Detection of Historical Alarm Subsequences Using Alarm Events and a Coactivation Constraint

GIANLUCA MANCA AND ALEXANDER FAY, (Senior Member, IEEE)

Institute of Automation Technology, Helmut Schmidt University, 22043 Hamburg, Germany

Corresponding author: Gianluca Manca (gianluca.manca89@gmail.com)

ABSTRACT This paper aims to provide an in-depth study of the detection of historical alarm subsequences, which are frequently used as an initial step for alarm flood analysis methods. Therefore, state-of-the-art approaches are comprehensively examined, evaluated, and compared. To overcome the limitations of these methods, a novel approach is presented, which uses outlier detection in time distances between alarm events (activation and return to normal) and an alarm coactivation constraint. The effectiveness and performance of the examined methods are illustrated by means of an openly accessible dataset, which is introduced in this paper. It is based on the ''Tennessee-Eastman-Process'', a benchmark in process automation. The intent is to provide a suitable dataset for the development and evaluation of alarm management methods in complex industrial processes using both quantitative and qualitative information from different sources. It is shown that the integration of supplementary information is beneficial for the overall performance and robustness of the detection method proposed here. This method allows for a more accurate detection of coherent historical abnormal situations, including phases with active root-cause disturbances and the normalization phases that follow their termination. Furthermore, the proposed method has the advantage that the detection results are less influenced by the alarm count, the propagation velocity, the duration of the situation, and the time distance between two causally independent situations in comparison to state-of-the-art approaches.

INDEX TERMS Abnormal situations, alarm analysis, alarm floods, alarm management, alarm systems.

NOMENCLATURE

ACT	Alarm activation.
ARC	Active root-cause.
ESD	Emergency shutdown.
MV	Manipulated variable.
P&ID	Piping and instrumentation diagram.
PV	Process variable.
RCD	Root-cause disturbance.
RTN	Return to normal.
TEP	Tennessee-Eastman-Process.
A, A_i	Alarm, ith alarm instance.
AS, AS_i	Alarm subsequence, ith alarm subsequence.
K	Total number of detected AS.
M	Total number of alarm variables.
MAD	Median absolute deviation.
MAD^{dist}	MAD-denominated distance.
N	Total number of alarm instances.
$PSI, aPSI$	Pair sets index, adapted pair sets index.
T_i^e	Starting time of AS_i.
T_i^s	End time of AS_i.
d	Time distance between two alarm events.
t, t_i	Time, ith time stamp.
t^A, t_i^A	Alarm event time, ith alarm event time stamp.
t^{ACT}, t_i^{ACT}	ACT time, ACT time stamp of A_i.
t^{RTN}, t_i^{RTN}	RTN time, RTN time stamp of A_i.
x, x_i	Process variable, ith process variable.
α, α_i	Alarm variable, ith alarm variable.
$\alpha^{ACT}, \alpha_i^{ACT}$	Alarm activation variable, ith alarm activation variable.
$\alpha^{bi}, \alpha_i^{bi}$	Binary alarm variable, ith binary alarm variable.
$\alpha^{RTN}, \alpha_i^{RTN}$	Alarm deactivation variable, ith alarm deactivation variable.
$\tilde{\alpha}, \tilde{\alpha}_i$	Indicator variable, ith indicator variable.
δ	Sampling interval.
ρ	Alarm rate.
τ^c	Coactivation threshold.
τ^e	ρ- or φ-threshold for the end of an AS.
τ^{MAD}	MAD^{dist}-threshold.

τ^s	ρ- or φ-threshold for the beginning of an AS.
$\tau^{s,e}$	ρ-threshold pair, $= [\tau^s, \tau^e]$.
φ	Number of active alarm variables.
ω	Window length.

I. INTRODUCTION

Both the complexity of modern process plants and the dominance of resource and energy prices as competitive factors increase the requirements for efficient and safe production that runs as automatically as possible. This objective can only be achieved with the help of suitable process automation. A human expert is called upon as a decision-maker when manual intervention is necessary due to malfunctions or deviations. This requires suitable interfaces between the plant operator and the plant so that the human operator can be made aware of abnormal situations. This function is fulfilled by alarms and alarm systems that are part of the process control system [32, p. 22], [58]. With the help of alarms, the plant operator should be able to intervene in a timely and targeted manner to transfer the process to a desired state and to prevent an automatic ESD of the system or inefficient process states [23, p. 1], [58]. Alarm management supports the plant operator in avoiding and controlling abnormal situations, e.g., by providing suitable interpretation support or preprocessing of raw data and information [58]. So-called alarm floods are among the most frequent and at the same time greatest challenges in alarm management. They overload the plant operator with alarms, thus limiting the ability to operate and monitor the assigned plant section, which can ultimately result in critical alarms or process states being overlooked [58]. Alarm floods can be one of the reasons for the poor performance of an alarm system and have sometimes been the cause of serious industrial accidents [13], [29], [30], [32, pp. 19-24].

According to [75], one cause that can trigger or promote an alarm flood is the propagation of abnormal situations due to causal dependencies throughout the process. These dependencies require the existence of a connection, i.e., a material, energy, or information connection [82, pp. 1-6]. Abnormalities in the form of process variable changes can propagate along these connections and thus trigger further deviations in other components. Abnormal situations are therefore not necessarily limited to their place of origin [7], [10], [19], [70], [75], [82, pp. 1-6]. If the resulting deviations in the respective underlying process variables exceed the defined alarm threshold values, a large number of consecutive alarms can be activated, which are the symptoms of a common RCD [7], [39], [69, pp. 47-74].

In recent years, many publications have dealt with the analysis and reduction of alarm floods. A comprehensive overview of alarm systems in general and open research challenges regarding alarm floods is given in [29] and [75]. A review and categorization of alarm data analysis approaches for alarm rationalization, online operator support, and root-cause analysis by using either time series analysis

or sequence mining techniques can be found in [51]. Furthermore, [42] studied and compared different methods regarding the similarity analysis of alarm sequences.

Methods such as those presented in [3], [25], [50], [66], [77], and [81] focus on the detection of historical alarm flood situations as a primary step in alarm sequence analysis. More generally, the detection results are referred to as alarm subsequences [3], [73], [81], which are shorter sequences that are contained in an original alarm sequence and are derived by using an appropriate analysis method. As an alternative, data mining approaches can be used to identify alarms that frequently occur together, as described in [26], [41], [73], and [79]. These methods are restricted to alarm clusters that are statistically relevant, and thus, they show limitations when an abnormal situation occurs only once. Alternatively, human experts can be used to find potential causal dependencies between process variables [1]. This is a time-consuming preprocessing step that requires extensive expert knowledge.

This paper focuses on methods for the detection of historical alarm subsequences. A systematic analysis of criteria for the offline detection of alarm floods was presented in [77]. Their limitations were analyzed by using a small selection of industrial examples. Since [77] was published, some novel alarm subsequence detection methods have been proposed, e.g., in [25] and [50]. Furthermore, [77] does not consider different parameter settings in the analysis of the methods. The relevance and not yet comprehensively examined characteristics of methods for the detection of historical alarm subsequences justify an updated and extended study.

The contributions of this paper are the following:

1) It provides an examination and performance assessment of existing approaches for the detection of historical alarm subsequences and alarm floods.
2) It develops a new method for the detection of historical alarm subsequences.
3) It introduces a novel and openly accessible alarm management dataset based on the TEP.
4) This dataset is used for the evaluation of the method proposed here to show its advantages.

The rest of the paper is organized as follows: Section II describes and analyzes the state-of-the-art methods for the detection of historical alarm subsequences with regard to existing limitations. It then derives suitable requirements. Section III describes the development of a novel approach based on the findings and requirements described in Section II. This proposed approach uses outlier detection in the time distances between alarm events (ACT and RTN) and an alarm coactivation constraint for the detection of historical alarm subsequences. Section IV introduces a novel alarm management dataset based on the TEP. This dataset is used in Section V for an in-depth evaluation and comparison of the methods in Sections II and III. Finally, this paper concludes with a comprehensive discussion of the evaluation results and an outlook on potential future work in Section VI.

II. STATE OF THE ART

A. TERMINOLOGY

With the aim of facilitating understanding of the terms used here, definitions are given below.

1) ABNORMAL SITUATION AND DISTURBANCE

An abnormal situation is characterized by a disturbed process [23, p. 237], where a disturbance is an unwanted deviation from a desired, normal, or defined state of at least one characteristic property, physical quantity, or parameter of a system [61]. In addition, the control system is unable to cope with these deviations so that operator intervention is required [8]. According to [8], it can be necessary to bring the process back to a safe state, establish the control of the process, and return the process to a normal operating state. Effectual and timely measures to counteract the RCD in an abnormal situation are the key to achieving these goals [35], [67], [75], [83]. Based on the alarm timeline described in [6], an abnormal situation is divided into the following phases:

1) An active root-cause (ARC) phase.
2) A normalization phase.

2) ROOT-CAUSE DISTURBANCE

The RCD is defined as the underlying cause in a complex cause-and-effect relation, on which all effects of the corresponding propagating abnormal situation, i.e., the other disturbances, are directly or indirectly dependent [75]. An abnormal situation can be initiated by either one or multiple RCDs [26]. The latter is a situation in which two or more simultaneously occurring RCDs can be classified as interrelated; i.e., it is not possible to distinguish between their respective effects on the process. Instead, a combined effect is observed.

3) ACTIVE ROOT-CAUSE PHASE

An alarm variable α_j is the unique identifier of a specific configured alarm; it can be represented, e.g., using the *"RTN time-stamped sequence"* or the *"multivalued alarm series"* [51]. The former consists of a sequence of chronologically ordered alarm instances A_i, where each A_i is defined by an ACT time (t_i^{ACT}) and an RTN time (t_i^{RTN}). The latter uses a time series representation for each α_j at time t, where j identifies the underlying process variable (following [51]):

$$\alpha_j(t) = \begin{cases} 2, & x_j(t) \geq \tau_j^{HIHI} \\ 1, & x_j(t) \geq \tau_j^{HI} \text{ and } x_j(t) < \tau_j^{HIHI} \\ -1, & x_j(t) \leq \tau_j^{LO} \text{ and } x_j(t) > \tau_j^{LOLO} \\ -2, & x_j(t) \leq \tau_j^{LOLO} \\ 0, & else \end{cases} \quad (1)$$

where τ_j^{HIHI}, τ_j^{HI}, τ_j^{LO}, and τ_j^{LOLO} represent alarm activation thresholds for high-high- (HIHI), high- (HI), low- (LO), and low-low alarms (LOLO), respectively. This alarm series allows for an unambiguous definition of the effective alarm

state so that only one of the conditions can be active at any given time. Fig. 1 shows an abnormal situation that triggers alarms in four different alarm variables. High- and low-alarm responses are represented with higher and lower levels, respectively. The total duration of the depicted abnormal situation is represented using a black arrow, whereas the corresponding ARC phase is marked with a red arrow. The latter begins with the RCD being initiated (red dotted line), which initially affects the process and begins propagation along causal dependencies. Eventually, the ARC phase ends with the termination of this RCD, which can be due to, among other factors, the operator taking effectual and timely action against it [8].
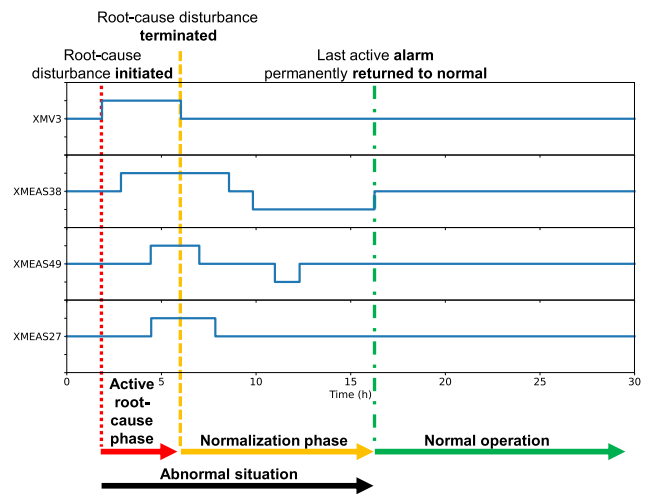

FIGURE 1. Illustration of the terminology used here. Solid blue lines represent the time trends of four alarm variables that experience an example abnormal situation. The lower level for each alarm variable represents a low alarm, and the higher level represents a high alarm.

4) NORMALIZATION PHASE

In Fig. 1, the normalization phase directly follows the ARC phase, as indicated by the successive yellow arrow and dashed line. However, this behavior does not always occur, as abnormal situations can escalate, eventually resulting in an ESD or major incident [6], [8]. In the case of a successful termination of the RCD, the normalization of a single process variable is characterized by a two-stage process response [6], which also applies to the process in its entirety:

1) A process deadtime.
2) A process response delay.

The first stage describes the phenomenon in which, despite effectual and timely operator intervention, the process continues to deviate further from its desired state. The second stage describes a time delay during which the process is responding to the operator intervention and the process variables are approaching their desired state [6]. The authors of this study have observed that plantwide normalization propagates throughout the process in a similar way as initial disturbance propagation does. Furthermore, the response of the implemented control system can cause different process variables
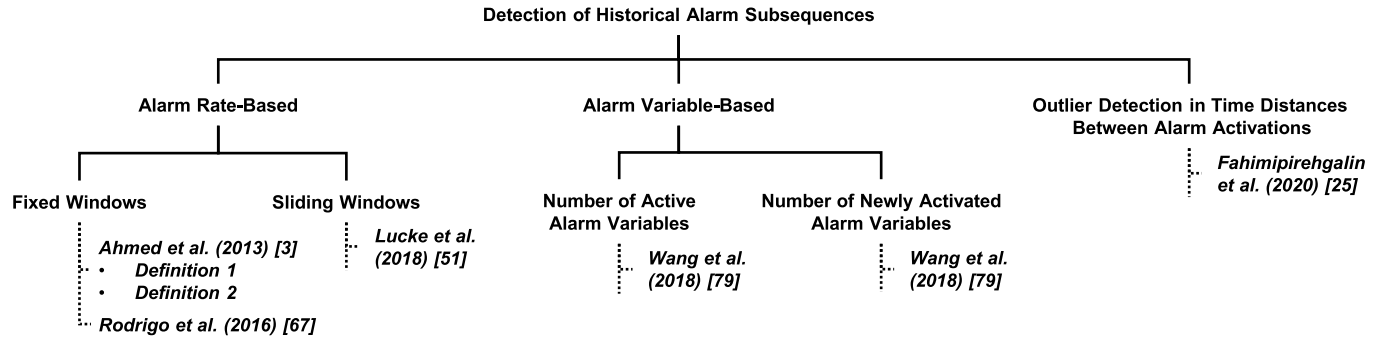
FIGURE 2. Categorization of existing approaches for the detection of historical alarm subsequences.

to show oscillating behavior, thus resulting in ACTs during the normalization phase. Due to complex interconnections and causal dependencies, it is possible for the normalization phase to take longer and raise more alarms than the actual ARC phase, which can confuse the operators, despite their reasonable primary intervention, and lead to misguided follow-up actions.

The difference between the ARC and normalization phases lies in the fact that the process is able to cope with the remaining deviations in the normalization phase so that additional operator intervention is not required. Eventually, the normalization phase results in all activated alarms switching back to an inactive alarm state for as long as no new abnormal situation arises. Fig. 1 depicts this successive normal operation with a green arrow and dashed-dotted line.

B. OVERVIEW

Various standards, guidelines, and publications have determined different quantitative alarm flood definitions by describing characteristic alarm rate-thresholds. According to [37] and [58], ρ is defined as the number of newly activated alarms per specified time window with length ω and operator. The following formula can be used to calculate ρ for a specific window starting at time t_i (adapted from [77]):

$$\rho\left(t_i, t_i + \omega - \delta\right) = \sum_{j=1}^{M} \sum_{t=t_i}^{t_i+\omega-\delta} \alpha_j^{ACT}(t) \qquad (2)$$

with the alarm activation variable α_j^{ACT}:

$$\alpha_j^{ACT}(t) = \begin{cases} 1, & \alpha_j(t-\delta) \neq \alpha_j(t) \ and \ \alpha_j(t) \neq 0 \\ 0, & else. \end{cases} \qquad (3)$$

Here, α_j^{ACT} considers the transition from an inactive to an active alarm state or between any two active alarm states as the ACT of a new alarm instance.

The following publications suggest using consecutive and nonoverlapping fixed windows with $\omega = 10$ min. In [6] and [32, p. 74], an alarm flood starts when $\rho \geq 10$ and ends when $\rho \leq 5$. Another characterization can be found in [20] and [23, p. 104], where an alarm flood is initially detected when $\rho > 10$. Aligning with the latter definition, [3] describes the end of an alarm flood as occurring when $\rho < 10$.

Additional divergent definitions can be found in [8] and [69, pp. 118-125]. Despite standardization efforts, no generally accepted quantitative alarm flood definition has been given. However, some of the approaches presented below use these thresholds.

Fig. 2 depicts a potential categorization of existing approaches for the detection of historical alarm subsequences. ρ-based approaches (left branch) use either fixed or sliding windows. In the latter case, ρ is calculated using overlapping windows with length ω computed for each t_i [50]. Fig. 3 shows four alarm variables (a), which are used for the calculation of ρ for fixed (b) and sliding (c) windows.

A common precondition of the approaches illustrated in Fig. 2, except for the α-based approaches (middle branch), is the assumption that chattering alarms are eliminated beforehand as far as possible using state-of-the-art methods, e.g., those described in [23, pp. 151-159], [32, pp. 107-125], [46], and [54]. Chattering alarms are alarms that frequently toggle between active and inactive alarm states over a certain period [23, p. 151]. They are a type of nuisance alarm, which are alarms that report excessively or unnecessarily; i.e., there is no abnormal situation or no operator response is required [20].

An alarm that is permanently active over a long period (e.g., more than eight [69, pp. 118-125] or 24 hours [6], [23, p. 239], [37]) is called a long-standing alarm. Different alarms of this type are described in [23, p. 215]. One type is long-standing alarms that are a nuisance to the operator, as they are either spurious and do not require any corrective action to be taken or can only be resolved in the long term. Example reasons for this kind of long-standing alarm are a faulty sensor [54], poorly tuned alarm thresholds, and long-lasting equipment maintenance [76]. Some of these alarms can be resolved by using state-of-the-art alarm management techniques, e.g., dynamic state-based alarming [76] or reconfiguration of the alarm thresholds [20], [23, pp. 51-58]. In addition, alarm shelving techniques can be used [23, p. 215], which allow alarms to be temporarily suppressed by using process knowledge [32, pp. 165-167]. In contrast, another type of long-standing alarm is defined by alarms that are in fact indicating an abnormal situation, which should be dealt with accordingly and which the operator needs to be
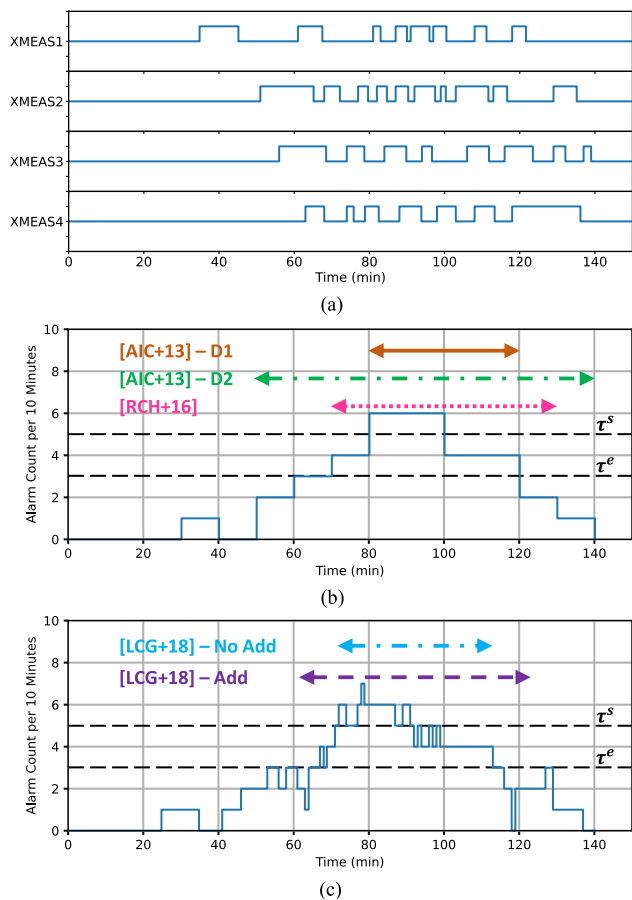
FIGURE 3. Alarm rate-based approaches. (a) Solid blue lines represent the time trends of four example alarm variables. The lower level for each alarm variable represents a low alarm, and the higher level represents a high alarm. (b) Fixed windows. (c) Sliding windows.

aware of. It is not recommended to ignore or suppress these true long-standing alarms [23, p. 215].

Alarm systems with a high number of nuisance alarms are not suitable for the application of most of the approaches shown in Fig. 2 or any other advanced alarm analysis method because their alarm behavior does not necessarily represent an abnormal situation from the process point of view. The resulting input would be a mix of flawed and reasonable alarm data, possibly leading to false and unreliable conclusions (in computer science, this is commonly referred to as the "*garbage-in, garbage-out*" concept [68]).

C. ALARM RATE-BASED APPROACHES
Due to the mismatch of alarm flood definitions, the evaluation in Section V considers different ρ-thresholds and ω, which have not been examined in depth in scientific publications. Furthermore, to provide a suitable framework for the comparison of the approaches, it is necessary to harmonize the representation of the alarm data input and the utilization of inequations in threshold design. Therefore, generic versions of the ρ-based approaches with three parameters, namely, τ^s, τ^e, and ω, are described here. These versions use a multivalued alarm series as an input. The mathematical

formulations given below are based on the original publications, which do not always specify the approaches in detail.

Reference [3] uses "*alarm flood detection and flood data isolation*" as a primary step in the similarity analysis of alarm floods. This type of analysis aims to find similar alarm patterns that can be used to group highly correlated alarms or to identify known patterns in an online manner to suggest appropriate corrective actions. Two definitions that describe the beginning and duration of an alarm flood subsequence are given in [3]. Henceforth, the generic versions of these definitions are referred to as "*[AIC+13]-D1*" and "*[AIC+13]-D2*".

For "*[AIC+13]-D1*", the starting times of the fixed windows can be described using the following equation:

$$t = [t_1, t_1 + \omega, t_1 + 2 * \omega, \ldots, t_1 + (P^\omega - 1) * \omega] \quad (4)$$

where P^ω is the total number of fixed windows with length ω. The beginning of an AS_k is then detected using the following inequation [3]:

$$\rho\left(t_i, t_i + \omega - \delta\right) \geq \tau^s. \quad (5)$$

The end of AS_k is detected by finding the first t_j in t, with $t_i < t_j$, that satisfies the following inequation:

$$\rho\left(t_j, t_j + \omega - \delta\right) \leq \tau^e. \quad (6)$$

Subsequently, AS_k can be defined by T_k^s and T_k^e:

$$T_k^s = t_i \quad and \quad T_k^e = t_j - \delta. \quad (7)$$

where the term $(t_j - \delta)$ represents the end of the fixed window that starts at $(t_j - \omega)$. The detected subsequences are modified in "*[AIC+13]-D2*"; i.e., the fixed windows before T_k^s, in which ρ increases from zero to τ^s, and after T_k^e, in which ρ decreases from τ^e to zero again, are added to AS_k. This definition was included due to the possibility that these windows may contain causally relevant alarms [3].

Reference [3] recommends using the following parameter settings: $\tau^s = 11$, $\tau^e = 10$, and $\omega = 10$ min. "*[AIC+13]-D1*" was also used in [14], [16], [34], and [48] for the identification of alarm flood subsequences.

Another ρ-based approach is proposed in [66]. The basic principle of this method is identical to "*[AIC+13]-D1*". On the grounds that the detected alarm subsequence has to be extended to include otherwise eliminated relevant alarms, the authors proposed attaching an additional fixed window both before and after it. Using (4), (5), and (6), the resulting T_k^s and T_k^e of an AS_k can be calculated as follows:

$$T_k^s = t_i - \omega \quad and \quad T_k^e = t_j - \delta + \omega. \quad (8)$$

However, these extensions can overlap and lead to individual alarms being contained in more than one alarm subsequence. Reference [66] does not give any indication as to whether such overlaps are to be merged; therefore, it is assumed that overlaps are a legitimate phenomenon. This must be appropriately considered during evaluation, as

in Section V. Henceforth, the generic version of the method proposed in [66] is referred to as *"[RCH+16]"*. Fig. 3 (b) shows an example of the resulting alarm subsequences using *"[AIC+13]-D1"*, *"[AIC+13]-D2"*, and *"[RCH+16]"* with example ρ-thresholds.

ρ-based methods with fixed windows have the advantage that they are easy to implement and have a low computational complexity. This benefit comes with several limitations: The overall structure of the fixed windows is determined by the examined dataset in terms of characteristics such as the start and end times. The specific ρ can be significantly altered if ACTs, already marginally enclosed by the window's boundaries, are deferred to adjacent fixed windows, thus possibly leading to missed alarm subsequences. Furthermore, the performance of these approaches depends greatly on suitable settings for ω and the ρ-thresholds, which are interrelated and should be tuned accordingly. For instance, an oversized ω could result in the overestimation of the detected subsequences' durations, which in turn could cause misleading interpretations in subsequent analyses insofar as the underlying abnormal situations would be poorly captured. This applies analogously to an undersized ω. A drawback of the recommended parameter settings arises from the sole emphasis on historical alarm flood situations, whereas all other abnormal situations and their inherent knowledge potential are neglected. On the other hand, manual parameter tuning demands substantial process knowledge since different criteria, e.g., the number of configured alarms, the complexity of the monitored plant, the alarm management techniques that are already implemented, and the overall alarm system quality, have to be considered.

A ρ-based approach that uses sliding windows is presented in [50]. The starting times of these windows can be described using the following equation:

$$t = [t_1, t_1 + \delta, t_1 + 2 * \delta, \ldots, t_1 + (S - 1) * \delta] \quad (9)$$

where S is the total number of samples. Similar to *"[RCH+16]"*, the preceding and following intervals with length ω, which differ from the overlapping sliding windows, are also included. To evaluate the effect of the extension, two generic versions of the method proposed in [50] are evaluated and compared in Section V. *"[LCG+18]-Add"* with additional windows, as originally intended, and *"[LCG+18]-No Add"* without. Using (9), (5), and (6), the resulting T_k^s and T_k^e of an AS_k can be calculated as follows: for *"[LCG+18]-No Add"*:

$$T_k^s = t_i \ and \ T_k^e = t_j + \omega - 2 * \delta \quad (10)$$

where the term $(t_j + \omega - 2 * \delta)$ represents the end of the sliding window that starts at $(t_j - \delta)$, and for *"[LCG+18]-Add"*:

$$T_k^s = t_i - \omega \ and \ T_k^e = t_j + 2 * (\omega - \delta). \quad (11)$$

Reference [50] recommends adapting τ^s, τ^e, and ω to the underlying process dynamics.

Fig. 3 (c) shows example alarm subsequences generated by *"[LCG+18]-Add"* and *"[LCG+18]-No Add"*. Comparing

Fig. 3 (b) and (c), it becomes apparent that methods using sliding windows tend to generate AS_k that show an earlier T_k^s as well as T_k^e and are therefore shifted to the left. Compared to fixed windows, the calculation of ρ using sliding windows has a higher computational complexity; e.g., one 10-min window with a δ of 10 s features 60 sliding windows. On the other hand, *"[LCG+18]-Add"* is less influenced by the global window structure, making it more robust against the assignment of individual alarms to specific windows. Other limitations already described for fixed window approaches also apply to *"[LCG+18]-Add"* and *"[LCG+18]-No Add"*.

D. ALARM VARIABLE-BASED APPROACHES

Because of the limitations of ρ-based approaches, two additional methods for the detection of historical alarm subsequences are presented in [77]. These methods utilize the number of active alarm variables in fixed windows. Similar to Subsection II.C, a generic version of the first method is described here. For each t_i in t, computed by using (4), the number of active alarm variables φ can be calculated using the following formula (following [77]):

$$\varphi(t_i, t_i + \omega - \delta) = \sum_{j=1}^{M} \tilde{\alpha}_j(t_i, t_i + \omega - \delta) \quad (12)$$

with the indicator variable $\tilde{\alpha}_j$ that specifies whether α_j has been active in the window starting at t_i and with length ω:

$$\tilde{\alpha}_j(t_i, t_i + \omega - \delta) = \begin{cases} 1, & \sum_{t=t_i}^{t_i + \omega - \delta} \alpha_j^{bi}(t) > 0 \\ 0, & else \end{cases} \quad (13)$$

and the binary variant α_j^{bi} of α_j:

$$\alpha_j^{bi}(t) = \begin{cases} 1, & \alpha_j(t) \neq 0 \\ 0, & else. \end{cases} \quad (14)$$

An AS_k is then detected by using (5) and (6) with φ instead of ρ and with τ^s and τ^e representing φ-thresholds. Subsequently, T_k^s and T_k^e can be described by using (7).

This method is robust against chattering alarms, although it is noted by the authors that nuisance long-standing alarms can have a negative impact on the detection of alarm subsequences since they increase φ. As a solution, a second detection criterion was proposed in [77]. It uses the number of newly activated alarm variables in a fixed 10-min window and adds the number of alarm variables that are still active in the current window but have not been active longer than a certain time limit parameter; e.g., the authors suggest using 30 min. This disregard for long-standing alarms is based on the assumption that they are not providing any valuable support for the operators [77]. However, as [23, p. 215] states, this cannot be true for all long-standing alarms (s. Subsection II.B).

The methods proposed in [77] are a suitable alternative to ρ-based approaches in cases where nuisance alarms cannot be reduced or eliminated using state-of-the-art alarm management techniques. The evaluation dataset used here, s. Section IV, shows no chattering or nuisance long-standing

alarms. Thus, regarding the examination conducted here, the φ-based approaches proposed in [77] have no advantage over the ρ-based approaches. φ-based approaches are therefore not considered in the evaluation in Section V.

E. UTILIZING OUTLIER DETECTION IN THE TIME DISTANCES BETWEEN ALARM ACTIVATIONS

A less parameter-dependent method is proposed in [25], henceforth referred to as *"[FWV20]"*. It utilizes outlier detection in the time distances between ACTs to cluster alarms that are close in time. It is based on the observation that abnormal situations in industrial plants tend to generate alarm sequences with alarm instances showing both short and long time distances between their respective ACTs. Two possible reasons for this behavior are given in [25]:

1) Disturbances within one part of the plant propagate more rapidly than between different parts of the plant.
2) Independent and nonoverlapping disturbances arise with a significant time distance.

Fig. 4 shows three example consecutive abnormal situations that trigger alarms in several alarm variables. The ARC phases of the first and second abnormal situations are each followed by a corresponding normalization phase and normal operation. The last ARC phase results in an ESD of the plant. In particular, the first and last abnormal situations show ACTs that are close in time. Furthermore, the transition from one abnormal situation to the following one can be identified by the significant distance between their respective last and first alarms.

As a first step in *"[FWV20]"*, the t_i^{ACT} of each alarm instance A_i must be extracted, e.g., by calculating the ACT transitions of each α_j using (3). Next, these time stamps must be sorted in chronological order [25]:

$$\boldsymbol{t} = [t_1^{ACT}, t_2^{ACT}, \ldots, t_N^{ACT}], \quad t_1^{ACT} < t_2^{ACT} < \ldots < t_N^{ACT}. \tag{15}$$

The vector $\boldsymbol{t}$ is then used to calculate a time distance vector $\boldsymbol{d}$ of neighboring alarms:

$$\boldsymbol{d} = [d_1, d_2, \ldots, d_{N-1}] \tag{16}$$

with the distance d_i of two adjacent alarms A_i and A_{i+1} [25]:

$$d_i = t_{i+1}^{ACT} - t_i^{ACT}. \tag{17}$$

The diagram in Fig. 5 illustrates the values of $\boldsymbol{d}$ for the example in Fig. 4 (left axis label). The peak values suggest the existence of statistical outliers. Instead of an absolute distance threshold, the *"[FWV20]"* uses a median absolute deviation (MAD)-denominated distance MAD^{dist} as an outlier detection method. In this way, the statistical characteristics of $\boldsymbol{d}$ are utilized. The value of MAD^{dist} for a specific d_i can be computed using the following equation [25]:

$$MAD^{dist}(d_i) = (|d_i - median(\boldsymbol{d})|)/(MAD(\boldsymbol{d})) \tag{18}$$

with:

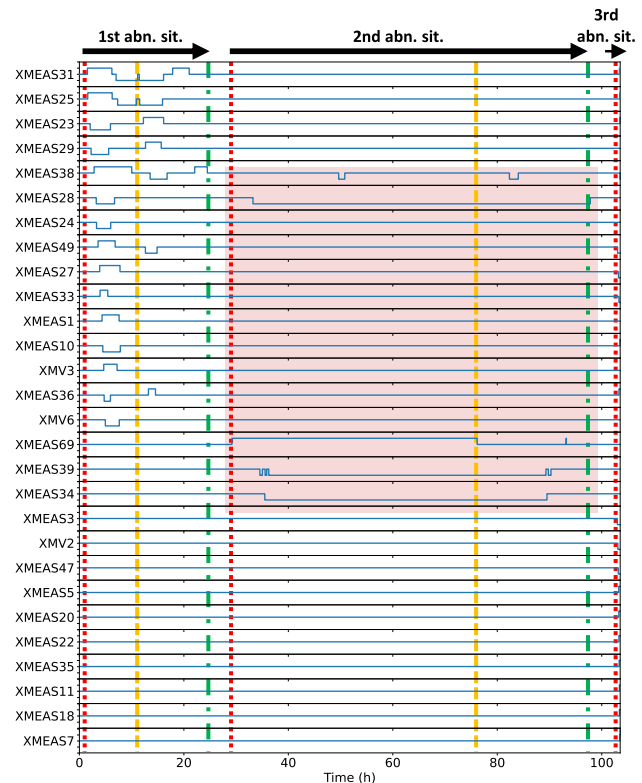$$MAD(\boldsymbol{d}) = median(|d_i - median(\boldsymbol{d})|). \tag{19}$$

FIGURE 4. Three example consecutive abnormal situations (abn. sit.). Solid blue lines represent the time trends of alarm variables. The lower level for each alarm variable represents a low alarm, and the higher level represents a high alarm. Red dotted lines represent the beginning of ARC phases. Yellow dashed lines represent the beginning of normalization phases. Green dashed-dotted lines represent the beginning of normal operation. The red area marks a period with high time distances between alarms located within one abnormal situation.

The resulting vector for all d_i in $\boldsymbol{d}$ can be described using the following formula:

$$\boldsymbol{MAD^{dist}}(\boldsymbol{d}) = \left[MAD^{dist}(d_i), \ldots, MAD^{dist}(d_{N-1})\right]. \tag{20}$$

The axis label on the right side of the diagram in Fig. 5 specifies the values of $\boldsymbol{MAD^{dist}}(\boldsymbol{d})$ for the above-mentioned example. Outliers are identified by the MAD^{dist}-threshold τ^{MAD} if the following inequation is satisfied [25]:

$$MAD^{dist}(d_i) > \tau^{MAD}. \tag{21}$$

The authors of [25] use a τ^{MAD} of 400 but recommend adjusting it to generate the desired cluster sizes. In the event of an outlier, the corresponding index i is then used as a cutoff point for generating timed clusters, which are equivalent to the notion of alarm subsequences. Namely, one AS_k ends with A_i and the next AS_{k+1} starts with A_{i+1}. Moreover, AS_1 starts with A_1 and AS_K ends with A_N. Further alarms are assigned to the respective subsequences according to their indices [25]:

$$AS_k = \left\{A_{k_1}, A_{k_1+1}, \ldots, A_{k_1+n_k-1}\right\}, \quad k = 1, \ldots, K \tag{22}$$

where k_1 is the index of the first alarm in AS_k and n_k is the number of alarms in AS_k. After this first step, follow-up steps are described by the authors: the identification of similar
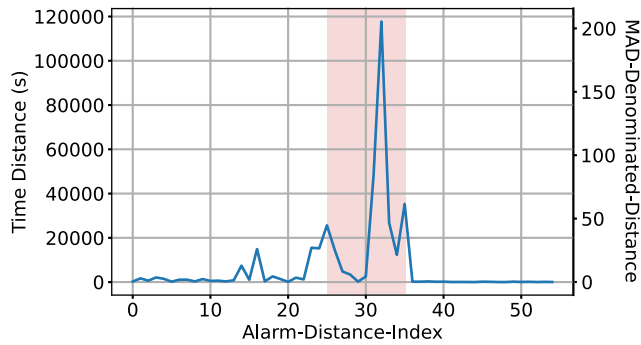
FIGURE 5. Time distance (left axis label) and MAD-denominated distance (right axis label) diagram for the example in Fig. 4 generated by using *"[FWV20]"*. The red area marks a period with high time distances between alarms located within one abnormal situation.

clusters by means of the Jaccard distance and the analysis of their potential root-causes by using transfer entropy.

On the one hand, *"[FWV20]"* shows some advantages compared to the ρ-based approaches. τ^{MAD} describes a relative threshold and is therefore more flexible regarding different process dynamics. Furthermore, all activated alarms are considered and unambiguously assigned to subsequences, making their inherited knowledge available for subsequent analysis steps. Ambiguous overlaps between subsequences are therefore fully avoided. On the other hand, some of the abovementioned limitations persist, and new limitations arise.

One major drawback of *"[FWV20]"* derives from different propagation velocities, initially described as a possible reason for the alarm behavior observed in [25]. If a disturbance propagates slowly through the process, the resulting time distances between alarms located within one abnormal situation can be higher than those between two individual situations. In this case, *"[FWV20]"* shows a limitation regarding the identification of coherent abnormal situations. An example of this phenomenon is illustrated in Fig. 4 and Fig. 5, marked by the two red areas. ACTs during the second abnormal situation occur with a higher time distance compared to the time distances between the first and second abnormal situations as well as between the second and third. This renders it impossible to define a τ^{MAD} for *"[FWV20]"* that is capable of correctly identifying all relevant alarm subsequences without falsely splitting the second abnormal situation.

Another drawback derives from transitions between ARC phases and normalization phases, which can have a complex time distance structure despite their causal coherence. *"[FWV20]"* is limited in effectively linking these two phases in one comprehensive alarm subsequence if, e.g., a long-lasting and mostly stable ARC phase is followed by a normalization phase that is accompanied by ACTs. In that case, *"[FWV20]"* would wrongly classify the stable, but still disturbed, period as a cutoff point.

In addition, the τ^{MAD} parameter still needs to be tuned, which requires suitable process knowledge. The evaluation in Section V examines the performance of different τ^{MAD} values.

F. REQUIREMENTS
The examination of already established approaches has demonstrated the existence of several limitations. Based on this examination, the following requirements (R1 to R3), which should be met by methods intended for the detection of historical alarm subsequences, can be expressed:

R1: An alarm subsequence should contain all alarms that arise during an abnormal situation, irrespective of the alarm count, alarm frequency, duration of the situation, and propagation velocity.

R2: An alarm subsequence should include both the ARC phase and normalization phase of the corresponding abnormal situation.

R3: A detection method should separate two causally independent and consecutive situations into two alarm subsequences irrespective of their distance in time.

III. PROPOSED APPROACH
A. OVERVIEW OF THE PROPOSED APPROACH
Based on the promising approach *"[FWV20]"*, this paper proposes an improvement to this method that aims at meeting the stated requirements, thus overcoming its limitations. The improvement is achieved by using additional information. In contrast to *"[FWV20]"*, which considers only ACTs, the proposed approach uses both alarm event types, namely, ACT and RTN, thus facilitating the detection of the normalization phases of abnormal situations. Furthermore, domain knowledge is used to define a general coactivation constraint, which is applied for the reasonable merging of detected alarm subsequences, thus expanding the view to periods in which alarms are active rather than single transition points from an inactive to an active alarm state. Fig. 6 shows the general structure of the proposed *"alarm coactivation and event detection method"* (ACEDM) using the *"formalized process description"* given in [72]. Process operators (green rectangles) as well as processed and generated information (blue hexagons) are described in detail below.

B. DETAILS OF THE PROPOSED APPROACH
The ACEDM starts with O1.1, *"Extraction and sorting of alarm events"*. The input to this first step is a set of preprocessed historical alarm data (Information I1.1). It is necessary to reduce chattering and nuisance long-standing alarms, as they pose an obstacle to the detection of proper alarm subsequences. Furthermore, historical data must be represented in a way that allows for ACT and RTN times of alarm instances to be extracted. Here, the multivalued alarm series described in (1) is used. The t_i^{ACT} of each A_i can be obtained similarly to *"[FWV20]"*. The corresponding t_i^{RTN} can be extracted using the following formula:

$$\alpha_j^{RTN}(t) = \begin{cases} 1, & \alpha_j(t-\delta) \neq 0 \text{ and } \alpha_j(t) \neq \alpha_j(t-\delta) \\ 0, & else. \end{cases}$$

(23)

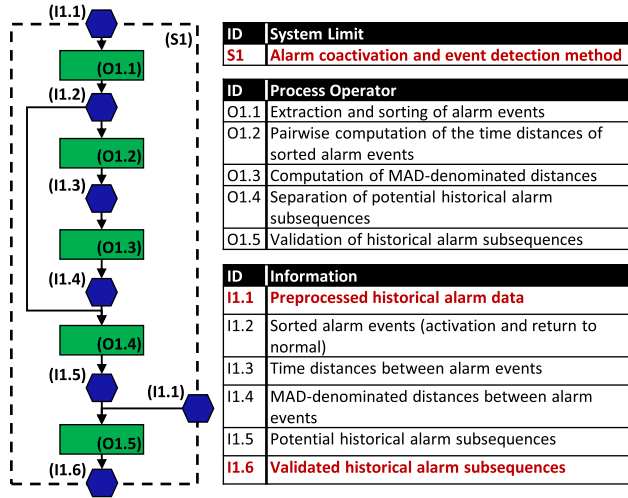| ID	System Limit
S1	Alarm coactivation and event detection method

ID	Process Operator
O1.1	Extraction and sorting of alarm events
O1.2	Pairwise computation of the time distances of sorted alarm events
O1.3	Computation of MAD-denominated distances
O1.4	Separation of potential historical alarm subsequences
O1.5	Validation of historical alarm subsequences

ID	Information
I1.1	Preprocessed historical alarm data
I1.2	Sorted alarm events (activation and return to normal)
I1.3	Time distances between alarm events
I1.4	MAD-denominated distances between alarm events
I1.5	Potential historical alarm subsequences
I1.6	Validated historical alarm subsequences

FIGURE 6. "Formalized process description" of the proposed ACEDM.

Here, α_j^{RTN} considers the transition from an active to an inactive alarm state or between any two active alarm states as the RTN of an alarm instance. It is assumed that no alarms are active at the beginning of the dataset. In addition, all alarms that are still active at the end of the dataset are considered to RTN at that time. Subsequently, the extracted time stamps must be sorted in chronological order (Information I1.2) (adapted from [25]):

$$\boldsymbol{t} = [t_1^A, t_2^A, \ldots, t_{2N}^A], \quad t_1^A < t_2^A < \ldots < t_{2N}^A \qquad (24)$$

where t_i^A is the time stamp of an alarm event (t^{ACT} or t^{RTN}) and $2N$ is the number of alarm events. Vector $\boldsymbol{t}$ is then used as an input to O1.2, *"Pairwise computation of the time distances of sorted alarm events"*. Here, d_i of two adjacent alarm events at times t_i^A and t_{i+1}^A is calculated using (17). A total of $(2N - 1)$ time distances are calculated since the last event has no successor and therefore no time distance. The time distances are subsequently described using the time distance vector (Information I1.3) (adapted from [25]):

$$\boldsymbol{d} = [d_1, d_2, \ldots, d_{2N-1}]. \qquad (25)$$

In O1.3, $\boldsymbol{MAD}^{dist}(\boldsymbol{d})$ is computed (Information I1.4) using (18) and (19). It is then described using (20), whereas the highest index of $\boldsymbol{d}$ is $(2N - 1)$ instead of $(N - 1)$. Afterwards, I1.4 is provided as an input to O1.4 for the separation and subsequent detection of potential historical alarm subsequences using (21). In the event of an outlier, the corresponding times t_i^A and t_{i+1}^A are then used as cutoff points for generating potential alarm subsequences. Namely, one potential subsequence pAS_k ends with t_i^A, and the next pAS_{k+1} starts with t_{i+1}^A. Moreover, pAS_1 starts with t_1^A and pAS_{K^p} ends with t_{2N}^A, where K^p is the number of detected potential alarm subsequences. Hence, each pAS_k can be subsequently defined by T_k^s and T_k^e, which are the time stamps of the first and last alarm events in pAS_k, respectively. Both as well as $\boldsymbol{t}$ (Information I1.2) are then utilized to assign the alarm instances to the potential alarm subsequences (Information I1.5). To determine if A_i with t_i^{ACT} is part of pAS_k, the following formula can be used:

$$A_i \in pAS_k, \quad if \ T_k^s \leq t_i^{ACT} \leq T_k^e. \qquad (26)$$

Finally, in O1.5, the detected potential alarm subsequences are validated based on an alarm coactivation constraint; here, a method presented in [50] and [51] regarding the application of alarm coactivations to online alarm flood classification is adopted as inspiration for this proposal. Furthermore, [41] and [79] showed that overlapping alarms can be an indicator of a causal connection between them. The principle of the constraint proposed here is as follows: two potential subsequences pAS_k and pAS_{k+1} are merged into a new validated alarm subsequence AS_i if the number of active alarms, which can be calculated for any point in time between pAS_k and pAS_{k+1} using (14), satisfies the following in equation:

$$\sum_{j=1}^{M} \alpha_j^{bi} \left(T_k^e + \delta \right) \geq \tau^c. \qquad (27)$$

The threshold τ^c allows for the consideration of known or assumed nuisance long-standing alarms by adjusting the sensitivity of the coactivation constraint. The approach used in O1.5 assumes that two distinct alarm subsequences are most likely part of the same abnormal situation and thus dependent if they share common alarms. A reason for this assumption is that a fault propagating along connections throughout the process can result in a multitude of consecutive and coactive alarms [7], [39], [69, pp. 47-74]. Furthermore, even if these overlapping subsequences turn out to be independent, e.g., because they are the result of independent disturbances in different parts of the process with no common effect on the plant, clustering methods are not suitable for the analysis of causal relations [25]. Hence, the subsequences in question should be merged and thereby made available for further processing using a suitable causal analysis method. The resulting validated alarm subsequences (Information I1.6) can be represented using (22).

C. ILLUSTRATION OF THE ADVANTAGES OF THE PROPOSED APPROACH

Fig. 7 illustrates an example application of the proposed alarm coactivation constraint. Here, two RCDs, which are initiated after one hour and 79 hours, are able to propagate throughout the process. The first RCD is terminated after 45 hours and is followed by a normalization phase, which transitions into normal operation after 20 more hours. The second RCD is terminated after being active for approximately nine hours and is followed by a normalization phase, which transitions into normal operation after 11 more hours. Now, let the red areas be potential alarm subsequences, which have been detected since they are composed of alarm events that are very close in time. These subsequences divide the original underlying abnormal situations into ten distinct parts. Between each pair of adjacent potential subsequences,
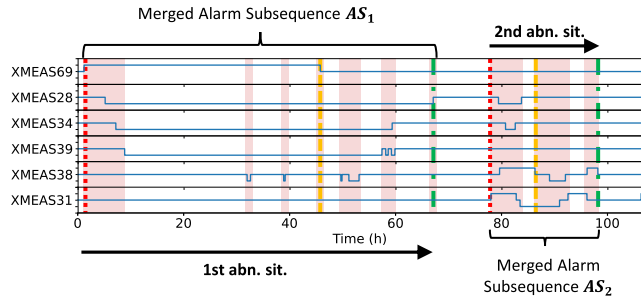
IEEE Access

G. Manca, A. Fay: Detection of Historical Alarm Subsequences Using Alarm Events and a Coactivation Constraint

FIGURE 7. Illustration of the application of the proposed alarm coactivation constraint to two example abnormal situations (abn. sit.). Solid blue lines represent the time trends of six alarm variables. The lower level for each alarm variable represents a low alarm, and the higher level represents a high alarm. Red dotted lines represent the beginning of ARC phases. Yellow dashed lines represent the beginning of normalization phases. Green dashed-dotted lines represent the beginning of normal operation. Red areas represent detected potential historical alarm subsequences by using the ACEDM.

FIGURE 8. Time distance (left axis label) and MAD-denominated distance (right axis label) diagram for the example in Fig. 4 generated by using the ACEDM. The red dashed-dotted line represents a τ^{MAD} of 11.

which are part of the same abnormal situation, at least one alarm is active, rendering it possible for the proposed coactivation constraint to be successfully applied. In doing so, the potential subsequences are merged into AS_1 and AS_2, which embrace the full length of the two depicted abnormal situations. If this ought to be done using *"[FWV20]"*, no τ^{MAD} will allow the generation of two suitable timed clusters for the two abnormal situations. This is due to the time distances between ACTs within the first abnormal situation being greater than the time distance between the two independent situations.

To compare the ACEDM to *"[FWV20]"* and to examine whether the existing limitations can be successfully overcome, both are applied to the example introduced in Subsection II.E. Fig. 8 therefore provides a corresponding time distance (left axis label) as well as a MAD^{dist} (right axis label) diagram with values derived by applying the ACEDM to the example in Fig. 4. Due to the utilization of both alarm events, the maximum distance index is twice as high as that of *"[FWV20]"*, which allows for a more accurate detection of the start and end of an abnormal situation. Furthermore, the number of peaks in the diagram is increased, resulting in a higher number of cutoff points and thus more potential alarm subsequences. Tests with varying integer values between zero and 200 for τ^{MAD} reveal optimal results using a threshold of 11 (marked with a red dashed-dotted line). The application of this τ^{MAD} value initially results in 10 potential alarm subsequences, which are merged by using the proposed alarm coactivation constraint with a τ^c of one. The resulting validated subsequences represent the three example abnormal situations to the full extent with regard to their starting and ending alarm events. Hence, an examination of this example shows that it is possible to overcome the two major limitations of *"[FWV20]"*, namely, dealing with slowly propagating disturbances and normalization phases, by utilizing additional information and a coactivation constraint.
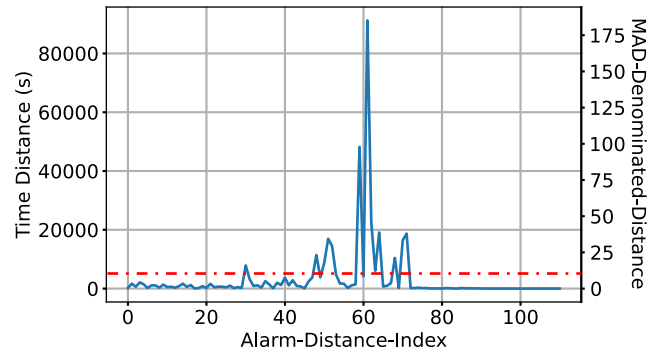
D. DISCUSSION OF THE LIMITATIONS OF THE PROPOSED APPROACH

One limitation of the ACEDM derives from its sensitivity to long-standing alarms when applying the coactivation constraint. For example, a nuisance long-standing alarm, which has not been dealt with by using τ^c, can cause all affected potential subsequences to be merged incorrectly into a single alarm subsequence. If, on the other hand, τ^c is set too high, the potential alarm subsequences are not merged, even if they are actually part of a coherent abnormal situation. Therefore, it is of great importance that nuisance long-standing alarms are either fully eliminated by using extensive process knowledge or state-of-the-art alarm management techniques (s. Subsection II.B), or by tuning τ^c appropriately.

Similar to *"[FWV20]"*, a τ^{MAD} needs to be set for the ACEDM. To avoid manual tuning, it can be set to a default value of zero. In this case, each alarm event will be put into a separate potential alarm subsequence, and these are then subsequently merged according to their overlapping alarms. Here, only the alarm coactivation constraint is used for the detection of historical alarm subsequences. The evaluation in Section V considers and examines different τ^{MAD}, including zero, to evaluate the effectiveness of the proposed coactivation constraint and additional alarm event input.

IV. EVALUATION DATASET AND SIMULATION MODEL

Reference [31] describes different evaluation methods for applications in information systems research. Among these are observational methods, e.g., case studies, and experimental methods, i.e., controlled experiments and simulations. Both case studies and controlled experiments use real plant data. However, [17] describes that acquiring suitable alarm and process data from industrial plants remains a significant challenge in developing and evaluating alarm management methods. This is due to the potential reservation of industrial companies as well as alarm systems that perform poorly, thus limiting the implementation of advanced alarm analysis methods. Therefore, this study uses a simulation model to generate a dataset of artificial process and alarm data. The following must be available for a suitable simulation model and dataset:

46860

VOLUME 9, 2021

1) Alarm and process data.
2) A piping and instrumentation diagram (P&ID) as well as additional connectivity information.
3) Relevant alarm management information, e.g., alarm thresholds and parameter settings of the implemented alarm management techniques.
4) Information about the induced abnormal situations, including their RCDs and process normalizations. This allows for an in-depth evaluation of the generated analysis results and is an advantage over industrial data.

Since the initial publication of the TEP in 1993 as "*a plant-wide industrial control problem*" [22], it has become accepted as a benchmark simulation model in the process automation of chemical plants [7], [9]. The TEP is based on the processes of an actual plant of the *Eastman Chemical Company* (Tennessee, United States of America). Its persistent academic relevance is sustained by different publications in the fields of fault detection and identification [7], [82, pp. 72-82] and the parameterization of control loops and structures [52] [57], [65]. Furthermore, TEP simulations have been used in different alarm management and causal analysis studies, e.g., [7], [36], [70], [81], and [83]. However, the systematic parameterization and implementation of suitable alarm thresholds and alarm management techniques for the TEP have mostly been disregarded to date. In reference [81], a set of 33 alarm thresholds is presented. A set of calculation formulas and corresponding alarm thresholds can be found in [56] and [44]. This work is used here as a basis for the systematic development and implementation of suitable alarm thresholds and state-of-the-art alarm management techniques.

The experimental evaluation method utilized here has the following two components:

1) A simulation model, which is used to generate an evaluation dataset. This MATLAB-Simulink model of the TEP was presented in [9] and can be accessed and downloaded online via the "*Tennessee Eastman Challenge Archive*".[1] The corresponding alarm thresholds and alarm management techniques are briefly described in Subsection IV.C. A detailed description of the systematic design process as well as the specific alarm threshold values and parameter settings of the alarm management techniques can be found in [53].
2) An evaluation dataset with several tests. Here, a test is a simulation run with specific abnormal situations. This dataset is used for the evaluation of the methods examined and proposed in Sections II and III (s. Section V). The dataset itself as well as a supplementary technical report can be accessed and downloaded via the

"*IEEE DataPort*" [53].[2] Subsection IV.D describes in detail the content of this dataset and how it was created.

Both of the above components are openly accessible, so the tests introduced here can be reused by other researchers. Alternatively, the described simulation model can be utilized to generate further tests.

A. TENNESSEE-EASTMAN-PROCESS

Fig. 9 shows the P&ID of the TEP according to [7], [9], and [22]. Considering the plant and process hierarchy described in [4] and [5], the TEP can be separated into five main equipment modules: a two-phase chemical reactor (*R003*), a condenser (*C115*), a vapor-liquid-separator (*S102*), a stripper (*E005*), and a reboiler (*G111*). Furthermore, the following control and equipment modules, placed along material connections (pipes) between the components, are part of the TEP: 11 automatic pneumatic control valves (*V160-V170*), two pumps (*P101* and *P102*), and one compressor (*K100*) [7], [22].

The P&ID in Fig. 9 displays the extended process of [9] and the tags adapted by [7]. These tags show 36 PVs of types flow (F), pressure (P), temperature (T), level (L), and chemical component concentration (A) that are indicated (I) and registered (R). In addition, the work of compressor *K100* is measured. Eight different chemical components (*A-H*) are part of the TEP, and each of the chemical component concentration PVs measures at least five of these components as a separate value so that a total of 73 PVs are measured and recorded. Furthermore, 12 MVs, namely, the 11 control valves and the agitator speed of *R003*, are measured and recorded. The TEP simulation model utilized here uses an *XMEAS-No.* (*XMEAS 1* to *73*) and an *XMV-No.* (*XMV 1 to 12*) to identify the corresponding PV and MV signals, respectively. A detailed overview and mapping of all PVs and MVs can be found in [9]. A detailed description of the process operations and process steps can be found in [7], [9], and [22].

B. CONTROL STRUCTURE

According to the simulation results presented in [55], the TEP is highly unstable and tends to exceed or go below the set ESD thresholds after a runtime of approximately one hour. Hence, a suitable control structure has to be implemented. Reference [65] describes a control structure consisting of 17 control loops, which is implemented in the TEP simulation model used here. Its unique features are as follows:

1) The reactor pressure is controlled by using the relatively low purge flow.
2) The production rate controller sets the ratio controller set points on all feeds, purge, and liquid flows.
3) The reactor level is controlled by setting the separator temperature controller, which controls the cooling water supply to the condenser [52].

---

[1]http://depts.washington.edu/control/LARRY/TE/download.html - Version: January 23, 2015 (MultiLoop_mode1.mdl Version 1.3.3)

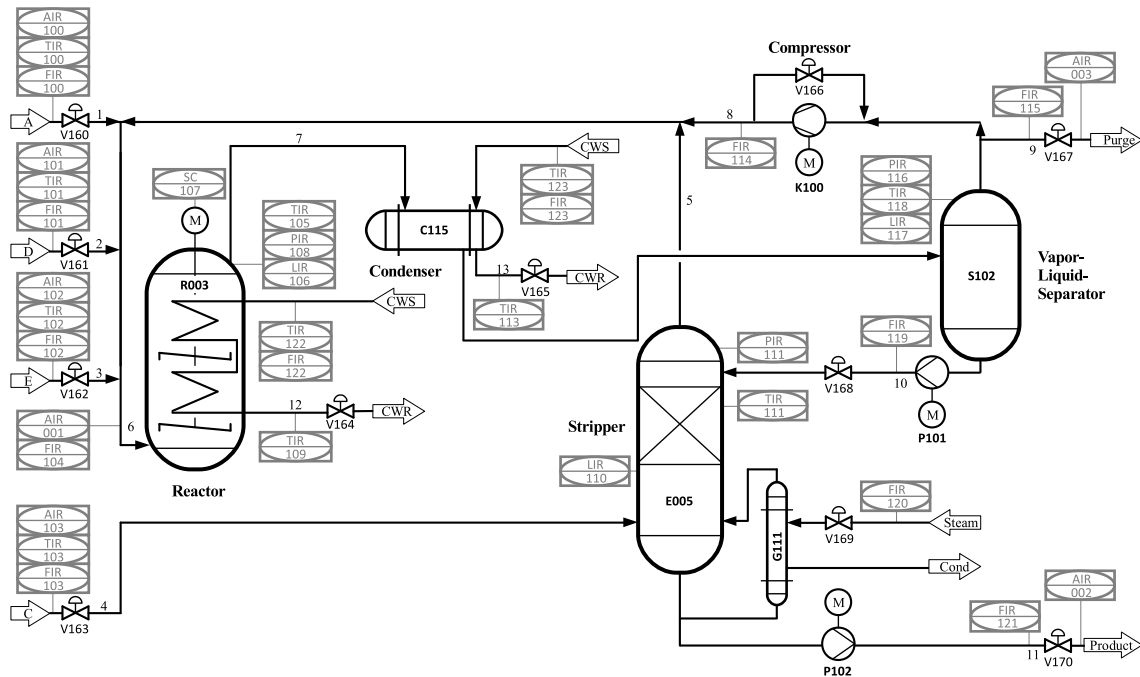[2]https://ieee-dataport.org/open-access/tennessee-eastman-process-alarm-management-dataset

FIGURE 9. P&ID of the TEP ([7], [9], and [22]).

C. ALARM DESIGN AND MANAGEMENT

The set of alarm thresholds presented in [81] is partly set inside the normal operating limits as defined by [22], which also apply to the adjusted normal operating mode described in [65]. Hence, corresponding alarms do not necessarily indicate an abnormal situation. The set of calculation formulas and corresponding alarm thresholds for the TEP, which was initially presented in [56] and adapted in [44], contains a total of 60 HI- and LO-Alarm thresholds. Due to the differences of the TEP model used here and the one used in [44] and [56], a high average ρ can be observed even during normal operation. Thus, the alarm thresholds from [44] and [56] are inadequate and cannot be used with this TEP simulation model and controller. Hence, the existing calculation formulas are used to design a novel set of alarm thresholds. This set is adapted to the adjusted normal operating mode and includes the additional process measurements from [9] to meet the target of no ACTs during normal operations and to provide an appropriate response during abnormal situations [37].

Based on the iterative rationalization step of the alarm management life cycle described in [6], several revision steps were conducted utilizing a selection of relevant tests. Thus, suitable alarm thresholds for all relevant PVs and MVs with 81 LO- and 81 HI-Alarm as well as five HIHI- and three LOLO-Alarm thresholds were obtained. Furthermore, two alarm management techniques, an exponential-weighted moving-average filter (EWMA filter) and alarm deadbands, were implemented and parameterized according to recommendations from [2], [23, pp. 151-159], [32, pp. 107-125], [37], [38], and [39]. The results of the

implemented alarm thresholds and alarm management techniques are that there are no ACTs during normal operations and an overall reduced alarm rate. Furthermore, no chattering or nuisance long-standing alarms arise during abnormal situations.

D. TEST DESIGN AND ANALYSIS

References [9] and [22] describe 28 different RCDs (*IDV 1* to *28*), each caused by manipulating one or more PVs. These disturbances lead to local or plantwide abnormal situations. Reference [81] describes that the RCDs of the *step* type, which are described as a sudden variation in a single PV or a combination of PVs, are the only *IDVs* that can propagate throughout the TEP and cause alarm floods. Hence, the dataset presented and used here is limited to RCDs of the *step* type. Furthermore, eleven additional RCDs of the *step* type, which are each initiated by a full closure of a single control valve (*XMV 1 to 11*), are described in [7].

In the first step, RCDs that are suitable for the application of alarm analysis methods must be selected. The following requirements need to be met:

1) The RCD and its corresponding ARC phase show a sufficient number (>1) of alarm variables in the alarm condition, which allows for the detection of nontrivial alarm subsequences. For this purpose, each ARC phase is simulated for 24 hours, except for situations where an ESD occurs.
2) The RCD does not cause an ESD of the TEP earlier than 15 min after its initiation in order to generate an

ARC phase that is long enough for the application of alarm prediction methods.

Only four IDV RCDs, namely, *IDV 1* (manipulation of the chemical component concentration of materials A and C in feed C), *IDV 2* (manipulation of the chemical component concentration of material B in feed C), *IDV 5* (manipulation of the temperature of the cooling water inlet of the condenser), and *IDV 6* (manipulation of the flow of feed A), and four XMV RCDs, namely, *XMV 2* (manipulation of *V162*), *XMV 3* (manipulation of *V160*), *XMV 4* (manipulation of *V163*), and *XMV 6* (manipulation of *V167*), satisfy these requirements. These RCDs are selected for further test design.

Next, suitable ARC phase durations and disturbance scaling must be selected for each of the selected RCDs. The latter have not yet been examined in alarm analysis publications. Preliminary tests have shown that disturbance scaling affects the number of active alarm variables, as well as the order of alarm instances and their dynamic behavior. For this dataset, the following scaling was used, which was selected by applying the same requirements as in the selection of the RCDs:

- *IDVs 1, 2*, and *5*: 100% and 95%.
- *IDV 6*: 100%, 90%, 80%, and 75%.
- *XMVs 2, 3, 4*, and *6*: 100%, 97.5%, and 95% (with 100% representing a full closure of the valve).

Reference [22] recommends a simulation time of 24 to 48 hours regarding the RCDs of the *step* type. In [44], an RCD duration of 7.5 hours was used. Other relevant publications, e.g., [7], [9], [56], and [81], do not provide any suggestions regarding ARC phase durations. All selected *XMV* RCDs as well as *IDV 6*, with a scaling of 90% and 100%, result in an ESD after a certain amount of time. These ESDs were used for further test design. For the remaining *IDVs*, the following disturbance durations were selected:

- *IDV 1:* 10 hours.
- *IDVs 2* and *5:* 48 hours.
- *IDV 6* with 80% and 75% scaling: 15 hours.

To facilitate the evaluation of the detection performance for RCDs with a short ARC phase, which occurs in the cases of prompt and effective operator intervention, some of the selected RCDs are also tested with a duration of one hour.

In addition to individual RCDs, [7] recommends examining abnormal situations that are initiated by more than one concurrent RCD. Due to the high interconnections and causal dependencies in the TEP simulation model used here, concurrent RCDs must be classified as causally associated in their respective effects on the process; i.e., they are part of one common abnormal situation. In the dataset presented here, *IDVs 1, 2,* and *5* are used to generate concurrent RCDs with variations regarding their combination and temporal overlap.

To date, no publication on the TEP has taken normalization phases into account. Example tests of the selected RCDs have shown that some normalization phases take longer and raise more alarms than the corresponding ARC phases. Therefore, it is important to take normalization phases into consideration for the test design. For this reason, each test consists of three consecutive abnormal situations. Consequently, the first two abnormal situations must have a normalization phase, in which the plant returns to normal operation, before the next RCD can be initiated. Therefore, consecutive abnormal situations have no direct causal effect on each other. The third abnormal situation of each test results in an ESD of the TEP. Each of the previously selected RCDs is used in several tests with variations in their respective durations. These variations have an influence on the consecutive normalization phases and therefore ensure statistical variation in alarm behavior and temporal distances between abnormal situations. Altogether, a total of 300 specific abnormal situations, including 100 that cause an ESD, were conducted in 100 tests. Each test starts with a normal operation period of either 30 min or one hour. Fig. 4 shows the time trends of alarm variables for a typical test. The first and second abnormal situations are *IDV 2* and *IDV 1*, respectively, each with a 100% scaling. The third abnormal situation is *XMV 2* with a 100% scaling.

To facilitate understanding of the dataset presented here, some relevant characteristics are given in this paragraph. Overall, a total of 7343 alarms were activated, 4860 of which arose during ARC phases and 2483 during normalization phases. Fig. 10 gives a detailed view of the ACT count distribution by using a boxplot diagram. It shows values for the following phases and situations:

1) Abnormal situations without (w/o) an ESD.
2) ARC phases w/o an ESD.
3) Normalization phases.
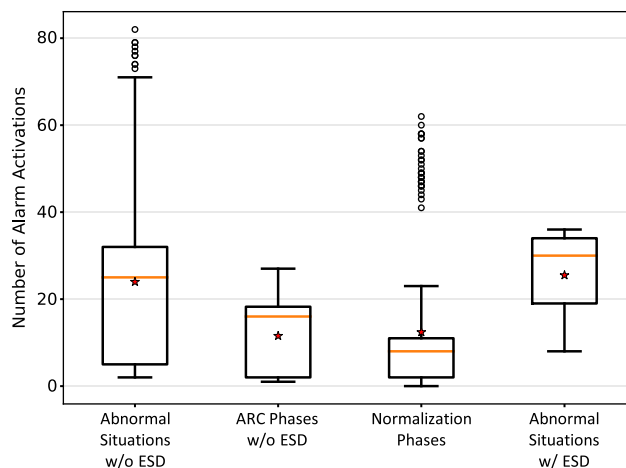4) Abnormal situations with (w/) an ESD.

FIGURE 10. Boxplot diagram of alarm activation counts for the TEP dataset.

Here, the first and third quartiles are represented by a box with the median illustrated by using a yellow line and the mean illustrated by using a red star. The whiskers are at most 1.5 times the length of the box, with any data point further away considered an outlier and marked by a black dot [80]. With an average test duration of 97.30 hours, each abnormal situation without an ESD includes an average of 23.97 ACTs. Each of the corresponding ARC phases features an average

of 11.55 alarms, and each normalization phase includes an average of 12.42 ACTs. In the case of an ESD, an average of 25.50 alarms are activated. The maximum and minimum numbers of ACTs are 36 and two for ARC phases and 62 and zero for normalization phases. The latter represents an immediate RTN with little to no impact on plantwide behavior.

Table 1 shows the availability of data sources for the dataset presented here. Each of the 100 tests includes four files, which contain the recorded PV and MV signals, the calculated alarm variables, and the corresponding time stamps. The sampling rate for the conducted simulation runs was set to 0.1 Hz (approximately 0.0028 h or 10 s). The alarm data are made available as a multivalued alarm series by using (1).

TABLE 1. Data availability.

Measured Variable	Sampling Rate	Availability	File Name
Process Variables (XMEAS)	0.1 Hz	Time Series	SIMOUT.xlsx
Manipulated Variables (XMV)	0.1 Hz	Time Series	XMV.xlsx
Time	0.1 Hz	Time Series	TOUT.xlsx
Alarm Variables	0.1 Hz	Multi-Valued Alarm Series	ALARMS.xlsx

V. EVALUATION AND PARAMETER OPTIMIZATION

This section evaluates and compares the performances and characteristics of the methods described in Section II in comparison with the ACEDM proposed in Section III. Subsection V.A deals with choosing a suitable evaluation measure, which will be used in Subsection V.B. All examined methods are applied to the TEP dataset that was presented in Section IV by using a comprehensive grid search for parameter tuning, thus gaining insights into the overall alarm subsequence detection performance.

A. EXTERNAL VALIDITY INDICES

For evaluation, a suitable measure needs to be chosen, which facilitates an appropriate performance comparison of different parameter settings and methods in terms of which detected subsequences best fit with the TEP dataset [64]. Therefore, validity indices are commonly used that provide insight into how well a method reflects the underlying reality [27]. Cluster validity indices are applicable to the methods analyzed here, as the mutual characteristic of grouping similar objects suggests. A variety of measures with different characteristics have been proposed in the scientific literature and can be categorized into internal and external cluster validity indices [64], [78]. The latter assess agreement between a ground-truth partition and a trial partition [21], [78]. Here, the ground-truth consists of the historical alarm subsequences, which truly represent the abnormal situations in the tests. These are available for the TEP dataset; thus, external validity indices are deemed applicable.

According to [49], [60], and [64], these indices can be classified into three subtypes:

1) Pair-counting indices.
2) Information theoretic indices.
3) Set-matching indices.

Pair-counting indices, which describe the agreement or disagreement of two partitions on the pairs of objects in the dataset [64], have already been applied to the alarm analysis domain. References [50] and [51] used the Jaccard similarity coefficient (introduced in [40]), whereas [3] and [25] used the corresponding distance variant of this index. References [3], [50], and [51] aimed to calculate the similarity between a pair of binary alarm series, whereas [25] used the Jaccard index for computing the similarity between a pair of timed-clusters. These approaches have in common that similarity is analyzed only for a pair of alarm subsequences, and they are thus different from comparing two partitions of a dataset; i.e., the comparison of two sequence sets for the desired purpose of performance evaluation. Both sets are generated from the same test, each including at least one alarm subsequence and containing alarm instances characterized by a unique identifier. In this way, it is possible to compare two partitions regarding their common alarm instances. Other frequently used pair-counting indices are the Rand index [63] and adjusted Rand index [33].

When used for the comparison of a ground-truth partition and a generated trial partition, pair-counting indices are mainly determined by the agreement and disagreement on larger clusters; hence, smaller clusters have only a limited impact on the overall index value [78]. This cluster size imbalance sensitivity has been analyzed in detail for different indices, e.g., in [28], [59], [60], and [62]. References [64] and [78] both point out that this behavior might be desirable for some applications, but it is subsequently assumed that in general, all clusters should be evaluated with the same relevance irrespective of their sizes. Cluster size imbalance can also be found throughout the TEP dataset in varying numbers of ACTs per abnormal situation. However, the number of alarms in an abnormal situation is not necessarily correlated with its severity; e.g., a single alarm can be critical and could possibly lead to a dangerous escalation, whereas numerous less critical alarms might be of relatively minor significance. Hence, the external validity index applied here should be invariant with respect to the size of the clusters. According to [71], this sensitivity to the size of clusters also applies to most information theoretic indices.

A set-matching validity index, which is insensitive to cluster size imbalances, was proposed in [64]. This PSI uses an overall similarity that is normalized separately for each cluster, thus assuring that the index is invariant with respect to the cluster sizes. It is based on the Braun-Blanquet formula (introduced in [12, p. 363] and among other similarity measures described and examined in [15] and [18]), which was used for alarm sequence similarity analysis in [66] and is used in the PSI to calculate the similarity between a pair of

clusters. To obtain an efficient and optimal pairing of clusters, the Hungarian algorithm is used (presented in [47]), which aims to maximize the overall similarity of two compared partitions under the restriction that clusters may only be matched once. The remaining unpaired clusters show a mismatch of the number of clusters between the two partitions [64]. The concomitant property of symmetric similarity, which is ensured by dividing the calculated overall similarity by the maximum number of clusters in either partition, needs to be adapted for the purpose of the evaluation conducted here. The following example illustrates this necessity: a ground-truth partition consisting of three subsequences is matched to three subsequences in a computed trial partition, resulting in a maximized but subpar overall similarity. Furthermore, three single alarms that are not part of the match are put into additional subsequences. In case *"A"*, these alarms are put into three individual subsequences, whereas in case *"B"*, all three alarms are part of one joint subsequence. In both cases, these unmatched alarms are equivalent regarding their negligible utility for further alarm analysis. The application of the original matching measure proposed in [64] would result in a relatively low PSI value for case ''A'', as the maximum number of subsequences adds up to six, whereas the overall similarity is only calculated using the three matched subsequences. The corresponding PSI value for case ''B'' would be significantly higher, although a similar situation would be described. This example highlights an undesired characteristic of the PSI for the purpose of externally validating the detection of historical alarm subsequences. Hence, it is proposed that the number of unmatched clusters should be of no relevance for the performance assessment of the detection methods examined here. Therefore, only the number of clusters in the ground-truth partition is used for the calculation of the aPSI. Other properties, such as normalization to the range 0 to 1, still apply. The aPSI for a test with a ground-truth partition X and a set of detected subsequences AS is calculated by using the following equation (adapted from [64]):

$$
\begin{aligned}
&aPSI(X, AS) \\
&= \begin{cases} (\sigma(X, AS) - 1)/(K_X - 1), & \sigma \geq 1, K_X > 1 \\ 0, & \sigma < 1 \\ 1, & K_X = K_{AS} = 1 \end{cases} \quad (28)
\end{aligned}
$$

with the overall similarity $\sigma(X, AS)$ between the two partitions that is to be maximized [64]:

$$
\sigma(X, AS) = \sum_i (\theta_{i,j}) \quad (29)
$$

and the Braun-Blanquet formula for the calculation of the similarity $\theta_{i,j}$ between two alarm subsequences X_i and AS_j [64]:

$$
\theta_{i,j} = (n_{i,j})/(\max(|X_i|, |AS_j|)) \quad (30)
$$

where K_X and K_{AS} are the number of subsequences in X and AS, respectively. For two alarm subsequences X_i and AS_j, $|X_i|$ and $|AS_j|$ represent their respective alarm counts, and n_{ij} denotes the number of shared alarm instances.

For the purpose of comparing the performance by using only a single index value per detection method and parameter setting, the average aPSI over all 100 tests of the dataset presented in Section IV is used.

B. PERFORMANCE EVALUATION AND PARAMETER OPTIMIZATION USING GRID SEARCH

This section evaluates the performance of the alarm subsequence detection methods described in Section II and proposed in Section III. In the first step, an adequate approach for the definition of suitable parameter settings is selected. These settings are then used to examine and compare the ρ-based approaches, namely, *"[AIC+13]-D1"*, *"[AIC+13]-D2"*, *"[RCH+16]"*, *"[LCG+18]-Add"*, and *"[LCG+18]-No Add"*, as well as the MAD-based approaches, namely, *"[FWV20]"* and the ACEDM. The optimal parameter settings determined hereby are eventually used to compare the performance over all methods. In addition, the ACEDM is compared to another version of it that does not use the coactivation constraint but only the process operators O1.1 to O1.4 (s. Fig. 6). Henceforth, this version is referred to as the *"alarm event detection method"* (AEDM). This evaluation approach allows for a systematic and in-depth examination of the effectiveness of the ACEDM and its components, namely, the proposed coactivation constraint and outlier detection by using the alarm event input.

Following the definitions given in [74], the targeted maximization of the average aPSI over all considered tests can be characterized as a mathematical optimization problem. A simple approach to solve this problem is to use a grid search, which is in the category of direct search methods. Here, a sequence of all possible parameter value combinations is constructed [11]. The optimal parameter setting is then considered to be the one with the highest average aPSI. To reduce the number of trials in search of the optimal parameter setting, a manual search can be conducted, thereby identifying promising intervals and suitable step sizes. Despite the existence of more efficient methods, e.g., random search [11] and simulated annealing [45], the combination of manual search and grid search is widely-used due to their beneficial properties, including simplicity, reliability when using only a low number of parameters, and ability to provide insights into the behavior of the analyzed methods [11].

For MAD-based approaches, the lower limit of τ^{MAD} is zero, which is the case for two alarm instances or events with a time distance equal to the median of d (s. (16) and (25)). The corresponding upper limit is set to 500 here, covering the recommendation of 400 given in [25]. Higher values are considered unpromising, as more than 90% of the TEP tests do not include any MAD^{dist} value greater than 500. The lower and upper limits are then used to generate an integer sequence of τ^{MAD} values. Furthermore, since no nuisance long-standing alarms occur in any of the TEP tests used, the coactivation threshold τ^c is set to one.

In the case of ρ-based methods, three parameters must be considered: τ^s, τ^e and ω. Regarding the latter, the lower

limit is defined manually by applying the 10-min window recommended in [3], [50], and [66]. Furthermore, the step size for ω is set to 10 min. The corresponding upper limit is set to 600 min, as preliminary tests have shown that values below 600 min are the most promising. The step size and lower limit for the ρ-thresholds are determined by ρ values that are limited to nonnegative integers and therefore set to one and zero, respectively. The upper limit is set to 20 alarms per window, as preliminary tests have shown that most tests do not include alarm rates above 20 alarms for any of the considered window lengths. In addition, the ρ-thresholds are merged into a single threshold pair parameter $\tau^{s,e} = [\tau^s, \tau^e]$ since τ^e is conditional upon τ^s; i.e., τ^e has to be smaller than τ^s. Table 2 shows all $\tau^{s,e}$ values used and their corresponding indices. The latter are used for the purpose of clarity in the subsequent diagrams.

TABLE 2. Mapping of alarm rate-threshold indices to alarm rate-threshold pairs.

Alarm Rate-Threshold Indices	Alarm Rate-Threshold Pairs $[\tau^s, \tau^e]$	Alarm Rate-Threshold Indices	Alarm Rate-Threshold Pairs $[\tau^s, \tau^e]$
1	[1, 0]	56 to 66	[11, 0] to [11, 10]
2, 3	[2, 0], [2, 1]	67 to 78	[12, 0] to [12, 11]
4 to 6	[3, 0] to [3, 2]	79 to 91	[13, 0] to [13, 12]
7 to 10	[4, 0] to [4, 3]	92 to 105	[14, 0] to [14, 13]
11 to 15	[5, 0] to [5, 4]	106 to 120	[15, 0] to [15, 14]
16 to 21	[6, 0] to [6, 5]	121 to 136	[16, 0] to [16, 15]
22 to 28	[7, 0] to [7, 6]	137 to 153	[17, 0] to [17, 16]
29 to 36	[8, 0] to [8, 7]	154 to 171	[18, 0] to [18, 17]
37 to 45	[9, 0] to [9, 8]	172 to 190	[19, 0] to [19, 18]
46 to 55	[10, 0] to [10, 9]	191 to 210	[20, 0] to [20, 19]

Fig. 11 depicts the average aPSI values for all five ρ-based approaches over the determined window lengths and all TEP tests. The respective data points are computed by determining

FIGURE 11. Performance of alarm rate-based approaches for optimal alarm rate-threshold parameter settings over all window lengths and tests.

the optimal $\tau^{s,e}$ for the specific window length and calculating the corresponding aPSI value. At first glance, it becomes apparent that all five methods show a similar minimum average aPSI for a ω of 10 min. Short window lengths consistently show low average aPSI values, although these values steeply increase in a range from 10 min to 80 min, with all five methods showing a similar rate of change. This demonstrates that the underlying propagation dynamics of abnormal situations in the TEP model analyzed here can only be reflected using greater window lengths; otherwise, the trial partitions show highly fragmented alarm subsequences, which rarely fit the given ground-truth. Moreover, the diagram shows that all five ρ-based approaches have rather similar performances up to a ω of 200 min, with fixed window methods performing identically and *"[LCG+18]-No Add"* consistently showing the lowest average aPSI values. In contrast, *"[LCG+18]-No Add"* achieves the highest average aPSI values for ω greater than 310 min compared to other ρ-based approaches. Furthermore, Fig. 11 shows that each depicted method has one distinct maximum average aPSI over all window lengths. The corresponding values can be obtained from the legend in Fig. 11. All three fixed window methods feature the same aPSI value of 0.794 for a ω of 180 min. The highest average aPSI for all ρ-based approaches is obtained by the sliding window method *"[LCG+18]-Add"* with a value of 0.809 and a ω of 280 min, whereas for *"[LCG+18]-No Add"*, the peak is not reached until a ω as great as 300 min. The graphs of all five methods show an increasing trend for values smaller than the respective maximum average aPSI and a decreasing trend for values greater than it. Another significant phenomenon revealed in Fig. 11 is that all three fixed window methods show a more variant trend for ω greater than 200 min compared to both sliding window methods. This is due to the high dependency of *"[AIC+13]-D1"*, *"[AIC+13]-D2"*, and *"[RCH+16]"* on the underlying structure of the calculated fixed windows, which can alter significantly in terms of the assigned alarms even when applying only small changes to the used window length. The optimal $\tau^{s,e}$ over all window lengths and for all methods were either [1, 0], [2, 0], or [2, 1]. Except for *"[LCG+18]-Add"*, which has an optimal $\tau^{s,e}$ of [2, 0] for ω in a range from 170 min to 190 min, all other methods have an optimal $\tau^{s,e}$ of [1, 0] for ω in a range from 10 min to 190 min. This makes the underlying concept of *"[AIC+13]-D2"* equivalent to that of *"[AIC+13]-D1"*. For ω greater than 270 min, all approaches adopt an optimal $\tau^{s,e}$ of [2, 1], except for *"[AIC+13]-D2"*, which has a threshold pair of [1, 0] for all window lengths. As illustrated in Fig. 11, this steady threshold value is accompanied by *"[AIC+13]-D2"*, presenting the lowest performance for greater window lengths.

An in-depth analysis of the resulting alarm subsequences reveals that additional fixed windows, as used in *"[RCH+16]"*, have little to no impact on the overall performance when compared to the straightforward method of *"[AIC+13]-D1"*. This demonstrates that the overlapping of subsequences is hardly an issue regarding
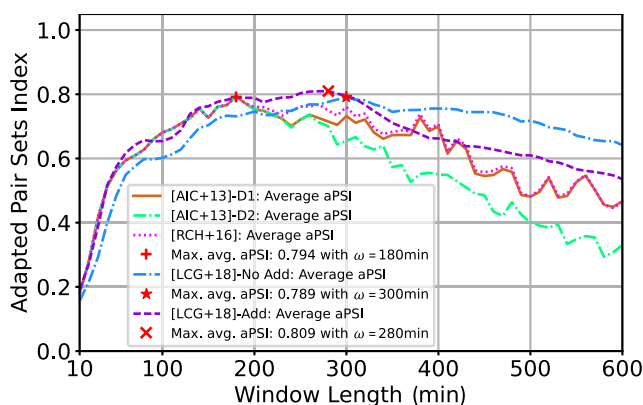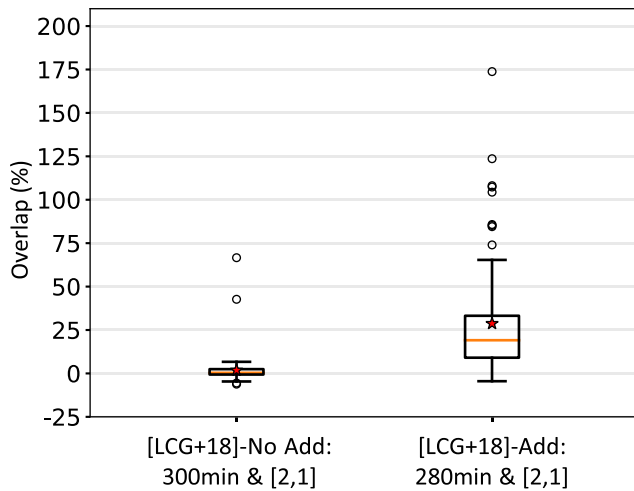
FIGURE 12. Boxplot diagram of alarm subsequence overlap per test for the optimal parameter settings of "[LCG+18]-No Add" and "[LCG+18]-Add".

"[RCH+16]", which holds true for the full range of considered window lengths, as indicated by the constant adjacency of "[AIC+13]-D1" and "[RCH+16]" in Fig. 11. However, with greater window lengths, overlapping increasingly becomes an issue for "[LCG+18]-Add", which negatively affects the similarity of the matched ground-truth and detected subsequences. Fig. 12 illustrates this overlapping phenomenon of detected alarm subsequences using optimal parameter settings for "[LCG+18]-No Add" and "[LCG+18]-Add". Each data point represents overlapping for an individual test and is generated by calculating the sum of alarm instances over all detected alarm subsequences, thereby also allowing for double counting and comparing it to the total number of distinct alarm instances in the particular test. For example, an overlap of 100% means that twice as many alarms are contained in the detected subsequences than alarm instances exist for this test. Negative overlapping values indicate that not all alarm instances are detected. It becomes apparent that even when using a smaller window length, "[LCG+18]-Add" shows an overall higher overlap, with an average of 26.8%, than "[LCG+18]-No Add", with an average of 1.0%.

Fig. 13 and Fig. 14 illustrate the determination process regarding the optimal $\tau^{s,e}$ per ω for each of the ρ-based approaches using two example window lengths, namely, 10 min and the respective optimal window length. The red dotted and light blue dash-dotted lines in each of the subfigures show the highest and lowest aPSI values, respectively, for any of the 100 tests. The dark blue solid line represents the average aPSI over all tests. The optimum $\tau^{s,e}$ and the corresponding maximum average aPSI value for this specific window length are marked according to the diagram legend.

Fig. 13 (a), (c), and (e) show the corresponding diagrams for a ω of 10 min for all three fixed window approaches. These methods only have average aPSI values greater than zero for ρ-threshold indices smaller than seven; i.e., all fixed

10-min windows in the dataset have an alarm rate smaller than four. Both sliding window approaches show a similar behavior in Fig. 14 (a) and (c); however, they show average aPSI values greater than zero up to $\tau^{s,e}$ with an index of 15.

Fig. 13 (b), (d), and (f) as well as Fig. 14 (b) and (d) show the corresponding diagrams for the optimal window lengths of all ρ-based methods. They reveal that for all the approaches, only smaller $\tau^{s,e}$ values perform well, as emphasized by the optimal threshold pair being [1, 0] for fixed window methods and [2, 1] for sliding window methods. The recommended $\tau^{s,e}$ values with $10 \leq \tau^s \leq 11$ (indices 46 to 66) consistently show relatively low aPSI values for all window lengths and approaches. Furthermore, "[AIC+13]-D1", "[RCH+16]", "[LCG+18]-Add", and "[LCG+18]-No Add" show a characteristic sawtooth behavior with local peaks for $\tau^e = 0$ or $\tau^e = 1$. Both of these phenomena trace back to a moderate number of tests that contain abnormal situations with only a small impact on plantwide ACTs. In contrast, "[AIC+13]-D2" shows a progression consisting of several plateaus, since only τ^e and ω determine the performance of this method, whereas τ^e is overruled by the second alarm flood definition given in [3]. Furthermore, both sliding window approaches show aPSI values greater than zero for higher ρ-threshold indices compared to fixed window approaches. This illustrates that abnormal situations with a relatively high number of ACTs can be detected more accurately by applying sliding windows instead of fixed windows.

This paragraph examines the performance of MAD-based approaches. Fig. 15 depicts the average aPSI values for all three MAD-based approaches over the determined τ^{MAD} values and all TEP tests. At first glance, it becomes apparent that "[FWV20]" and the AEDM show similar behavior, with the former approach performing marginally better for thresholds up to 14 and in a range from 32 to 37, whereas the latter presents higher average aPSI values for thresholds in a range from 15 to 31 and greater than 37. With an initial τ^{MAD} of zero, both approaches generate highly fragmented alarm subsequences, each consisting of a single alarm event, resulting in an average aPSI value of zero. An increase in the τ^{MAD} towards values of approximately 20 results in a substantial progression for both methods. Regarding their maximum performance, the AEDM has a slightly higher average aPSI of 0.764 when using an optimal threshold of 20 compared to 0.743 with a τ^{MAD} of 29 for "[FWV20]". Furthermore, Fig. 15 reveals that thresholds greater than the respective optimum values induce a change in direction in terms of a slow but persistent performance decline, which arises from a tendency towards longer and fewer alarm subsequences. Regarding the ACEDM, it becomes apparent that in contrast to the other two methods, it already shows an initially high performance with an average aPSI value of 0.923 using a τ^{MAD} of zero, which is equivalent to the exclusive application of the proposed alarm coactivation constraint. This is followed by a short increase up to a threshold of eight, where the ACEDM, now utilizing the MAD-based outlier
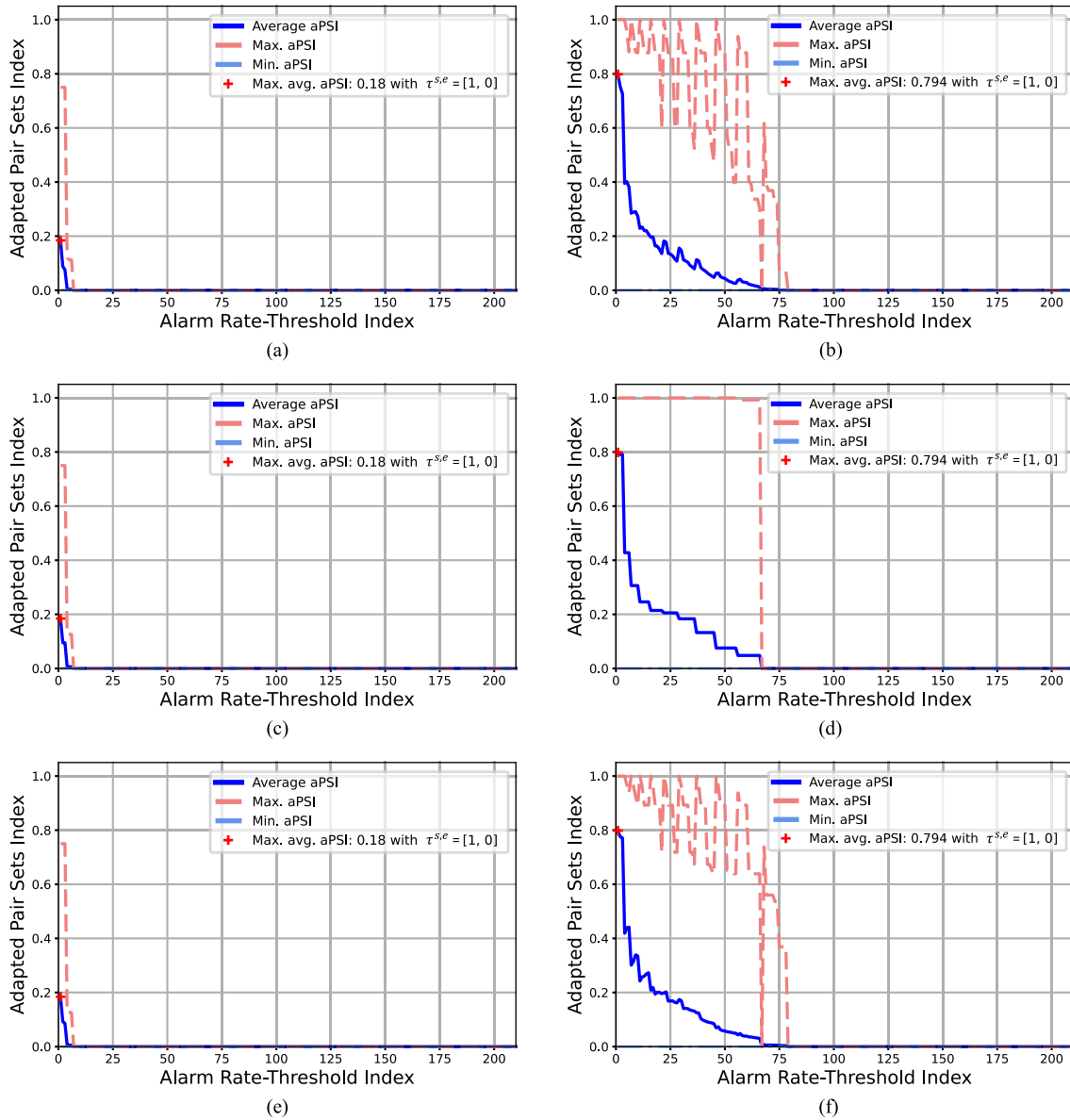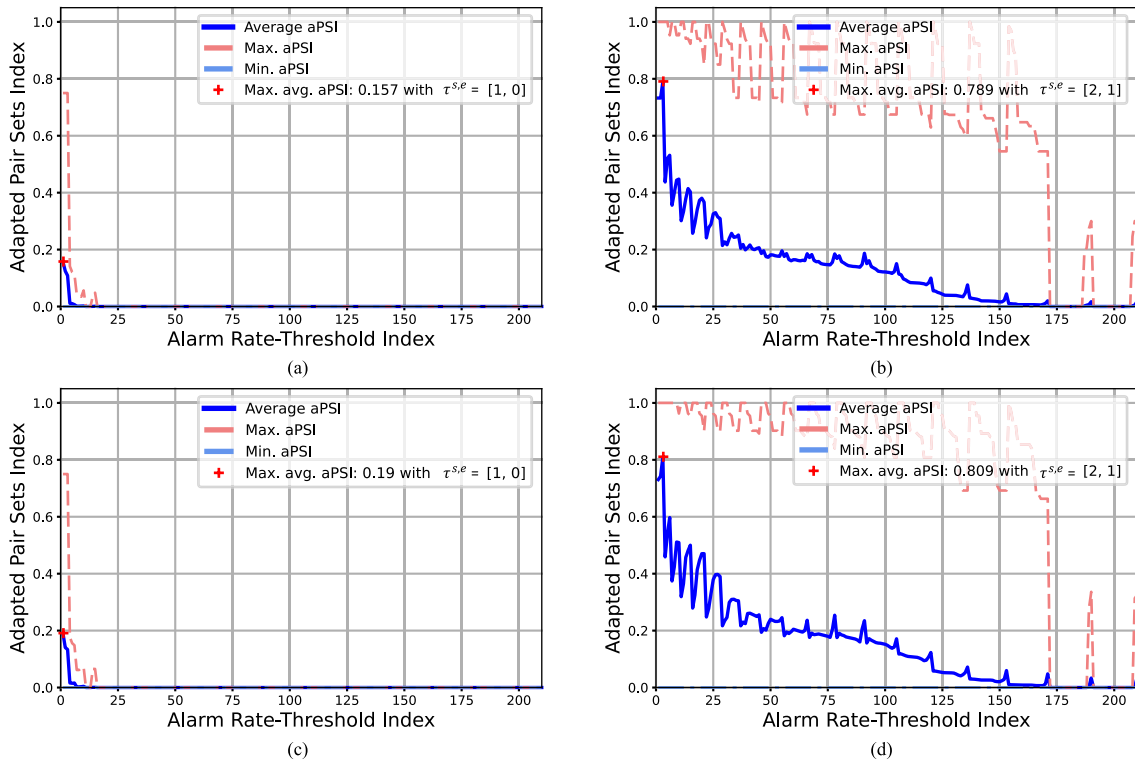
FIGURE 13. Maximum, minimum, and average performance of alarm rate-based approaches with fixed windows over all alarm rate-threshold pairs ($\tau^{s,e}$) and tests for a selection of window lengths. (a) "[AIC+13]-D1" with ω of 10 min. (b) "[AIC+13]-D1" with ω of 180 min. (c) "[AIC+13]-D2" with ω of 10 min. (d) "[AIC+13]-D2" with ω of 180 min. (e) "[RCH+16]" with ω of 10 min. (f) "[RCH+16]" with ω of 180 min.

detection method as well as the constraint, reaches its maximum average aPSI value of 0.957. This is also the highest average aPSI value over all eight alarm subsequence detection approaches. However, a further increase in τ^{MAD} results in a significant performance decrease, showing lower aPSI values for thresholds greater than 29 when compared to the other two MAD-based approaches. This negative sensitivity for higher τ^{MAD} has its seeds in an extended merging of spurious alarm subsequences by applying the proposed alarm coactivation constraint.

To compare and further assess the performance of all eight approaches, Fig. 16 shows boxplot diagrams that depict the

average aPSI values for each of the 100 tests by using the previously identified optimal parameter settings. One characteristic all methods have in common is that each of them has at least one test with an aPSI value of zero. *"[FWV20]"* and the three fixed window approaches have the highest number of these tests, with nine and eight tests, respectively, whereas *"[LCG+18]-Add"* and the ACEDM show the lowest number of tests with an aPSI of zero, with one and two tests, respectively. As the parameter optimization determines an identical setting for fixed window approaches, they are characterized by identical boxplots. Two additional groups of similar performing methods are the two solely MAD-based

FIGURE 14. Maximum, minimum, and average performance of alarm rate-based approaches with sliding windows over all alarm rate-threshold pairs ($\tau^{s,e}$) and tests for a selection of window lengths. (a) "[LCG+18]-No Add" with ω of 10 min. (b) "[LCG+18]-No Add" with ω of 300 min. (c) "[LCG+18]-Add" with ω of 10 min. (d) "[LCG+18]-Add" with ω of 280 min.

approaches and the two sliding window approaches. It can be seen that ρ-based approaches tend to show less variation, as indicated by the smaller interquartile range, i.e., the lengths of the respective boxes, compared to *"[FWV20]"* and the AEDM. Furthermore, Fig. 16 illustrates that the ACEDM, in addition to showing a high average performance, achieves a median aPSI of 0.997; i.e., 50% of the tests have a similar or better aPSI. As indicated by the median being at the upper edge of the box, the resulting aPSI distribution shows a significant positive skewness towards the value of 1.0. Only 13% of all tests show aPSI values below the average of 0.957.

An additional in-depth analysis of the generated alarm subsequences revealed significant differences between the approaches regarding the number of subsequences detected in each test. *"[FWV20]"* and the AEDM show a median of four and six subsequences, respectively. This difference, although there is similar performance overall, can be explained by the fact that most of the additional subsequences generated by the AEDM contain only RTNs. All three fixed-window approaches and *"[LCG+18]-Add"* show a median of five detected alarm subsequences, whereas *"[LCG+18]-No Add"* has a median of four. Only the ACEDM shows a median of three and an average of approximately 3.5 subsequences, which resembles the ground-truth most accurately. Analyzing the alarm subsequences generated by applying the ACEDM, it becomes apparent that the excessive subsequences mostly contain only a single alarm, which can be characterized
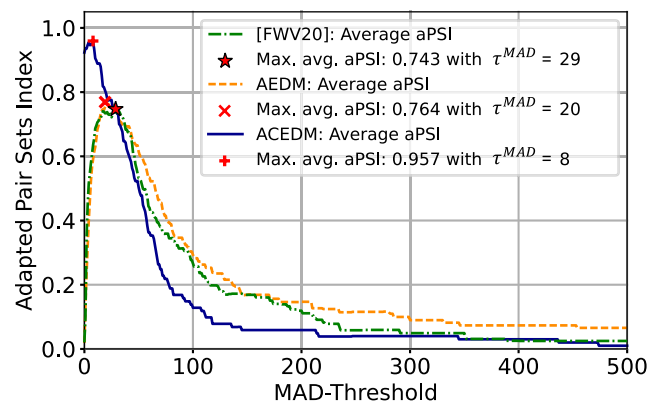
FIGURE 15. Performance of MAD-based approaches over all τ^{MAD} and tests.

as being the final alarm in a particular normalization phase.

Fig. 17 shows a heatmap diagram of the performances of all examined approaches using optimal parameter settings over the top eleven tests with the lowest average aPSI over all eight methods. Tests no. 61, 66, 78, and 98 are analyzed in detail in the following, as they allow us to gain insight into the characteristic behavior of the evaluated detection methods. Fig. 17 reveals that neither ρ-based approach can correctly detect the three abnormal situations in test no. 61. Moreover,
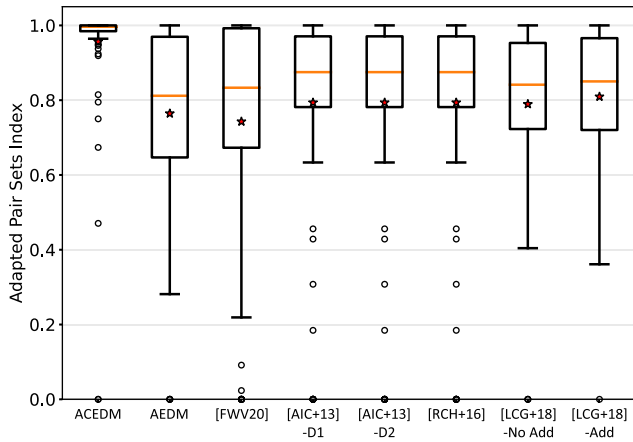
FIGURE 16. Boxplot diagram of the performance of MAD-based and alarm rate-based approaches using optimal parameter settings over all tests.
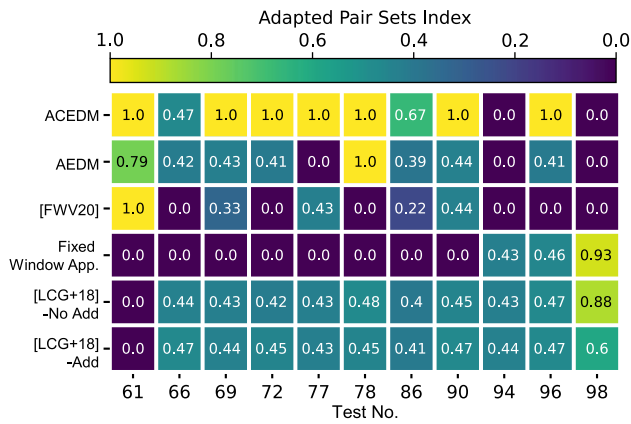

FIGURE 17. Heatmap diagram of the performances of MAD-based and alarm rate-based approaches (app.) using optimal parameter settings over the top eleven tests with the lowest average aPSI over all methods. The color bar at the top indicates the mapping of the colors used and the corresponding aPSI values.

this is the only test in which *"[LCG+18]-Add"* shows an aPSI of zero, with an overlap of 81.19%. Due to a low overall duration of 13.575 hours and the distances between individual situations being rather short, only *"[FWV20]"* and the proposed ACEDM are able to perfectly identify all three abnormal situations in test no. 61. Test no. 66 presents another challenging detection task. It includes a combination of short distances between the normalization phase of one abnormal situation and the initial ACTs of the next one, i.e., merely 0.5 hours, as well as exceedingly slowly propagating disturbances and normalizations. Except for *"[FWV20]"* and the fixed window approaches, with an aPSI of zero, all other methods show a similar performance on this test, with an aPSI in the range from 0.42 to 0.47. Despite their common fundamentals, *"[FWV20]"* is not able to properly detect the abnormal situations in test no. 78, whereas the AEDM shows an aPSI value of 1.0 for the same test. This is mainly due to the additional consideration of alarm events, thus emphasizing gaps between individual abnormal situations to a greater

extent. This also holds true for similar tests with quickly propagating RCDs and slowly propagating normalizations. On the other hand, the AEDM has significantly lower performance in tests where the final RTNs of one situation and the first ACTs of another situation are too close to each other, as in test 77. For the ACEDM, an application of the alarm coactivation constraint allows for a smaller τ^{MAD} to be used, substantially reducing this undesired effect. Two tests, namely, tests no. 94 and 98, show the lowest aPSI value of zero for the ACEDM. Although test no. 94 remains a challenge for other methods, fixed window approaches and *"[LCG+18]-No Add"* show better performance regarding test no. 98, with aPSI values of 0.926 and 0.879, respectively. This is mainly due to two characteristics of this test. First, there is exactly one fixed window, coincidently lying between the second and third abnormal situations, which does not include any ACTs, although it shows several RTNs. Second, both the second and third abnormal situations show a duration of approximately ten hours, which is twice the size of the optimal window length of *"[LCG+18]-No Add"*. Moreover, the distance between the final RTN of the second abnormal situation and the first ACT of the third abnormal situation is only 27.5 min, which negatively affects the performance of the ACEDM; i.e., the overall similarity between the ground-truth partition and the trial partition is smaller than 1. However, the fixed window approaches and *"[LCG+18]-No Add"* were not able to correctly detect the normalization period of the first abnormal situation, which was achieved by the ACEDM.

VI. DISCUSSION AND CONCLUSION

Subsection V.B showed that existing ρ-based and MAD-based approaches are not able to meet the requirements described in Section II to the fullest extent. The examination of fixed and sliding window approaches showed that they achieve relatively good performance results despite their rigid time windows. However, both types of approaches depend closely on the selected parameter settings, thus necessitating the cumbersome tuning of three highly interrelated parameters. It was further revealed that the proposed setting of 10 alarms in a 10-min window does not lead to valuable results when applied to the TEP dataset used here. In addition, some ρ-based approaches show a strong tendency towards overlapping; this may or may not be desired depending on the application and has to be taken into account when selecting a suitable method. The MAD-based approach proposed in [25] shows good performance results in tests where abnormal situations occur with longer time intervals. However, this method shows distinct limitations in the case of slowly propagating RCDs and normalization phases. It was further demonstrated that the proposed ACEDM is able to fulfil all given requirements and shows the best performance of all considered alarm subsequence detection methods. Its limitations result from more extensive data preprocessing (elimination of erroneous long-standing alarms), the necessity of tuning both the τ^{MAD} and the τ^c parameter, and a remaining but reduced negative

sensitivity to abnormal situations with short distances in time. Regarding the latter, a comprehensive assessment showed that the ACEDM allows a smaller τ^{MAD} to be used and is thus able to distinguish independent situations with smaller time distances than the method proposed in [25]. The individual components of the ACEDM were also examined, and it was demonstrated that each of them facilitated an improvement in the detection performance. In this context, it was proven that the ACEDM can also be used effectively without tuning the parameters, that is, when only the proposed alarm coactivation constraint is used. Nevertheless, the best performance was achieved when applying the ACEDM as intended, thus affirming the assumption that there is a potential advantage when using additional input information and process knowledge.

In future work, additional plant and process information can be included to improve the detection of relevant alarm subsequences in larger and more complex plants. According to [24] and [43], it is likely that purely data-driven methods wrongly group alarms that stem from separate and causally independent subprocesses and plant sections into the same alarm subsequence if they are close in time. Engineering documents, such as P&IDs and device documentation, as well as sensor data can be useful for the determination of the relevant plant and process hierarchy. This hierarchy information can then be used to preprocess the historical alarm data by splitting it into subsets that represent the hierarchy of subprocesses and plant sections.

This paper also introduced a novel and openly accessible alarm management dataset based on a TEP simulation model. In future work, the 100 tests and 300 individual abnormal situations included here can be further analyzed and used for the application of different alarm management techniques and methods. Furthermore, the proposed alarm subsequence detection approach can be implemented as a primary step for advanced alarm analysis methods, such as alarm flood root-cause analysis, the identification of recurring abnormal situations, and the prediction of upcoming alarm events.

REFERENCES

[1] L. Abele, M. Anic, T. Gutmann, J. Folmer, M. Kleinsteuber, and B. Vogel-Heuser, "Combining knowledge modeling and machine learning for alarm root cause analysis," *IFAC Proc. Volumes*, vol. 46, no. 9, pp. 1843–1848, 2013, doi: 10.3182/20130619-3-RU-3018.00057.

[2] N. A. Adnan, "Performance assessment and systematic design of industrial alarm systems," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Alberta, Edmonton, AB, Canada, 2013. Accessed: Dec. 4, 2020. [Online]. Available: https://era.library.ualberta.ca/items/2c7a4695-d97b-4f13-ab9e-6307bc12f7a0

[3] K. Ahmed, I. Izadi, T. Chen, D. Joe, and T. Burton, "Similarity analysis of industrial alarm flood data," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 2, pp. 452–457, Apr. 2013, doi: 10.1109/TASE.2012.2230627.

[4] *Batch Control Part 1: Models and Terminology*, Standard 88.01-1995, ANSI/ISA, 1995.

[5] *Enterprise-Control System Integration—Part 1: Models and Terminology*, Standard 95.00.01-2010, ANSI/ISA, 2010.

[6] *Management of Alarm Systems for the Process Industries*, Standard 18.2-2016, ANSI/ISA, 2016.

[7] E. Arroyo, "Capturing and exploiting plant topology and process information as a basis to support engineering and operational activities in process plants," Ph.D. dissertation, Inst. Automobile Technol., Helmut Schmidt Univ., Hamburg, Germany, 2017. Accessed: Dec. 4, 2020. [Online]. Available: https://edoc.sub.uni-hamburg.de/hsu/volltexte/2017/3162/pdf/PhDDissertation_Arroyo_Final.pdf

[8] A. J. R. Consortium, J. Errington, D. V. Reising, and C. Burns, "Appendix 3: Preventing and responding to abnormal situations," in *ASM Consortium Guidelines—Effective Alarm Management Practices*, Phoenix, AZ, USA: ASM Consortium, 2009, ch. A3, pp. 84–90.

[9] A. Bathelt, N. L. Ricker, and M. Jelali, "Revision of the Tennessee Eastman process model," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 309–314, Jun. 2015, doi: 10.1016/j.ifacol.2015.08.199.

[10] M. Bauer, J. W. Cox, M. H. Caveness, J. J. Downs, and N. F. Thornhill, "Finding the direction of disturbance propagation in a chemical process using transfer entropy," *IEEE Trans. Control Syst. Technol.*, vol. 15, no. 1, pp. 12–21, Jan. 2007, doi: 10.1109/TCST.2006.883234.

[11] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. 10, pp. 281–305, Feb. 2012, doi: 10.5555/2503308.2188395.

[12] J. Braun-Blanquet, H. S. Conard, and G. D. Fuller, Eds, *Plant Sociology; the Study of Plant Communities*. New York, NY, USA: McGraw-Hill, 1932.

[13] D. C. Brown and M. O'Donnell, "Too much of a good thing? Alarm management experience in BP oil. Part 1: Generic problems with DCS alarm systems," in *Proc. IEE Colloq. Stemming Alarm Flood*, London, U.K., Jun. 1997, doi: 10.1049/ic:19970743.

[14] S. Charbonnier, N. Bouchair, and P. Gayet, "Fault template extraction to assist operators during industrial alarm floods," *Eng. Appl. Artif. Intell.*, vol. 50, pp. 32–44, Apr. 2016, doi: 10.1016/j.engappai.2015.12.007.

[15] A. H. Cheetham and J. E. Hazel, "Binary (presence-absence) similarity coefficients," *J. Paleontol.*, vol. 43, no. 5, pp. 1130–1136, Sep. 1969.

[16] Y. Cheng, I. Izadi, and T. Chen, "Pattern matching of alarm flood sequences by a modified Smith–Waterman algorithm," *Chem. Eng. Res. Des.*, vol. 91, no. 6, pp. 1085–1094, Jun. 2013, doi: 10.1016/j.cherd.2012.11.001.

[17] M. Chioua, M. Hollender, and V. Rodrigo. (Dec. 2019). *Systematic Alarm Flood Reduction*. Accessed: Jan. 7, 2021. [Online]. Available: https://zenodo.org/record/3530676/files/Chioua%20Hollender.pdf?download=1, doi: 10.5281/zenodo.3530676.

[18] S.-S. Choi, S.-H. Cha, and C. C. Tappert, "A survey of binary similarity and distance measures," *J. Systemics, Cybern. Inform.*, vol. 8, no. 1, pp. 43–48, 2010.

[19] L. Christiansen, A. Fay, B. Opgenoorth, and J. Neidig, "Improved diagnosis by combining structural and process knowledge," in *Proc. Int. Conf. Emerg. Technol. Factory Automat.*, Toulouse, France, Sep. 2011, pp. 1–8, doi: 10.1109/ETFA.2011.6059056.

[20] *Management of Alarms Systems for the Process Industries (IEC 62682:2014); German Version EN 62682:2015*, DIN EN Standard 62682:2016-02, 2016.

[21] B. E. Dom, "An information-theoretic external cluster-validity measure," in *Proc. 18th Conf. Uncertainty Artif. Intell.*, Edmonton, AB, Canada, Aug. 2002, pp. 137–145, doi: 10.5555/2073876.2073893.

[22] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Comput. Chem. Eng.*, vol. 17, no. 3, pp. 245–255, Mar. 1993, doi: 10.1016/0098-1354(93)80018-I.

[23] *Alarm Systems—A Guide to Design, Management and Procurement*, EEMUA, London, U.K., 2013.

[24] F. El Sakka, H. Bloch, J. Kinghorst, M. F. Pirehgalin, A. Fay, and B. Vogel-Heuser, "Improved alarm flood analysis by cluster identification and alarm assignment," *Automatisierungstechnik*, vol. 66, no. 8, pp. 647–655, Aug. 2018, doi: 10.1515/auto-2018-0071.

[25] M. Fahimipirehgalin, I. Weiss, and B. Vogel-Heuser, "Causal inference in industrial alarm data by timely clustered alarms and transfer entropy," in *Proc. Eur. Control Conf. (ECC)*, Saint Petersburg, Russia, May 2020, pp. 2056–2061, doi: 10.23919/ECC51009.2020.9143823.

[26] J. Folmer and B. Vogel-Heuser, "Computing dependent industrial alarms for alarm flood reduction," in *Proc. Int. Multi-Conf. Syst., Sygnals Devices*, Chemnitz, Germany, Mar. 2012, pp. 1–6, doi: 10.1109/SSD.2012.6198008.

[27] V. G. D. Fonseca and N. R. J. Fieller, "Distortion in statistical inference: The distinction between data contamination and model deviation," *Metrika*, vol. 63, no. 2, pp. 169–190, Apr. 2006, doi: 10.1007/s00184-005-0010-2.

[28] P. Fränti, M. Rezaei, and Q. Zhao, "Centroid index: Cluster level similarity measure," *Pattern Recognit.*, vol. 47, no. 9, pp. 3034–3045, Sep. 2014, doi: 10.1016/j.patcog.2014.03.017.

[29] P. Goel, A. Datta, and M. S. Mannan, "Industrial alarm systems: Challenges and opportunities," *J. Loss Prevention Process Industries*, vol. 50, pp. 23–36, Nov. 2017, doi: 10.1016/j.jlp.2017.09.001.

[30] (Jul. 1994). *The Explosion and Fires at the Texaco Refinery*. Milford Haven, Health Safety Executive, Suffolk, U.K., 1997. Accessed: Dec. 4, 2020. [Online]. Available: https://www.jesip.org.uk/uploads/media/incident_reports_and_inquiries/Texaco%20Refinery%20Explosion.pdf

[31] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, Mar. 2004, doi: 10.5555/2017212.2017217.

[32] B. Hollifield and E. Habibi *Alarm Management Handbook—A Comprehensive Guide*, 2nd ed. Research Triangle Park, NC, USA: ISA, 2011.

[33] L. Hubert and P. Arabie, "Comparing partitions," *J. Classification*, vol. 2, no. 1, pp. 193–218, Dec. 1985, doi: 10.1007/BF01908075.

[34] W. Hu, J. Wang, and T. Chen, "A local alignment approach to similarity analysis of industrial alarm flood sequences," *Control Eng. Pract.*, vol. 55, pp. 13–25, Oct. 2016, doi: 10.1016/j.conengprac.2016.05.021.

[35] W. Hu, T. Chen, S. L. Shah, and M. Hollender, "Cause and effect analysis for decision support in alarm floods," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 13940–13945, Jul. 2017, doi: 10.1016/j.ifacol.2017.08.2215.

[36] W. Hu, S. L. Shah, and T. Chen, "Framework for a smart data analytics platform towards process monitoring and alarm management," *Comput. Chem. Eng.*, vol. 114, pp. 225–244, Jun. 2018, doi: 10.1016/j.compchemeng.2017.10.010.

[37] *Management of Alarm Systems for the Process Industries*, IEC Standard 62682, 2014.

[38] I. Izadi, S. L. Shah, D. S. Shook, S. R. Kondaveeti, and T. Chen, "A framework for optimal design of alarm systems," *IFAC Proc. Volumes*, vol. 42, no. 8, pp. 651–656, Jun. 2009, doi: 10.3182/20090630-4-ES-2003.00108.

[39] I. Izadi, S. L. Shah, D. S. Shook, and T. Chen, "An introduction to alarm analysis and design," *IFAC Proc. Volumes*, vol. 42, no. 8, pp. 645–650, Jun. 2009, doi: 10.3182/20090630-4-ES-2003.00107.

[40] P. Jaccard, "The distribution of the flora in the alpine zone," *New Phytol.*, vol. 11, no. 2, pp. 37–50, Feb. 1912, doi: 10.1111/j.1469-8137.1912.tb05611.x.

[41] S. A. Jacobs and A. Dagnino, "Large-scale industrial alarm reduction and critical events mining using graph analytics on Spark," in *Proc. 2nd Int. Conf. Big Data Comput. Service Appl. (BigDataService)*, Oxford, U.K., Mar. Apr. 2016, pp. 66–71, doi: 10.1109/BigDataService.2016.21.

[42] A. H. Kargaran, A. Neshastegaran, I. Izadi, and E. Yazdian, "Analytical derivation and comparison of alarm similarity analysis methods," 2020, *arXiv:2003.10600*. Accessed: Jan. 6, 2021. [Online]. Available: http://arxiv.org/abs/2003.10600

[43] J. Kinghorst, H. Bloch, A. Fay, and B. Vogel-Heuser, "Integration of additional information sources for improved alarm flood detection," in *Proc. IEEE 21st Int. Conf. Intell. Eng. Syst.*, Larnaca, Cyprus, Oct. 2017, pp. 19–26, doi: 10.1109/INES.2017.8118568.

[44] D. Kirchhübel, "Combining functional modelling and reasoning with online event analytics," Ph.D. dissertation, Dept. Elect. Eng. Automat. Control, Tech. Univ. Denmark, Lyngby, Denmark, 2019.

[45] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, May 1983, doi: 10.1126/science.220.4598.671.

[46] S. R. Kondaveeti, "Advanced analysis and redesign of industrial alarm systems," Ph.D. dissertation, Dept. Chem. Mater. Eng., Univ. Alberta, Edmonton, AB, Canada, 2013. Accessed: Dec. 4, 2020. [Online]. Available: https://era.library.ualberta.ca/items/b8d5ec6e-f7c1-453e-9f52-62e10069b0f4

[47] H. W. Kuhn, "The Hungarian method for the assignment problem," in *50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art*, M. Jünger, Ed. Berlin, Germany: Springer, 2010, pp. 29–47.

[48] S. Lai and T. Chen, "Methodology and application of pattern mining in multiple alarm flood sequences," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 657–662, Jun. 2015, doi: 10.1016/j.ifacol.2015.09.043.

[49] Y. Lei, J. C. Bezdek, J. Chan, N. X. Vinh, S. Romano, and J. Bailey, "Extending information-theoretic validity indices for fuzzy clustering," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 4, pp. 1013–1018, Aug. 2017, doi: 10.1109/TFUZZ.2016.2584644.

[50] M. Lucke, M. Chioua, C. Grimholt, M. Hollender, and N. F. Thornhill, "On improving fault detection and diagnosis using alarm-range normalisation," *IFAC-PapersOnLine*, vol. 51, no. 24, pp. 1227–1232, Aug. 2018, doi: 10.1016/j.ifacol.2018.09.695.

[51] M. Lucke, M. Chioua, C. Grimholt, M. Hollender, and N. F. Thornhill, "Advances in alarm data analysis with a practical application to online alarm flood classification," *J. Process Control*, vol. 79, pp. 56–71, Jul. 2019, doi: 10.1016/j.jprocont.2019.04.010.

[52] W. L. Luyben, "Simple regulatory control of the eastman process," *Ind. Eng. Chem. Res.*, vol. 35, no. 10, pp. 3280–3289, Jan. 1996, doi: 10.1021/ie960162x.

[53] G. Manca, "'Tennessee-eastman-process' alarm management dataset," *IEEE DataPort*, Nov. 2020, doi: 10.21227/326k-qr90.

[54] Z. Mannani, I. Izadi, and N. Ghadiri, "Preprocessing of alarm data for data mining," *Ind. Eng. Chem. Res.*, vol. 58, no. 26, pp. 11261–11274, Feb. 2019, doi: 10.1021/acs.iecr.8b05955.

[55] C. Martin-Villalba, A. Urquia, and G. Shao, "Implementations of the Tennessee eastman process in modelica," *IFAC-PapersOnLine*, vol. 51, no. 2, pp. 619–624, Feb. 2018, doi: 10.1016/j.ifacol.2018.03.105.

[56] X. Ma and D. Li, "A hybrid fault diagnosis method based on fuzzy signed directed graph and neighborhood rough set," in *Proc. 6th Data Driven Control Learn. Syst. (DDCLS)*, Chongqing, China, May 2017, pp. 253–258, doi: 10.1109/DDCLS.2017.8068078.

[57] T. J. McAvoy and N. Ye, "Base control for the Tennessee Eastman problem," *Comput. Chem. Eng.*, vol. 18, no. 5, pp. 383–413, May 1994, doi: 10.1016/0098-1354(94)88019-0.

[58] *Alarm Management*, NAMUR, Leverkusen, Germany, 2008.

[59] X. V. Nguyen, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Is a correction for chance necessary?" in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, Montreal, QC, Canada, Jun. 2009, pp. 1073–1080, doi: 10.1145/1553374.1553511.

[60] X. V. Nguyen, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *J. Mach. Learn. Res.*, vol. 11, no. 95, pp. 2837–2854, Oct. 2010, doi: 10.5555/1756006.1953024.

[61] T.P. Omdahl, *Reliability, Availability and Maintainability Dictionary*. Milwaukee, WI, USA: ASQC Quality Press, 1988.

[62] D. Pfitzner, R. Leibbrandt, and D. Powers, "Characterization and evaluation of similarity measures for pairs of clusterings," *Knowl. Inf. Syst.*, vol. 19, no. 3, pp. 361–394, Jun. 2009, doi: 10.1007/s10115-008-0150-6.

[63] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *J. Amer. Stat. Assoc.*, vol. 66, no. 336, pp. 846–850, Dec. 1971, doi: 10.2307/2284239.

[64] M. Rezaei and P. Franti, "Set matching measures for external cluster validity," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 8, pp. 2173–2186, Aug. 2016, doi: 10.1109/TKDE.2016.2551240.

[65] N. L. Ricker, "Decentralized control of the Tennessee Eastman challenge process," *J. Process Control*, vol. 6, no. 4, pp. 205–221, Aug. 1996, doi: 10.1016/0959-1524(96)00031-5.

[66] V. Rodrigo, M. Chioua, T. Hagglund, and M. Hollender, "Causal analysis for alarm flood reduction," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 723–728, Jun. 2016, doi: 10.1016/j.ifacol.2016.07.269.

[67] J. J. Rooney and L. N. V. Heuvel, "Root cause analysis for beginners," *Qual. Prog.*, vol. 37, no. 7, pp. 45–46, Jul. 2004.

[68] L. T. Rose and K. W. Fischer, "Garbage in, garbage out: Having useful data is everything," *Meas., Interdiscipl. Res. Perspective*, vol. 9, no. 4, pp. 222–226, Oct. 2011, doi: 10.1080/15366367.2011.632338.

[69] D. H. Rothenberg, *Alarm Management for Process Control: A Best-Practice Guide for Design, Implementation, and Use of Industrial Alarm Systems*. 1st ed. New York, NY, USA: Momentum Press, 2009.

[70] J. Thambirajah, L. Benabbas, M. Bauer, and N. F. Thornhill, "Cause-and-effect analysis in chemical processes utilizing XML, plant connectivity and quantitative process history," *Comput. Chem. Eng.*, vol. 33, no. 2, pp. 503–512, Feb. 2009, doi: 10.1016/j.compchemeng.2008.10.002.

[71] H. van der Hoef and M. J. Warrens, "Understanding information theoretic measures for comparing clusterings," *Behaviormetrika*, vol. 46, no. 2, pp. 353–370, Oct. 2019, doi: 10.1007/s41237-018-0075-7.

[72] *Formalised Process Descriptions—Concept and Graphic Representation*, VDI/VDE, Berlin, Germany, 2015.

[73] B. Vogel-Heuser, D. Schütz, and J. Folmer, "Criteria-based alarm flood pattern recognition using historical data from automated production systems (APS)," *Mechatronics*, vol. 31, pp. 89–100, Oct. 2015, doi: 10.1016/j.mechatronics.2015.02.004.

[74] G. R. Walsh, "Search methods for unconstrained optimization," in *Methods of Optimization*, London, U.K.: Wiley, 1975, ch. 3, pp. 74–75.

[75] J. Wang, F. Yang, T. Chen, and S. L. Shah, "An overview of industrial alarm systems: Main causes for alarm overloading, research status, and open problems," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 1045–1061, Apr. 2016, doi: 10.1109/TASE.2015.2464234.

[76] J. Wang and T. Chen, "Main causes of long-standing alarms and their removal by dynamic state-based alarm systems," *J. Loss Prevention Process Industries*, vol. 43, pp. 106–119, Sep. 2016, doi: 10.1016/j.jlp.2016.05.006.

[77] J. Wang, Y. Zhao, and Z. Bi, "Criteria and algorithms for online and offline detections of industrial alarm floods," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 5, pp. 1722–1731, Sep. 2018, doi: 10.1109/TCST.2017.2723578.

[78] M. J. Warrens and H. van der Hoef, "Understanding partition comparison indices based on counting object pairs," 2019, *arXiv:1901.01777*. Accessed: Jan. 6, 2021. [Online]. Available: http://arxiv.org/abs/1901.01777

[79] I. Weiss, J. Kinghorst, T. Kroger, M. F. Pirehgalin, and B. Vogel-Heuser, "Alarm flood analysis by hierarchical clustering of the probabilistic dependency between alarms," in *Proc. IEEE 16th Int. Conf. Ind. Informat. (INDIN)*, Porto, Portugal, Jul. 2018, pp. 227–232, doi: 10.1109/INDIN.2018.8471973.

[80] H. Wickham and L. Stryjewski. *40 Years of Boxplots*. Accessed: Dec. 4, 2020. [Online]. Available: http://byrneslab.net/classes/biol607/readings/wickham_boxplots.pdf

[81] Y. Xu, J. Wang, and Y. Yu, "Alarm event prediction from historical alarm flood sequences based on Bayesian estimators," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 2, pp. 1070–1075, Apr. 2020, doi: 10.1109/TASE.2019.2935629.

[82] F. Yang, P. Duan, S. L. Shah. and T. Chen, *Capturing Connectivity and Causality in Complex Industrial Processes*. Cham, Switzerland: Springer, 2014.

[83] F. Yang and C. Guo, "Survey on advanced alarm strategies based on multivariate analysis," in *Proc. 6th Int. Symp. Adv. Control Ind. Processes (AdCONIP)*, Taipei, Taiwan, May 2017, pp. 612–617, doi: 10.1109/ADCONIP.2017.7983850.

GIANLUCA MANCA was born in Buchholz in der Nordheide, Germany, in 1989. He received the B.Sc. degree in mechanical engineering and the M.Sc. degree in mechatronics from Helmut Schmidt University, Hamburg, Germany, in 2013 and 2014, respectively, and the B.Sc. degree in information systems from the University of Hagen, Germany, in 2019. He is currently pursuing the Ph.D. degree in engineering with Helmut Schmidt University.

From 2016 to 2018, he was a Platoon Leader with the 6th Reconnaissance Battalion of the German Federal Armed Forces. Since 2018, he has been an Officer of Operations Research and Simulation with the Army Warfighting Simulation Centre in Wildflecken, Germany. His research interests include the analysis and reduction of alarm floods in industrial alarm systems as well as the application of data mining and machine learning methods to industrial processes.

ALEXANDER FAY (Senior Member, IEEE) was born in 1970. He received the Diploma and the Ph.D. degree in electrical engineering from the Technical University of Braunschweig, Germany, in 1995 and 1999, respectively.

He had worked five years at the ABB Corporate Research laboratories in Heidelberg and Ladenburg. Before, he was appointed as a Full Professor with the Institute of Automation Technology, Helmut Schmidt University, Hamburg, Germany, in 2004. His research interests include models and methods for the engineering of large automated systems, especially in the process and manufacturing industries, in buildings and transportation systems. He has served as a member for the AdCom of the IEEE Industrial Electronics Society. He was a Program Co-Chair of the IEEE International Conference on Automation Science and Engineering (CASE) in 2018. From 2011 to 2017, he has also served as an Associate Editor for the IEEE Transactions on Industrial Informatics.

● ● ●