

Applications

Birgit Vogel-Heuser, Alexander Fay, Bernhard Rupprecht*, Franz C. Kunze, Victoria Hankemeier, Tom Westermann and Gianluca Manca

Exploring challenges of alarm root-cause analysis across varying production process types

Herausforderungen der Alarm Root-Cause Analyse in verschiedenen Arten von Produktionsprozessen

<https://doi.org/10.1515/auto-2023-0180>

Received September 27, 2023; accepted November 2, 2023

Abstract: Alarm management systems in the process industry use root-cause analysis methods to reduce alarm logs. To enable the application of these methods in different plant types, the alarm characteristics of a continuous, two discrete, and a hybrid plant are examined. The main contribution is threefold. First, root-cause analysis requirements, posed by different plant types, are revealed. Next, existing approaches are assessed against the requirements. Since the root-cause is not necessarily the first alarm in time, its justification requires further plant knowledge. Thus, engineering documents and the necessary formalized knowledge to justify root-causes are identified.

***Corresponding author: Bernhard Rupprecht**, Institute of Automation and Information Systems, Department of Mechanical Engineering, TUM School of Engineering and Design, Technical University of Munich, Boltzmannstr. 15, 85748 Garching, Germany, E-mail: bernhard.rupprecht@tum.de

Birgit Vogel-Heuser, Institute of Automation and Information Systems, Department of Mechanical Engineering, TUM School of Engineering and Design, Technical University of Munich, Boltzmannstr. 15, 85748 Garching, Germany; Core Member of MDSI and Lead of Sector Work of MIRMI, Technical University of Munich, Munich, Germany; and Munich Institute of Integrated Materials, Energy and Process Engineering, Technical University of Munich, Garching, Germany, E-mail: vogel-heuser@tum.de

Alexander Fay, Franz C. Kunze and Tom Westermann, Institute of Automation Technology, Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany, E-mail: alexander.fay@hsu-hh.de (A. Fay), franz.kunze@hsu-hh.de (F.C. Kunze), tom.westermann@hsu-hh.de (T. Westermann)

Victoria Hankemeier, Institute of Automation and Information Systems, Department of Mechanical Engineering, TUM School of Engineering and Design, Technical University of Munich, Boltzmannstr. 15, 85748 Garching, Germany, E-mail: victoria.hankemeier@tum.de

Gianluca Manca, Industrial AI, ABB Corporate Research Center Germany, Wallstadter Straße 59, 68526 Ladenburg, Germany; and Institute of Automation Technology, Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany, E-mail: gianluca.manca@de.abb.com

Keywords: alarm floods; alarm management; engineering knowledge; hybrid plants; root-cause analysis

Zusammenfassung: Alarmmanagementsysteme in der Prozessindustrie begrenzen dargestellte Alarmer durch Root-Cause-Analyse. Die Herausforderungen bei der Übertragung dieser Methoden auf andere Produktionsanlagen werden durch die Untersuchung der Alarmcharakteristika von kontinuierlichen, diskreten und hybriden Anlagen aufgezeigt. Dieser Beitrag zeigt die Anforderungen an die Alarmursachenanalyse für verschiedene Anlagentypen auf und bewertet vorhandene Ansätze entsprechend. Da die Alarmursache nicht zwingend der zeitlich erste Alarm ist, erfordert ihre Plausibilisierung weiteres Anlagenwissen. Deshalb werden notwendige technische Dokumente und formalisiertes Wissen zur Plausibilisierung identifiziert.

Schlagwörter: Alarmfluten; Alarmmanagement; Engineering-Wissen; Hybride Anlagen; Root-Cause Analyse

1 Introduction

The need for highly efficient production has led to an extensive use of automated production systems [1], which are managed by supervisory and control systems. As a constituent, alarm management systems (AMS) raise alarms whenever a process deviation or a fault is detected [2]. They inform the operator about sensor and actuator values exceeding or falling below a certain threshold or if some internal system conditions, like pre- and postconditions in a process chain, are violated [3]. In undesirable conditions, an initial disturbance or fault might propagate throughout the plant causing additional deviations in other components, worsening the existing abnormal situation and possibly leading to emergency shutdowns [3, 4]. To avoid such a shutdown, a quick reaction to abnormal behavior is crucial. This generally requires manual intervention by the plant

operator [2], which becomes a bottleneck if alarm floods (AF) appear.

AF pose a major obstacle in AMS design [5, 6]. They occur when a large number of alarms are activated within a short period of time (e.g., more than 10 alarms in 10 min) [3, 5]. During AF, operators struggle to accurately determine the underlying reasons for the issue and respond to faults effectively [6], e.g. by identifying and addressing the root-cause, which is the underlying fault or trigger leading to an AF. Missing crucial alarms can yield costly downtimes and industrial accidents [3]. Often, a straightforward handling of successively occurring alarms, called an alarm sequence, may not be feasible, since the root-cause alarms are not necessarily first in the sequence [7, 8]. Thus, AMS employ advanced methods and tools to enhance the operator's decision-making process and situational awareness [3, 5] by preprocessing raw alarm data and outputting easy-to-interpret information about the ongoing situation [6]. Alarm flood root-cause analysis (AF-RCA) methods determine the cause-and-effect relationships between disturbed process variables and help to identify the most likely root-cause [4, 5, 9].

Given the significance and benefits of AF-RCA, this paper examines existing approaches. Research to date focuses mostly on continuous processes within the process industry and has limited transferability to discrete ones found in manufacturing and intralogistics. However, also in discrete processes, a vast number of alarms caused by a single root-cause can be a challenge for the plant operator. For instance, a stuck workpiece or carrier can block subsequent parts causing a chain of building up alarms with changing order depending on the desired transport route [9].

To bridge this gap between industries, we aim to provide a first contribution to the development of comprehensive AF-RCA methods applicable to a wide range of process types and use cases. Four exemplary plants covering discrete, continuous, and hybrid processes are presented and compared to derive a set of requirements for AF-RCA development based on their commonalities and differences. Furthermore, relevant engineering knowledge and documents for the respective process types to validate potential dependencies between alarms in a sequence are suggested. Knowing the root-cause alarm and dependencies between alarms, unnecessary ones can be filtered to reduce the operator's workload [5, 9, 10].

Related work in data-driven and knowledge-based AF-RCA for different process types is examined in Section 2, showing the limitations of purely data-driven methods. Section 3 presents and compares four illustrative plants with distinct characteristics. Therefore, suitable requirements for a comprehensive AF-RCA in those plants are

derived in Section 4. In Section 5, the AF-RCA methods, which incorporate additional knowledge and overcome limits of purely data-driven ones, are assessed against the requirements from Section 4. In Section 6, engineering documents and expert knowledge useful for the advancement of AF-RCA are proposed. Section 7 concludes with a summary and an outlook on future research.

2 Related work in alarm flood root-cause analysis for varying production process types

Automated production systems are grouped into four categories based on their type of production process [11]. Continuous processes are characterized by a continuous product stream and consistent output [12]. In contrast, discrete ones [13] have production steps with defined start and endpoints, which are repeated for each individually traceable product [14]. Additionally, batch processes are continuous subprocesses that are completed before being reinitialized [14]. If a process is composed of more than one type, it is a hybrid process [15].

The following subsections analyze AF-RCA approaches, regarding their suitability for the previously mentioned process types. They are categorized by data-driven methods that solely rely on alarm data and methods incorporating further engineering knowledge, representing any knowledge about the process and its dependencies. This comprehensive examination aims to uncover shared characteristics and distinctions among the various approaches.

2.1 Data-driven AF-RCA methods

Since similar sequences will likely have the same root-cause, historical alarm sequences with an identified root-cause, were considered. These sequences with a known root-cause are often found in operator annotations in the alarm log [16] or in incident reports [17]. However, to discover new, unknown root-causes, a causal analysis of the AF must be conducted, e.g., by data-driven AF-RCA methods. Since the process industry plays a leading role in the development of AF-RCA methods most research is conducted in this domain. For an extensive overview of state of the art methods for data-driven AF-RCA in the process industry, please refer to [11]. Moreover, Al-Dabbagh et al. [18] rank and reorder AF by a criticality index, based on a set of selected metrics associated with the alarms in the AF. The index is updated every time as a new alarm is detected. Besides, Guo and Guo [19] use a bag-of-words feature extraction to predict faults

from multivariate time series and extract critical patterns associated with the underlying root-cause in a continuous paper manufacturing process as well as in discrete self-pierce riveting in the automotive industry.

In contrast, there are still comparably few data-driven AF-RCA methods designed for discrete and hybrid processes. Fan et al. [20] show a semiconductor manufacturing fault detection approach by identifying key fault producing operations and associated process parameters using machine learning techniques. Kinghorst et al. [21] split the alarm data into groups of statistically dependent alarms using a graph-based approach. They assume that alarms are statistically dependent if their activation time overlaps. Fahimipirehgalin et al. [22] used a clustering algorithm based on transfer entropy and time trends. Pezze et al. [23] forecast future alarms of a dairy product packaging process by a multi-label classification task with past alarm sequences as input. Folmer et al. [24] developed a method that bases causality on the temporal relationship of alarm appearance considering cyclic occurring patterns. Most of these approaches were evaluated on real world industrial data sets originating from continuous, discrete, or hybrid process plants. However, they require dependable knowledge about activation and deactivation timestamps, which is not always available. Moreover, in [25], different similarity measures for the detection of comparable alarm sequences were evaluated.

However, purely data-driven AF-RCA methods are prone to identifying false causality [26] and one alarm sequence may consist of alarms originating from multiple root-causes [27]. Therefore, data-driven AF-RCA methods need to rely on additional reasoning, such as human experts, to verify and increase trust of the detected causal relations [28].

2.2 AF-RCA methods incorporating engineering knowledge

The incorporation of engineering knowledge further extends the AF-RCA methods to justify causal dependencies (justification). In [29] the data-driven identification of possible root-causes was extended by the comparison of alarm sequences. They capture the connections between plant components to verify that the root-cause alarm can in fact be the cause for consecutive alarms in the sequence. Moreover, mass and energy flows using Multilevel Flow Modelling (MFM) were considered in [30]. Early detectability and traceability of errors in chemical systems was established in [31, 32] focusing on the spatial flow of materials. Connectivity information can be extracted from Piping and Instrumentation Diagrams (P & ID). The approach in [33] used this connectivity

information to create Dynamic Causal Digraphs (DCDG) for fault detection, fault diagnosis and alarm management. DCDG are cause-effect graphs extended by additional information and so-called fault propagation carriers which determine the distribution paths of different fault types along the DCDG. A related approach is presented in [4]. Here, P & IDs are used to detect possible propagation paths between the alarm sources combining them with an AF-RCA method based on temporal dependencies. Both approaches were evaluated with data from a simulated continuous production process. For an application on real-world plants, Schlegel et al. [10] used additional knowledge about the process from Computer Aided Engineering Exchange (CAEX) files to support a rule-based approach for alarm grouping. The causal relationships between process variables were obtained using nonparametric multiplicative regression models with the incorporation of CAEX connectivity information [34], to decide whether the found causal dependencies are justifiable.

Notably, all approaches mentioned in this subsection were designed and evaluated within the process industry. Some of the concepts are only applicable to continuous processes and an adaption of those methods for other process types is likely to require extensive changes or is even impossible. One example is fault propagation carriers as temperature or pressure, that describe how different fault types propagate along a continuous product flow [33].

Alarms in manufacturing and intralogistics are usually generated instantly by discrete failure events, without a warning state. Those faults have the potential to propagate to other interconnected machines, by starving them of materials or blocking the output of finished products, leading to prolonged downtimes and reducing the overall plant efficiency [35]. Vuković and Thalmann [36] provide a review on causal discovery in manufacturing and integration of engineering knowledge. The similarity of causal discovery to AF-RCA methods makes an adaption to alarm analysis conceivable. Besides, in manufacturing, model-based AF-RCA approaches were enriched by knowledge obtained from interviews [37, 38].

In [39], explicit models of process knowledge for the automated prediction of quality parameters were created using the example of a hybrid timber particle board production presented in Section 3.4. A statistical process optimization and control approach was extended by expert knowledge from interviews and card sorting which was formalized in cause-effect graphs [40]. Another approach considering interview-based expert knowledge is presented in [38] with a hybrid bottle filling module used for evaluation. Considering a hybrid hot strip mill process, Dong et al.

[41] use a hierarchical causal graph to divide the production process into sub-blocks based on process knowledge. They perform fault detection, causal and propagation analysis using causal feature long short-term memory (CF-LSTM) networks.

Most approaches incorporating further knowledge solely rely on expert interviews which are costly, time consuming and prone to mistakes or operator bias. Despite the limited availability of AF-RCA methods in other domains than the process industry, the growing complexity of today's intralogistics and manufacturing systems leads to a rising demand of such approaches [42]. However, regarding the development and applicability of AF-RCA methods in discrete and hybrid processes, the significant gap in research underlines the importance of developing and evaluating novel interdisciplinary AF-RCA methods [43].

3 Plants for varying production process types

Highly variable production process types are presented below: continuous chemical production, discrete manufacturing with buffers, discrete intralogistics, and hybrid timber particle board production, which includes batch, continuous, and discrete processes. These examples were chosen to identify and illustrate different production processes' characteristics and their effects on alarm sequences (Section 3.5 and Table 5). This forms the base for deriving interdisciplinary AF-RCA requirements in Section 4.

3.1 Tennessee-Eastman Process (TEP)

The Tennessee-Eastman Process (TEP) has been recognized as a benchmark for chemical plant simulation, in alarm management and for AF-RCA [4, 33, 44–46]. As a continuous process based on a real-world chemical plant, the TEP's flat process hierarchy consists of a two-phase chemical reactor, a condenser, a vapor-liquid-separator, a stripper, and a steam reboiler. Pipes, two pumps, and one compressor transport material between above-mentioned components [47]. Necessitated by the TEP's unstable nature, automatic control valves in 17 control loops are used for process stabilization [48]. A full overview of the process operations and phases is provided in [47, 49]. The considered revised TEP model [47], which enables extended monitoring of process variables, including flow, pressure, temperature, level, and chemical component concentration, is considered within this paper. Among these variables, the chemical component concentration has a deadline and a delay. As a result, the

current value of the concentration is available with a time lag from 10 to 15 min.

Regarding alarm sequences, a sudden variation in at least one variable can propagate throughout the TEP and cause AFs [46]. As an example, two sequences "A" and "B" (Table 1), were generated using some random variations of an identical root-cause, namely a step change in the inlet flow of material feed A, by manipulating "XMV3" and raising alarm "XMV3_H" (XMV: manipulated variable; H: high alarm). The sequences show the activated alarms and their respective timestamps of activation during the first 3 h of the propagating abnormal situation. For better understandability, the start of both sequences is shifted to time zero. The red-marked alarm variable "XMEAS1_L" (XMEAS: process variable, L: low alarm) directly corresponds to the underlying root-cause. As indicated in both sequences, this disturbance propagates and affects numerous other process variables. However, due to set alarm thresholds and random effects on corresponding process values, the alarm variable "XMEAS1_L" is only activated in alarm sequence "A".

Relevant phenomena are demonstrated: The first alarm does not necessarily represent the underlying root-cause and the real causal alarm is not necessarily included in the sequence. This can lead to ambiguous alarm sequences for the same root-cause.

For example, in sequence "B", alarm variable "XMEAS10_L" corresponds to a low flow in the TEP's purge, which is located on the other side of the plant and is sensitive to changes in the reactor's inlet.

Secondly, despite having the same root-cause, the sequences "A" and "B" have considerable differences in their alarm orders (indicated by black lines connecting pairs of identical alarm variables in Table 1). Thus, data-driven causality analysis methods that solely rely on the alarm order may yield erroneous causal relationships [4].

3.2 Modular manufacturing plant MPS500

The MPS500 is a highly modular, discrete manufacturing plant consisting of seven stations, which can perform independent manufacturing steps (Figure 1). All stations are connected via a conveyor system which transports parts on workpiece carriers. Two product types – thermostats and cylinders – each with multiple subvariants, are made from a cylindrical base part that is processed at the drilling station according to the desired product type. At the assembly station, an industrial robot mounts additional components to the base part. The order of necessary manufacturing steps is planned via a Manufacturing Execution System (MES), which tracks the products using RFID tags on the workpiece carrier.

Table 1: Two TEP alarm sequences “A/B” from [49]. Red-marked variables: correspond to underlying root-cause.

Alarm Sequence “A”		Alarm Sequence “B”	
Alarm Variable	Activation Time	Alarm Variable	Activation Time
XMEAS1_L	00:00:00	XMEAS1_L	Not Activated
XMV3_H	00:35:20	XMEAS10_L	00:00:00
XMEAS38_H	01:26:00	XMV3_H	00:02:40
XMEAS31_H	01:44:20	XMEAS38_H	00:57:30
XMEAS49_H	01:46:40	XMEAS49_H	01:18:20
XMEAS23_L	01:54:20	XMEAS23_L	01:48:10
XMEAS27_H	01:58:10	XMEAS33_H	01:55:00
XMEAS25_H	02:15:40	XMEAS29_L	01:57:00
XMEAS29_L	02:20:30	XMEAS31_H	01:57:20
XMEAS33_H	02:35:00	XMEAS27_H	01:57:40
XMEAS36_L	02:37:30	XMEAS36_L	02:13:50
XMEAS32_H	03:10:50	XMEAS25_H	02:15:50

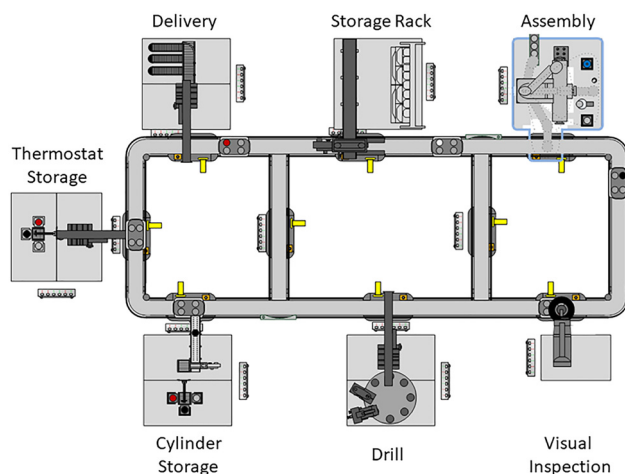


Figure 1: Layout of the modular manufacturing plant with seven stations.

Alarms during manufacturing are raised if product characteristics do not meet the expected ones, i.e., product type (thermostat or cylinder) at the assembly station, height, and material (metal or plastic) at the drilling station, and color of the base part at the inspection station. Unlike in the TEP, these alarms are based on discrete events, usually with multiple software conditions using input from binary sensors instead of checks against a limit of a continuous variable. Alarm sequences (Table 2) can occur, for example, if two base parts – one for a thermostat and one for a cylinder – are accidentally swapped. As both parts traverse their manufacturing sequences, alarms are successively raised at different stations. For each base part, the height does not

match the expected value (Height Check). If the color differs from the expectation, a color mismatch is detected (Color Check). These two alarms inform the operator about the deviation without aborting the manufacturing process. At the assembly station, the base parts are checked for installation notches, which identify the base part type. If the type does not match, another alarm is activated (Ass. Check) and the production is aborted.

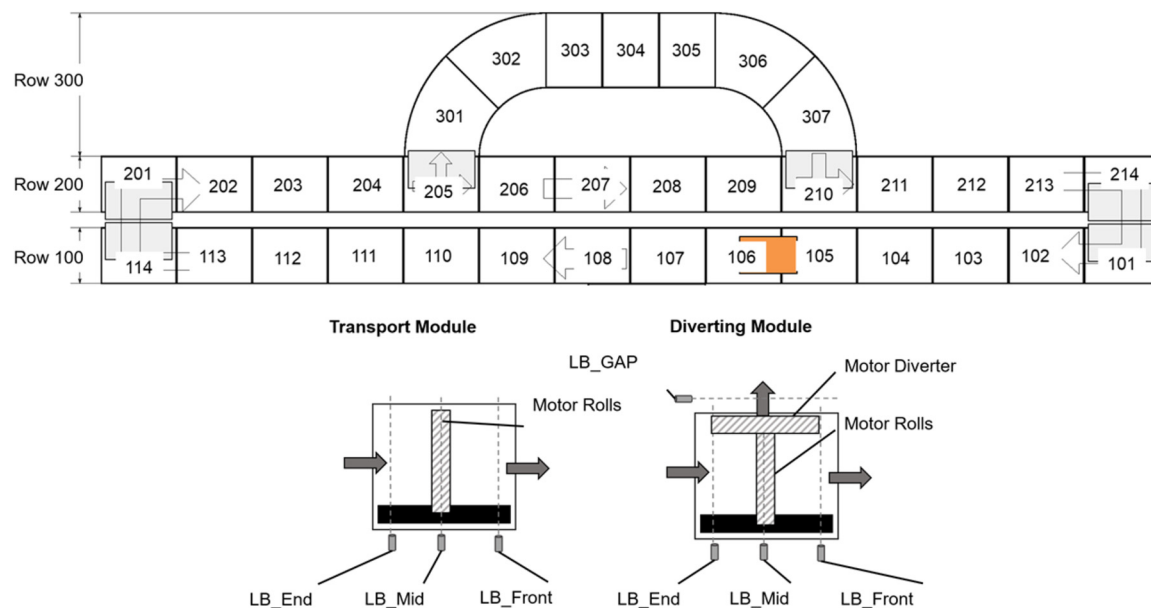
The resulting alarm sequences’ (“A” and “B” in Table 2) different characteristics pose challenges to their interpretation. Although both sequences have the same root-cause, they show variations in their order. The occurrence of some alarms, like the color mismatch, is influenced by statistical effects depending on the (random) color of the base parts. While the order of alarms for a single workpiece is almost completely determined by the fixed order of manufacturing steps, the alarm order can vary considerably across multiple products because the conveyor system allows round trips of a product if a station is busy. Alarm timings are influenced by the possibility to temporarily store the unfinished product in a storage rack. A shortest plausible time span between two alarms, corresponding to a direct, nonstop route from one station to the next, exists. However, deviations from this timespan occur due to wait times at busy stations and storage with no upper time limit.

3.3 Intralogistics small load carrier transport plant

The industrial small load carrier (SLC) transport plant (Figure 2) represents common intralogistics properties,

Table 2: Two exemplary alarm sequences of the MPS500 with the same root-cause (marked red).

Alarm Sequence “A”			Alarm Sequence “B”		
Alarm Variable	Workpiece	Activation Time	Alarm Variable	Workpiece	Activation Time
Root Cause	Both	Unobserved	Root Cause	Both	Unobserved
Height Check	Thermostat	00:00:00	Height Check	Thermostat	00:00:00
Color Check	Thermostat	00:00:32	Height Check	Cylinder	00:00:48
Height Check	Cylinder	00:00:45	Ass. Check	Cylinder	00:01:13
Color Check	Cylinder	00:01:17	Color Check	Thermostat	00:01:16
Ass. Check	Thermostat	00:01:49	Ass. Check	Thermostat	00:02:01
Ass. Check	Cylinder	00:02:23			

**Figure 2:** Layout and basic modules with sensors and actuators of the intralogistics plant.

including multiple available material flow paths and a high degree of modularity [13]. The plant can transport SLCs on modular roller conveyors by making use of 35 basic modules organized in two straight (row 100/200) and one curved row (row 300). Determined by a scanned QR code, the SLC is either transported straight down the row 200 or transposed onto the curved section (row 300). The basic modules are of two basic types (Figure 2): The transport modules (102–113, 202–204, 206–209, 211–213, 301–307) can transport SLCs in one spatial direction, the diverting modules (114, 201, 205, 210, 214, 101) can move SLCs in a second spatial direction using an ejection mechanism. Both module types are equipped with three light-barriers (LB_End, LB_Mid, LB_Front) to detect the SLC position and a motor to power movement in the first spatial direction. The diverting module has an additional motor and light-barrier (LB_Gap).

The transport process of the intralogistics plant is loosely coupled because interfaces between standardized modules are simple with limited signal flow (e.g., just material in/out or conveyor speed). Hardware and software composition follow the ISA-88 standard [50], as the plant is divided into units (rows), equipment modules (e.g., transport and diverting modules), and control modules (actuators and sensors). The basic module names indicate both the location and material flow direction, e.g., 103 and 105 are neighboring modules of 104 in row 100, whereby SLCs pass these modules in ascending or descending numerical order.

Alarm messages are raised by individual hardware modules, rows, or the entire intralogistics plant. Conceptually, an upward and downward propagation of alarms by the control software is observed as, e.g., a module shutdown is followed by a row shutdown and vice versa.

Table 3: Two exemplary alarm sequences of the intralogistics plant. Red-marked variables: correspond to underlying root-cause.

Alarm Sequence “A”		Alarm Sequence “B”	
Alarm Variable	Activation Time	Alarm Variable	Activation Time
Motor Rolls 205 Torque low	00:00:00	Motor Rolls 205 Torque low	00:00:00
LB_205_end not reached	00:03:20	LB_205_end not reached	00:03:20
LB_206_front not reached	00:05:00	LB_301_front not reached	00:05:30
Modul 206: Handover Error	00:09:20	Modul 301: Handover Error	00:12:25
LB_205_mid permanently interrupted	00:15:50	LB_205_mid permanently interrupted	00:16:50
Modul 205: SLC jammed	00:20:25	Modul 205: SLC jammed	00:21:34
FG 200: „Shutting Down“	00:27:40	FG 200: „Shutting Down“	00:22:45
„Shutting Down“ error, SLC in Lightbarrier	00:34:30	„Shutting Down“ error, SLC in Lightbarrier	00:34:56

Alarm sequence “A” (Table 3) is caused by lubricant placed on the rolls at module 205 (e.g., after maintenance work). The reduced friction leads to the first alarm of “Motor Rolls 205 Torque Low”, which corresponds to the underlying root-cause. If the remaining friction is too low, the SLC does not move, resulting in subsequent alarms after software-defined wait times (LB_205_end not reached, LB_206_front not reached, Module 206 Handover Error, LB_205_mid permanently interrupted, and Modul 205: SLC jammed). Finally, the control software recognizes a non-recoverable failure in row 200, leading to a complete shutdown of the respective row (FG200: “Shutting down”), which triggers the automated removal of all remaining SLCs on this row. However, due to the stuck SLC at module 205, the automatic shutdown fails (“Shutting Down” error, SLC in Lightbarrier). The respective row is now in an error state waiting for operator response. In the case of sequence “B”, the ejection mechanism of module 205 is activated to transport the SLC to the curved section (module 301). The modular plant structure with variable routes, however, leads to similar alarm sequences raised by varying modules for the same root-cause. High modularity often results in similar sequences despite differing transport routes.

3.4 Hybrid timber particle board production

The hybrid timber particle board production process [15] consists of a batch material preparation (drying and glue

preparation), a continuous forming and press process with up to 80 press frames for pressure and temperature control, and a discrete board splitting and handling. For simplification, this paper focuses on the batch and the continuous process parts (Figure 3). A time-correct assignment of process data, warnings, and alarms to one specific fiber board segment using the durations of production steps is necessary. In Figure 4 exemplary process parameters and corresponding timestamps of one board are depicted. The material’s density, humidity, and mass depend on its retention time in the buffers. This retention time can be estimated for a known filling level, buffer volume and throughput. The throughput again is indicated by measurements of the outfeed velocity using colored material at the buffer infeed. Understanding these parameter relations, the retention time can be determined for different filling levels, buffer volumes and throughputs. The duration of the batch processes I and II need to be estimated by experts. At the conveyor belt (Figure 3), time-correct assignment depends on the thickness of the board which influences the conveyor speed, i.e., thinner boards are produced with higher velocity. An inline quality assessment continuously measures the thickness at five different positions across the board near the press outfeed. Moreover, a cut of the fiber board is examined using destructive material lab testing. These lab results are obtained with a time delay.

For time-correct assignment of the material weight measured by the scales w_{B1} and w_{B2} (Figure 3) to one fiber

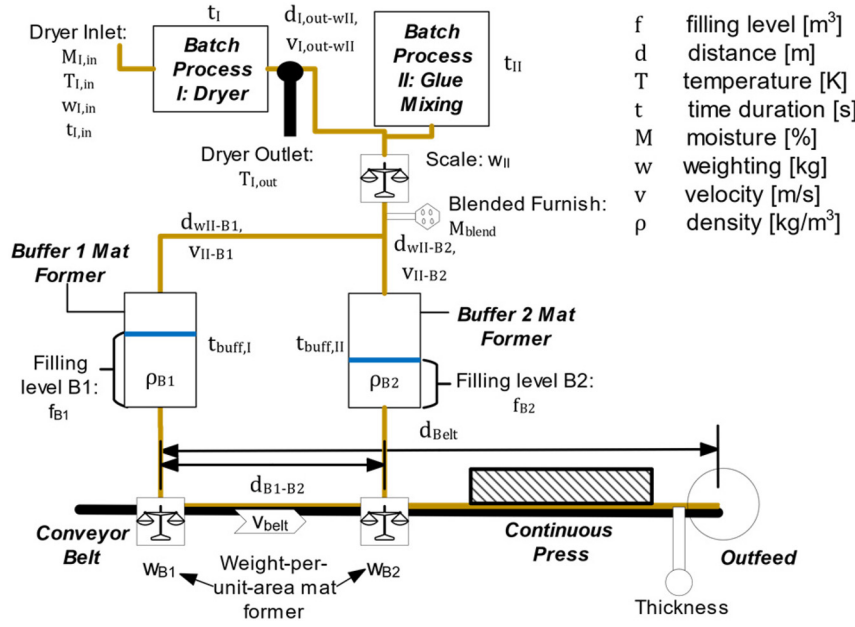


Figure 3: Highly simplified schematics for hybrid timber particle board production (not to scale). Units based on ISO3511.

Temperature dryer outlet	...	Moisture of blended furnish	Weight-per-unit-area mat former	Prepress pressure infeed	Temperature main heating x	...	Temperature main heating y	Pressure frame z middle	Thickness	Lab sample
160.4 °C	...	8.12 %	23.70 kg/m ²	126.8 bar	269.4 °C	...	196.4 °C	51.31 bar	20.20 mm	
9:45:00	...	11:56:26	11:58:01	11:59:34	11:59:55	...	12:01:00	12:01:09	12:01:23	12:10:00
Continuous press										
Alarms need to be assigned properly		Material might be rejected (=alarm) in case deviation from expected values too high								

Figure 4: Exemplary times and process parameters for hybrid timber particle board production to project lab quality (inspired by [15]).

board at the outfeed, knowledge about the distances and velocities of the conveyor belt is required. The weights at w_{B1} and w_{B2} in turn depend on storage processes in the buffers, material properties, and the weight measured at w_{II} . This results in the following dependencies:

$$w_{Bi} = f\left(f(f_{Bi}(\rho_{Bi}(t_{buff,i}))), f(M_{blend}), w_{II}\left(\frac{d_{wII-Bi}}{v_{II-Bi}}\right)\right) \quad (1)$$

with $i = 1, 2$ and

$$t_{buff,i} = f(\dot{V}_{buff,i}) \quad (2)$$

denoting the estimated retention time of the material in the buffer based on the throughput. These dependencies must be considered for time-correct assignment between w_{II} and w_{B1} or w_{B2} . Finally, the weight w_{II} is a function of dryer infeed parameters, the durations of batch process I and II, which depend on the Moisture ($M_{I,in}$) of the infeed:

$$w_{II} = f\left(t_{II}, \frac{d_{I,out-wII}}{v_{I,out-wII}}, t_I(M_{I,in}), T_{I,in}, t_{I,in}, w_{I,in}\right). \quad (3)$$

Table 4: Two exemplary alarm sequences of hybrid timber particle board production with the same root-cause (marked red).

Alarm Sequence “A”		Alarm Sequence “B”	
Alarm Variable	Activation Time	Alarm Variable	Activation Time
Moisture $M_{I,in}$ too high	00:00:00	Moisture $M_{I,in}$ too high	00:00:00
Weight w_{II} too high	02:00:00	Weight w_{II} too high	02:01:10
Density ρ_{B1} too high	02:02:10	Density ρ_{B2} too high	02:04:10
Density ρ_{B2} too high	02:02:30	Density ρ_{B1} too high	02:04:37
Weight w_{B1} too high	02:05:55	Weight w_{B2} too high	02:06:33
Weight w_{B2} too high	02:06:45	Weight w_{B1} too high	02:08:52

The influences in (3) have to be taken into account for time-correct assignment in the batch processes an exemplary alarm sequence “A” of a faulty product segment traversing through the plant is shown in Table 4. A weight above the tolerated range is first recognized by the scale w_{B2} , raising an alarm (Weight w_{B2} too high). Through time-correct assignment as described above, previously raised alarms are mapped to the faulty product segment, identifying a too-high moisture at the dryer inlet (Moisture $M_{I,in}$ too high) as a possible root-cause.

After a varying duration of the batch process I in the dryer, the scale w_{II} indicates an exceeding weight (Weight w_{II} too high). Due to material transport to buffer 1 and 2 respectively, alarms indicating an elevated density in each buffer are raised (Density $\rho_{B1/2}$ too high). After a varying retention time in the buffers 1 and 2 (formula (2)), the scales w_{B1} and w_{B2} recognize a weight (Weight $w_{B1/2}$ too high) above the threshold as well. Alarm messages in and after the buffers highly depend on the filling level and furnish material properties in the buffer. The alarm sequences “A” and “B” (Table 4) show that the alarm order of buffers and scales w_{B1} or w_{B2} may differ within sequences having the same root-cause and are thus ambiguous.

Within the continuous press itself, process parameters at the press frames depend on pressure and distances of any preceding and succeeding frame [15], leading to a tightly coupled process with high interconnection of process parameters.

Within one frame, material properties are recorded by analog sensors transverse to the working direction. Due to locally varying material properties (e.g., higher, inhomogeneous moisture caused by a short-time malfunction of the dryer motor, interrupting the tumble dry process) faulty particle board segments might remain undetected. No alarm during production is raised, because the averaged material properties (e.g., weight) of the segment stay closely

below the alarm raising threshold in the subsequent process. The fault is, however, detected in the lab cut. In this case large non-constant timespans between related alarms (e.g., lab cut and alarm of the dryer motor) make AF-RCA a challenging task.

3.5 Properties of the considered plants

Plant properties influencing alarm patterns are grouped in four categories (Table 5, marked bold). The first category “Structure” describes the overall structure of the plant differentiating between modular plants, and the plants’ hierarchical structure. The second category concerns “Connections” within the plant which influence alarm propagation. It is distinguished between tightly and loosely coupled processes, types of material flow, physical relations, and energy flow. Thirdly, the plants’ sensor types influence the properties of the alarm log (“Alarm Type”). In contrast to binary sensors, a continuous measurement of process variables enables multiple alarm states, logging alarm deactivations (RTN: Return-to-Normal) and is prone to chattering (fluctuating variables that cause multiple alarm activations). The last category considers influence on the alarm “Timing” since timespans between alarms can change depending on material-flow speed, and varying time delays.

4 Requirements for AF-RCA for varying production process types

Based on the previously identified influencing properties on alarm sequences (Table 5), this section derives relevant requirements for interdisciplinary AF-RCA methods. A summary of the requirements is shown in Table 6.

Table 5: Overview of influencing properties on alarm sequences for varying production process types.

Process type	Structure	Connection			Alarm type		Timing	
		Material flow/coupling	Energy flow	Physical relations	L/H alarms	RTN & chattering	Timing influences	Buffer
Continuous (TEP)	Individual sensors	Pipes/tightly coupled	Yes	Yes, e.g., temperature-pressure, flow-level	Yes	Yes	Material-flow speed, control relations, buffer	Yes, reactor tanks
Discrete manufacturing (MPS)	Different modules	Conveyor/loosely coupled	No	No	No	No	Conveyor speed, buffer, wait times	Yes, on conveyor
Discrete intralogistics (SelfX)	Different modules	Conveyor/loosely coupled	No	Yes, e.g., SLC weight-speed (via friction)	No	No	Conveyor speed, wait times	No
Hybrid (particle board production)	Individual sensors	Pipes and press frames/tightly coupled	Environmental influence	Yes, e.g., density-weight; moisture-time, temperature-moisture	Only in material preparation	Yes	Conveyor speed, buffer, moisture, density material	Only in material preparation

$R_{\text{hierarchy}}$: Alarms from closely related plant components are more likely to be causally related [10]. Thus, the multi-level hierarchical composition of the structure of a plant (see column ‘Modularity & Hierarchy’ in Table 5) can provide valuable information for alarm propagation and should be considered in AF-RCA. All sensors of the TEP and the particle board production are individual units that together comprise the full plant [50] and consequently form a flat hierarchy. The hierarchy of the MPS500 and the intralogistics plant is further divided, as both comprise process cells incorporating fully separable modules that again entail sensors.

$R_{\text{modularity}}$: The level of modularity should also be considered. The degree of modularity defines how encapsulated software and hardware are and thus reconfigurability of the plant layout [1]. Both the MPS500 and the intralogistics plant consist of different modules, while the modules differ in their degree of linkage. For example, knowing the number, type, and sequence of basic modules in the intralogistics plant can assist the justification of found alarm sequences. The TEP and the hybrid particle board production follow a more monolithic design. Modular design, e.g. by using module type packages (MTPs) [51], is growing in importance in the process industry and should be considered in the design of AMS.

$R_{\text{material-flow}}$: Alarms can propagate through the production process, e.g., by a SCL blocking the conveyor in the intralogistics plant. The consideration of material flow can thus support finding causally coherent alarm propagation paths [10]. Applications in intralogistics pose the challenge

of bidirectional material flow due to two conveying directions. For the MPS500, the material flow is unidirectional. In discrete applications like manufacturing and intralogistics, alarm sequences are often closely tied to a specific product and its flow through the production process. In the TEP and the hybrid particle board production, most alarms are based on physical material properties exceeding safe operation limits. As the material moves through the pipes to subsequent sensors an alarm propagation is likely. The material flow is therefore relevant in all production processes (see column ‘Material Flow/Coupling’ in Table 5).

$R_{\text{info-flow}}$: Overcompensation of control loops can cause propagation of alarm sequences leading to a chain reaction in which one controller tries to compensate the overcompensation of the other. The information flow can therefore provide valuable information about the propagation and causal relations of alarms due to the underlying software structure [1]. Both the TEP and the hybrid particle board production feature multiple control loops and corresponding information can enrich AF-RCA methods. The underlying software structure directly influences alarm propagation since an alarm in one module can propagate to a software triggered shutdown of the whole row (Table 3).

$R_{\text{energy-flow}}$: Especially in process industry, AF-RCA relevant properties are influenced by energy flows caused by spatial proximity or environmental interactions (see column ‘Energy-flow’ in Table 5) [10]. A change in environmental conditions may lead to varying heat transfer between components and their environment. For instance, parts of

a plant that are close to an open window or gate are significantly impacted by outside temperature conditions. As heat transfer between components with spatial proximity is more likely than between components that are far apart, the physical component location provides valuable information about energy flows and consequent alarm propagation. In the TEP an exothermic reaction exchanges energy with two cooling water supplies and a steam heater, which makes energy flow without the existence of material flow possible. Regarding hybrid particle board production, a change in environmental temperature may result in a too fast or too slow cooling process of the hybrid particle board production in turn yields to diminished stability causing alarms. Since the different modules of the MPS500 and the intralogistics plant operate independently, energy flow between these modules is limited to common power supply.

$R_{physical}$: Alarms can also be related by physical laws [33], i.e., pressure and temperature inside a tank in the TEP (Section 3.1), and the continuous part of the hybrid particle board production (Section 3.4). Furthermore, within the intralogistics plant and the MPS500, physical laws are employed to elucidate the causal relationships of alarms. For example, the input current of the drive and the velocity of the conveyor are directly coupled by a physical equation (see column ‘Physical relations’ in Table 5). Known physical relations should be considered in the AF-RCA directly, as they can reduce the degrees of freedom in alarm analysis and directly coupled alarms.

$R_{alarm-limit}/R_{RTN}/R_{chatter}$: Alarms in the process industry are usually raised based on the upper or lower limits of process variables. Further alarm types (e.g. high-high alarms) are applied to more extreme values, usually corresponding to a more severe alarm [52]. Recorded RTNs can assist in identifying the process state as well as related alarm messages [5]. Alarm logs thus provide valuable insight into the historic production process. However, for threshold-based alarms, fluctuating variables can cause chattering, hindering the analysis of alarm sequences due to changes in count and order of raised alarms. AF-RCA thus need chattering countermeasures. When analyzing the TEP and the hybrid particle board production, all three requirements should be considered (see columns ‘L/H Alarms’/‘RTN & Chattering’ in Table 5). Abnormal events of the MPS500 and the intralogistics plant are discrete and detected by a logic which is set in the software and triggered by multiple binary sensors. Chattering describes the fluctuation around the threshold and a continuous instability. Discrete processes like the MPS500 and the intralogistics plant are thus not susceptible to continuous alarm chattering. RTN timestamps are not known for the MPS500 and the intralogistics plant.

$R_{time-simple}/R_{time-complex}$: The MPS500, the intralogistics plant, and the hybrid particle board production, comprise conveyor belts with a set speed, defining the velocity of alarm propagation along with a conveyed product. Thus, the simple and analytically resolvable relationship ($R_{time-simple}$) between conveyor speed and the length of the conveyor belt should be considered to exclude alarm sequences with impossible time-deltas [53]. In continuous processes, however, these interrelations are more complex. Since the time of an alarm activation does not necessarily correspond to the point in time of the disturbance, not analytically describable time relations ($R_{time-complex}$) can occur. For example, material flow is interrupted by batch and buffer processes or wait times, which also cause an interruption to the material-bound alarm propagation. Effects like these are observed in the TEP, the MPS500 and the hybrid particle board production (see columns ‘Timing Influences’ and ‘Buffer’ in Table 5). Additionally, changing process speeds and varying data sampling rates increase the difficulty of time-correct assignment of alarms to process data in the TEP and the hybrid particle board production. While some variables in the TEP are checked once per second, other values are only sampled every 15 min, with an equal measurement delay, and therefore lag significantly behind the current value of the process variable. This phenomenon can also be observed in the hybrid particle board production, considering the assessment of a lab cut. In general, delays due to lab procedures and slow sampling rates are a type of retention time and should be incorporated into time-correct assignment, including the calculation and estimation of complex timing relationships ($R_{time-complex}$), which is crucial for AF-RCA methods in continuous and hybrid plants (Section 3.4).

$R_{ambiguous-order}$: Even equivalent alarm sequences with a common root-cause may exhibit an ambiguous order (see Section 3), which AF-RCA methods must consider during analysis [22]. The mechanisms that cause such order switching differ between the process types. In the process industry, alarm thresholds together with stochastic variations in sensor measurements influence the activation order of alarms [29, 54]. As a result, the root-cause alarm may be raised after the subsequent influenced alarm. In a manufacturing plant comprising multiple disjunct subprocesses (like MPS500), product specifications are fulfilled sequentially by the subprocesses. Thus, the specification dictates the order of possible alarms. However, order switching may still occur, if independent or redundant subprocesses are reordered to improve the overall process efficiency. Alarm orders of the hybrid particle board production are ambiguous, especially

Table 6: Overview of AF-RCA requirements for varying production process types and assessment of AF-RCA methods against requirements (REQ.: Requirement; RTN: Return to normal).
 Top: Plants (x: AF-RCA requires REQ; (x): AF-RCA partly requires REQ; -: AF-RCA does not require REQ) Bottom: AF-RCA Methods (x: supports REQ; (x): partly supports REQ; -: does not support REQ; in round brackets: intended use of the method).

Process types	$R_{modularity}$	$R_{hierarchy}$	$R_{material-flow}$	$R_{info-flow}$	$R_{energy-flow}$	$R_{physical}$	$R_{alarm-limit}$	R_{RTN}	$R_{chatter}$	$R_{time-simple}$	$R_{time-complex}$	$R_{ambiguous-order}$	$R_{dynamic-causality}$	$R_{variable-processes}$
Continuous (TEP)	-	(x)	x	x	x	x	x	x	x	-	x	x	x	-
Discrete manufacturing (MPS)	x	x	x	-	-	x	-	-	x	-	-	(x)	-	x
Discrete intralogistics (SelfX)	x	x	x	x	-	x	-	-	-	x	-	-	-	x
Hybrid (particle board production)	-	(x)	x	x	x	x	x	x	x	x	x	x	x	x
AF-RCA methods														
Arroyo [33] (continuous process)	-	-	x	x	x	x	x	-	x	-	-	-	-	-
Kirchhübel et al. [30] (continuous process)	-	-	x	x	x	x	x	x	(x)	-	-	x	x	x
Manca and fay [4] (continuous process)	-	-	x	x	(x)	-	x	x	x	(x)	(x)	x	x	-
Rodrigo et al. [29] (continuous process)	-	-	x	x	(x)	-	-	-	x	-	-	(x)	-	-
Schleburg et al. [10] (continuous process)	-	-	x	x	(x)	x	-	-	-	x	-	-	-	-
Wunderlich and Niggemann [38] (hybrid process)	-	-	x	x	(x)	x	-	-	-	-	-	-	-	-
Kottre et al. [37] (no information)	-	-	-	-	-	-	-	x	-	(x)	-	x	(x)	(x)
Dong et al. [41] (hybrid process)	x	x	-	-	-	-	-	-	-	(x)	-	-	-	-

in the material preparation section, caused by changing material retention times in the buffers.

R_{dynamic-causality}: Propagating abnormal situations can change over time and thus, cause-and-effect relationships do as well [4]. Disturbing influences or state changes may increase, decrease, or even invert the causal impact of certain variables on other process variables. Thus, the dynamically emerging causal behavior must be tracked. Especially for tightly coupled processes changing causality should be considered in AF-RCA.

R_{variable-processes}: While some systems are designed for one specific production task, others allow the production products with varying characteristics. The induced changes in the production process can have an influence on the alarm sequences and timings. While the TEP was designed to perform a single production process, the MPS500, the intralogistics plant and the hybrid particle board production show different behavior based on manufacturing or transport instructions. The effects on the timing and order of the alarm sequences ought to be assessed in AF-RCA.

To fulfill the above-mentioned requirements additional information about the plant is needed. It can originate from the operators running the plant or from engineers who designed it. Operator knowledge is usually compiled in time-consuming expert interviews and is often limited, as even an expert will likely not be aware of every causal relation in a complex production plant [55]. Engineering knowledge, in contrast, is usually already available in the form of engineering documents, with the possibility of their automatic analysis saving time and cost.

5 Assessment of requirements satisfaction by selected AF-RCA methods

In this section, the presented methods for AF-RCA incorporating engineering knowledge (Section 2.2) are evaluated for compliance with the previously defined requirements defined in Section 4. This selection was made to pay tribute to the previously identified need for further knowledge.

R_{modularity}/R_{hierarchy}: Information about the plant structure and hierarchy are considered using engineering knowledge of the production process and of the plant's composition in [41]. That information was used to divide process variables into sub-groups which were evaluated separately.

R_{material-flow}/R_{info-flow}/R_{energy-flow}: To incorporate material and information flows, connectivity information is provided in P & IDs or Control Logic Diagrams (CLDs) [33], supporting the assessment of causal relations [29] or

grouping alarms with the same root-cause [10]. Manca and Fay [4] assigned knowledge about the material flow to process variables and in [30, 38] this knowledge is included in MFM and a first principle model, respectively. Energy flows were considered as connection type in [30, 33].

R_{physical}: Physics-based relations between different process variables are established using first order physical principles [33, 38] and included in an MFM [30]. Furthermore, Schlegel et al. [10] define characteristics and limitations of alarm propagations in a plant based on physical effects.

R_{alarm-limit}: In the presented methods, the incorporation of alarm limits was achieved by considering different alarm types directly [33], by normalizing high and low alarms using known thresholds [4], or by creating directed causal trees including the alarm state [30].

R_{RTN}/R_{chatter}: The approaches in [4, 30, 37] contemplate the active time of alarms, by keeping track of the alarm types and/or the active alarm periods. Disruption in AF-RCA by chattering alarms are avoided by removing them out of the alarm log as a preprocessing step [29, 33], by using other data-sources instead of the alarm log [4], or by comparing causal relations on currently active alarms [30].

R_{time-simple}/R_{time-complex}: Time relations were considered by setting expected delays between alarms depending on the process type in [33, 41]. Schlegel et al. [10] considered a maximum possible time between alarms during grouping. Besides, inconsistent timing behavior is described by an uncertainty score when matching alarm patterns [37]. Based on Taken's delay embedding theorem, which provides a way to reconstruct the underlying dynamics of a system, the AF-RCA method by Manca and Fay [4] dynamically adjusts for varying time delays in order to account for complex time dependencies between pairs of process variables.

R_{ambiguous-order}/R_{dynamic-causality}: Some of the presented methods are robust to changes in the alarm order [4, 30]. The approach in [4] determines whether the prediction of an alarm improves with the knowledge about past alarms and thus does not rely on the specific alarm order. Kirchhübel et al. [30] use MFM of different operating situations in the plant to investigate alternative behaviors.

Kottre et al. [37] compare sequences with the same root-cause applying a decision function which allows order switching of alarms. Rodrigo et al. [29] allow small differences in the alarm order. However, only the methods in [4, 30, 37] keep track of dynamical changes in causality, using sliding windows for change tracking, dynamic generation of propagation paths, and adaptive decision functions, respectively.

R_{variable-processes}: Varying production processes were considered by multiple predefined states in an MFM [30], and

by use of preselected pattern descriptions matching the current process [37].

Altogether, no method fulfills all formulated requirements (see lower part of Table 6). Most methods were designed towards a certain process type without consideration of transferability.

6 Engineering knowledge for AF-RCA methods in varying production process types

To meet the requirements (Section 4), AF-RCA methods need to be enriched with additional process information. Therefore, this section provides useful engineering knowledge and corresponding domain-specific documents (Table 7) – without claim to completeness.

Information about the modularity ($R_{modularity}$) of the intralogistics plant and the MPS500, which includes the number and types of basic building modules as well as their arrangement next to each other, is obtained from a layout plan. The hierarchy description ($R_{hierarchy}$) should describe the different hardware modules and submodules. The hardware hierarchy is extracted from a CAD model. Such information for continuous processes like the TEP, as well as the association of sensors to units of machinery, is structured in the P & IDs.

Table 7: Exemplary overview of useful engineering knowledge and documents for AF-RCA.

Requirements	Documents/knowledge
$R_{modularity}$	Layout plan
$R_{hierarchy}$	Labeling conventions (IEC 81346-2), CAD, P & ID
$R_{material-flow}$	Labeling conventions (IEC 81346-2), layout plan, P & ID
$R_{info-flow}$	PLC-SW, P & ID
$R_{energy-flow}$	CAD, layout plan, P & ID
$R_{physical}$	Analytical formulae, heuristics, process data data-driven, P & ID
$R_{alarm-limit}/R_{RTN}$	Alarm log, AMS documentation, PLC-SW, P & ID
$R_{chatter}$	Alarm log, process data, sensor metadata
$R_{time-simple}$	Analytical formulae, PLC variables, CAD, device documentation, process data
$R_{time-complex}$	Process data, expert knowledge
$R_{ambiguous-order}$	Material flow models, product specification, process data, alarm thresholds
$R_{variable-processes}$	MES, PLC-SW, PLC variables, layout plans
$R_{dynamic-causality}$	Process data, expert knowledge

Predecessor-successor relations of components result directly from the material flow directions ($R_{material-flow}$). From equipment labeling according to classification schemes with defined object classes, as standardized in ISO/IEC 81346-2 [56], material flow relations of a system's components are derived (e.g., module 102 is a neighboring module of module 103 in the intralogistics plant). Limits of labeling conventions (like information about the cross-row connections at the diverting modules in the intralogistics plant) are overcome by examining a layout plan. Furthermore, P & IDs are used for hybrid and continuous processes to find material flow connections and differentiate between different piping systems, e.g., material and cooling water systems in the TEP [57], as well as branching paths of material flow in the buffers (Figure 3).

To fulfill the requirement $R_{info-flow}$ an analysis of call sequences and implemented control loops in PLC source code can display information about the information flow [1]. Furthermore, P & IDs often show the information flow as a dashed line [10].

Closely located components or components interacting with the environment are influenced by energy flows ($R_{energy-flow}$). The exact location, including distances, are extracted from a CAD file, where for example close components might exchange thermal energy. Layout plans may provide a rough location and allow to distinguish between components near and far apart while thermal contact equipment like condensers in P & IDs indicate possible energy transfer between connected pipes. To distinguish material transporting connections from purely thermal connections, e.g., cooling water pipes connected to the reactor, CAD files are evaluated.

Physical relations ($R_{physical}$) can either be described as an analytical formula, obtained by expert knowledge in form of known heuristics, or can be learned using process data. In P & IDs, the type of sensor gives information about the type of physical relation. The type can be retrieved from standardized naming conventions [58, 59], e.g., TIR100 being a temperature sensor.

Information on whether alarm limits ($R_{alarm-limit}$) exist and if alarm deactivations (R_{RTN}) are tracked can either be obtained directly from the alarm log, or from the AMS documentation. Furthermore, the number of implemented alarm limits and their thresholds can be obtained from PLC software. Moreover, instrument abbreviations in P & IDs may provide knowledge about the type of alarms implemented for a sensor, e.g., high, or high-high.

The detection of chattering alarms ($R_{chatter}$) can be supported by identifying sensor types from alarm data, sensor metadata or process data, to decide if a sensor is prone to

cause chattering alarms. For instance, a continuous temperature sensor is more likely to generate chattering than a discrete light barrier on a conveyor.

Simple, analytically resolvable timing relations ($R_{\text{time-simple}}$) are calculated using formulae considering e.g., speed, distances, and software-defined timeouts. Distances are given in an accurate CAD model. The material transport speed by PLC variables, process data or the drive device's documentation.

For complex timing behavior ($R_{\text{time-complex}}$), a direct analytic relationship usually cannot be formulated. Timing deviations between the alarm sequences can be high (see Section 3.4). Thus, a combination of documents, like the production recipe, process data, engineering information, expert knowledge, and process data support an accurate estimation of time constraints.

The causes of changing alarm orders ($R_{\text{ambiguous-order}}$) depend on the process type. For continuous process variables, such changes in order can be understood from process data combined with alarm thresholds. In discrete industries such as manufacturing or intralogistics, this variability can arise due to diverse production recipes or routings. These variations can be identified by leveraging process data or material flow models, like MFM, which provide insights into the currently produced product.

To deal with varying production processes ($R_{\text{variable-processes}}$) detailed information about the current process is necessary. For manufacturing plants, the bill of processes, which details the production steps and involved machinery, can often be extracted from the MES. For the intralogistics plant and hybrid particle board production, different routing options can be extracted from the PLC software/variables or layout plans.

Dynamic tracking of causal relations ($R_{\text{dynamic-causality}}$) is not bound to a single document, but rather requires constant reevaluation of existing causal relations. However, access to timing information, e.g., process data, and information about an alarm's RTN are important for this task. Expert knowledge can help to decide if dynamic relations exists and if so, identify boundary conditions between different dynamics.

Overall, available documents vary in type and information content for different plants and process types. Since those documents contain the necessary knowledge to fulfill each requirement, they can be used to enrich existing AF-RCA methods. To avoid excessive manual work, engineering knowledge should be automatically parsed from the documents, if possible. Intermediary formats may be needed (e.g., export standardized data formats instead of

proprietary project files) and manual analysis can serve to verify the automatic procedure.

7 Conclusion and future work

This paper presents challenges of AF-RCA to support the continued improvement of AMS systems, highlighting that most current methods are limited to a single process type. AF-RCA methods for continuous processes have been extensively studied. In contrast, other process types are mostly neglected in AF-RCA research such that their requirements are not yet well-understood. Four comprehensive case studies are conducted to reveal needs for future research: A continuous chemical, a discrete manufacturing, a discrete intralogistics, and a hybrid process. Based on the analysis of distinctive characteristics in the alarm dynamics of these plants, requirements for AF-RCA methods are derived. These requirements can serve as a guideline for the development of novel AF-RCA approaches for varying production process types in future contributions. Notably, AF-RCA methods for hybrid processes must address the requirements of its combined process types. Thus, the development of hybrid AF-RCA methods is an important challenge as such methods are likely to be applicable in both discrete and continuous production processes.

Furthermore, the usage of formalized engineering knowledge in novel AF-RCA methods is suggested to increase their performance and accuracy. However, the integration of such knowledge is still an ongoing challenge, as knowledge extraction, formalization, and usage pose obstacles to overcome. Promising advancements in the domain of engineering knowledge formalization were made in [60, 61]. This paper suggests a selection of relevant knowledge and corresponding documents. Furthermore, engineering information already available in machine readable formats, such as AutomationML, MTPs, or via an Asset Administration Shell in the context of Digital Twins, can also be integrated into the knowledge formalization process by skipping the information extraction step [60]. Thereby, due to the heterogeneity of the considered plants, engineering documents heavily differ in informational content and availability. Since real-world plants can additionally impose the challenge of missing, outdated, and inconsistent information, the combination of different documents is a promising approach. Consequently, methods for AF-RCA should be designed flexibly enough to deal with varying input data and be robust to missing information.

For future contributions towards interdisciplinary AF-RCA using formalized engineering knowledge, novel methods meeting the presented requirements should be

developed. Subsequently, the effectiveness of those methods should be assessed in comparison to traditional, highly customized solutions revealing remaining shortcomings. Operator trust in the findings of AF-RCA methods may be increased by presenting excerpts of the engineering knowledge that underly the analysis [62].

Research ethics: Not applicable.

Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: The authors state no conflict of interest.

Research funding: This research is part of the project “Causal Alarm pattern analysis by the Integration of Technical Information from Engineering Documents (CausALITI)”, funded by the German Research Foundation (DFG) under the project number 455823267.

Data availability: Not applicable.

References

- [1] B. Vogel-Heuser, J. Fischer, S. Feldmann, S. Ulewicz, and S. Rösch, “Modularity and architecture of PLC-based software for automated production Systems: an analysis in industrial companies,” *J. Syst. Softw.*, vol. 131, pp. 35–62, 2017.
- [2] B. Vogel-Heuser, D. Schütz, and J. Folmer, “Criteria-based alarm flood pattern recognition using historical data from automated production systems,” *Mechatronics*, vol. 31, pp. 89–100, 2015.
- [3] *Management of Alarm Systems for the Process Industries, ANSI/ISA 18.2*, 2016, Research Triangle Park, NC, ANSI/ISA: International Society of Automation, 2016. Available at: <https://www.isa.org/products/ansi-isa-18-2-2016-management-of-alarm-systems-for>.
- [4] G. Manca and A. Fay, “Off-line analysis of dynamic causal dependencies in evolving industrial alarm floods,” in *2022 IEEE ICPS*, UK, 2022, pp. 1–8.
- [5] 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, 2019.
- [6] J. Folmer, D. Pantförder, and B. Vogel-Heuser, “An analytical alarm flood reduction to reduce operator’s workload,” in *14th HCI International*, vol. 6764, Springer, 2011, pp. 297–306.
- [7] G. Manca, M. Dix, and A. Fay, “Clustering of similar historical alarm subsequences in industrial control systems using alarm series and characteristic coactivations,” *IEEE Access*, vol. 9, pp. 154965–154974, 2021.
- [8] K. Ahmed, I. Izadi, T. Chen, D. Joe, and T. Burton, “Similarity analysis of industrial alarm flood data,” *IEEE Trans. Automat. Sci. Eng.*, vol. 10, no. 2, pp. 452–457, 2013.
- [9] J. Folmer and B. Vogel-Heuser, “Computing dependent industrial alarms for alarm flood reduction,” in *IEEE SSD*, Chemnitz, Germany, 2012, pp. 1–6.
- [10] M. Schleburg, L. Christiansen, N. F. Thornhill, and A. Fay, “A combined analysis of plant connectivity and alarm logs to reduce the number of alerts in an automation system,” *J. Process Control*, vol. 23, no. 6, pp. 839–851, 2013.
- [11] H. S. Alinezhad, H. M. Roohi, and T. Chen, “A review of alarm root cause analysis in process industries: common methods, recent research status and challenges,” *Chem. Eng. Res. Des.*, vol. 188, pp. 846–860, 2022.
- [12] *Batch Control — Part 1: Models and Terminology, IEC61512-1*, Geneva, Switzerland, International Electrotechnical Commission (IEC), 1997.
- [13] L. Overmeyer, K. Ventz, S. Falkenberg, and T. Krühn, “Interfaced multidirectional small-scaled modules for intralogistics operations,” *Logist. Res.*, vol. 2, nos. 3–4, pp. 123–133, 2010.
- [14] M. Barker and J. Rawtani, *Practical Batch Process Management*, 1st ed. Oxford, Elsevier, 2005.
- [15] B. Vogel-Heuser, “Automation in the wood and paper industry,” in *Springer Handbook of Automation*, Berlin, Heidelberg, Springer, 2009, pp. 1015–1026.
- [16] S. Charbonnier, N. Bouchair, and P. Gayet, “A weighted dissimilarity index to isolate faults during alarm floods,” *Control Eng. Pract.*, vol. 45, pp. 110–122, 2015.
- [17] A. Noroozifar and I. Izadi, “Root cause analysis of process faults using alarm data,” in *27th ICEE*, 2019.
- [18] A. W. Al-Dabbagh, W. Hu, S. Lai, T. Chen, and S. L. Shah, “Toward the advancement of decision support tools for industrial facilities: addressing operation metrics, visualization plots, and alarm floods,” *IEEE Trans. Automat. Sci. Eng.*, vol. 15, no. 4, pp. 1883–1896, 2018.
- [19] S. Guo and W. Guo, “Process monitoring and fault prediction in multivariate time series using bag-of-words,” *IEEE Trans. Automat. Sci. Eng.*, vol. 19, no. 1, pp. 230–242, 2022.
- [20] S.-K. S. Fan, C.-W. Cheng, and D.-M. Tsai, “Fault diagnosis of wafer acceptance test and chip probing between front-end-of-line and back-end-of-line processes,” *IEEE Trans. Automat. Sci. Eng.*, vol. 19, no. 4, pp. 3068–3082, 2022.
- [21] J. Kinghorst, M. F. Pirehgalin, and B. Vogel-Heuser, “Graph-based grouping of statistical dependent alarms in automated production systems,” *IFAC-PapersOnLine*, vol. 51, no. 24, pp. 395–400, 2018.
- [22] M. Fahimipirehgalin, I. Weiss, and B. Vogel-Heuser, “Causal inference in industrial alarm data by timely clustered alarms and transfer entropy,” in *IEEE ECC*, Saint Petersburg, Russia, 2020, pp. 2056–2061.
- [23] D. D. Pezze, C. Masiero, D. Tosato, A. Beghi, and G. A. Susto, “FORMULA: a deep learning approach for rare alarms predictions in industrial equipment,” *IEEE Trans. Automat. Sci. Eng.*, vol. 19, no. 3, pp. 1491–1502, 2022.
- [24] J. Folmer, F. Schuricht, and B. Vogel-Heuser, “Detection of temporal dependencies in alarm time series of industrial plants,” *IFAC Proc.*, vol. 47, pp. 1802–1807, 2014.
- [25] M. Fullen, P. Schüller, and O. Niggemann, “Validation of similarity measures for industrial alarm flood analysis?” in *Technologies for Intelligent Automation, IMPROVE*, Berlin, Heidelberg, Springer, 2018, pp. 93–109.
- [26] F. Yang, P. Duan, S. L. Shah, and T. Chen, *Capturing Connectivity and Causality in Complex Industrial Processes, [Online]*, 1st ed. Cham, Springer International Publishing, 2014.

- [27] J. Kinghorst, H. Bloch, A. Fay, and B. Vogel-Heuser, "Integration of additional information sources for improved alarm flood detection," in *IEEE 21st INES*, Larnaca, 2017, pp. 19–26.
- [28] Y. Laumonier, J.-M. Faure, J.-J. Lesage, and H. Sabot, "Towards alarm flood reduction," in *22nd IEEE ETFA*, 2017, pp. 1–6.
- [29] V. Rodrigo, M. Chioua, T. Hagglund, and M. Hollender, "Causal analysis for alarm flood reduction," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 723–728, 2016.
- [30] D. Kirchhübel, X. Zhang, M. Lind, and O. Ravn, "Identifying causality from alarm observations," in *International Symposium on Future Instrumentation and Control for Nuclear Power Plants*, 2017 [Online]. Available at: https://www.researchgate.net/publication/329389596_Identifying_Causality_from_Alarm_Observations.
- [31] S. Sierla, B. M. O'Halloran, T. Karhela, N. Papakonstantinou, and I. Y. Tumer, "Common cause failure analysis of cyber–physical systems situated in constructed environments," *Res. Eng. Des.*, vol. 24, no. 4, pp. 375–394, 2013.
- [32] N. Papakonstantinou, S. Proper, B. O'Halloran, and I. Y. Tumer, "Simulation Based Machine Learning for Fault Detection in Complex Systems Using the Functional Failure Identification and Propagation Framework," in *ASME CIE*, Buffalo, New York, USA, 2014.
- [33] E. Arroyo, *Capturing and Exploiting Plant Topology and Process Information as a Basis to Support Engineering and Operational Activities in Process Plants*, Dissertation, Helmut-Schmidt-Universität Hamburg, 2017.
- [34] R. Landman and S.-L. Jämsä-Jounela, "Hybrid causal analysis combining a nonparametric multiplicative regression causality estimator with process connectivity information," *Control Eng. Pract.*, vol. 93, p. 104140, 2019.
- [35] Z. Guo, Y. Zhang, X. Zhao, and X. Song, "CPS-based self-adaptive collaborative control for smart production-logistics systems," *IEEE Trans. Cybern.*, vol. 51, no. 1, pp. 188–198, 2021.
- [36] M. Vuković and S. Thalmann, "Causal discovery in manufacturing: a structured literature review," *JMMP*, vol. 6, no. 1, p. 10, 2022.
- [37] A. Kottre, T. Schöler, and C. Legat, "Applying engineering knowledge in alarm flood reduction to reduce machine downtime," *IFAC-PapersOnLine*, vol. 55, no. 2, pp. 54–59, 2022.
- [38] P. Wunderlich and O. Niggemann, "Structure learning methods for Bayesian networks to reduce alarm floods by identifying the root cause," in *22nd IEEE ETFA*, Limassol, 2017, pp. 1–8.
- [39] G. Bemardy and B. Scherff (now: Vogel-Heuser), "SPOC-process modelling provides on-line quality control and predictive process control in particle and fibreboard production," in *24th IEEE IECON*, Aachen, Germany, 1998, pp. 1703–1707.
- [40] B. Vogel-Heuser, V. Karaseva, J. Folmer, and I. Kirchen, "Operator knowledge inclusion in data-mining approaches for product quality assurance using cause-effect graphs," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 1358–1365, 2017.
- [41] J. Dong, K. Cao, and K. Peng, "Hierarchical causal graph-based fault root cause diagnosis and propagation path identification for complex industrial process monitoring," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023.
- [42] J. Wilch, B. Vogel-Heuser, J. Mager, et al., "A distributed framework for knowledge-driven root-cause analysis on evolving alarm data – an industrial case study," *IEEE Robot. Autom. Lett.*, vol. 8, no. 6, pp. 3732–3739, 2023.
- [43] A. Vodencarevic and T. Fett, "Data analytics for manufacturing systems," in *20th IEEE ETFA*, Luxembourg, 2015, pp. 1–4.
- [44] G. Manca and A. Fay, "Identification of industrial alarm floods using time series classification and novelty detection," in *20th IEEE IES*, 2022, pp. 698–705.
- [45] 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, 2009.
- [46] Y. Xu, J. Wang, and Y. Yu, "Alarm event prediction from historical alarm flood sequences based on bayesian estimators," *IEEE Trans. Automat. Sci. Eng.*, vol. 17, no. 2, pp. 1070–1075, 2020.
- [47] A. Bathelt, N. L. Ricker, and M. Jelali, "Revision of the Tennessee eastman process model," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 309–314, 2015.
- [48] N. Lawrence Ricker, "Decentralized control of the Tennessee eastman challenge process," *J. Process Control*, vol. 6, no. 4, pp. 205–221, 1996.
- [49] G. Manca, "Tennessee-Eastman-Process alarm management case study," *IEEE DataPort*, 2020, <https://doi.org/10.21227/326K-QR90>.
- [50] ISA88, *Batch Control- ISA*, [Online], Available at: <https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa88> [accessed: Mar. 23, 2023].
- [51] L. Bittorf, J. Oeing, T. Kock, R. Garreis, and N. Kockmann, "Design of module type package services for modular downstream units and process analytic Technology," *Chem. Eng. Technol.*, vol. 46, pp. 1502–1510, 2023.
- [52] J. Liu, K. W. Lim, W. K. Ho, K. C. Tan, R. Srinivasan, and A. Tay, "The intelligent alarm management system," *IEEE Softw.*, vol. 20, no. 2, pp. 66–71, 2003.
- [53] J. W. Vásquez, L. Travé-Massuyès, A. Subias, F. Jimenez, and C. Agudelo, "Alarm management based on diagnosis," *IFAC-PapersOnLine*, vol. 49, no. 5, pp. 126–131, 2016.
- [54] M. H. Roohi, P. Ramazi, and T. Chen, "Towards accurate root-alarm identification: the causal bayesian network approach," in *IEEE Systol*, 2021.
- [55] 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, 2016.
- [56] DIN EN IEC 81346-2:2020-10, 2020, Berlin, Beuth Verlag GmbH. Available at: <https://dx.doi.org/10.31030/3146080>.
- [57] E. Arroyo, M. Hoernicke, P. Rodríguez, and A. Fay, "Automatic derivation of qualitative plant simulation models from legacy piping and instrumentation diagrams," *Comput. Chem. Eng.*, vol. 92, pp. 112–132, 2016.
- [58] *Graphical Symbols for Diagramms*, ISO 14617-1, 2005. [Online]. Available at: <https://www.iso.org/standard/41838.html>.
- [59] ISA5.1, *Instrumentation Symbols and Identification*, [Online], Available at: <https://www.isa.org/products/ansi-isa-5-1-2022-instrumentation-symbols-and-iden> [accessed: May 23, 2023].
- [60] F. Ocker, B. Vogel-Heuser, and C. J. J. Paredis, "Applying semantic web technologies to provide feasibility feedback in early design phases," *J. Comput. Inf. Sci. Eng.*, vol. 19, no. 4, p. 12, 2019.
- [61] A. Kocher, C. Hildebrandt, L. M. Da Vieira Silva, and A. Fay, "A formal capability and skill model for use in plug and produce scenarios," in *25th IEEE ETFA*, Vienna, Austria, 2020.
- [62] A. Kotriwala, B. Klöpper, M. Dix, G. Gopalakrishnan, D. Ziobro, and A. Potschka, "XAI for operations in the process industry – applications, theses, and research directions," in *AAAI Spring Symposium Combining Machine Learning with Knowledge Engineering*, 2021.

Bionotes



Birgit Vogel-Heuser

Institute of Automation and Information Systems, Department of Mechanical Engineering, TUM School of Engineering and Design, Technical University of Munich, Boltzmannstr. 15, 85748 Garching, Germany; Core Member of MDSI and Lead of Sector Work of MIRMI, Technical University of Munich, Munich, Germany; and Munich Institute of Integrated Materials, Energy and Process Engineering, Technical University of Munich, Garching, Germany
vogel-heuser@tum.de

Prof. Dr.-Ing. Birgit Vogel-Heuser received a Diploma degree in Electrical Engineering and a Ph. D. degree in Mechanical Engineering from RWTH Aachen. Since 2009, she is a full professor and director of the Institute of Automation and Information Systems, Department of Mechanical Engineering, TUM School of Engineering and Design, Technical University of Munich. Her current research focuses on systems and software engineering. She is member of the acatech (German National Academy of Science and Engineering), fellow of IEEE, editor of IEEE T-ASE, and member of the science board of MIRMI at TUM.



Alexander Fay

Institute of Automation Technology, Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany
alexander.fay@hsu-hh.de

Prof. Dr.-Ing. Alexander Fay (born 1970) is Director of the Institute of Automation Technology at Helmut Schmidt University Hamburg. His main research interests are models, methods, and tools for the efficient engineering of distributed automation systems. Prof. Fay also heads the division “Methods of automation” and the Technical Committee “Engineering and operation of automated systems” in the German association for Measurement and Automation (VDI-/VDE-GMA) and is member of acatech — National Academy of Science and Engineering and of the Scientific Advisory Board of the German Platform “Industrie 4.0”.



Bernhard Rupprecht

Institute of Automation and Information Systems, Department of Mechanical Engineering, TUM School of Engineering and Design, Technical University of Munich, Boltzmannstr. 15, 85748 Garching, Germany
gianluca.manca@de.abb.com

Bernhard Rupprecht received an M.Sc. in Automotive Engineering from Technical University of Munich (TUM), Munich, Germany in 2021. He is currently pursuing a Ph.D. at the Institute of Automation and Information Systems, Department of Mechanical Engineering, TUM School of Engineering and Design, Technical University of Munich. His main research interests are data-driven fault detection and algorithm performance benchmarking with focus on low power embedded and edge devices.



Franz Kunze

Institute of Automation Technology, Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany
franz.kunze@hsu-hh.de

Franz Kunze was born 1992 in Frankfurt a. M., Germany. He received his M.Sc. degree in physics from the Carl von Ossietzky University of Oldenburg. Currently he is a research fellow and is pursuing a Ph.D. in engineering at the Helmut-Schmidt-University, Hamburg, Germany. His main research interest is the use of engineering knowledge for causal analysis and root-cause detection for industrial alarm floods.



Tom Westermann

Institute of Automation Technology, Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany
tom.westermann@hsu-hh.de

Tom Westermann M.Sc. (born 1992) is a Research Associate at the Institute of Automation Technology at Helmut Schmidt University Hamburg. His research interests include the semantic description of engineering information as well as methods for the analysis of process data, especially with regard to time.