An autonomous valve stiction detection system based on data characterization

Published in:
CONTROL ENGINEERING PRACTICE

Published: 01/01/2013

Document Version
Peer reviewed version

Please cite the original version:
An autonomous valve stiction detection system based on data characterization

Alexey Zakharov\textsuperscript{a}, Elena Zattoni\textsuperscript{c}, Lei Xie\textsuperscript{b}, Octavio Pozo\textsuperscript{a}, Sirkka-Liisa Jämsä-Jounela\textsuperscript{a}

\textsuperscript{a}Aalto University School of Chemical Technology, Department of Biotechnology and Chemical Technology, P.O. Box 16100, 00076 Aalto, Finland
\textsuperscript{b}Department of Control Science & Technology, Zhejiang University, Hangzhou, Zhejiang Province, 310027, P.R. China
\textsuperscript{c}Department of Electrical, Electronic and Information Engineering “G. Marconi”, Alma Mater Studiorum · University of Bologna, 40136 Bologna, Italy

Abstract

This paper proposes a valve stiction detection system which selects valve stiction detection algorithms based on characterizations of the data. For this purpose, novel data feature indexes are proposed, which quantify the presence of oscillations, mean-nonstationarity, noise and nonlinearities in a given data sequence. The selection is then performed according to the conditions on the index values in which each method can be applied successfully. Finally, the stiction detection decision is given by combining the detection decisions made by the selected methods. The paper ends demonstrating the effectiveness of the proposed valve stiction detection system with benchmark industrial data.

Key words: Valves; Stiction; Oscillations; Control Loops; Fault Detection; Diagnosis; Industrial Applications.

\* Corresponding author O. Pozo. Fax +358 9 470 23846.
Email addresses: alexey.zakharov@aalto.fi (Alexey Zakharov), elena.zattoni@unibo.it (Elena Zattoni), leix@iipc.zju.edu.cn (Lei Xie), octavioado@gmail.com (Octavio Pozo), sirkka-liisa.jamsa-jounela@aalto.fi (Sirkka-Liisa Jämsä-Jounela).
1 Introduction

Since most control loops in process industry have operational concerns, the last two decades have witnessed a rapid development of research activities in control system performance monitoring and assessment (e.g., Jelali and Huang, 2010). An extensive survey (Desborough, 2003) of over twenty-six thousand PID controllers collected from a wide variety of process industries reveals that only 16% of the loops own excellent performance, while 32% actually work in open loop, and those which are classified as fair or poor amount to 38%. A study by Desborough and Miller (2002) shows that nonlinear phenomena in control valves account for 32% of the fair and poor control loops and are the major contributors of control loop performance degradation. In particular, a survey by Yang and Clarke (1999) shows that 30% of all the control loops in Canadian paper mills are oscillating due to valve problems. Moreover, the comprehensive analysis of the faults of an industrial paperboard machine carried out by Jämsä-Jounela et al. (2012) shows that valves are the source of nearly 10% of the machine faults.

As is well-known from cost analysis done in process industry, oscillations in control loops result in high variability of product quality, increased energy consumption, and accelerated equipment aging, which bring down the plant profit significantly. In particular, valve static friction — or valve stiction, as is briefly called in the literature — is recognized as one of the main causes of valve malfunctions: see, e.g., (Pozo Garcia et al., 2013; Shoukat Choudhury et al., 2008, 2006; Thornhill, 2005). Since the selection of the right stiction detection method, in connection with the properties of the available data, often requires a huge investment of time and effort, also because of the large number (i.e., hundreds or even thousands) of control loops that must be processed, the development of an autonomous valve stiction detection system, based on routine operation data, is essential to improve process efficiency. To be more specific, characterizing the properties of the available data sequences can lead to an automatic selection of the applicable methods among a wide array of them. Moreover, the various degree of suitability of the methods selected, with respect to the properties of the data processed, can be exploited in order to obtain the final stiction detection decision as a weighted combination of the decisions provided by the single methods, thus enhancing the reliability of the whole detection procedure.

The contribution of this work consists in providing an autonomous system for valve stiction detection based on the aforementioned ideas. In particular, deep investigations have been carried out in order to accomplish the following objectives: i) define appropriate indexes aimed at characterizing the main features of the available data; ii) determine the relations between the features of the available data and the properties of the stiction detection methods;
iii) devise an algorithm for the automatic selection of the stiction detection methods applicable to the given sets of data; iv) devise an algorithm for achieving the final detection decision on the basis of the decisions provided by the single detection methods applied, by exploiting information on their reliability in connection with the features of the given data sequences.

In particular, the autonomous stiction detection system presented in this work is based on the introduction of five indexes for quantifying the relevant properties of the given data sequences, the implementation of four stiction detection methods, and the definition of two decision algorithms. The indexes are an oscillation index, a data-sampling index, a mean-nonstationarity index, a noise index, and a nonlinearity index. The methods are a curve-fitting method which integrates the triangular-fitting method (He et al., 2007) and the rectangular-fitting method (Hågglund, 2011), the cross-correlation method (Horch, 1999), the histogram method (Horch, 2006a), and the area-ratio method (Singhal and Salsbury, 2005). All these methods are encompassed in the so-called data-driven approach to valve stiction detection and, more specifically, the former two are known as shape-based methods. Data-driven methods are herein preferred to the so-called model-based approaches (e.g. Qi and Huang, 2011; Nallasivam et al., 2010; Jelali, 2008; Srinivasan et al., 2005) not only because, as mentioned above, they do not require a detailed model of the process, but also because their implementation is less demanding from the computational point of view.

The paper is organized as follows. An outline of the autonomous stiction detection system is presented in Section 2. Data characterization through the definition of appropriate indexes is discussed in Section 3. A review of the stiction detection methods implemented in the autonomous system is presented in Section 4. The impact of the data characteristics on applicability of the considered detection methods is analyzed in Section 5. The algorithm for selecting the applicable detection methods, given the data to be processed, and the algorithm for computing the final detection decision from the single decisions provided by each of the methods applied are shown in Section 6. Validation of the proposed autonomous stiction detection system with benchmark industrial data is discussed in Section 7. Some concluding remarks are set forth in Section 8.

2 Outline of the autonomous stiction detection system

The aim of this section is to introduce the overall concept of the proposed valve stiction detection system. The system processes the available data sequences and computes the values of a set of indexes characterizing the data, in order to establish whether each of the four valve stiction detection methods considered
is applicable or not. On this basis, a decision algorithm operates the selection of the applicable methods. Then, the selected methods are run in parallel and each of them provides its own detection decision. The definitive detection decision is finally provided by an algorithm that weights the single decisions with information on the reliability of each of the selected methods with respect to the given data sequences. A schematic diagram of the autonomous stiction detection system proposed in this work is shown in Fig. 1.

![Flow chart of the autonomous stiction detection system](image)
In more detail, the proposed autonomous stiction detection system works as described in the following. At the first stage, the recorded data sequences are processed to calculate the indexes characterizing those data features which most affect the correct operation of the stiction detection methods. The presence of oscillations is ascertained first, since all the methods implemented require oscillating data sequences. The number of available samples in a certain amount of time is also determined at this stage. Then, the data sequences are further characterized by quantifying the presence of noise and the possible nonstationarity of the mean value. Finally, possible functional nonlinearities in the valve input-output characteristic are evaluated, since some stiction detection methods presuppose that data are generated by control loops only including valves with linear input-output characteristics. The quantification of the data features by appropriate indexes leads to the definition of applicability thresholds, which, however, must be properly tuned depending on the diverse nature of the loops and the signals considered. At the second stage, the detection methods which can be applied are selected from an array of four among the most widely used in stiction detection practice: these are the area-ratio method, the histogram method, the curve-fitting method, and the cross-correlation method. The selection is performed by an algorithm that compares the values of the data indexes with the respective thresholds, defined for the specific method, and allows its application if and only if all the constraints are satisfied. Then, the selected methods are run and provide their individual decisions. At the third and last stage, the decisions individually provided by each of the selected methods contribute to the final decision according to an algorithm that weights the single decisions by means of coefficients which depend on the degree of suitability of the run methods to the specific data sequences considered.

3 Data characterization and definition of the indexes used in the autonomous system

This section concerns the determination of the features suitable for characterizing process data prior the selection of the applicable stiction detection methods. A preliminary study of the performance of the selected valve stiction detection methods on a benchmark set of seventy-six industrial control loops (Jelali and Huang, 2010) was carried out to reveal the typical causes of deteriorated detection results. The outcome of the study is reported below.

First, it was observed that, in some control loops, the oscillation shape can change significantly from one period to another. On the other hand, the selected stiction detection methods assume the presence of the steadfast oscillation pattern, so that, if the oscillation shape is not stable enough, the results obtained can be misleading. Therefore, stability of the oscillation pattern in
the signals is the first feature selected to characterize the data.

Next, it was considered that most of the selected stiction detection methods use zero-crossings of the signal to determine the half-periods of the oscillation. However, the mean value of the signal may deviate from zero (mean-nonstationarity), which can lead to incorrect identification of the half-periods and, therefore, to the wrong stiction detection results. In the case the data exhibit such variations, either the low-frequency disturbances are filtered out, or the half-periods are identified with alternative algorithms, or other stiction detection methods, with the property of being insensitive to the fluctuation of the mean value, are employed. Thus, mean-nonstationarity is the second data feature considered in this work.

Thirdly, it was pointed out that high-frequency noise complicates the determination of the oscillation patterns and disturbs the results of all stiction detection methods. However, some methods are more robust to the presence of noise than others. Thus, the presence of high-frequency noise is the third data feature to take into account.

Moreover, many methods assume the presence of a specific oscillation pattern in the signals for sticky and non-sticky cases. However, the pattern can vary because of the valve functional nonlinearity. In particular, nonlinear valve operating curves may lead to a noticeable asymmetry in the oscillation shape and, therefore, may corrupt the detection results. Hence, the nonlinearity of the valve operating curve is assumed as the fourth feature of process data to be checked.

Finally, a too long sampling time results in a relatively small number of data samples over an oscillation period and this, in turn, may not be sufficient to describe the oscillation shape precisely. On these conditions, the performance of all stiction detection methods suffers. Thus, the knowledge of the data resolution, which is defined as the number of samples in the oscillation period, is essential for a comprehensive characterization of the data.

The indexes quantifying the data features listed above are introduced in the remainder of this section.

3.1 Oscillation index

The presence of oscillations caused by valve stiction in control loops is a common problem in process industry. Many stiction detection methods are based on the presence of oscillations and provide reliable detection results only for signals with a steadfast oscillation pattern. Hence, the autonomous system for stiction detection proposed in this work requires an oscillation index which
should be robust to all most frequent types of uncertainties in the signals to be processed: those concerning the mean value, the length of the oscillation period, the amplitude of the oscillation, the presence of high-frequency noise, and the occurrence of non-steadfast oscillation patterns. The zero-crossing method does not define the oscillation half-periods properly if the signal is mean-nonstationary. Automatically filtering out the low-frequency noise is not always possible and, therefore, some manual actions are still needed. Therefore, in this work, in order to achieve an autonomous system, the peak-based approach is proposed to define the half-periods.

The oscillation index introduced herein identifies the periods of the oscillation by considering various half-periods and defining the oscillation peaks. Namely, given a signal $x$ and a possible integer upper-bound for the half-period length $d$, the proposed method determines the periods of the oscillations by defining the locations of the peaks. The symbol $m_i^+$ will be used for the location of the maximum in the $i$-period, while $m_i^-$ will be used for the location of the minimum in the same period. The search for the values $m_i^+$ and $m_i^-$, with $i = 1, 2, \ldots$, is made according to the following formulas:

\[
\begin{align*}
    m_i^+ &= \max_{k=1, \ldots, 2d} x(k), \\
    m_i^- &= m_i^+ + \min_{k=(d/2), \ldots, d} x(m_i^+ + k), \quad i = 1, 2, \ldots, \\
    m_i^+ &= m_{i-1}^- + \max_{k=(d/2), \ldots, d} x(m_{i-1}^- + k), \quad i = 2, \ldots,
\end{align*}
\]

where $\{ \cdot \}$ denotes the integer part of the argument. The intervals between two subsequent maxima are considered as periods, and the period search continues according to (1)–(3), while there are at least $d$ samples available after the last maximum or minimum found. At the next stage, similarity of two subsequent periods is evaluated using the correlation coefficient. Since the lengths of the identified periods are usually different due to deviations of the oscillation periods from the pattern and noise, resampling is required to obtain two data sets of the same length. The correlation coefficients $C(i)$, with $i = 1, 2, \ldots$, between these two data sets are computed to characterize the similarity of the subsequent periods $i$ and $i + 1$. The presence of oscillations in the signal is recognized if most of the coefficients $C(i)$ are close to 1, and thus it can be measured with 80% quantile the aforementioned coefficients:

\[
I_{OSC}(d) = \max_{\theta} \{ \theta : \theta \leq C(i) \text{ for 80\% of } i \}. \tag{4}
\]

Finally, all possible lengths of the oscillation period are considered and the oscillation index is defined as the maximum value of the quantile in (4):

\[
I_{OSC} = \max_d I_{OSC}(d). \tag{5}
\]

The proposed method proved to perform satisfactorily in all cases, including the square-shape signals, which do not have a single peak in an oscillation.
half-period. Nonetheless, in this case, performance can be further improved, for instance, by considering a narrower search interval in (2) and (3), which will allow finding similar peak locations in the oscillation periods.

The oscillation in a signal is very steadfast when the oscillation index is above 0.85. When the index is between 0.70 and 0.85, the signal is usually noisy and the oscillation periods frequently differ from the pattern. If the index is below 0.70, it is typically impossible to recognize any oscillations in the signal plot and it can be concluded that the signal is non-oscillating.

### 3.2 Mean-nonstationarity index

In industrial processes, oscillating signals are not periodic, in general, due to several reasons, including fluctuations in the mean value, oscillation amplitude, period length, and oscillation shape. Variations in the mean value (mean-nonstationarity), in particular, seem to have a significant effect on the performance of stiction detection methods. Hence, an index related to the behavior of the mean value of the signals to be processed is also considered in this work. However, since the mean value, as is usually defined, is very sensitive to signal outliers, a more robust ‘middle level’ of the signal is proposed and used herein to quantify the mean variations. The so-called ‘middle level’ of the signal in an oscillation period is defined as the point at which the respective areas of the upper and lower half-periods are equal, as is shown in Fig. 2.

This approach requires the determination of the oscillation periods, which is achieved by computing the oscillation index according to the technique presented in Section 3.1. In particular, the time values \( m_i^+ \) and \( m_i^- \), with \( i = 1, 2, \ldots \), respectively corresponding to the maximum and the minimum of the signal over each period are used. Given an initial approximation of the middle level \( L_i \) of the signal in the \( i \)-th period, the lower and the upper half-periods are respectively defined as the subsets of the intervals between two subsequent maxima and minima, according to the following formulas:

\[
P_i^-(L_i) = [m_i^+, m_{i+1}^+] \cap \{k \mid x(k) < L_i \},
\]

\[
P_i^+(L_i) = [m_i^-, m_{i+1}^-] \cap \{k \mid x(k) > L_i \}.
\]

An iterative procedure, like, e.g., bisection, is then employed to determine the middle level as that particular value of \( L_i \) such that

\[
\sum_{i \in P_i^-(L_i)} (L_i - x(i)) = \sum_{i \in P_i^+(L_i)} (x(i) - L_i)
\]

is satisfied. The mean-value variation index, or mean non-stationarity index as
it will be called henceforth, is finally computed as the ratio of the estimations of the standard deviation of the middle levels and the mean magnitude in the oscillation periods: i.e.,

\[
I_{NS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (L_i - \bar{L})^2} / \left( \frac{1}{n} \sum_{i=1}^{n} (x(m_i^+) - x(m_i^-)) \right),
\]

(6)

where

\[
\bar{L} = \frac{1}{n} \sum_{i=1}^{n} L_i
\]

is the mean value of the middle level.

For stationary signals, the index value is below 0.1. The index value exceeding 0.25 indicates significant mean-value variations. In particular, for control loops with very strong nonstationary behavior, the index value ranges between 0.7 and 1.

### 3.3 Noise index

Even though high-frequency noise deteriorates the results of most valve stiction detection methods, no indexes quantifying the presence of noise in the signals to be processed by stiction detection methods are used in the literature. The index proposed in this work relies on the fact that the behavior of the signal over the half-periods of the oscillation patterns located between two peaks is typically monotonous. However, in signals with high-frequency
noise, the behavior of the signal over the half-periods is no longer monotonous. This fact is therefore exploited to evaluate the level of noise according to the following reasoning. The total variation of a signal in the interval \([m_1^+, m_n^+]\) is defined by

\[
V(x, m_1^+, m_n^+) = \sum_{i=m_1^+}^{m_n^+-1} |x(i + 1) - x(i)|.
\]

where the set of values \(m_i^+, \text{ with } i = 1, \ldots, n,\) are determined according to the technique presented in Section 3.1. The noise index is defined as an increase of the total variation of the signal caused by high-frequency noise: i.e.,

\[
I_{NO} = \left( \frac{V(x, m_1^+, m_n^+)}{\sum_{i=1}^{n-1} (|x(m_i^-) - x(m_i^+)| + |x(m_{i+1}^+) - x(m_i^-)|)} - 1 \right) \sqrt{d}, \quad (7)
\]

where \(d\) is the average number of samples in the oscillation half-periods. In fact, scaling is introduced to correctly reflect the ability of noise to corrupt the signal shape, depending on the number of samples in the half-periods. Namely, availability of only few samples in a half-period frequently makes an oscillation pattern impossible to recognize even in an almost noiseless signal. Instead, availability of many samples in a half-period makes the oscillation pattern easy to identify also in noisy signals. For almost noiseless signals, the value of the noise index is below 0.1. For higher values of the noise index, it is difficult to precisely distinguish the oscillation pattern, particularly when the noise index becomes greater than 0.25. It is worth noting that, in the case of a noiseless signal, the total variation of the signal in the interval \([m_1^+, m_n^+]\) is given by

\[
V(x, m_1^+, m_n^+) = \sum_{i=1}^{n-1} (|x(m_i^-) - x(m_i^+)| + |x(m_{i+1}^+) - x(m_i^-)|).
\]

Hence, \(I_{NO} = 0,\) according to (7).

### 3.4 Nonlinearity index

As previously mentioned, many valve stiction detection methods are only applicable to process data with clear oscillations. A sticky valve, however, is not the only contributor to process oscillations. Other possible causes include controller overtuning and nonlinear valve characteristics. In particular, a nonlinear valve characteristic represents a nonlinear relation between the position of the valve flow-controlling element (e.g., the valve stem) and the resulting flow rate. In contrast to valve characteristic, valve stiction is used to describe
the nonlinear relation between the controller output (i.e., the valve input) and the position of the valve stem. These two different nonlinear valve features may result in similar oscillation patterns and, thus, may render many valve stiction detection methods unable to discriminate between them. As depicted in Fig. 3, where the flow rate is plotted versus the valve stem position, there are three basic valve characteristic curve shapes: linear (solid line), equal-percentage (dash-dot line), and fast-opening (dash line). The linear characteristic follows a straight line, the equal-percentage one shows the flow rate increasing exponentially, and the fast-opening one shows a quick flow increment at the beginning of the stem travel off the seat.

Fig. 3. Common valve characteristic curves — Flow rate vs stem position

Figs. 4 and 5 present simulation results concerning a control loop with a sticky valve. The figures just aim to illustrate the effect of the nonlinear valve curve on the control loop signals and to demonstrate how the nonlinearity index reflects the valve nonlinearity. Stiction in a control loop with a linear valve results in an oscillating signal which has symmetric lower and upper half-periods as is shown in Fig. 4. Instead, stiction in a control loop with a nonlinear valve gives rise to an oscillating signal with asymmetric lower and upper half-periods, as is shown in Fig. 5.

Hence, the index proposed to evaluate the nonlinearity in the valve characteristic is defined as follows:

\[ I_{NL} = \frac{1}{n} \sum_{i=1}^{n} \log \left| \frac{m_i^+ - L_i}{L_i - m_i^-} \right|, \]  

(8)

where \( m_i^+ \), \( m_i^- \), and \( L_i \), with \( i = 1, \ldots, n \), are defined as in Sections 3.1 and 3.2. A nonlinearity index close to 0 corresponds to a linear valve characteri-
istic, whereas positive and negative values indicate fast-opening and equal-
percentage valve characteristics, respectively.

![Symmetric control error in a control loop with a linear valve — Amplitude vs time (number of samples)](image)

**Fig. 4.** Symmetric control error in a control loop with a linear valve — Amplitude vs time (number of samples)

### 4 A review of the stiction detection methods implemented in the autonomous system

A significant number of stiction detection methods are available in the literature. However, successful application of some of them is conditioned by very strict requirements, such as, in particular, detailed knowledge of the process physical model. Hence, only data-driven stiction detection methods have been implemented in the autonomous stiction detection system presented in this work. For the reader’s convenience, these methods are reviewed below.

The curve fitting technique proposed by He et al. (2007) evaluates the fitting of a triangular signal and a sinusoidal signal to a process measurement. According to this method, if the triangular signal has a lower fitting error, stiction is detected. The rectangular fitting algorithm presented by Hägglund (2011) compares a rectangular signal and a sinusoidal signal with a process measurement and calculates the so-called loss function, based on the integral of the absolute value of the error (IAE, or integrated absolute error). If the loss function for the rectangular matching is lower than that for the sinusoidal matching,
stiction is assumed to be the source of oscillations. In this work, the so called curve-fitting method encompasses both the techniques respectively introduced by He and Hägglund. As mentioned above, both techniques base their detection decision on the oscillation pattern of a measurement, but they analyse different signals. If the loop is self-regulating, He’s method uses the controller output (also denoted as $op$ signal), while Hägglund’s method employs the process output (or, $pv$ signal). As to integrating loops, He’s method employs the $pv$ signal, while Hägglund’s method can be applied provided that the $pv$ signal is filtered with an appropriate low-pass filter. In this work, Hägglund’s method is not applied to integrating loops, since a precise determination of the filter parameters is a task not suited for automatic implementation. In case of self-regulating loops, both methods are used separately and their results contribute individually to the final decision of the system.

The cross correlation method proposed by Horch (1999) determines the presence of stiction by evaluating the symmetry of the cross correlation between the controller output ($op$ signal) and the process output ($pv$ signal). Thus, unlike the aforementioned curve-shape methods, it processes two different signals instead of only one. The cross correlation method has been designed to operate only on self-regulating oscillating control loops.

The histogram method, also developed by Horch (2006b), constructs a data histogram using the filtered second derivative of the process output ($pv$ signal) for integrating processes or the first derivative for self-regulating processes.
However, in this work, the filtered second derivative of the controller output (op signal) is used for diagnosing self-regulating processes. This is possible since the controller output of a self-regulating loop produces approximately the same stiction shape of the controller error in an integrating loop. The next step in the algorithm is to match the data histogram with two probability density functions: the first is a gaussian distribution while the second is a so-called camel distribution. If the first probability distribution shows a better fit, stiction is determined as the source of the oscillations in the loop.

The method presented by Singhal and Salsbury (2005) analyses the symmetry of the process variable signal in self-regulating processes by calculating the ratio of the areas before and after the oscillation peak in each half-period. If the area preceding the peak is significantly larger than that following the peak and this circumstance frequently occurs in the sequence of data considered, stiction can be assumed as the cause of the oscillations in the loop. This method is only suitable for self-regulating loops.

5 Data features and detection methods: requisite set-up

A detailed analysis of how the data features examined in Section 3 affect successful application of the stiction detection methods reviewed in Section 4 is presented in this section. To this aim, the stiction detection methods are tested on a set of benchmark industrial data borrowed from (Jelali and Huang, 2010). Nineteen loops from chemical industry, concerning flow, temperature, level, and pressure control have been considered. In particular, these control loops have oscillating signals. The testing procedure consists of the following steps. First, the indexes are calculated for each data sequence. Next, the stiction detection methods are run. Finally, the detection decision provided by each of the methods applied is compared with the actual — either faulty or non-faulty — condition known from the literature and the outcome is analysed in connection with the values taken by the feature indexes. The test results are synthetically reported in Table 1.

The following notation is used in Table 1. The loop number refers to the source of the processed data in (Jelali and Huang, 2010). The symbols $pv$ and $op$ respectively identify the process output and the controller output. The columns denoted as oscillation, non-stationarity, noise, and nonlinearity show the values of the indexes $I_{OSC}$, $I_{NS}$, $I_{NO}$, and $I_{NL}$, as respectively defined in (5)–(8), for the processed data. The column denoted as samples per half-period shows the average number of samples available in half a period of the oscillating signal, where the half-period is computed according to the technique described in Section 3.1. The column denoted as condition shows the actual condition — either faulty or non-faulty — of the considered loop as is stated in (Jelali and
Huang, 2010): the symbol 0 means that the loop is not affected by valve stiction, while the symbol 1 means the opposite. The column labelled as loop type distinguishes between self-regulating loops and integrating loops: the symbol 0 indicates a self-regulating loop, while the symbol 1 is used for an integrating loop. The last four columns show the values of the stiction detection index provided by each of the methods applied for the specific data sequence. The curve-fitting column only displays the stiction index of He’s method on integrating loops. On self-regulating loops, both methods frequently provided similar results and the average of these two indexes is displayed. The symbol NA means that the method does not apply. Values of the stiction detection index greater than 0.59 indicate the detection of stiction. Values of the stiction detection index lower than 0.45 indicate that the loop is not affected by stiction. Values of the stiction detection index ranging between 0.45 and 0.59 means that no decisions can be made.

In the light of the tests summarized in Table 1, a detailed analysis of the relations between the values of the feature indexes for the processed data and successful application of the stiction detection methods has been performed. The result, consisting in the set-up of threshold values of the indexes for the methods to be applicable with a reasonable expectation of success, are presented in the remainder of this section.

5.1 Presence of oscillations

The fact that all the considered stiction detection methods only apply to oscillating signals is a clear sign of the influence of this feature. In particular, the proposed oscillation index reveals not only the presence of oscillations, but also the consistency of oscillation parameters such as, shape, frequency, and amplitude: the higher is the index, the more consistent are the parameters. Thus, applicability of the curve-fitting method, the cross-correlation method, and the area-ratio method can be related to an oscillation index at least equal to 0.7. Lower values of the index point out non-steadfast oscillations, which can affect the correct operation of all these methods. The histogram method is especially sensitive to inconsistencies in the oscillation parameters. Inaccurate data histograms are produced and matching to a specific shape is interdicted when the oscillation index is lower than 0.85. For instance, a similar circumstance occurs in Loop 22, which is a self-regulating loop and where \( I_{OSC} = 0.8 \) for the \( op \) signal. The \( op \) signal and the histogram of the filtered second derivative of the \( op \) signal are shown in Fig. 6. The shape of the histogram does not fit any of the distributions, gaussian and camel, assumed for the comparison. Hence, no detection decision is given.
Table 1
Data features and detection indexes for the tested control loops

<table>
<thead>
<tr>
<th>Loop No</th>
<th>Oscillation</th>
<th>Non-stationarity</th>
<th>Noise</th>
<th>Samples per half-period</th>
<th>Non-linearity</th>
<th>Condition</th>
<th>Loop-type</th>
<th>Curve-fitting</th>
<th>Cross-correlation</th>
<th>Histogram</th>
<th>Area ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pv</td>
<td>0.92</td>
<td>0.04</td>
<td>0.13</td>
<td>-0.13</td>
<td>1</td>
<td>0</td>
<td>0.7</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.96</td>
<td>0.24</td>
<td>0.01</td>
<td>0.83</td>
<td>0</td>
<td>1</td>
<td>0.45</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>5</td>
<td>pv</td>
<td>0.96</td>
<td>0.39</td>
<td>0.02</td>
<td>-0.13</td>
<td>0</td>
<td>1</td>
<td>0.56</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.97</td>
<td>0.55</td>
<td>0</td>
<td>0.8</td>
<td>1</td>
<td>0</td>
<td>0.56</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>pv</td>
<td>0.97</td>
<td>0.01</td>
<td>0.13</td>
<td>0.00</td>
<td>1</td>
<td>0</td>
<td>0.87</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.99</td>
<td>0</td>
<td>0</td>
<td>0.75</td>
<td>0</td>
<td>1</td>
<td>0.87</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>pv</td>
<td>0.95</td>
<td>0.01</td>
<td>0.26</td>
<td>0.03</td>
<td>1</td>
<td>0</td>
<td>0.37</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.98</td>
<td>0.01</td>
<td>0.11</td>
<td>0.64</td>
<td>1</td>
<td>0</td>
<td>0.37</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>12</td>
<td>pv</td>
<td>0.72</td>
<td>0.02</td>
<td>0.28</td>
<td>233</td>
<td>1</td>
<td>0</td>
<td>0.83</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.91</td>
<td>0.03</td>
<td>0.01</td>
<td>228</td>
<td>1</td>
<td>0</td>
<td>0.83</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>pv</td>
<td>0.98</td>
<td>0.18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.31</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.97</td>
<td>0.76</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0.31</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>pv</td>
<td>0.81</td>
<td>0.26</td>
<td>0.21</td>
<td>0.18</td>
<td>0</td>
<td>0</td>
<td>0.68</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.97</td>
<td>0.31</td>
<td>0.01</td>
<td>0.10</td>
<td>0</td>
<td>0</td>
<td>0.68</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>pv</td>
<td>0.7</td>
<td>0.11</td>
<td>0.21</td>
<td>67</td>
<td>1</td>
<td>0</td>
<td>0.49</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.8</td>
<td>0.3</td>
<td>0.02</td>
<td>68</td>
<td>1</td>
<td>0</td>
<td>0.49</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>pv</td>
<td>0.93</td>
<td>1.07</td>
<td>0.01</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.89</td>
<td>10.81</td>
<td>0.01</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0.5</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>22</td>
<td>pv</td>
<td>0.74</td>
<td>0.01</td>
<td>0.14</td>
<td>73</td>
<td>1</td>
<td>0</td>
<td>0.32</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.8</td>
<td>0.03</td>
<td>0.02</td>
<td>65</td>
<td>1</td>
<td>0</td>
<td>0.32</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>23</td>
<td>pv</td>
<td>0.83</td>
<td>0.01</td>
<td>0.48</td>
<td>64</td>
<td>1</td>
<td>0</td>
<td>0.92</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.98</td>
<td>0.02</td>
<td>0</td>
<td>64</td>
<td>1</td>
<td>0</td>
<td>0.92</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>pv</td>
<td>0.91</td>
<td>0.06</td>
<td>0.02</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0.45</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.94</td>
<td>0.18</td>
<td>0.01</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0.45</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>26</td>
<td>pv</td>
<td>0.9</td>
<td>0.1</td>
<td>0.37</td>
<td>66</td>
<td>1</td>
<td>1</td>
<td>0.58</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.9</td>
<td>0.13</td>
<td>0.3</td>
<td>66</td>
<td>1</td>
<td>1</td>
<td>0.58</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>28</td>
<td>pv</td>
<td>0.99</td>
<td>0.02</td>
<td>0.03</td>
<td>65</td>
<td>1</td>
<td>1</td>
<td>0.55</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.98</td>
<td>0.03</td>
<td>0</td>
<td>64</td>
<td>1</td>
<td>1</td>
<td>0.55</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>32</td>
<td>pv</td>
<td>0.99</td>
<td>0.03</td>
<td>0.05</td>
<td>45</td>
<td>1</td>
<td>0</td>
<td>0.77</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.99</td>
<td>0.06</td>
<td>0.04</td>
<td>45</td>
<td>1</td>
<td>0</td>
<td>0.77</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>35</td>
<td>pv</td>
<td>0.93</td>
<td>0.31</td>
<td>0.03</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0.6</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.86</td>
<td>5.03</td>
<td>0.02</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0.6</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>36</td>
<td>pv</td>
<td>0.77</td>
<td>0.21</td>
<td>0.19</td>
<td>70</td>
<td>0</td>
<td>1</td>
<td>0.59</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.76</td>
<td>0.31</td>
<td>0.18</td>
<td>70</td>
<td>0</td>
<td>1</td>
<td>0.59</td>
<td>NA</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>38</td>
<td>pv</td>
<td>0.84</td>
<td>0.02</td>
<td>0.07</td>
<td>92</td>
<td>0</td>
<td>0</td>
<td>0.81</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.99</td>
<td>0</td>
<td>0.03</td>
<td>92</td>
<td>0</td>
<td>0</td>
<td>0.81</td>
<td>0.5</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>
5.2 Nonstationary mean-value

The mean-nonstationarity index indicates variations in the mean value of the oscillating signal. A low value of the index indicates a stationary signal, which can be analysed by all the methods. Conversely, a high value of the index can affect the methods which employ zero-crossings to determine the oscillation periods. Nevertheless, an alternative procedure is used in this work for calculating the oscillation period. Thus, mean-nonstationarity has little effect on the curve-fitting and the area-ratio methods. On the contrary, methods like cross-correlation and histogram can successfully be applied only to clearly mean-stationary signals. Thus, the requirement for the latter two methods is a value of the mean-nonstationarity index lower than 0.25, while, for the former two methods, the boundary value is set at 0.5. Fig. 7 shows a typical effect occurring when the histogram method is applied to data with weak characteristics of mean-stationarity: the mean-nonstationary signal produces a skewed histogram.

5.3 High-frequency noise

In general, high-frequency noise seriously affects successful application of any stiction detection method. For instance, Fig. 8 shows that a signal corrupted by noise does not lend itself to straightforward curve fitting. However, the investigation carried out in this work has brought to the following consider-
vations, which allow some thresholds for the value of the noise index to be defined in connection with each of the considered methods. As to the area-ratio method, Singhal and Salsbury (2005) point out that noise can affect the performance of their method and propose to use a low pass filter, otherwise it is impossible for the method to accurately locate the peak of the signal in each half-period (as it happens, for instance, in Loops 11, 12, 23). As a consequence, unreliable or false detection decisions are caused by the frequent occurrence of inconsistent area ratios between the half-periods. Similar reasoning also holds for the curve-fitting method. So that, a maximal acceptable value for the noise index can be set to 0.15 for both methods. As to the cross-correlation method, one of its main advantages is its intrinsic robustness with respect to high-frequency noise, mainly due to the processing of two different sequences of data. Also for the histogram method, the filtering action implicit in the method alleviates the issues related to high-frequency noise. However, since the presence of high-frequency noise can be combined with the presence of other unfavourable conditions, like, e.g., unclear oscillation patterns or low data resolution, a conservative threshold for the noise index can be set at 0.3 for the latter two methods.
Fig. 8. Triangular and sinusoidal fitting in Loop 11

5.4 Nonlinear valve characteristics

The presence of stiction in valves with a nonlinear characteristic generates oscillating signals whose shapes are different from those assumed in stiction detection methods devised to detect stiction in valves with linear characteristics. Thus, the nonlinearity index is used to define some tolerance of the stiction detection methods to variations in the shape of the oscillating signal. For instance, an irregular oscillation pattern is shown in Fig. 9, where the $pv$ signal does not resemble any of the shapes assumed by the curve-fitting methods. Thus, for the curve-fitting method, the absolute value of the nonlinearity index is required to be lower than 0.15. Similar arguments hold for the area-ratio method, which has approximately the same tolerance to oscillation shape variations, so that the requirement is still an absolute value of the nonlinearity index lower than 0.15. The cross-correlation method gives ambiguous detection results when applied to data characterized by a highly nonlinear behavior (like, e.g., those of Loops 5 and 20). The cause of the unreliable detection result is a phase shift between the $op$ signal and the $pe$ signal, which, in turn, derives from the erratic oscillation shape of both signals, as is shown in Fig. 10, with reference to the data of Loop 5 in particular. Similarly, nonlinearity affects the expected stiction shape of the data histogram. Therefore, for both methods, the maximal admissible absolute value for the nonlinearity index is 0.3.
Fig. 9. Rectangular fitting for $pv$ signal in Loop 19

Fig. 10. Cross-correlation function (top) and $pv$, $op$ signals (bottom) in Loop 5
5.5 Data resolution

As it is intuitive, low data resolution may compromise effective applicability of any stiction detection method. In particular, uncertain detection decisions obtained by running the curve-fitting method with the data of Loops 4, 13, 14, 35 can be explained by the poor data resolution. In fact, as is shown in Fig. 11, a small number of available samples makes the evaluation of any shape fitting an impossible task. The performance of the histogram method can be affected by a small amount of oscillation periods in the data, if this is coupled with systematic deviations from the oscillation pattern. This circumstance is indicated by relatively low values of the oscillation index, which causes the method to construct an inaccurate data histogram and classify it wrongly, as it occurs in Loops 22 and 36. Moreover, low data resolution limits the cut-off frequency of the filter. In fact, the cut-off frequency has to be such that the oscillating behavior of the signal and its derivatives till the second order must be preserved. Figure 12 shows that, in the presence of a filtering action, the higher is the order of the derivative, the less clear is the oscillation pattern in the signal. Consequently, in loops with low data resolution the high-frequency noise cannot be removed completely and the histogram may be strongly corrupted by the noise. Similarly, for the cross correlation method, when the data resolution is low, the recorded points are not consistent from one period to another, which can lead to indefinite or misleading results. The area-ratio method also produces unreliable results in control loops with low data resolution (Loops 5, 20, 35). In these cases, the poor data resolution hin-
ders accurate identification of the half-periods and the level of the starting and finishing point of each half period becomes highly variable. This produces unreliable area calculation, which leads to indefinite or wrong stiction detection. Thus, for the selected stiction detection methods, the data resolution should be of at least 20 samples per half-period.

5.6 A summary of the requisites

The results derived in Sections 5.1 to 5.5 are summarized in Table 2. The first column reports all the stiction detection methods which have been considered. The first row shows all the data feature indexes which have been defined to characterize the data to be processed. All indexes are followed by a sign, either ≥ or ≤, which indicates the inequality with respect to the values presented in the corresponding column. It is worth noticing that as far as the nonlinearity index is concerned, the bound is given on the absolute value since the index can also assume negative values. The average number of samples per half period is also considered and expressed through its inverse: the so-called resolution index, or $I_{RES}$. In conclusion, it is worth mentioning that these index thresholds can be used as a guideline, although they are not intended to be definitive under all circumstances.
Table 2
Summary of requisites on data feature indexes for applicability of detection methods

| Detection Method       | $I_{OSC}$ | $I_{NS}$  | $I_{NO}$  | $|I_{NL}|$ | $I_{RES}$ |
|------------------------|-----------|-----------|-----------|-----------|-----------|
| curve fitting          | 0.7       | 0.5       | 0.15      | 0.15      | 0.05      |
| cross correlation      | 0.7       | 0.25      | 0.3       | 0.3       | 0.05      |
| histogram              | 0.85      | 0.25      | 0.3       | 0.3       | 0.05      |
| area-ratio             | 0.7       | 0.5       | 0.15      | 0.15      | 0.05      |

6 Decision algorithms

Two algorithms are presented in this section. The first algorithm is devised for selecting the applicable detection methods for a given set of values of the data feature indexes (Section 6.1). The second algorithm is aimed at computing the final detection decision on the basis of the decisions respectively made by the selected methods (Section 6.2). The algorithms are presented with reference to a generic set of data feature indexes and a generic set of detection methods, in order to stress their flexibility: the possibility of applying the same decision algorithms to different sets of data feature indexes and different sets of detection methods, or, even, more generally, to different detection problems.

The following notation is used. $\mathbb{R}$ stands for the set of real numbers. $\mathbb{B}$ stands for the set of Boolean elements: $\mathbb{B} = \{0, 1\}$. Matrices are denoted by uppercase letters, vectors and scalars by lower-case letters. The symbol $\bar{a}$, where $a$ is a boolean variable, denotes the complement of $a$ (or, not $a$).

6.1 Algorithm for selecting the detection methods based on the data features

Let $\mathcal{I} = \{1, 2, \ldots, n\}$ be the finite index set associated with the set of the considered detection methods, so that, for any $i \in \mathcal{I}$, $M_i$ denotes the $i$-th detection method considered. Let $\mathcal{J} = \{1, 2, \ldots, m\}$ be the finite index set associated with the features considered for the data sequences, so that, for any $j \in \mathcal{J}$, $P_j$ denotes the $j$-th data feature considered. Let $\tilde{P} \in \mathbb{R}^{n \times m}$ be an $n \times m$ matrix of real values defined in such a way that, for any $i \in \mathcal{I}$ and for any $j \in \mathcal{J}$, $\tilde{P}_{ij}$ is the minimum value of the $j$-th feature index for the $i$-th method to be applicable. In particular, $\tilde{P}_{ij} = 0$ means that the $j$-th feature is irrelevant for applicability of the $i$-th method. Let $p \in \mathbb{R}^m$ denote the vector of the feature indexes computed for a given data sequence. Let $S \in \mathbb{B}^{n \times m}$ denote an $n \times m$
matrix whose entries are boolean values computed as follows:

\[
S_{i,j} = \begin{cases} 
0, & \text{if } p_i \geq \bar{P}_{i,j}, \\
1, & \text{otherwise}, 
\end{cases} i \in \mathcal{I}, j \in \mathcal{J}.
\]

Let \(a \in \mathbb{B}^n\) denote an \(n\)-dimensional vector of boolean variables computed as follows:

\[
a_i = \begin{cases} 
1, & \text{if } \bigvee_{j=1}^m S_{i,j} = 0, \\
0, & \text{otherwise}, 
\end{cases} i \in \mathcal{I}.
\]

Then, the set of the methods that can be applied for the given data sequence is defined as follows:

\[
\mathcal{M} = \{ M_i : a_i = 1, i \in \mathcal{I} \}.
\] (9)

6.2 Algorithm for computing the final detection decision

Let \(\tilde{\mathcal{I}} = \{1, 2, \ldots, \tilde{n}\}\), with \(\tilde{n} \leq n\), denote the index set associated with the set of the applicable algorithms for a given data sequence, so that the set \(\tilde{\mathcal{M}}\), defined by (9), can also be written as

\[
\tilde{\mathcal{M}} = \{ \tilde{M}_i, i \in \tilde{\mathcal{I}} \}.
\]

Let \(R_i\), with \(i \in \tilde{\mathcal{I}}\), denote the reliability of the applicable method \(\tilde{M}_i\), with \(i \in \tilde{\mathcal{I}}\). Let \(R_i\) be defined as the product of the so-called reliability of the applicable method \(\tilde{M}_i\), with \(i \in \tilde{\mathcal{I}}\), with respect to each of the considered data features \(P_j\), with \(j \in \mathcal{J}\): i.e.,

\[
R_i = \Pi_{j \in \mathcal{J}} R_{i,j}, \quad i \in \tilde{\mathcal{I}}.
\]

Moreover, let us assume that, for each of the considered features, the reliability \(R_{i,j}\) be a function of the feature index \(p_j\) defined by the following rule:

\[
R_{i,j} = \begin{cases} 
1, & \text{if } p_j \leq \alpha_{i,j}, \\
1 - \frac{p_j - \alpha_{i,j}}{\beta_{i,j} - \alpha_{i,j}}, & \text{if } \alpha_{i,j} < p_j < \beta_{i,j}, \\
0, & \text{if } \beta_{i,j} \leq p_j,
\end{cases}
\]

where \(\alpha_{i,j}, \beta_{i,j}\) are nonnegative real numbers with \(\alpha_{i,j}\) equal to or greater than the minimum value of the admissible range for the \(j\)-th feature index \(p_j\) and \(\beta_{i,j}\) less than or equal to the maximum value of the admissible range for the \(j\)-th feature index \(p_j\). Hence, \(0 \leq R_{i,j} \leq 1\) for all \(j \in \mathcal{J}\) and \(i \in \tilde{\mathcal{I}}\). Consequently, \(0 \leq R_i \leq 1\) for all \(i \in \tilde{\mathcal{I}}\). For any \(i \in \tilde{\mathcal{I}}\), let \(F_i \in \mathbb{B}\) denote the decision obtained by
running the method $M_i$ for the given data sequence, according to the following rule:

$$F_i = \begin{cases} 
1, & \text{if fault is detected,} \\
0, & \text{otherwise.}
\end{cases}$$

Let $W_F, W_{\bar{F}} \in \mathbb{R}$ be defined as follows:

$$W_F = \sum_{i=1}^{\hat{n}} R_i F_i, \quad W_{\bar{F}} = \sum_{i=1}^{\hat{n}} R_i \bar{F}_i.$$

If

$$\frac{|W_F - W_{\bar{F}}|}{W_F + W_{\bar{F}}} \geq \Delta,$$

where $\Delta$ is assigned on the basis of the reliability required for the decision to be made, then the final decision $F$ is determined according to the following rule:

$$F = \begin{cases} 
1, & \text{if } W_F - W_{\bar{F}} > 0, \\
0, & \text{otherwise,}
\end{cases}$$

where $F = 1$ means that the fault is detected, while $F = 0$ means that the fault is not detected.

A suitable value of the threshold parameter $\Delta$ for the sets of data considered in this work is around 0.7. Indeed, this value is used for all the tests considered in this manuscript. A more precise determination of the threshold requires that the proposed approach be applied to a bigger amount of industrial data.

7 Validation of the autonomous stiction detection system

In this section, the autonomous stiction detection system devised in this work is tested on a set of eight control loops still taken from (Jelali and Huang, 2010), but not used for the preliminary analysis presented in Section 5. It is worth mentioning that reducing all the applicability conditions pointed out in Section 5 to the syntax assumed in Section 6 is a question of mere technicalities and, therefore, will be skipped herein. The results of the validation process are shown in Table 3. In this table, the symbol NA in the column denoted as stiction decision indicates that the method is not applicable, the symbol 1 means the stiction is detected, the symbol 0 means that no stiction is detected. The system provides the correct results for the tested control loops that fulfill the applicability conditions for at least one method. Conversely, the system is able to correctly classify the control loops in which no method can be implemented. This is the case of the two first validation control loops. In the first control loop, the data resolution is too low for any of the methods to correctly analyse the data. While, in the second control loop, the high
Table 3
Data characteristics and stiction detection results for the validation data

<table>
<thead>
<tr>
<th>Process variable</th>
<th>Oscillation</th>
<th>Non-stationarity</th>
<th>Noise</th>
<th>Data resolution</th>
<th>Nonlinearity</th>
<th>Condition</th>
<th>Loop type</th>
<th>Applicable methods</th>
<th>Individual Diagnosis</th>
<th>Stiction decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>pv</td>
<td>0.94</td>
<td>0.09</td>
<td>0.01</td>
<td>1.6</td>
<td>0.01</td>
<td>1</td>
<td>None</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.95</td>
<td>0.28</td>
<td>0.01</td>
<td>1.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>pv</td>
<td>0.70</td>
<td>0.05</td>
<td>0.18</td>
<td>0.3</td>
<td>0.83</td>
<td>1</td>
<td>None</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.83</td>
<td>0.10</td>
<td>0.01</td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow</td>
<td>pv</td>
<td>0.73</td>
<td>0.02</td>
<td>0.02</td>
<td>0.47</td>
<td>0.18</td>
<td>1</td>
<td>Cross correlation</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.91</td>
<td>0.13</td>
<td>0.01</td>
<td>0.47</td>
<td></td>
<td></td>
<td>Histogram</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>pv</td>
<td>0.89</td>
<td>0.14</td>
<td>0.26</td>
<td>0.29</td>
<td>0.34</td>
<td>1</td>
<td>Histogram</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.86</td>
<td>0.19</td>
<td>0.21</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concentration</td>
<td>pv</td>
<td>0.99</td>
<td>0.03</td>
<td>0.02</td>
<td>0.47</td>
<td>0.14</td>
<td>1</td>
<td>Curve fitting</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.99</td>
<td>0.06</td>
<td>0.01</td>
<td>0.47</td>
<td></td>
<td></td>
<td>Histogram</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Temperature</td>
<td>pv</td>
<td>0.98</td>
<td>0.04</td>
<td>0.01</td>
<td>0.14</td>
<td>0.02</td>
<td>0</td>
<td>Curve fitting</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.98</td>
<td>0.04</td>
<td>0.01</td>
<td>0.14</td>
<td></td>
<td></td>
<td>Cross correlation</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>pv</td>
<td>0.99</td>
<td>0.02</td>
<td>0.00</td>
<td>0.34</td>
<td>0.19</td>
<td>1</td>
<td>Area ratio</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.99</td>
<td>0.04</td>
<td>0.01</td>
<td>0.34</td>
<td></td>
<td></td>
<td>Histogram</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Level</td>
<td>pv</td>
<td>0.96</td>
<td>0.07</td>
<td>0.05</td>
<td>0.4</td>
<td>0.43</td>
<td>1</td>
<td>Histogram</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>op</td>
<td>0.95</td>
<td>0.14</td>
<td>0.01</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The nonlinearity index indicates that none of the methods is applicable. In the control loops satisfying the applicability conditions for at least one method, the proposed system is able to select the methods which provide the correct diagnosis. However, the parameters concerning the reliability of each detection method in connection with the values of the feature indexes must be carefully tuned to avoid possible misdiagnoses.

8 Conclusions

The article presents an autonomous system for detection of valve stiction in control loops of industrial processes. Indexes have been introduced in order to quantify the features of processed data that most affect the correct application
of some among the most common, data-driven, stiction detection methods. A preliminary analysis has been carried out in order to identify minimal requisites for the detection methods to be successfully applicable. A first decision algorithm has been devised for the automatic selection of the applicable stiction detection methods. A second decision algorithm has been designed in order to compute the final detection decision as a weighted combination of the detection decisions provided by the single methods, based on the reliability of each method with respect to the processed data sequence. The system has been validated by testing on a set of benchmark industrial data.

References

URL http://www.prior-ip.com/patent/26565842/
URL http://dx.doi.org/10.1007/s00170-012-4296-8