

DATA-DRIVEN ALARM PARAMETER OPTIMIZATION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

TAYFUN EYLEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

MAY 2025

DATA-DRIVEN ALARM PARAMETER OPTIMIZATION

Submitted by Tayfun Eylon in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Information Systems Department, Middle East Technical University** by,



Date:

21.05.2025



I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Tayfun Eylen

Signature : _____

ABSTRACT

DATA-DRIVEN ALARM PARAMETER OPTIMIZATION

Eylen, Tayfun

Ph.D, Department of Information Systems

Supervisor: Prof. Dr. Erhan Eren

May 2025, 62 pages

Most manufacturing sector businesses utilize advanced control mechanisms to sustain their ongoing operations. An alarm management system is one of these control mechanisms that works as a safety barrier, and it contains alarm messages indicating abnormal situations to operators. The causes of alarms mainly result in a harmful state of operations that should be eliminated as quickly as possible to minimize possible negative results. However, the size of the system, lack of people directing the system, and process-dependent peak conditions may lead operators to miss some critical alarms. Quality and quantity of products, job safety, and operational costs are some of the features negatively affected by these missing alarms. The proposed work aims to combine a well-established alarm management philosophy with advanced data analytics techniques to optimize decision variables in alarm management processes. This study introduces a novel data-driven optimization method that leverages the Tennessee Eastman Process as a benchmark to validate its effectiveness. The proposed method aims to ensure continuous alarm system health by contributing to the automation of the parameter optimization process in the life cycles of alarm management systems. Key contributions include the development of a method to associate disturbances with alarms, the creation of an alarm simulation platform, and the improvement of alarm parameters through a unique optimization approach. The results show that there is a trade-off between alarm reaction delay, which refers to the time between disturbances and the first relevant alarm and number of alarms and alarm on times. This trade-off can be evaluated in the desired direction by taking into account the priorities of the process.

Keywords: Alarm Management, Process Safety, Alarm Parameter Optimization, Alarm KPI's, Tennessee Eastman Process

ÖZ

VERİ TEMELLİ ALARM PARAMETRE OPTİMİZASYONU

Eylen, Tayfun

Doktora, Bilişim Sistemleri Bölümü

Tez Yöneticisi: Prof. Dr. Erhan Eren

Mayıs 2025, 62 sayfa

Üretim sektöründeki işletmelerinin çoğu, devam eden faaliyetlerini sürdürmek için gelişmiş kontrol mekanizmalarından yararlanmaktadır. Alarm yönetim sistemi de güvenlik bariyeri görevi gören bu kontrol mekanizmalarından biri olup, operatörlere anormal durumları bildiren alarm mesajlarını içerir. Alarmların nedenleri esas olarak, olası olumsuz sonuçları en aza indirmek için mümkün olduğunca çabuk ortadan kaldırılması gereken zararlı bir çalışma durumuna neden olur. Ancak sistemin büyüklüğü, sistemi yönlendirecek kişilerin bulunmaması ve prosese bağlı yoğunluklar, operatörlerin bazı kritik alarmları gözden kaçırmalarına neden olabilir. Ürünlerin kalitesi ve miktarı, iş güvenliği ve operasyonel maliyetler bu kaçırılan alarmlardan olumsuz etkilenen özelliklerden bazılarıdır. Önerilen çalışma, alarm yönetimi süreçlerinde karar değişkenlerini optimize etmek için köklü bir alarm yönetimi felsefesini gelişmiş veri analitiği teknikleriyle birleştirmeyi amaçlamaktadır. Bu çalışma, etkinliğini doğrulamak için Tennessee Eastman Süreci'nden yararlanan yeni bir veriye dayalı optimizasyon yöntemini tanıtmaktadır. Önerilen yöntem, alarm yönetim sistemlerinin yaşam döngülerinde parametre optimizasyon sürecinin otomasyonuna katkıda bulunarak sürekli alarm sistemi sağlığını sağlamayı amaçlamaktadır. Ana katkılar arasında, bozulmaları alarmlarla ilişkilendirmek için bir yöntemin geliştirilmesi, bir alarm simülasyon platformunun oluşturulması ve benzersiz bir optimizasyon yaklaşımıyla alarm parametrelerinin iyileştirilmesi yer almaktadır. Sonuçlar, bozulmalar ile ilk ilgili alarm arasındaki zamanı ifade eden alarm tepki gecikmesi ile alarm sayısı ve alarmda kalma süreleri arasında bir denge olduğunu göstermektedir. Bu denge, sürecin öncelikleri dikkate alınarak istenen yönde değerlendirilebilir.

Anahtar Sözcükler: Alarm Yönetimi, Süreç Güvenliği, Alarm Parametre Optimizasyonu, Alarm Performans Metrikleri, Tennessee Eastman Süreci



ACKNOWLEDGMENTS

First of all, I would like to thank my supervisor Prof. Dr. Erhan Eren for his support throughout my long PhD process. He was there for me with his valuable suggestions and comments, sometimes about academic life, sometimes about business life, and sometimes about social life.

Besides my supervisor, I would like to thank Prof. Dr. Altan Koçyiğit for his contributions to all processes. I would also like to point out that he is a role model by using his technical and academic perspectives in every field and does not spare his support even in the busiest periods.

I would also like to thank my wife for always supporting me, adding color to my life, and always thinking of our well-being. Since the thesis period takes place together with other critical stages of life, it can sometimes be very challenging, and no matter how much effort is put in beforehand, it is one of the first areas to be waived during difficult periods. No matter how tired and busy my wife was during these periods, she always created space for me to complete my thesis. I can say that I have been able to reach this stage thanks to her.

TABLE OF CONTENTS

ABSTRACT	iv
ÖZ.....	v
DEDICATION	vi
ACKNOWLEDGMENTS.....	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	x
LIST OF FIGURES.....	xi
LIST OF ABBREVIATIONS	xii
CHAPTER	
1. INTRODUCTION.....	1
2. BACKGROUND.....	7
3. LITERATURE REVIEW.....	9
3.1. Alarm Parameter Design	9
3.2. Alarm System Evaluation.....	11
3.3. Alarm Chattering	12
3.4. Sequence Analysis.....	13
3.5. New Alarm Types.....	16
3.6. Indicator Indexes	17
3.7. State Based Alarming.....	18
3.8. Operator Action Analyses	18
3.9. Literature Gaps	18
4. DATA-DRIVEN ALARM PARAMETER OPTIMIZATION	21
4.1. KPI's.....	22
4.1.1. KPI Definitions	22
4.2. Tennessee Eastman Process (TEP).....	24
4.2.1. TEP Components.....	24
4.2.2. Chemical Equations.....	25
4.2.3. Process Summary	26

4.2.4. Implementation	26
4.2.5. Control Strategy	27
4.2.6. TEP Dataset.....	28
4.2.7. Alarm System Design	29
4.3. Alarm – Disturbance Relation.....	30
4.4. KPI Implementation	33
4.4.1. KPI Normalization	33
4.4.2. KPI Weighting & Cost Function.....	34
4.4.3. Parameter Boundaries	35
5. RESULTS AND DISCUSSIONS.....	37
6. CONCLUSION AND FUTURE WORK	49
6.1. Validity Threats.....	50
6.1.1 Internal Validity Threats	50
6.1.2 External Validity Threats	50
6.1.3. Mitigation Strategies	50
6.2. Answers to Research Questions	51
6.2.1. Research Question 1: How is the performance of an alarm management system or an alarm measured?	51
6.2.2. Research Question 2: Which key performance indicators should be utilized to measure the performance of an alarm management system?.....	51
6.2.3. Research Question 3: How should alarm parameters be optimized for better performance? What is required for the optimization of alarm parameters?	52
6.3. Future Work	52
REFERENCES.....	55
CURRICULUM VITAE	61

LIST OF TABLES

Table 1: Relation between alarm parameters and KPI's	24
Table 2: TEP disturbances.....	28
Table 3: Example Alarm – Disturbance Labeling for s1	31
Table 4: Example optimization results for s1 low alarm.	36
Table 5: Sample table performance comparison	38
Table 6: Minimum reaction delays.....	40
Table 7: Minimum reaction delays (2 nd scenario).....	43

LIST OF FIGURES

Figure 1: Layers of protection, ISA 18.2 (International Society of Automation, 2016)....	2
Figure 2: ISA 18.2 Alarm Management Lifecycle (International Society of Automation, 2016)	3
Figure 3: Block diagram summarizing the proposed work	21
Figure 4: Tennessee Eastman Process (Bathelt et al. 2015).....	25
Figure 6: Alarm state transition diagram	29
Figure 7: Disturbance Labeling.....	30
Figure 8: Disturbance process value graph for s1. 1, 5, 6, and 8 disturbances are determined as valid disturbances for low alarm, and they are shown with vertical lines on the graph.....	32
Figure 9: : Disturbance process value graph for s1. 1, 2, 6, and 8 disturbances are determined as valid disturbances for high alarm, and they are shown with vertical lines on the graph.....	32
Figure 10: KPI Calculation & Optimization	33
Figure 11: Example optimization visualization for s1 low alarm. Red circles indicate entering an alarm state, and blue circles indicate existing from the alarm state.	35
Figure 12: Delay comparison	39
Figure 13: Number of alarms comparison	39
Figure 14: Alarm on time comparison	40
Figure 15: Number of alarms comparison (2 nd scenario).....	41
Figure 16: Filtered number of alarms comparison (2 nd scenario)	41
Figure 17: Delay comparison (2 nd scenario)	42
Figure 18: Alarm on time comparison (2 nd scenario).....	42
Figure 19: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV1	43
Figure 20: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV2	44
Figure 21: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV3	44
Figure 22: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV4	45
Figure 23: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV5	46
Figure 24: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV6	46
Figure 25: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV7	47
Figure 26: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV8	48

LIST OF ABBREVIATIONS

DCS	Distributed Control Systems
DPLS	Discriminant Partial Least Squares
ESD	Emergency Shut Down
EWMA	Exponentially Weighted Moving Average
FAR	False Alarm Rate
FDA	Fisher Discriminant Analysis
FFS	Forward-Flux Sampling
HBM	Hierarchical Bayesian Model
HMI	Human Machine Interface
IDS	Intrusion Detection Systems
ISA	International Society of Automation
KPI	Key Performance Indicator
MAR	Missed Alarm Rate
MV	Manipulated Variable
PCA	Principle Component Analysis
PID	Proportional Integrator Derivative
PV	Process Variable
TEP	Tennessee Eastman Process

CHAPTER 1

INTRODUCTION

As a result of increased attention to robotics and computer sciences, automation in manufacturing became a key factor. There are many examples in the literature showing applications of automated systems. Some of the benefits of an automated system are summarized in Stephanou (1995) as an increase in product quality, an increase in the safety of the operational environment, and a quick response to changing consumer demands for new products. These features can be accepted as a direct result of less human interaction in operations.

Complete automation of a system is a complex task. All interrelated components need to be modeled. Moreover, interactions between components should also be defined. The high number of components and interactions between them increases the system's complexity. As a result, large systems like nuclear reactors and oil refinery plants are complex and very hard to automate fully.

Human reactions are necessary for these large systems, especially for abnormal conditions. Operators should monitor deviations from steady states or transitions between them to react in time in the case of unwanted situations. However, the number of locations or sensors where continuous monitoring is crucial mostly exceeds the operators' capacity. Alarm messages are used to deal with this problem. These messages are created when a sensor reading exceeds a predefined threshold limit in predefined conditions.

An alarm system is a layer of protection. As shown in Figure 1, it is used for prevention. At the base layer (the first layer, Process Design), all process boundary values should be specified. Then the process control layer exists. Process control is a method used to monitor and manage industrial processes to ensure consistent quality and achieve desirable outcomes. The process control layer should do automated actions to prevent process values from exceeding their boundaries. If process control cannot keep values within limits, then operator action is required. The layer where operator intervention occurs is called alarm layer. After this point, which means that operator intervention is not enough or there is no alarm to warn operators, in order to prevent greater damage, the ESD (Emergency Shut Down) equipment is activated to continue production less profitably but

safely. However, this equipment may not be enough to stop the problem, and incidents occur. In the passive protection layer, partial shutdowns can be applied to reduce the effect. It is guaranteed there is a big financial loss in this layer. If the problem gets bigger, there will be a unit shutdown at a higher level, as human life will be endangered. If the last two layers are reached, we can say a catastrophic incident happened.

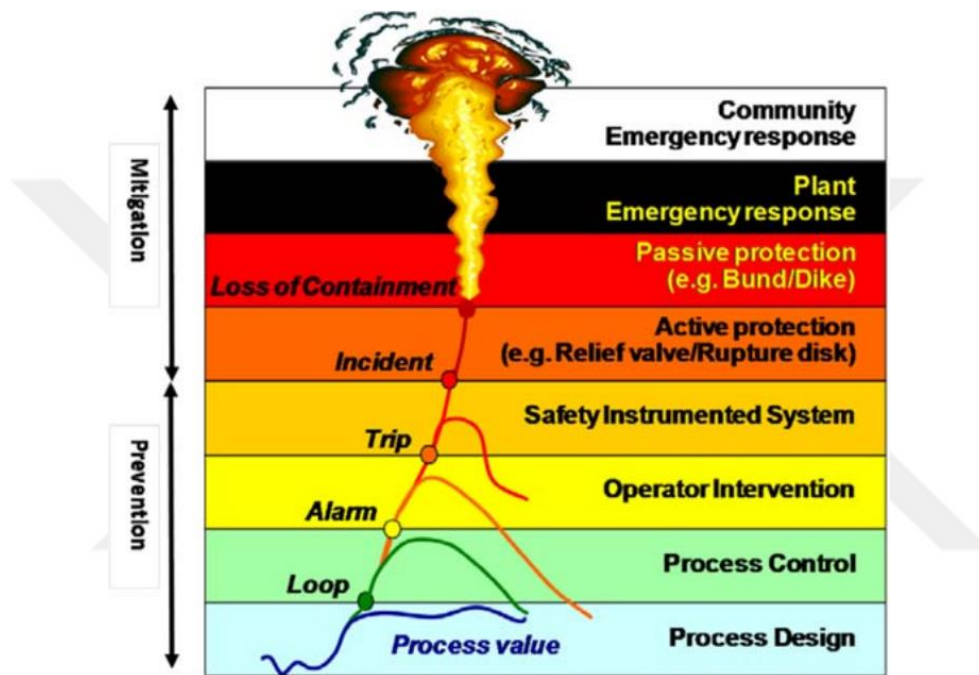


Figure 1: Layers of protection, ISA 18.2 (International Society of Automation, 2016)

Unfortunately, there are many examples of catastrophic incidents where the alarm system fails, such as Three Mile Island (Walker, J. S., 2004), Milford Haven Explosion (HSE Books., 1995), and BP Texas (Hopkins, A., 2008). Three Mile Island is the most significant U.S. commercial nuclear power plant accident. Some design flaws caused it, but it can be prevented. At the time of the accident, 100 alarms appeared simultaneously because there were duplicate alarms. So, operators could not realize the root cause or take any action. Cleanup took about 14 years and cost \$1 billion in 1993, approximately equivalent to \$2 billion today. In the case of the Texaco Refinery in Milford Haven, the accident started with a lightning strike. A fire started in the crude distillation unit. Then, related plants were shut down as normal. However, at that point, the debutanizer discharge valve should have been open, but it remained closed even though the control system indicated it was open. So, the drum was filled beyond its design capacity, resulting in overflow and an explosion. There was one alarm for the drum's high level, but there were many irrelevant alarms, and the operator missed the most important one. The BP Texas City accident happened during a start-up (initialization of operations). Similar to the Milford Haven case, there were wrong signals. The distillation tower level seemed to be declining when it was overfilling. Moreover, alarms were ignored because of a startup,

and some were undefined. As a result, explosions happened, 15 people were killed, 180 people were injured, and it cost about \$1.5 billion.

If the alarm system was healthy, the above-mentioned accidents could be prevented (Walker, J. S., 2004, HSE Books., 1995, Hopkins, A., 2008). Then, the important question is, “How can we have a healthy alarm system?” For this purpose, ISA-18.2 (International Society of Automation., 2016) standards exist that define the alarm management lifecycle. This lifecycle describes steps to improve the alarm system continually. It starts with creating an alarm management philosophy and includes identification, rationalization, design, implementation, operation, and maintenance steps. From identification to maintenance, steps are in a loop. Each step takes the output of the previous step as input.

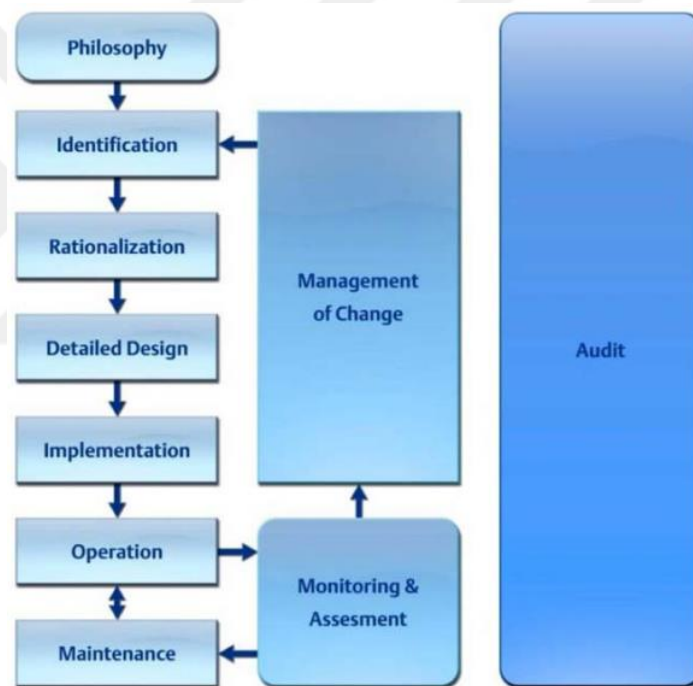


Figure 2: ISA 18.2 Alarm Management Lifecycle (International Society of Automation, 2016)

When generating an alarm philosophy document, each step in the lifecycle is defined. The first definition is about alarms. These definitions are made to be consistent. All steps must be perceived in the same way by all stakeholders. Moreover, roles and requirements should be defined.

For the identification and rationalization steps, actions for each alarm should be studied. Consequences of not taking any action should also be defined. Alarms with no necessity to take any action should be eliminated. Calibration of alarm parameters is also a subject of rationalization process. In this step, documentation is also important. Type, class, priority and limits should be included in the report.

In detailed design and implementation steps, capabilities and limitations of the process are considered. Test runs are applied to see the validity of parameters. Moreover, suppression logics are implemented, which means internal filtering mechanisms can be applied in the design phase. HMI (Human Machine Interface) design is also considered in this step. Color conventions and icons can be defined.

Operation, maintenance, monitoring & assessment, and management of change are runtime operations. Operation and maintenance steps are self-defining. In the monitoring & assessment step, the performance of the alarm management system is considered. KPIs are generated in this step. Management of change step considers applied changes. It is also an important step, because there are many stakeholders. Individual changes should be shared, and this process should be managed not to have asynchronous operators.

ISA 18.2 is a well-documented lifecycle approach that is used by many companies for alarm systems. However, it is also very costly. It requires a multifunctional team, including process engineers, control engineers, lead operators, and safety engineers. Moreover, periodic meetings are required with this team to analyze failure mechanisms. These meetings are called alarm rationalization meetings and alarm parameters are calibrated in these meetings.

The process of setting alarm parameters is currently carried out as a manual process. At each iteration during these manual operations, the process is monitored, and alarm parameters are changed as necessary. Iterations may typically take at least one week. The process cannot be carried out properly due to prolonged iteration times and the need for extensive participation. The proposed work aims to facilitate the mentioned process.

The developments made in this field in literature mostly aim to reduce alarm floods, that is, to reduce the number of alarms with new filtering methods. However, one of the most important issues to be considered in the alarm parameter setting process is the reaction time for early detection of disturbances. Generally, publications in literature provide this issue by taking the opinions of domain-experts. However, this is also an obstacle to the automation of the process. Our proposed approach aims to automate and optimize alarm parameter settings. For this purpose, a new KPI, alarm reaction delay, is proposed. The key contributions of this study are as follows:

1. Developing a method to associate faults with alarms and thus calculate the alarm reaction delay value, a novel KPI which facilitates the automation of alarm parameters optimization
2. Creation of a comprehensive alarm simulation platform that includes the most common alarm parameters using the Tennessee Eastman Process (TEP)
3. Improvement of alarm parameters through a novel data-driven optimization method

Despite there are many studies on alarm management in the literature, there are still gaps that this study addresses:

1. Lack of automated optimization methods
2. Trade-off between alarm filtering and timely reactions

By addressing these gaps, the proposed work contributes to the field of alarm management and abnormal situation management, providing a novel perspective and practical solutions for improving alarm system performance.

To guide this study, the following research questions have been formulated:

1. How is the performance of an alarm management system or an alarm measured?
2. Which key performance indicators should be utilized to measure the performance of an alarm management system?
3. How should alarm parameters be optimized for better performance? What is required for the optimization of alarm parameters?

In Chapter 2, the background information is given, defining alarm parameters and explaining some of the critical definitions. In Chapter 3, the literature overview is given in four categories: alarm parameter design, alarm system evaluation, alarm chattering, sequence analyses, new alarm types, indicator indexes, state-based alarming, and operator action analysis. In Chapter 4, the proposed work is explained in 4 sections. In Section 4.1, selected KPIs are introduced. Tennessee Eastman Process (TEP) is utilized to show the validity of the proposed work. In Section 4.2, TEP and constructed alarm management systems are introduced. Section 4.3 discusses disturbances of TEP and disturbance–alarm relations. In Section 4.4, the optimization structure and implementation of the work are explained. In Chapter 5, results are given with comparisons from literature. Lastly, the conclusion and future work is provided in Chapter 6.



CHAPTER 2

BACKGROUND

Process engineers and plant manufacturers define the parameters for generating and filtering alarm messages. The most common parameters included in most alarm systems are;

- Threshold values,
- On/Off delay parameters,
- Hysteresis ratio.

A threshold is a limit value. As stated above, an alarm message is created when a sensor reading exceeds this value. An on-delay parameter is a delay value for an alarm to occur. No alarm is generated if a sensor reading exceeds the threshold and returns in less than on-delay time. Off-delay is a similar parameter to on-delay. The alarm condition is cleared when a sensor reading crosses the threshold line while returning to a normal state. If an off-delay parameter exists, the alarm condition is kept for off-delay time. A hysteresis ratio is a ratio of the maximum sensor reading value behaving like the off-delay parameter. From these parameters, the threshold value is used to generate alarms. Other parameters are utilized for filtering.

As a result of cascaded relations between processes in large plants, a single failure or disturbance may propagate and affect other parts. In these situations, sensor readings in healthy parts of a plant also deviate and generate alarm messages. This is called alarm flooding. In the case of floods, the number of alarms exceeds the limit to which operators can react in time. The ISA-18.2 standard suggests 10 alarms per 10 min for each operator (International Society of Automation, 2016). If the alarm rate exceeds this recommended value, the state of the alarm management system is accepted as “flooding.”

The causes of alarm floods stated in literature (Ahmed et al., 2013; Cheng et al., 2013; Hu et al., 2015; Lai & Chan 2017; Hu et al., 2018; Dorgo & Abonyi, 2018; Vogel-Heuser et al., 2015) can be summarized as a high correlation between sensor readings and sequential process structure. A high correlation between sensors can be exemplified by having multiple alarms and sensors for a single process point. On the other hand, if the system’s

physical structure causes a sequential relation between sensors, then alarms created by these sensors come successively. Similarly, a sensor in a process can be a predecessor of another sensor. In this case, the second sensor will create an alarm from the first one after some time. For example, the first and the second sensors may measure the temperature of a liquid at the start and end of a pipe, respectively. The time passing for the liquid to arrive from beginning to end will be the time lag between two separate alarms. In this case, a single abnormality in a process results in multiple alarms.

Another reason for alarm floods, evaluated in a different category, is chattering alarms. Multiple alarm occurrences are observed in a short time when alarm chattering happens. The reasons for chattering alarms have also been studied well. Some of them stated in the literature (Wang & Chen, 2014; Wang & Chen, 2016; Zang et al., 2015) are malfunctioning equipment, poorly selected alarm parameters, and missing operator responses.

Participant tags of alarm floods are named “Bad Actors.” Decreasing the number of alarms created by “Bad Actors” is accepted as a common solution to alarm flood problems. The ISA-18.2 standard defines bad actors as the top 10 alarms when an alarm flood exists (International Society of Automation, 2016). Most of the applications for alarm management systems focus on the top 10 alarms to solve flooding problems. These top 10 alarms may have different types of issues. Each one is evaluated individually to eliminate the majority of the problems.

Periodic analysis of alarm management systems to eliminate bad actors is called “Alarm Rationalization.” Processes and equipment conditions change with time. Worn-off equipment behaves differently compared to its initial conditions. As a result, rationalization meetings should be conducted when changes are detected or periodically. One of the aims of the alarm rationalization process is to eliminate bad actors that are previously observed. As stated before, solving alarm problems individually can be very time-consuming. Some reasons for that are manual analyses, complex processes, and crowded meetings. As a result, omitted rationalization meetings cause poorly designed alarm parameters and poorly optimized alarm management systems.

CHAPTER 3

LITERATURE REVIEW

This section provides an overview of related studies published in the literature. Identified papers are categorized into eight groups: alarm parameter design, alarm system evaluation, alarm chattering, sequence analyses, new alarm types, indicator indexes, state-based alarming, and operator action analysis. Each category is discussed in relation to how it informs and supports the objectives of this study, particularly in the context of automating and optimizing alarm parameter settings.

3.1. Alarm Parameter Design

The studies summarized in this section highlight various methods for designing alarm parameters, such as the use of exponentially weighted moving average (EWMA) filters, hysteresis ratios, and manual adjustments. These methods are foundational to our proposed work, which aims to automate and optimize alarm parameter settings. By leveraging data-driven techniques, our approach seeks to enhance the efficiency and accuracy of alarm parameter optimization, addressing the limitations of manual and semi-manual methods discussed in these studies.

In a technical report by Manca (2020), Tennessee Eastman Process data is semi-manually analyzed to detect alarm parameters. In five revisions, parameters were obtained. In the first revision, chattering analyses were conducted. In the second revision, EWMA filters were applied. In the third revision, manual threshold adjustments were applied. In the fourth revision, hysteresis ratios were utilized. In the last revision, some final manual adjustments were made. In the end, thresholds, filtering parameters of EWMA, and hysteresis ratios were presented. Moreover, performance comparisons with starting conditions were shown. The results of this work are compared in the results and comparisons chapter.

Kumari et al. (2020) proposes a new framework for root cause analysis of key process variable deviations for rare events in the chemical process industry. The framework combines relative information gain, Pearson correlation coefficient, and a Hierarchical Bayesian Model (HBM) to handle data scarcity and source-to-source variability.

Sudarshan et al. (2021) discusses the application of forward-flux sampling (FFS) to simulate and analyze rare safety and reliability events in chemical processes. It highlights

the importance of understanding these events, particularly in extreme operating conditions, and compares the brute-force (BF) method with the FFS method. The study uses an exothermic continuous stirred tank reactor (CSTR) model to demonstrate the FFS method's efficiency in tracing pathways to rare events. The results show that the FFS method is more effective and computationally efficient than the BF method, making it a valuable tool for improving alarm systems and ensuring plant safety and reliability.

Parsa et al. (2021) constructed a graph model of an alarm management system. They used the graph model to build process flow, piping, and instrumentation diagrams (P&ID). Alarm rationalization and workshop activities were discussed, and the safety integrity levels (SIL) were given to show how another perspective, like their graph solution, affects the whole process. The model was realized for causal relations between alarms. BP Texas Refinery explosion (2005) was chosen as a case study. The graph's root is specified as the ESD alarm that was bypassed when the explosion happened. The importance of alarms decreases when going from root to leaf nodes. The alarm graph model aims to have a framework for alarm rationalization and evaluation at the design stage. Parsa et al. (2022) created a similar alarm graph model for the “Tennessee Eastman Process” and utilized PageRank, degree centrality, betweenness centrality, closeness centrality, and Dijkstra algorithms to measure alarm importance.

Gyasi and Wang (2023) propose a method to jointly design alarm thresholds and delay timers to reduce missed and false alarm rates. To simultaneously calculate these two values, which are normally designed separately, the authors combined separate distributions. To combine them, they assumed that both variables were independent and in an identical distribution. Using the Bayesian estimation rule cumulative probabilities of alarm durations were estimated.

Gyasi et al. (2023) developed an adaptive method to update alarm deadbands for non-stationary process variables. To do this, they first identify temporal moments of significant changes. These changes are achieved when mean value of the process variable deviates by a predefined value. After detecting changes, the deadband width is updated with the Bayesian estimation rule. The method has been tested with numerical and industrial data and compared with fixed-width deadbands.

Sudarshan et al. (2023) discusses the development of multivariate alarm systems designed to recognize rare unpostulated abnormal events. The study highlights the importance of using advanced algorithms to improve the detection and management of such events.

Sudarshan et al. (2024) introduces a path-sampling algorithm called Branched-Growth Forward-Flux Sampling (BG-FFS) and the XGBoost machine learning algorithm to develop predictive models for committer probabilities. These models are used to create dynamic multivariate alarm systems capable of real-time response actions.

3.2. Alarm System Evaluation

The evaluation methodologies discussed in this section, such as the informativeness and effectiveness metrics proposed by Park et al. (1997), provide a basis for assessing alarm systems. Our proposed work builds on these evaluation techniques by introducing a novel KPI, alarm reaction delay, which measures the timeliness of alarm responses. This KPI, along with the number of alarms and alarm-on times, forms the core of our optimization framework, ensuring a comprehensive evaluation of alarm system performance.

Park et al. (1997) focused on evaluating an advanced alarm system. They stated that an advanced alarm system should be assessed according to its alarm reduction rate and informativeness. They proposed a methodology to measure informativeness that is the answer to the following research question. "How many important alarms are conducive to identifying process deviation provided?" Then, they combined these two metrics to measure effectiveness. For the measurement process, operators labeled alarms for their informativeness. Using analyses of true/false acceptance/rejection values, an advanced alarm system was evaluated. They tested and evaluated their system with data from the Korea Institute of Nuclear Safety simulator of Yonggwang Units 1 and 2.

Kondo et al. (2015) present an approach for alarm rationalization based on process mining techniques. The methodology involves systematically investigating the alarm database to ensure that all parameters are properly configured. Process mining techniques are used to analyze event logs from industrial processes, identifying patterns, anomalies, and significant structures. This approach improves the alarm system by reducing the number of unnecessary alarms and enhancing the reliability of the remaining alarms.

Gao et al. (2015) proposes a systematic rationalization method for multivariate correlated alarms to perform root cause analysis and alarm prioritization. The approach combines interpretive structural modeling (ISM) and a revised Likert scale (RLS) to identify and prioritize critical alarms. The Tennessee Eastman process is used as a case study to demonstrate the effectiveness of the proposed approach, showing that it can reduce alarm floods and improve alarm system performance.

Sudarshan et al. (2024) focuses on alarm rationalization and dynamic risk analyses for rare abnormal events. The key aim of alarm rationalization is to significantly reduce both the number of false alarms, and the total alarms annunciated, ensuring that every alarm is a quality alarm. The study claims that through rationalization, the total number of alarms can be reduced by 50%, coupled with reductions in nuisance alarms. Post-rationalization, the risk associated with multivariate alarm systems is evaluated using dynamic risk analysis (DRAn), a technique developed to achieve real-time proactive risk management against rare abnormal events.

3.3. Alarm Chattering

The issue of alarm chattering, as addressed by Wang et al. (2014) and others, is a critical challenge in alarm management. Our proposed work incorporates techniques to minimize chattering alarms by optimizing alarm parameters such as on-off delays and hysteresis ratios. By reducing the occurrence of chattering alarms, our approach aims to improve the overall reliability and effectiveness of the alarm system.

Wang et al. (2014) focused on detecting and removing chattering and repeating alarms. Chattering alarms are defined as alarms that stay very short in alarm conditions (Short means below the limit of an operation reaction time for an operator according to ISA standards). Repeating alarms are those with the same or close inter-arrival times between occurrences. They created two rules to detect chattering and repeating alarms.

- If the alarm duration or interval is less than 20s (minimum reaction time), the chattering alarm is present.
- If the alarm occurs multiple times regularly, the repeating alarm is present.

Then, they discussed shelving and delay timer operations for their effectiveness in removing alarms. They concluded that delay timer operations are more effective. They created an algorithmic method (m-sample delay timer) to remove detected chattering and repeating alarms online by utilizing their two rules and delay timers. They tested their algorithms with data from a large-scale thermal power plant in Shandong Province, China.

Wang et al. (2016) studied the problem of long-standing alarms in industrial plants. Long-standing alarms are separated from nuisance alarms. They summarized the causes of long-standing alarms. Four statements are presented for their causes:

- An operator should but do not take an action,
- Maintenance staff should but do not take an action,
- The problem cannot be resolved in a short time (for example, a material needs to be cool down),
- Wrong alarm parameters.

Using these causes, they created an algorithm to obtain a state diagram of a plant to detect reasons for a long-standing alarm. Using their state-defining algorithm, they proposed a dynamic state-based alarm system to remove long-standing alarms. They tested their methodology with industrial examples from a power generation unit. They stated that their algorithm is insufficient for the cases in which their third reason occurs. It is stated that when the third reason is valid, a dynamic alarm trip point design should be performed.

Zang et al. (2015) proposed a method to increase the efficiency of alarm delay timers. Alarm-delay timers reduce false and missed alarm rates (FAR and MAR). An alarm is generated when a process value crosses a predefined threshold value (alarm set point). When alarm delay-timers (on-off delay parameters) are used, an alarm is not generated immediately. It waits for a predefined time. After the predefined time, if the process value is still above (or below for low alarms) the set point, an alarm is generated. This process adds delays to alarms with decreasing FAR and MAR. The authors proposed a second threshold value that validates certain occurrences of an alarm. They stated that the drawbacks of alarm-delay timers can be removed by using this value. They tested their algorithm with a simulation of a single process value. They simulated a normal and an abnormal state for the process. The states are modeled as two overlapping Gaussian distributions with predefined set points. Results show that the benefits of the alarm-delay timer are kept with a decreased delay time.

Wang et al. (2015) studied deadbands (hysteresis ratio). They proposed an index derived from normalized alarm duration and deviations to measure performance for reducing nuisance alarms. They also proposed a method to specify deadband width for optimal elimination of nuisance alarms without having a large delay. Their work consists of analyses of an analog signal to specify which technique is more suitable: deadbands, delay timers, and filters. They tested their algorithm with four industrial signals from a large-scale 300-MW thermal power plant. They concluded that when alarm durations are large, and alarm deviations are small, deadbands are more suitable.

3.4. Sequence Analysis

The sequence analysis methods discussed in this section, including the use of clustering algorithms and similarity indices, are relevant to our work in identifying and analyzing alarm patterns. Our proposed method for associating disturbances with alarms leverages similar data-driven techniques to detect and label alarm sequences, facilitating the optimization of alarm parameters based on real-time process data.

Ahmed et al. (2013) dealt with the classification problem for alarm floods. Alarm floods are detected according to ISA standards 18.2. They used the Agglomerative Hierarchical clustering algorithm (Manning et al., 2008) to the unsupervised nonparametric alarm flood data that contains several occurrences of alarms. To apply the clustering algorithm, they created a similarity/distance metric to evaluate the similarity of two floods. A frequency matrix that stores consecutive alarm occurrences and common alarm percentages for two floods calculates the similarity index. Detected similar floods are validated using the Dynamic Time Warping algorithm (Müller, 2007). The proposed work was tested with an oil hydrotreater's real industrial alarm data set.

Cheng et al. (2013) dealt with the problem of detecting similar alarm sequences where alarm floods occur. They proposed a new similarity index between two sequences. Sequences include several alarms that can be defined as tag names and time stamp pairs.

The algorithm for similarity detection is a modified version of the Smith-Waterman algorithm (Smith & Waterman, 1981). The Smith-Waterman algorithm studied sequence alignment problems without having a time stamp. The authors proposed a new similarity metric that considers the equality of tags and the time distances of an alarm to other included alarm types (time distance vector) in a sequence. Using a Gaussian kernel function, they converted the time distance vector to a time weight vector used to calculate positional similarity. Then, an alignment cost matrix is built to capture the best alignment sequence. The similarity index is measured as the maximum value in the matrix for the best alignment sequence. After having a similarity metric, all captured alarm floods are compared, and a flood similarity map is constructed. Then, clustering for similarities is completed to detect similar alarm flood patterns. The proposed algorithm's results on actual alarm data were compared with the original Smith-Waterman algorithm's results. Better results, according to operators, were obtained with the proposed method.

Hu et al. (2015) focused on detecting lagged correlated alarms. Lag is defined as occurrence and correlation delays. Correlation delay is the time difference between two correlated process signals. Occurrence delay is the time difference between two alarms generated from separate correlated process signals. Before detecting the correlation between the two alarms, they cleared chattering or nuisance alarms with an m -sample delay timer. Then, they obtained lag values with the proposed occurrence delay-based estimation algorithm. After obtaining the lag value, they synchronized alarm times with shifting lagged alarms with obtained lagged values. Then, the alarm occurrences are converted to a Gaussian distribution with parameter h , which is optimized according to two rules: separation of distinct alarms and maximum value close to m (delay timer). After this conversion, alarm distributions are compared by their correlations. A list including highly correlated alarms is created to support operators. They gave the number of alarms a sequence should include for each extracted pattern to have a reliable correlation test. They tested their algorithms with data from a refinery plant in Alberta, Canada.

Lai & Chen (2017) dealt with the alignment problem for multiple alarm sequences that are time-stamped and tagged messages. It is an extension of the work proposed by Cheng et al. (2013). After detecting similar sequences, the authors stated that alignment should be performed to have summarized sequences. The first elements in a sequence can indicate the occurrence of a whole sequence. Cheng et al. modified the Smith-Waterman algorithm by adding time stamp information. As a result, the previous work could align two sequences; however, it does not apply to more than two alarm sequences. The authors generalized their algorithm to work with more than two alarm sequences in this work. They applied their algorithms to actual petrochemical plant data and obtained aligned sequences. They used operator approval to validate their results because there is no other work to compare in the literature.

Hu et al. (2018) focused on the detection of frequent alarm patterns as sets of alarms (not sequences). Firstly, they preprocessed alarm logs to get rid of nuisance alarms. Then, they marked flooding cases using ISA standards. Marked alarms were converted to alarm sets that are called alarm transactions. From these transactions, they detected closed alarm sets

with maximum support value compared to larger and covering sets. Finally, a representative alarm pattern algorithm was proposed and applied. Results were visualized using three plots: alarm burst plot, alarm flood spiral graph, and alarm pattern color map. The proposed methodology was implemented with MATLAB and applied to a real industrial case.

Dorgo & Abonyi (2018) proposed a methodology to create alarm suppression rules to reduce the number of alarms in an industrial system. Firstly, they created an event table from alarm logs. An event can be exemplified when the duration of a sensor reading is higher than the threshold value for a high alarm. Then, four relations between events are defined to create alarm chains: equal, before, during, and overlap. These chains are analyzed to state root causes with supervised and unsupervised machine learning methods. Decided chain starter alarms are used for suppression. Finally, they tested their algorithms by simulating a water treatment testbed in MATLAB.

Vougel-Heuser et al. (2015) dealt with the problem of detecting similar flood patterns from historical alarm data. Their novel approach is to combine predefined criteria with pattern recognition algorithms. They made one base and three derived hypotheses to define their problem. The base hypothesis is the existence of causalities between notifications. Their pattern recognition algorithm uses the frequency of occurrence to detect patterns. Their algorithm starts by combining two notifications and combining others iteratively controlling frequencies. The output of their algorithm is notification patterns occurring sequentially. They tested the proposed work with eight existing industrial automated production systems (aPS). Results were assessed by industrial experts and found successful.

Shang & Chen (2019) focused on the early classification of alarm floods. At first, they separated alarm flood data from alarm and event (A&E) logs. Then, two arrays are constructed to define a flood. The first array contains binary data to state whether an alarm type is in the flood. The second one contains relative time distance values of the first alarm of an alarm type to start the flood. These two arrays were used to classify a flood using a weighting algorithm with an exponential character. Exponentially weighted time distance values enabled a better early classification concerning k-NN and ACM-NN. They tested their algorithm with simulation data from the Tennessee Eastman process (TE).

Xu et al. (2020) dealt with the problem of alarm sequence prediction. They used the Smith-Waterman algorithm (Smith & Waterman, 1981) to calculate sequence similarities between two alarm floods. When a recent flood occurs, the current alarm sequence is extracted, and similarities are calculated. Then, they obtained a descending order of similarities for previously obtained floods. After that, sub-sequences of the current flood are extracted, and probabilities of upcoming alarms are calculated. A Bayesian estimator calculates the posterior probabilities and their confidence intervals. The authors stated that the usage of sub-sequences for the calculation of probabilities resolves the problems of varying lengths of sequences. They tested their algorithm using data from the Tennessee Eastman (TE) benchmark process.

Shirshahi et al. (2024) dealt with the problem of detecting similar alarm sequences to identify root causes of faults. They used transfer entropy as a criterion to detect similarity. They also used process variables for similarity detection. To be able to work with large-scale plants, they realized multi-sensor information fusion. They tested their algorithm using data from the Tennessee Eastman (TE) benchmark process.

3.5. New Alarm Types

The introduction of new alarm types through multivariate analyses, as explored by Kondaveeti et al. (2009) and others, aligns with our goal of enhancing alarm system performance. Our work extends these concepts by developing a comprehensive alarm simulation platform that includes various alarm parameters and types. This platform enables the testing and validation of our optimization methods under different scenarios.

Kondaveeti et al. (2009) focused on multivariate analyses of alarms. They introduced new alarm types extracted from multivariate analyses of multiple process variables. They measured T square and Q statistics of each pair of process variables to detect broken correlation case and false alarms. For validation of results obtained by mentioned techniques, they labelled faults in a given data and calculated a confusion matrix for each selected pair and for threshold values used for single process variable. Then, they built a ROC (Receiver Operating Characteristic) curve to detect ones that have the highest performance metrics (the lowest missed and false alarms). They provided two simulation examples to illustrate their methodology.

Chenaru et al. (2016) deals with the problem of extracting rules that identify faulty cases. The authors used simple statistical models to create alarms. For univariate analyses, they analyzed variance of a variable looking at first three sigma levels. For multivariate analyses, correlations of variables are analyzed. PCA, FDA and DPLS techniques are used for fault diagnosis in multivariate analyses. They extracted and tested their rules obtained from their algorithms using simulation data created with MATLAB.

Wang et al. (2016) studied a risk-based warning system that can be used for alarm generation. They seek operating points with looking at process value distribution according to frequency. Operating point can be defined as a set value set by process engineers. The set point is a target value specified for a process variable. When process value is plotted against its frequency, it can be seen that higher frequencies can be obtained near set points if process proceeds in its normal (desired) conditions. Moreover, if there exist state transitions between set points, on-the-way virtual state can be observed. The authors modeled these set points and on-the-way states using Gaussian Mixture Model. Then, they gave weights about criticality of an instant to regions on process value – frequency graph. The risk of a time instant equals to the weight of the region the time instant occurs. They evaluated labeled alarms created on a stirred tank reactor. Evaluation was done according to comparison between measured risk values and actual/false alarm identification.

Perdisci et al. (2006) dealt with the problem of alarm clustering and fusion for computer networks. Intrusion Detection Systems (IDSs) classify network state and produce alarms to network administrators when a predefined intrusion is detected. To be able to maximize the performance of detection, multiple IDSs are used together. When multiple IDSs are deployed, many alarms indicate the same intrusions. This case results in alarm flooding for computer networks. The authors proposed a method to describe different IDSs alarms in same features and classify them. Alarms belonging to the same classes are fused together and presented to the network administrator as a summary report. When an alarm cannot be put in a class, a new class is created. Classification is realized with nearest neighbor algorithm. The algorithm is tested with a live network containing dozens of hosts. Three attack types are analyzed: portscan, webscan and DoS.

3.6. Indicator Indexes

The use of indicator indexes to detect alarm floods and other anomalies, as proposed by Wang et al. (2018) and Charbonnier et al. (2015), informs our approach to KPI selection and optimization. Our work incorporates similar metrics, such as the number of alarms and alarm-on times, to evaluate and optimize alarm system performance. By integrating these indicators into our cost function, we aim to achieve a balanced and effective alarm management system.

Wang et al. (2018) dealt with the problem of detection of alarm floods. Two metrics are specified as previously used for detection of alarm floods: number of alarm occurrences and number of variables in alarm state. This paper proposes a third metric that indicates a newly encountered alarm. The reason for specifying this third metric is false detection of alarm floods. The authors stated that reasons for false detection are chattering and long-standing alarms. Existing metrics are not sufficient for getting rid of these cases. They proposed an algorithm using the proposed metric for detection of alarm floods. The proposed work is tested with data coming from a large-scale power plant at Shandong Province in China.

Charbonnier et al. (2015) proposed a dissimilarity index to compare alarm floods. The proposed index consists of weights of alarms that strongly indicate a specific fault. Extracting weights is realized by filtering out alarms that occur in separate fault types. Analyzing multiple occurrences of a fault type, an alarm is linked to a fault. If the link between the alarm and the fault is strong, the alarm is considered as an indicator of a fault and its weight increases. After having weights for each alarm, similarity is measured as scaled sum of common alarm weights. Time stamps in each alarm are not considered for the proposed work. They tested their weighted dissimilarity index with both artificial and simulated datasets. The simulated dataset is obtained from the CERN Large Hadron Collider process simulator (SCADA System). It was seen that the proposed method outperforms other indexes obtained from literature (Hamming Distance).

3.7. State Based Alarming

The concept of state-based alarming, as discussed by Hu et al. (2018), is relevant to our work in optimizing alarm parameters for different process states. Our proposed method considers the dynamic nature of process conditions and aims to optimize alarm parameters accordingly. This approach ensures that alarms are timely and relevant, reducing the likelihood of missed or false alarms.

Hu et al. (2018 - May) focused on relations between alarms and mode changes. Modes can be described as the resulting state of operations. The authors assumed the existence of event logs that have information about time and description of operational changes. Firstly, they defined time windows to analyze the relation between an alarm and an operational change. Then, they generalized this algorithm to analyze the relation between an alarm and multiple operational changes. If a captured relation's confidence, which is number of times same relation exists, is greater than a predefined number, then relation is accepted as valid. Algorithms to realize the above-mentioned process are given with mathematical formulations defining window lengths and threshold values. The proposed work was tested with data that is obtained from an oil plant. Results that were represented as relations were confirmed with process knowledge.

3.8. Operator Action Analyses

The analysis of operator actions in response to alarms, as explored by Hu et al. (2016), underscores the importance of timely and accurate alarm generation. Our proposed work enhances this aspect by optimizing alarm parameters to minimize reaction delays and improve operator response times. By providing operators with more reliable and actionable alarms, we aim to enhance overall process safety and efficiency.

Hu et al. (2016) focused on operator action discovery with respect to a univariate alarm. It is stated that operator actions are stored in alarm and event logs (A&E). The authors proposed a methodology to extract operator actions for an alarm by fetching these logs. Specifying actions for each alarm type helps record operator expertise and process knowledge. Obtained knowledge can be used for training new operators. They created a generic state transition table for alarms. With using alarm occurrence time, event logs are segmented and necessary operations combined with alarms. They used Petri net models to create state transition tables. The proposed methodology was tested with data coming from an oil pipeline system.

3.9. Literature Gaps

The gaps encountered while searching for answers to the research questions after the literature review can be summarized in the following items.

1. First of all, we can say that most of the studies are generally divided into two types. In the first type of study (New Alarm Types, Indicator Indexes, Alarm Chattering), the effects of new alarm parameters that are structured differently on the alarm system are examined. In the second type of study (Sequence Analysis and State Based Alarming), the focus is on alarm filtering and alarm flood or alarm pattern prediction in an existing alarm system with more advanced analytical methods. We can express this situation as the first type being quite simple and the second type being quite complex. It is quite difficult to find a systematic solution in the middle of these two groups of study. We see that the most convergent studies (Alarm Parameter Design) are presented using domain knowledge rather than being data-driven. The fact that the solution is based on domain knowledge is an obstacle to generalizability.
2. One of the biggest challenges faced by alarm management systems, as mentioned in the literature review, is the high number of alarms encountered by operators, and as a result, critical alarms are overlooked, and precautions cannot be taken. For this reason, most studies in literature are set out with the aim of reducing this number (All categories except Alarm System Evaluation and Operator Action Analysis). However, the factors that will balance the reduction of the number of alarms have usually been overlooked or have just been based on domain knowledge. This situation, similar to the previous item, prevents the generalizability of the proposed studies. In fact, we see that many studies perform the verification process in a real unit based on the comments of the operators (Park et al., 1997; Wang et al., 2014; Wang et al., 2015; Ahmed et al., 2013; Hu et al., 2015; Vougel-Heuser et al., 2015; Wang et al., 2018; Hu et al., 2018 - May). We can evaluate the lack of a secondary KPI that balances the number of alarms as the second item.
3. Finally, although the Tennessee Eastman Process simulation study is frequently used in literature, we have not encountered a publication describing the process of generating alarm data using the most commonly used alarm parameters from the obtained process data (Zu at al., 2017; Peng at al., 2017; Ragab at al., He at al., 2019; Shirshahi at al., 2024). Although many studies share their own alarm data, they do not explain how they generate this data.

Our solution suggestions for the above-mentioned literature gaps constitute the novel aspects of the proposed study, as stated in the introduction chapter.



CHAPTER 4

DATA-DRIVEN ALARM PARAMETER OPTIMIZATION

The proposed work is explained in 4 sections: KPI's (Key Performance Indicators), Tennessee Eastman Process, Alarm-Disturbance Relation and KPI Implementation. The flow diagram given in Figure 3 explains the relationship between these sections and also the modules that constitute the proposed approach. KPI Calculation and Optimization module is explained in KPI's and KPI Implementation subchapters. TEP Simulation and Alarm Generator modules are explained in Tennessee Eastman Process subchapter.

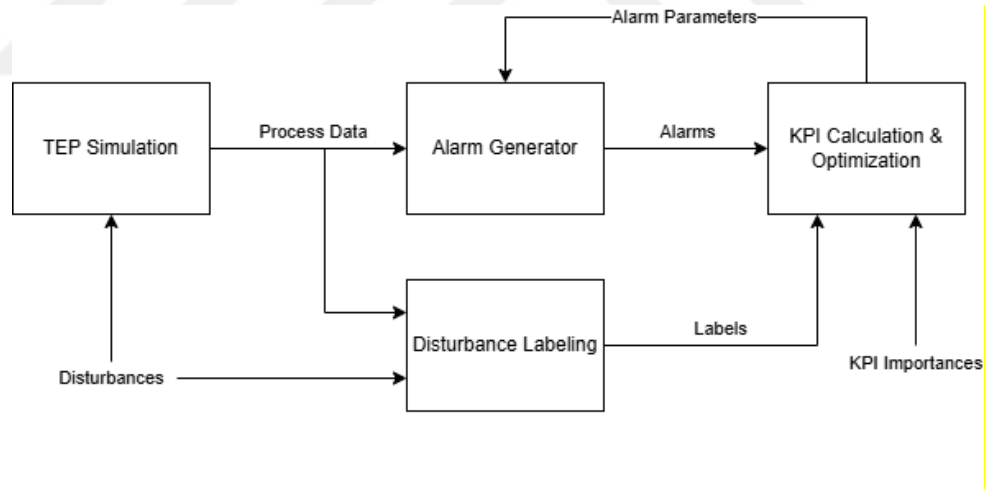


Figure 3: Block diagram summarizing the proposed work

In this study, data was collected under stable conditions and during disturbances. Alarms are then generated from this process data using predetermined sets of the most common alarm parameters. At the same time, the effects of known disturbances on process variables were examined and matching was made between them with a new technique. Then, the most optimal parameter set is determined by taking into account the generated alarms and disturbance matches.

4.1. KPI's

In this work, a novel data-driven alarm parameter optimization method is proposed. Selected parameters are the most common ones existing in almost all Distributed Control Systems (DCS). These parameters are introduced in Chapter 2: threshold value, on-off delay, and hysteresis ratio. The optimization function (or cost function) is constructed using well-known or novel KPIs introduced in this chapter.

Alarm parameter optimization is a routine part of the alarm rationalization process. It is mostly done with domain knowledge, including the opinions of process experts responsible for the systems and vendor specifications. The alarm rationalization process is conducted periodically. In these periodical checks, problematic cases are evaluated. On the other hand, parameter optimization is first done in the design phase, so parameters are kept constant if a problematic case is not observed after the design phase. Changing process conditions, worn-off equipment, and even changing operators (changing experience) are reasons why alarm parameters should be changed. However, continuous and exclusive parameter updates are very costly. The proposed work introduces an autonomous way of alarm parameter optimization.

As stated, alarm parameters are mostly changed to solve problems arising after process changes. However, every alarm management problem cannot be solved with changing parameters (Ahmed et al., 2013; Cheng et al., 2013; Hu et al., 2015; Lai & Chan 2017; Hu et al., 2018; Dorgo & Abonyi, 2018; Vogel-Heuser et al., 2015; Shirshahi et al., 2024). Some advanced techniques are proposed in the literature to capture alarm patterns. On the other hand, it is stated that parameter optimization is a prerequisite for the application of more advanced techniques. Namely, it should be noted that problems that can be solved with changing parameters should be solved with that.

4.1.1. KPI Definitions

Alarm systems can be evaluated with many metrics. Chattering and long-standing alarm ratios, number of alarm floods, number of alarms, and total alarm on times are some of these metrics. Apart from these metrics, more complex ones also exist (Park et al., 1997). Informativeness and effectiveness are some of these complex ones. These complex ones require labeling for every alarm, and as a result, they are costly. The selection of KPIs is very important for automating and optimizing alarm parameters. Using complex KPIs for optimization results in a better-performing alarm management system; however, the optimization cost prevents being completely autonomous. Using simple KPIs results in a low-performance alarm management system. Moreover, some of the KPIs have strong correlations. Considering these factors, the below aggregated values are selected for the proposed work;

- Number of alarms
- Alarm on times
- Alarm reaction delay

The number of alarms and alarm-on times are two metrics that can be easily measured by only having alarm messages. The number of alarms is one of the most basic KPIs. Alarm flooding is detected with the number of alarms for a specific time window. Chattering is one of the reasons for the high number of alarms (Wang & Chen, 2014; Wang & Chen, 2016; Perdisci et al., 2006). Moreover, cascaded relations in a process may increase the number of alarms with some alarm patterns (Ahmed et al., 2013; Cheng et al., 2013; Hu et al., 2015; Lai & Chan 2017; Hu et al., 2018; Dorgo & Abonyi, 2018; Vogel-Heuser et al., 2015; Shirshahi et al., 2024).

The alarm-on time is the time difference between an alarm's on and off state. Alarms that have longer alarm-on times than predefined limits are stated as long-standing alarms (Wang & Chen, 2014). These alarms unnecessarily occupy a place on DCS screens and distract operators. Moreover, this type of alarm is one of the reasons for alarm floods.

The delay time of alarm messages to disturbances is also a critical parameter. Early detection of a disturbance enables operators to take timely actions. This way, problems causing disturbances can be eliminated as quickly as possible without further problems. Alarm reaction delay can be generated by having disturbance flags. Disturbance flags are the indicators of disturbances on a process timeline. It can be obtained from process records or step tests in an actual scenario not simulation. However, both types of information are hard to obtain. Process records mostly have ambiguous starting and ending times of disturbances. Tests that can only be obtained for ongoing processes are simple ones. Unfortunately, complex tests cannot be conducted without stopping the process. This work obtains process data from TEP. Disturbances are manually triggered. So, disturbance flags already exist.

To better understand KPI selection and the relationships between them, we can examine how each KPI will be affected by alarm parameters. First of all, if we evaluate the number of alarms, expanding the threshold values (increasing for high alarm, decreasing for low) and increasing all other parameters (on-delay, off-delay and hysteresis) will reduce the number of alarms. Secondly, for the alarm duration, expanding the thresholds and decreasing off-delay and hysteresis will cause the alarm duration to decrease. Finally, the alarm reaction delay is inversely proportional to the threshold expansion and all other parameters. This situation is summarized in the table below.

Table 1: Relation between alarm parameters and KPI's ("+", "-", and "N" mean positive, negative and no correlation, respectively)

Parameters\KPI's	Number of alarms	Alarm on times	Alarm reaction delay
Threshold	-	-	+
On-delay	-	N	+
Off-delay	-	+	N
Hysteresis	-	+	+

As can be seen in Table 1, the alarm reaction delay balances the number of alarms and the alarm on times for the threshold. For the on-delay parameter, the alarm on times does not have any effect; but the number of alarms and the alarm reaction delay have a balance. Similarly, for off-delay, there is a balance between the number of alarms and the alarm on times. Lastly for the hysteresis, the number of alarms balances the alarm on times and the alarm reaction delay. As a result, the selected KPIs balance the selected parameters and this is necessary to create a healthy optimization structure.

4.2. Tennessee Eastman Process (TEP)

Tennessee Eastman Process is defined as: "Medium-large scale nonlinear dynamic simulator of an actual plant and processes of the Eastman Chemical Company. Actual TEP is a process at an Eastman Chemical Company plant located in Tennessee. The company was a subsidiary of Kodak, and it spun off in 1994. Their focus is on chemicals, fibers, and plastics. The first version of the simulation was written in FORTRAN in 1993 (Downs & Vogel, 1993). The current version used to validate the proposed work is a revised version written in Matlab/Simulink in 2015 (Bathelt et al., 2015).

4.2.1. TEP Components

The process shown in Figure 4 requires four inputs: A, C, D, and E. As a result of the reactions, two products (G and H) and one byproduct (F) are generated. There is also an inert gas B, which is included in inputs. After the reactions of TEP, G, and H are collected at the output layer. B and F are purged during the process. The components to realize this process are introduced below;

- One two-phase chemical reactor: Exothermic reactions take place in the reactor.
- One condenser: It condenses the outputs of the reactor.

- One vapor-liquid separator: It separates byproducts and inert gases from the condenser.
- One stripper: Some raw materials are not involved in any reaction at the reactor's output. These materials can be used to feed the reactor. Stripper separates them using C.
- One reboiler: It is utilized to reboil the stripper content when necessary.

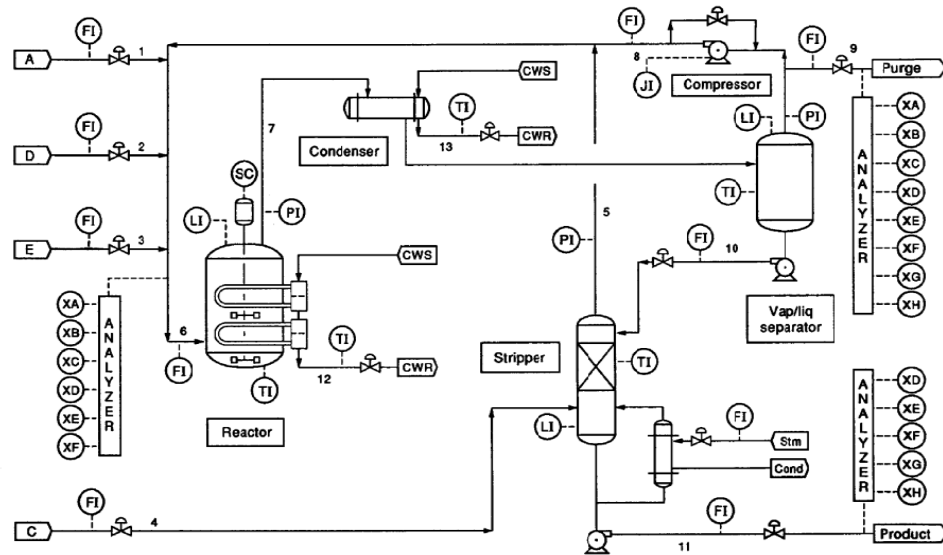
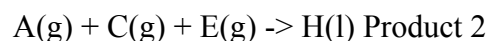
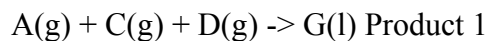


Figure 4: Tennessee Eastman Process (Bathelt et al. 2015)

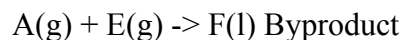
The whole process contains 73 process values (PV). These process values have the following types: flow, temperature, level, and chemical component concentration. The first version allows us to monitor only 41 of them. The current version enables us to see all 73 of them. Moreover, 12 manipulated variables exist (MV). These variables mostly represent valves, and they should be controlled in a closed loop to have TEP in a stable condition.

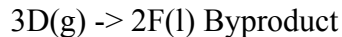
4.2.2. Chemical Equations

2 simultaneous gas-liquid exothermic reactions



2 additional byproduct reactions





F is not a waste or inert; but it is not used in the current process. So, it is behaved as inert or waste and separated at separator.

Also, there exists an inert gas B. It is also separated by separator.

TEP is highly unstable. Manipulated variables should be controlled dynamically to be in a safe state. An example scenario exists in the scripts where all manipulated variables are kept constant. In that scenario, an emergency shutdown happens at about 1 hour. An emergency shutdown state can be reached when some critical sensors' high or low limit values are crossed. For the mentioned case, the reactor pressure limit is reached (3000 kPc).

4.2.3. Process Summary

1. First, A, D, and E are fed to the reactor.
2. Then, an exothermic reaction takes place. Because it is exothermic, it should be cooled down. Cooling water supply is used for that purpose.
3. Reaction is realized with the help of catalysts. It is said that the catalyst is nonvolatile, which means that it behaves the same during process simulations.
4. After reactions, the reactor product stream is cooled in a condenser, and noncondensed components are recycled back to the reactor feed. At this stage, by-products and inert gas are purged by a vapor-liquid separator.
5. Remaining reactants and products go to a stripper. Stripper removes the remaining reactants using stream 4 (C).
6. It is assumed that products are separated. This final separation is not included in the process simulation.

4.2.4. Implementation

To be able model the above mentioned process, more than 30 ordinary differential equations are used. Two example equations are given below to clarify implementation.

- “The vapor-liquid equilibrium follows Raoult’s Law with the vapor pressure calculated using the Antoine equation”
- “Reaction rates are a function of temperature through an Arrhenius expression”

$$k = Ae^{\frac{-E_a}{RT}}, \quad k = Ae^{\frac{-E_a}{T_B T}},$$

The first expression uses Raoult's Law and Antoine equation to express vapor-liquid equilibrium point. The second expression is utilized for calculation of reaction rates. As can be seen from the equations, it is a function of temperature, and we can say that reaction rate increases with increasing temperature. Having exothermic reactions and increasing reaction rates with increasing temperature make process more unstable.

These examples are for explaining underlying chemistry. These two and about 30 more are used for calculating process states.

Inputs of the main simulation function are values of manipulated variables and all process variables (or states). At each iteration, simulation returns derivatives of the process variables.

4.2.5. Control Strategy

Processes like TEP have predefined modes. These modes are detected as steady state values of process values for targeted production types. TEP has 6 modes. There exist 3 different G/H mass ratios and 2 different production rates. A total of 6 combinations exists as modes. Most of the examples from literature use the first mode.

As it is stated before, TEP is highly instable. To be able to sustain the process in a selected mode, control loops are needed. Loops are designed according to a control strategy. The original paper (Bathelt et al. 2015) puts 5 objectives as a control strategy. We can interpret these objectives as constraints and pieces of objective function.

- Maintain process variables at desired values
- Keep process operating conditions within equipment constraints
- Minimize variability of product rate and product quality during disturbances
- Minimize movement of valves which affect other processes
- Recover quickly and smoothly from disturbances, production rate changes or product mix changes.

The main goal is definitely the first objective. The second objective gives us constraints as equipment limits. The third and fifth objectives are also objectives for disturbance cases. The fourth objective is an indirective one. When PID loops are utilized, most of the time one variable is manipulated to control another non-controllable one. However, an

indirect relation affects other non-controllable variables. As a result, minimum change in manipulated variables is preferred.

4.2.6. TEP Dataset

TEP data is used as a benchmark dataset in many publications (Zu at al., 2017; Peng at al., 2017; Ragab at al., He at al., 2019; Shirshahi at al., 2024).

To validate the proposed work, 1000 hours of steady state and 340 hours of disturbance process data are collected. In the disturbance phase, eight disturbances are introduced. The first disturbance is applied at the 10th hour of 340 hours. The disturbance window size is selected as 10 hours, meaning conditions causing a disturbance are kept unchanged for 10 hours (Some of the disturbances take up to 10 hours to take full effect). After each disturbance, 30 hours of stabilization time is given, because some of the disturbances take up to 30 hours to completely stable. After the last disturbance stabilization time, 10 hours of extra time is given. One hundred samples are collected per hour. 100000 and 34000 data points are collected for steady state and disturbance cases, respectively.

Twenty disturbances are declared in the first version of TEP. The first eight of them are used in the simulation. Disturbance types and changed process variables are given in Table 2.

Table 2: TEP disturbances

Disturbance Num.	Process Variable	Type
IDV1	A/C feed ratio, B composition constant (stream 4)	Step
IDV2	B composition, A/C ratio constant (stream 4)	Step
IDV3	D feed temperature (stream 2)	Step
IDV4	Reactor cooling water inlet temperature	Step
IDV5	Condenser cooling water inlet temperature	Step
IDV6	A feed loss (stream 1)	Step
IDV7	C header pressure loss, reduced availability (stream 4)	Step
IDV8	A, B, C feed composition (stream 4)	Random variation

4.2.7. Alarm System Design

Although many publications in literature use TEP to obtain alarm data, I have not encountered a clear method on how alarm data is generated from this data after obtaining process data. For this reason, the alarm data production process has been studied in detail in the proposed work and an alarm data production infrastructure has been setup in which the parameters we suggested could be used in the most appropriate way.

After collecting process data, an alarm system is implemented to have alarm logs. Each alarm log (or alarm message) consists of 4 parts: process tag name, alarm type, alarm state, and alarm time. These parts can be accepted as the most common parts of every alarm system log. Additional information can be available.

Each alarm log is not presented to operators. Alarms that are ‘on-state’ are shown on DCS screens. However, all alarm databases include other logs, and some of the KPIs, like alarm-on time, can be generated with this information.

Depending on alarm parameters, alarms have states. On-and-off states are the main states. Moreover, on/off delay states exist as transitions between main states. The state transition diagram is given in Figure 6.

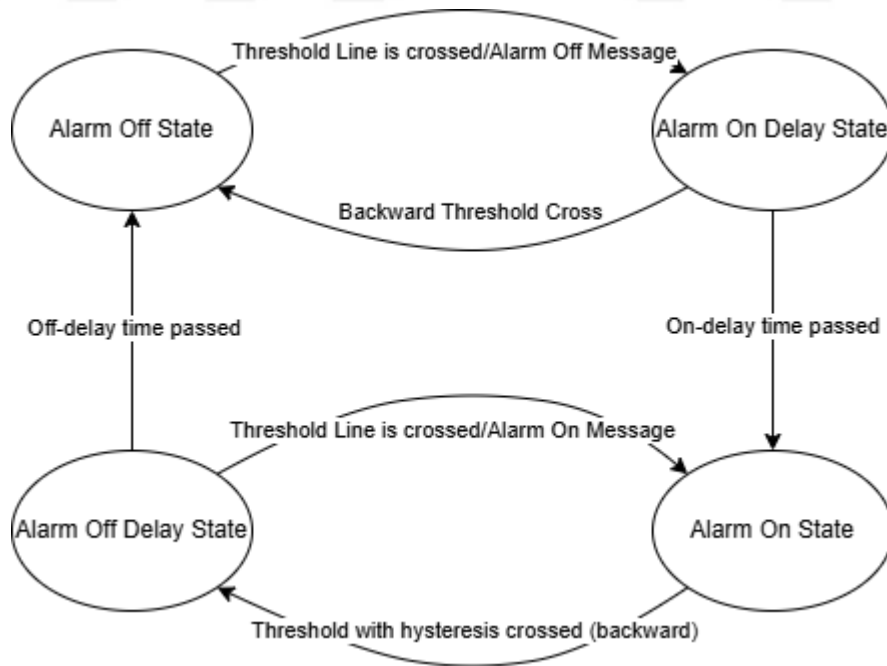


Figure 5: Alarm state transition diagram

The simulation starts from the alarm-off state for each alarm, and alarm logs concerning the given state-transition diagram are collected. Alarm logs are generated when entering alarm-on and alarm-off states.

4.3. Alarm – Disturbance Relation

In the proceeding subsection, alarm parameters are optimized with defined KPIs. However, one of the proposed KPIs (Alarm reaction delay) depends on reactions of process values to disturbances. Not every process value needs to be affected by every disturbance. A disturbance may be generated in an independent part of the process. Process values can be taken from a point the disturbance does not affect. So, process value & disturbance relations should be studied for correct optimization. For example, condenser cooling water disturbance (IDV5) does not affect the reactor. When IDV5 is applied, there will be no change in reactor temperature or pressure. As a result, alarm parameters for reactor temperature and pressure should not be optimized according to the performance of the IDV5 application. Moreover, a disturbance may be detectable only with high or low alarms. For example, if process values show only high spikes, optimization of low alarm parameters will be misleading.

A data-driven method is utilized to label disturbances for each alarm. First of all, steady-state statistics are calculated for each process value. These statistics include maximum, minimum and standard deviation values in a selected window. Then, comparisons are made for disturbance data with same-sized time windows. Window ending times are marked according to the diagram given in Figure 7.

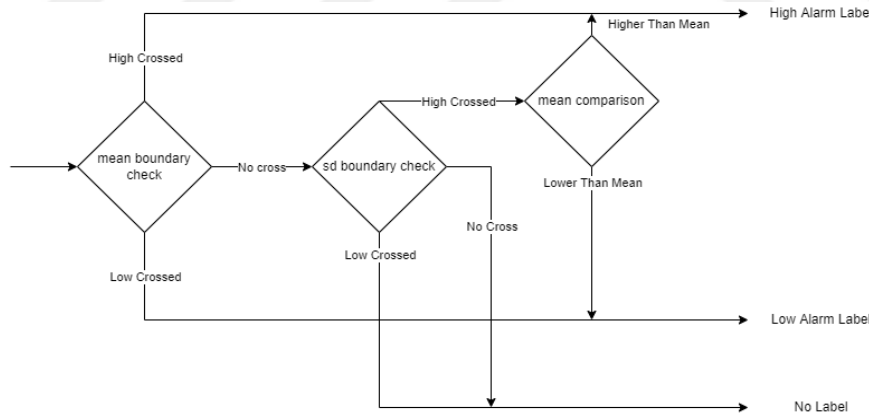


Figure 6: Disturbance Labeling

Boundary crosses are directly labeled as alarm labels. If the standard deviation increases and the window mean is higher or lower than the steady-state mean, the point gets high and low alarm labels, respectively. This is because PID loops may solve problems caused by the disturbance, and a steady-state mean can be achieved quickly (shorter than the selected time window). However, during stabilization, standard deviation boundaries will be crossed, and mean drift will indicate the type of alarm label.

According to the above rules, all disturbances are labeled with process value reaction delays for each process value. An example of labeling and process value graphs are given in Table 3, Figure 8 and Figure 9.

Table 3: Example Alarm – Disturbance Labeling for s1

IDV Index	IDV Delay - Low	IDV Delay - High
IDV1	46	56
IDV2	Inf	286
IDV3	Inf	Inf
IDV4	Inf	Inf
IDV5	3989	Inf
IDV6	1	1009
IDV7	Inf	Inf
IDV8	154	1303

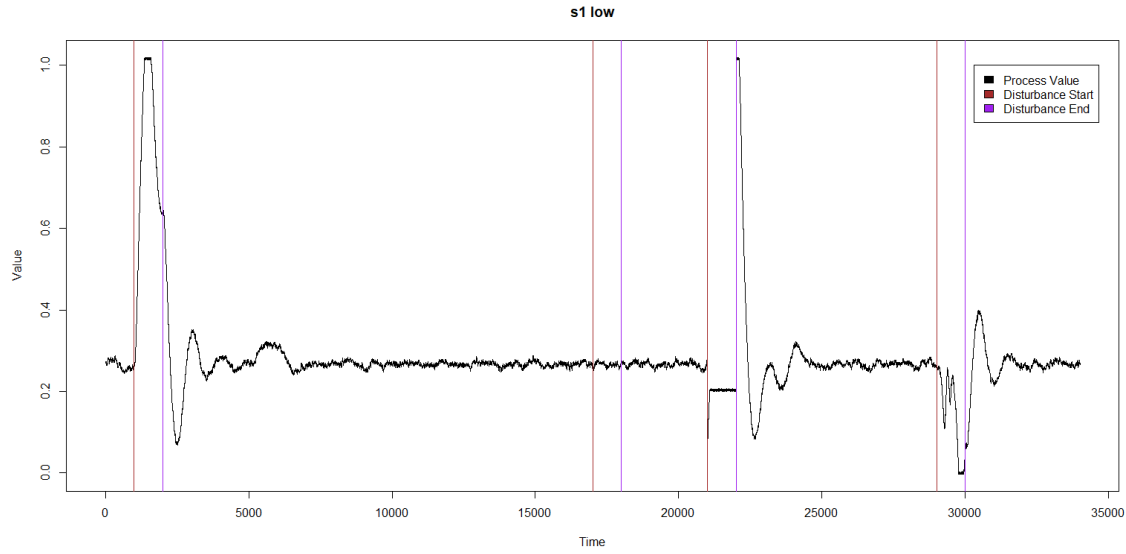


Figure 7: Disturbance process value graph for s1. 1, 5, 6, and 8 disturbances are determined as valid disturbances for low alarm, and they are shown with vertical lines on the graph.

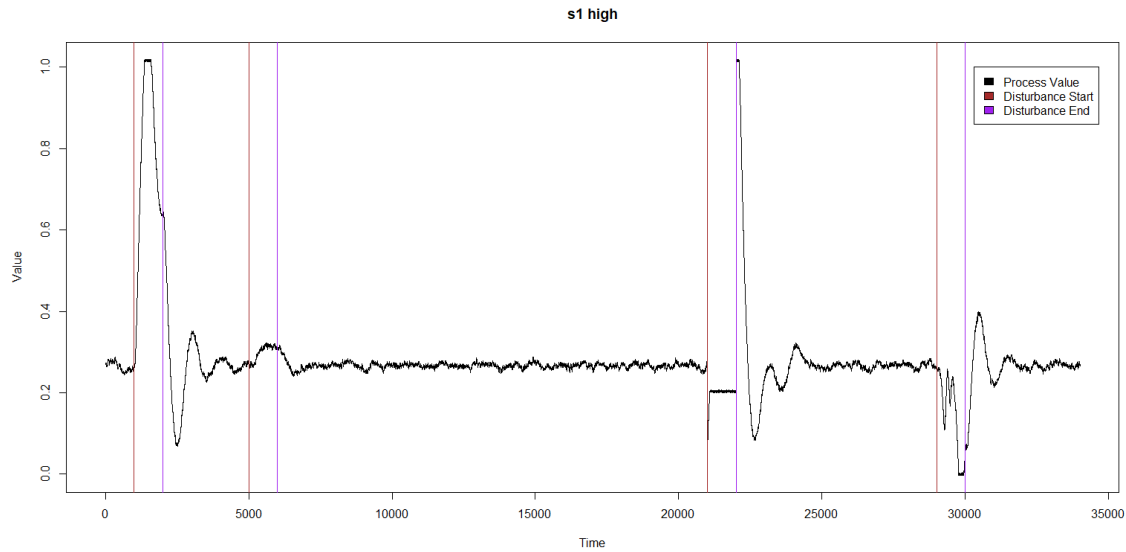


Figure 8: : Disturbance process value graph for s1. 1, 2, 6, and 8 disturbances are determined as valid disturbances for high alarm, and they are shown with vertical lines on the graph.

As a result of the alarm-disturbance labeling mechanism, some alarms have no relation to any disturbances. This may be caused by applying insufficient disturbances, or process value may not depend on process changes. For example, the outdoor temperature does not depend on any internal process condition. The number of alarms and alarm-on time KPIs are maximized for that cases, and alarm reaction delay is NA.

4.4. KPI Implementation

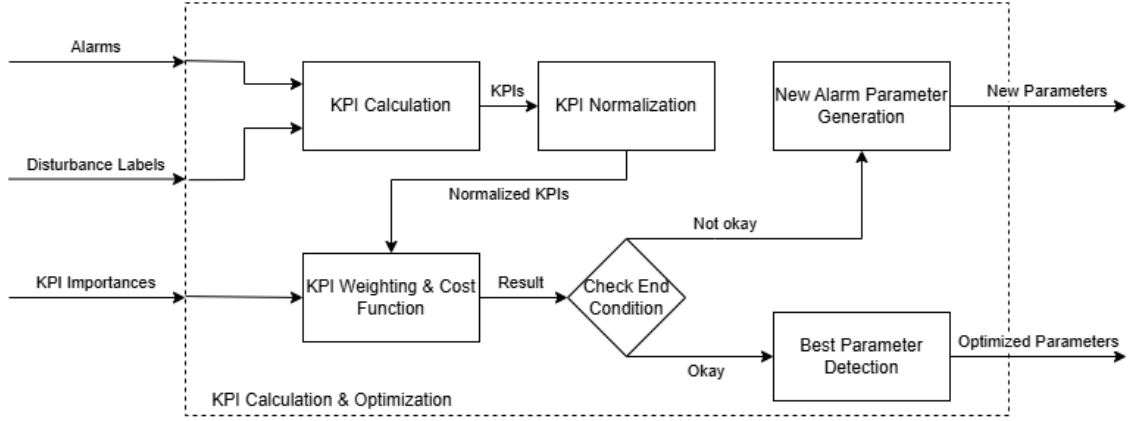


Figure 9: KPI Calculation & Optimization

The flow diagram for KPI implementation is given in Figure 10. KPIs are calculated after alarms and disturbance labels arrive. Normalization, which is explained in Section 4.4.1 below, is completed by including previously calculated KPIs. Then, the cost function is calculated by weighting it with importance values. The obtained values are compared in line with the acceptance conditions, and it is decided whether the next iteration will be performed or not. It checks whether all parameter sets with a certain similarity to the parameter set that produces the best result in the acceptance condition have been tried. In addition, reaching a predetermined percentage of all possible parameters is also an acceptance condition requirement. Parameter similarity for the results obtained includes one high and one low value for each parameter. In this way, it is guaranteed that the result is locally optimum. The completion percentage of all parameters was determined as 20% for the results at hand. If a new iteration is decided, the parameters are determined and sent to the alarm generator module. When determining the new parameter set, the parameter sets near the best result in the current iteration are checked. If all close possibilities have been tried and the specified percentage of all options has not been reached, the iteration continues by trying a random set of parameters. If the termination conditions are met, the run is completed with optimized parameters. Detailed explanations about the mentioned modules are given in the following sections.

4.4.1. KPI Normalization

Optimization of KPIs conducted concerning the disturbance dataset. Boundaries of KPIs are given below;

$$0 \leq \text{Number of alarms (AN)} \leq (\text{Number of points a process value is read})/2 \quad (1)$$

Where, $(\text{Number of points a process value is read})/2 = 17000$

$$0 \leq \text{Alarm on time (AT)} \leq \text{Total disturbance time} \quad (2)$$

Where, total disturbance time = 34000

$$0 \leq \text{Average alarm reaction delay(ARD)} \leq \text{Alarm reaction window size} \quad (3)$$

Where, alarm reaction window size = 4000

To be able to have a minimization formulation, KPIs are normalized according to the below formulation;

$$\text{norm}(x) = 1 + \frac{1}{1 + \frac{x}{\bar{x}}} \quad (4)$$

$\bar{x}$ is the mean of obtained values for each alarm parameter set. $\text{norm}(x)$ is the normalization function.

Using the above formulation, a normalized KPI's minimum and maximum values are equated to 1 and 2, respectively. The reason why the range is chosen as 1 to 2 instead of the standard 0 to 1 is that optimization is determined as the maximization of the cost function. In addition, since the importance values in the definition of the cost function are defined as an exponential function, the minimum value must be 1 or above. There is no particular reason why the maximum value is chosen as 2. As long as it is a constant and greater than 1, the use of the cost function is not affected.

4.4.2. KPI Weighting & Cost Function

After normalizing the KPI values, the cost function below is implemented.

$$\text{cost} = (AN')^a \times (AT')^b \times (ARD')^c \quad (5)$$

a, b, and c are exponential weights of KPIs. AN' , AT' and ARD' are normalized values of the KPIs.

The cost function can also be implemented in an additive manner. However, by doing that, one of the KPI costs may lose meaning by having very good values with the other two metrics. It is aimed at having the best results for each of the selected KPIs. So, a multiplicative equation is chosen.

Weights can be calibrated according to changing objectives. If the number of alarms is aimed to be low, selecting a higher value for 'a' would be more suitable. In the simulation of the proposed work, alarm reaction delay is accepted as the most important KPI. As a result, hence we give more weight to it by assigning weights to;

$$a = 1, b = 1 \text{ and } c = 2$$

4.4.3. Parameter Boundaries

Parameters are discretized within selected boundaries to select the best parameter set. After that, alarm logs are obtained for each parameter set. Higher values for on/off delay parameters and hysteresis ratio result in higher reaction delay, which is selected as the most important KPI for our simulation. So, the boundaries of these KPIs are kept low. Threshold values are started from the median value of the process value. Higher or lower boundary values are calculated as 10% higher or lower values of maximum or minimum process values for high and low alarms, respectively. This window is separated into N parts for the selection of threshold value. N is selected according to computation power of utilized hardware and performance increment. The utilized parameter set is given below;

On delays in [0, 1, 2, 3]

Off delays in [0, 1, 2, 3]

Hysteresis ratios in [0%, 1%, 2%, 3%, 4%]

For threshold boundaries $N = 21$

In total, 1680 parameter sets are compared for each alarm. There exist 73 PV and 12 MV values for TEP. For each process value, high and low alarms are considered. As a result, 170 alarms are considered with 1680 parameter sets. An example optimization result is visualized below.

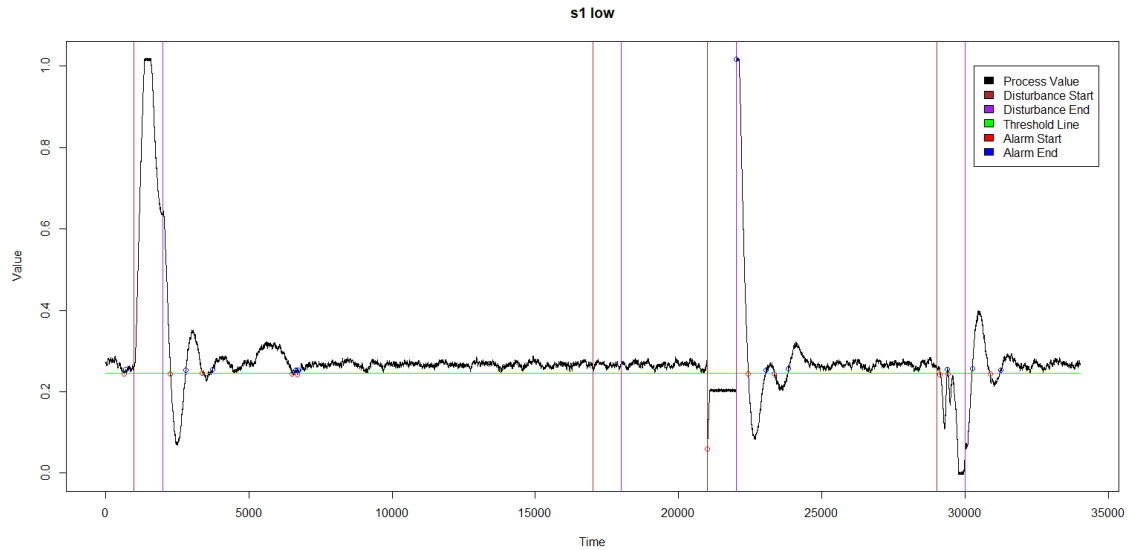


Figure 10: Example optimization visualization for s1 low alarm. Red circles indicate entering an alarm state, and blue circles indicate existing from the alarm state.

Table 4: Example optimization results for s1 low alarm.

Process Value	Alarm Type	Threshold	On Delay	Off Delay	Hysteresis Ratio
S1	Low	0.250772	0	1	0.03



CHAPTER 5

RESULTS AND DISCUSSIONS

For the comparison of the proposed work, alarm parameters for TEP were searched in the literature. However, a complete set of pertinent alarm parameters couldn't be found. The most inclusive alarm parameters are proposed by Manca (2020). Threshold values, EWMA algorithm with given parameters, and hysteresis ratios to generate and filter out alarms on TEP are given. Ma & Li (2017) and Kirchhübel (2020) proposed a subset of alarm parameters for TEP. Manca (2020) started with these parameter sets and manually enhanced them in five iterations using domain knowledge, as stated in the literature review. Although the parameter set is the most comprehensive, some process values do not have alarms.

As stated in Section 4.3, some process values do not depend on or react to disturbances. As a result, alarm parameter optimization could not be performed for those cases.

Comparisons are made for parameters that exist in Manca(2020) and ours. EWMA filters are utilized for the reference results, and ON-OFF delay values are utilized for the proposed work. Hysteresis ratios are used for both. 30 high and 27 low alarms are compared for the number of alarms, alarm-on time, and alarm reaction delay for the eight disturbances. A sample comparison is presented in Table 5 for s1, s9, and s25.

Table 5: Sample table performance comparison

Tag Name	Alarm Type	Source	IDV1	IDV2	IDV3	IDV4	IDV5	IDV6	IDV7	IDV8	Number of Alarms	Alarm On Time
S1	high	Opt. Result	44	NA	NA	NA	NA	NA	NA	NA	8	4292
S1	high	Ref. Result	346	NA	NA	NA	NA	NA	NA	NA	2	470
S9	high	Opt. Result	NA	NA	NA	1	NA	NA	26	NA	5	69
S9	high	Ref. Result	NA	NA	NA	Inf	NA	NA	Inf	NA	0	0
S25	high	Opt. Result	21	NA	NA	NA	NA	NA	NA	NA	14	3070
S25	high	Ref. Result	71	NA	NA	NA	NA	NA	NA	NA	4	1593

Numerical values on IDV and alarm on time columns are delay values from the start of the disturbances. Its unit is 1/100 hour. Opt. Results are optimized results of the proposed work, and Ref. Results are the results for reference. NA values mean that the IDV value is not valid for the alarm. Inf values mean that no alarm is observed during the disturbance.

The sample table shows that the reference parameters result in fewer alarms and fewer alarms-on time. However, alarm reaction delays are much larger compared to the proposed work. As an example, there exists about 3 hours between alarm activations for the s1 high alarm. Moreover, no alarm activation is observed for the s9 high alarm for the reference.

The average values of delays, number of alarms, and alarm-on time comparisons are shown below for a complete comparison. In these comparison figures (Figure 12, 13, 14), optimization is performed with the parameters specified in section 4.4.

The maximum delay boundary is set to 8 hours. Delays of more than 8 hours are set to 8 hours for visualization. The red line represents the reference, and the blue line represents the proposed work.

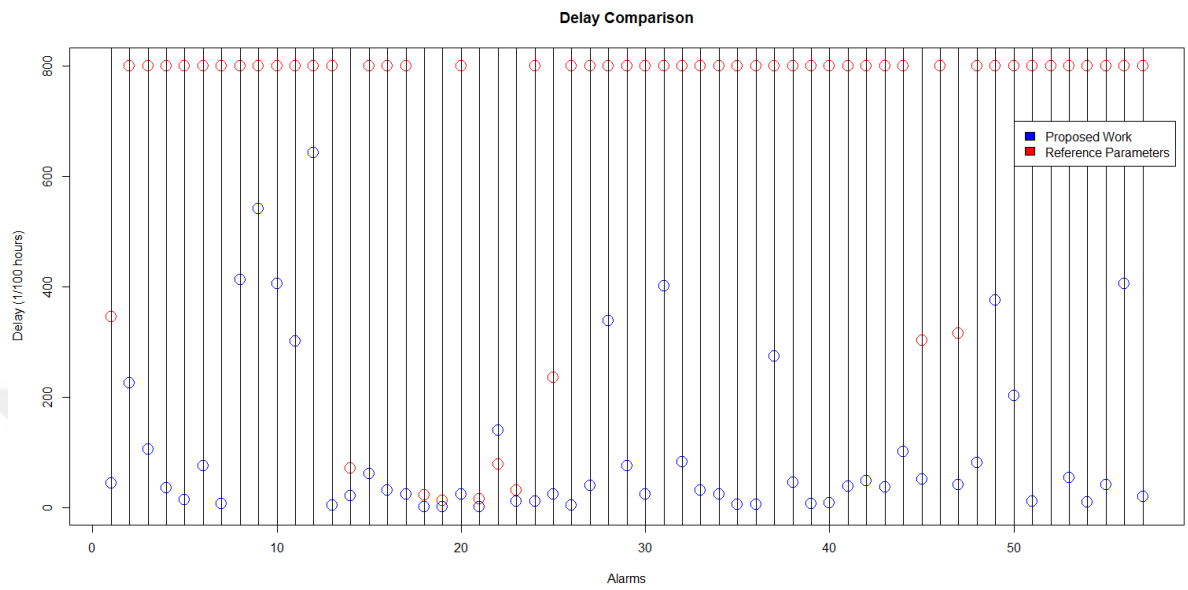


Figure 11: Delay comparison

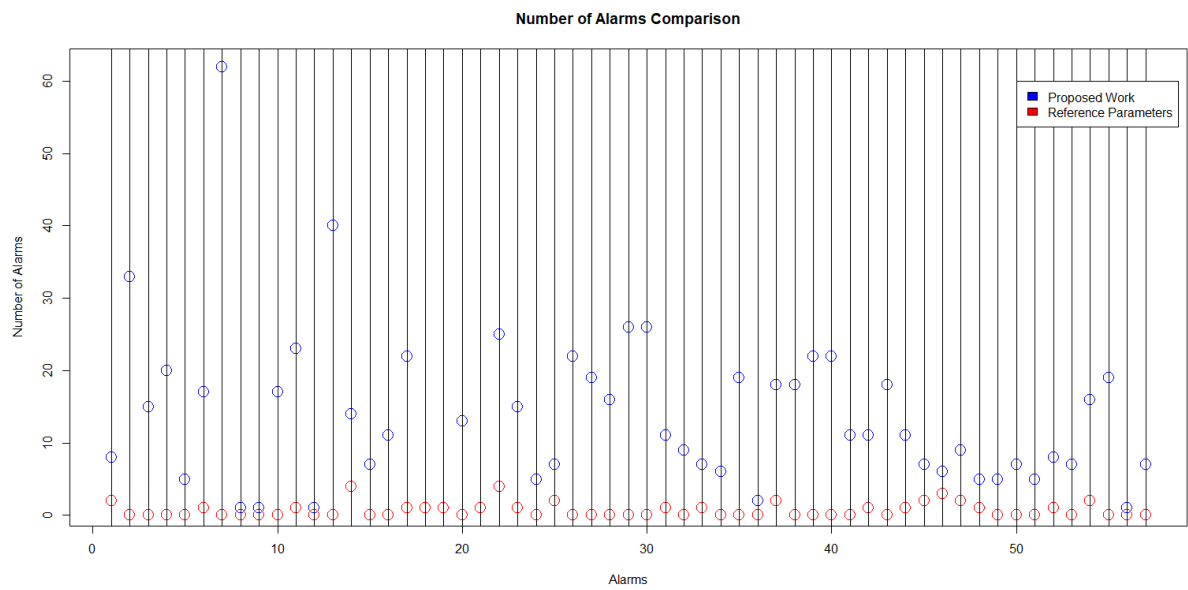


Figure 12: Number of alarms comparison

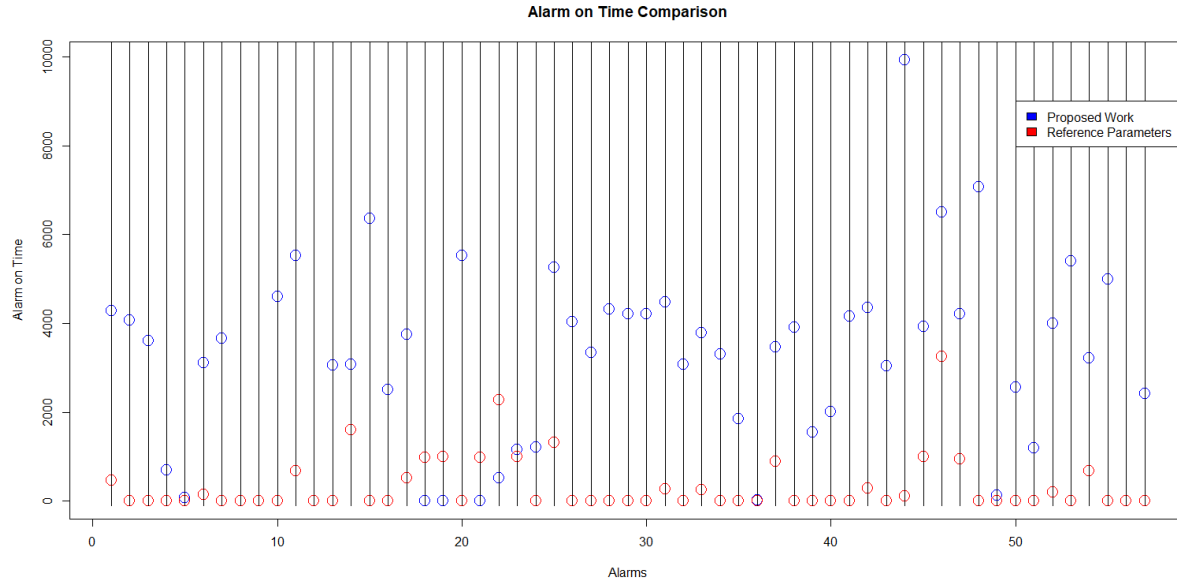


Figure 13: Alarm on time comparison

As can be seen from the figures, the sample table explanations are also valid for complete results. In the proposed work, alarm reaction delay is chosen as the most important parameter, as stated in the KPI weighting subsection. Having fewer alarms may solve the flooding problem in most cases. However, the introduction of large delays prevents operators from taking timely actions. These delays violate the main purpose of alarm systems. If larger delays become acceptable, then alarm reaction delay weight can be reduced for better optimization. To be able to comment on accepting larger delays, minimum delay values of all alarms are shown below for both works.

Table 6: Minimum reaction delays

Source	IDV1	IDV2	IDV3	IDV4	IDV5	IDV6	IDV7	IDV8
Opt. Result	7	11	1	1	1	1	2	76
Ref. Result	71	31	22	12	16	102	Inf	Inf

In order to compare the proposed study with the reference results from another perspective, figures 15, 16, 17 and 18 are given. In these comparisons, the optimization parameters specified in section 4.4 are changed and the weightings of the 3 KPI's were kept the same.

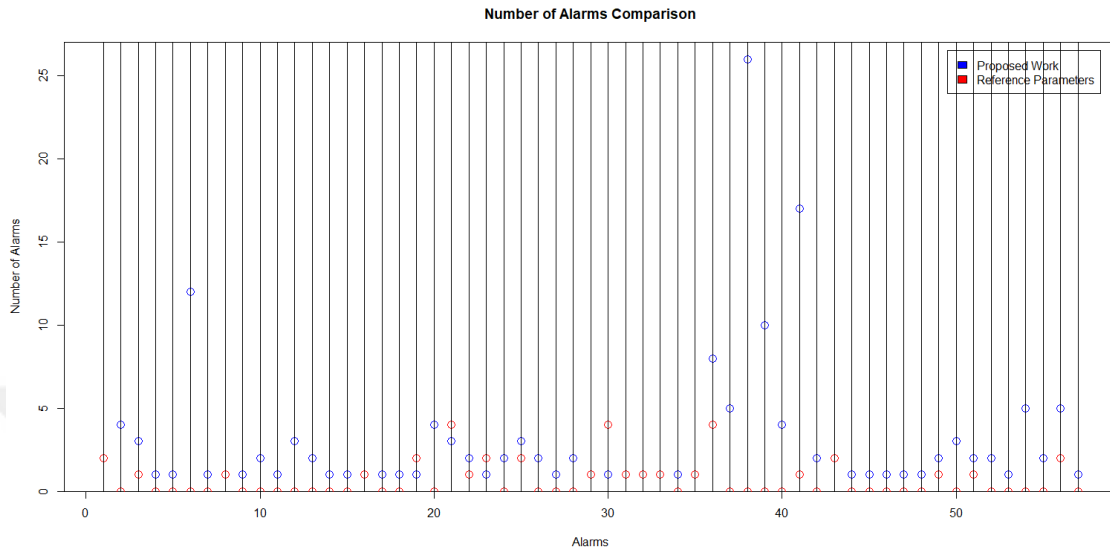


Figure 14: Number of alarms comparison (2nd scenario)

As can be seen in Figure 13 and 15, the reference results did not produce any alarms despite the disturbances on many process parameters. Although these results seem to be a good thing on the graph, in reality, the fact that these alarms are not produced is quite problematic. Of the 57 alarms whose parameters we could access, 36 did not produce any messages, so the operator was not warned about these disturbances. In Figures 16, 17 and 18, these values that did not produce any alarms were not evaluated, only the cases where the reference results created an alarm record were evaluated.

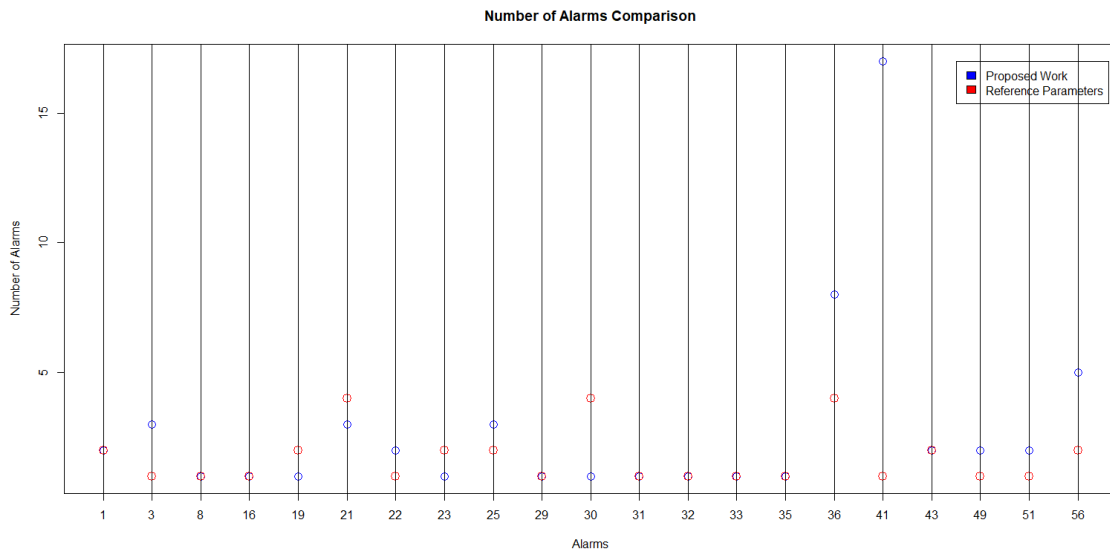


Figure 15: Filtered number of alarms comparison (2nd scenario)

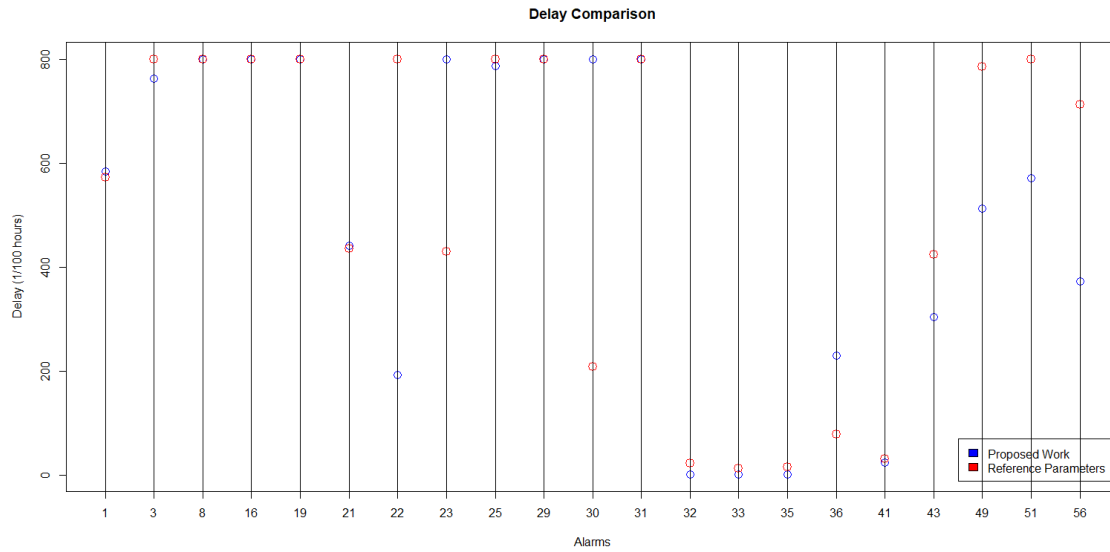


Figure 16: Delay comparison (2nd scenario)

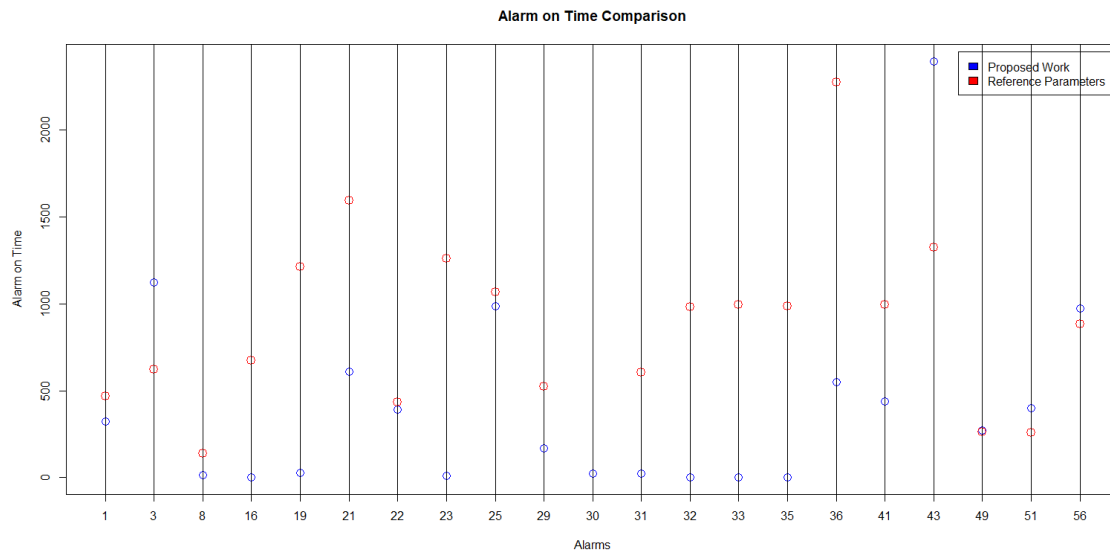


Figure 17: Alarm on time comparison (2nd scenario)

In the second scenario results, as can be seen in Figure 16, the alarm numbers are quite close to the reference results. However, alarm delays give quite good results compared to the reference results for many alarms (Figure 17). Similarly, as can be seen from Figure 18, the alarm on time values have decreased considerably compared to the reference results. Table 7 shows the minimum reactions delays for the second scenario.

Table 7: Minimum reaction delays (2nd scenario)

Source	IDV1	IDV2	IDV3	IDV4	IDV5	IDV6	IDV7	IDV8
Opt. Result	9	24	1	1	1	1	1	101
Ref. Result	71	31	22	12	16	102	Inf	Inf

For each disturbance, below graphs given for a comparison between the proposed work and reference parameters. As can be seen from the graphs, brown and purple vertical lines are given as start and end of the disturbances, respectively. Blue and red points indicate alarms for two parameter sets. The level of alarms shows alarm types. For example, s1 high alarms are plotted at 1 for the proposed work and at -1 for the reference parameter set.

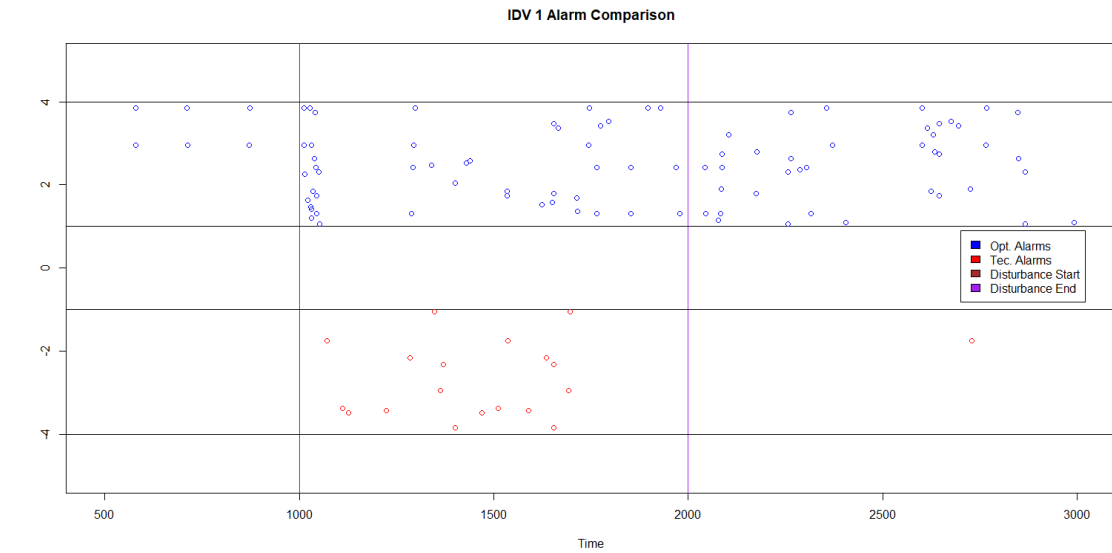


Figure 18: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV1

As can be seen from the graph, the alarms generated by the proposed work intensified immediately after the moment of disturbance. The first of these alarms occurred within 0.07 hours. The first of the alarms created with the help of the reference parameter set came after 0.71 hours. While the reference parameter set alarms come during the disturbance period, very few alarms come during the non-disturbance period. On the other hand, in the alarms created through the proposed work, after the intensity during the disturbance, the alarm intensity continued after the disturbance ended. Extra alarms during this period indicate that the post-disturbance period has not reached stable conditions.

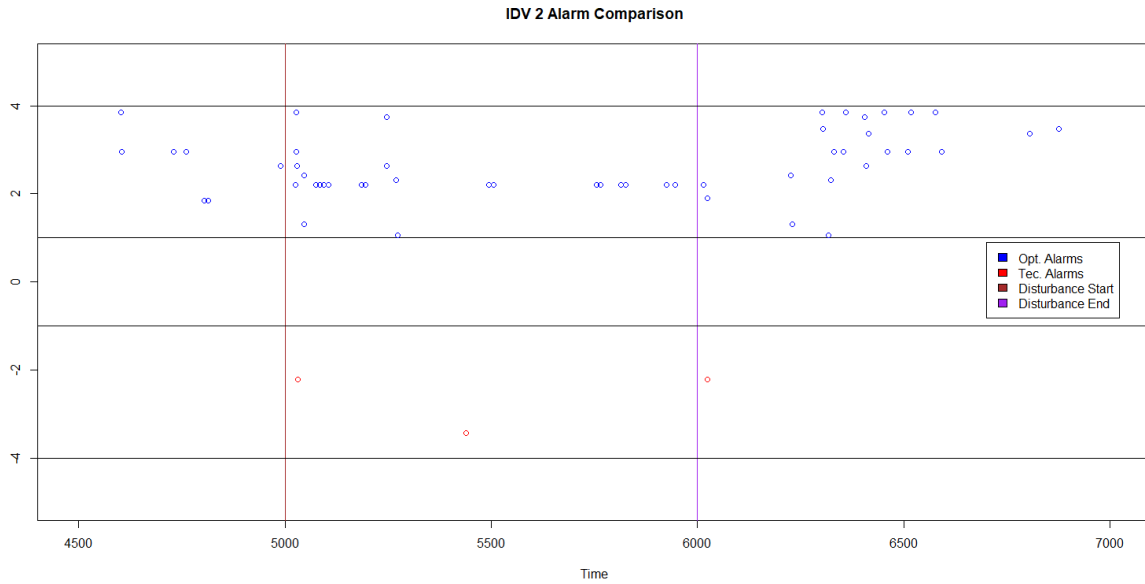


Figure 19: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV2

Similarly, after the second disturbance, alarms for the proposed job come first. Additionally, similar to the first disturbance, the intensity maintains itself after the disturbance, but the alarms that create the intensity change. This change can be detected with more advanced techniques and the alarms of the period that create extra intensity can be filtered with further studies.

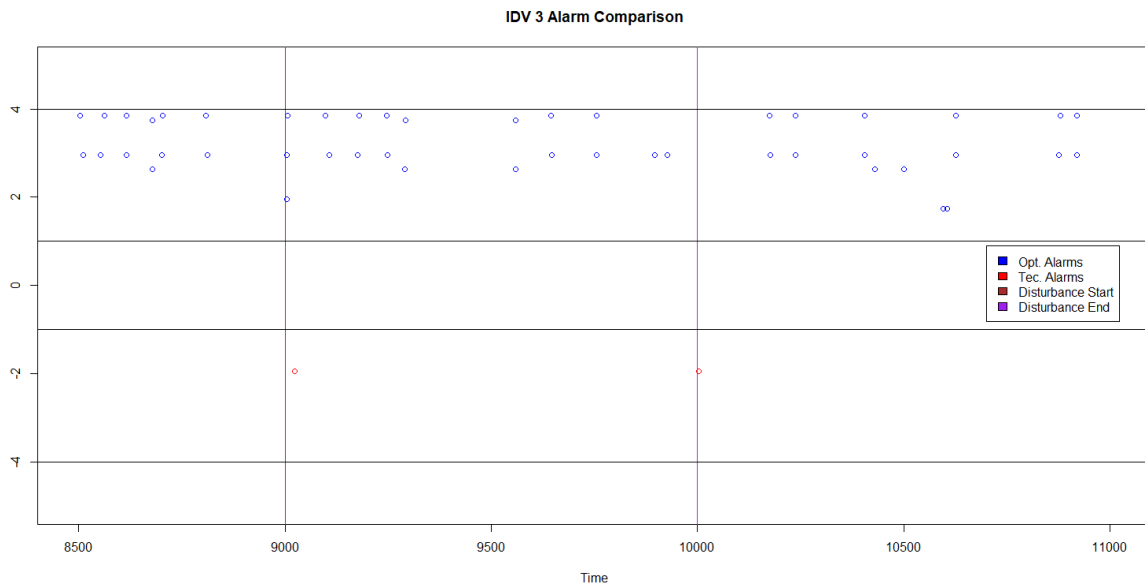


Figure 20: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV3

In the 3rd disturbance graph, there is no extra density visible after the disturbance. However, the first types of alarms that came for the first time and not before came exactly after the first moment of disturbance. Reference parameter set alarms give very good results, except for the delay (0.22 hours).

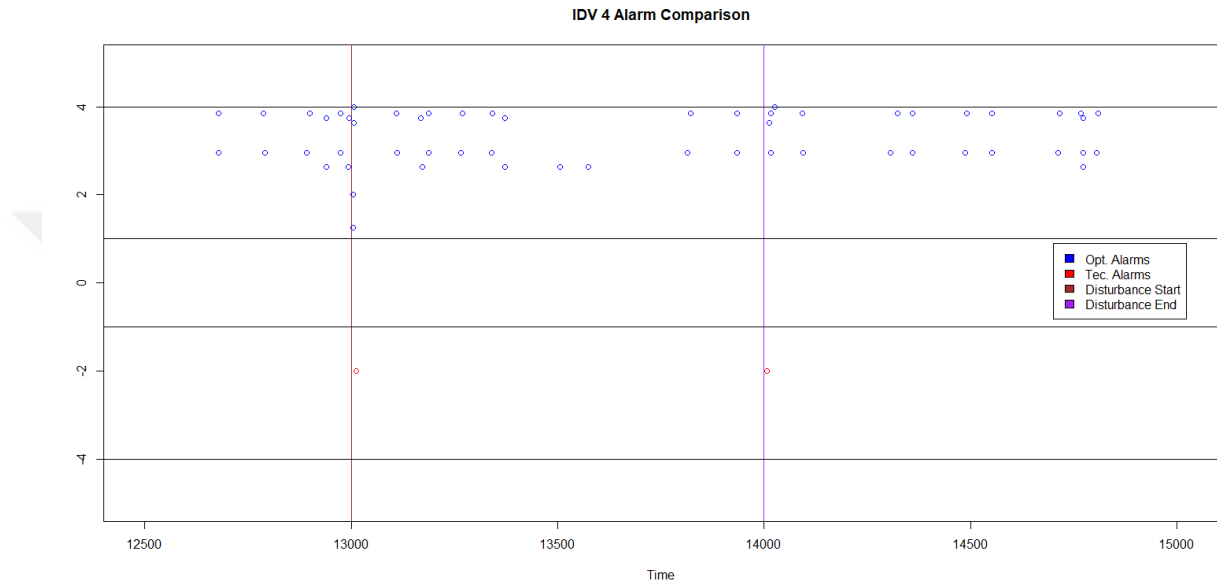


Figure 21: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV4

The results of the 4th disturbance are very similar to the 3rd disturbance. The only difference is that the delay in alarms of the reference parameter set has decreased.

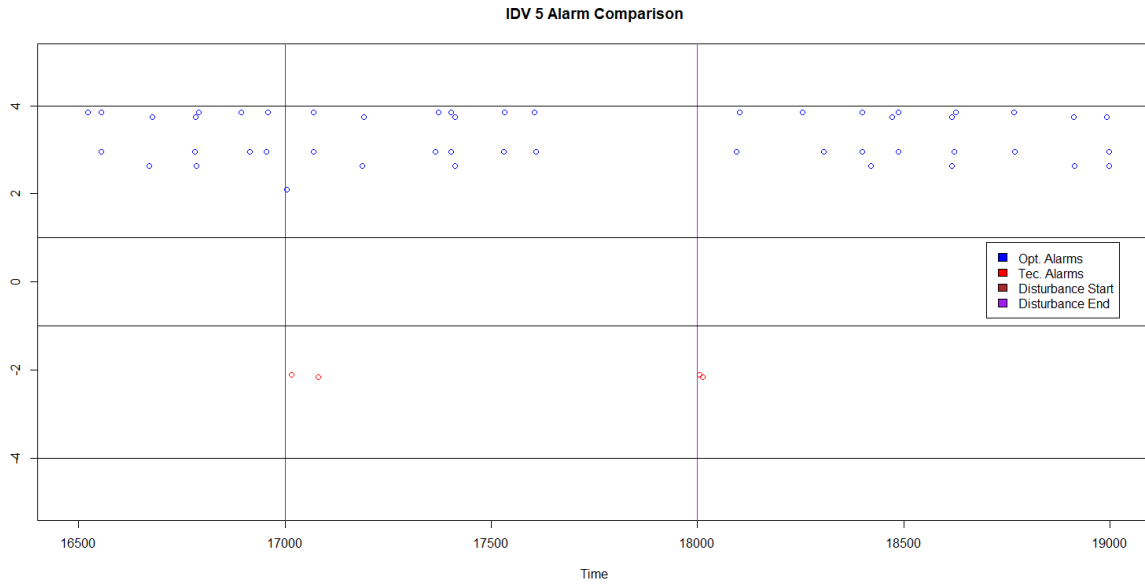


Figure 22: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV5

The 5th disturbance results are also largely similar to 3 and 4.

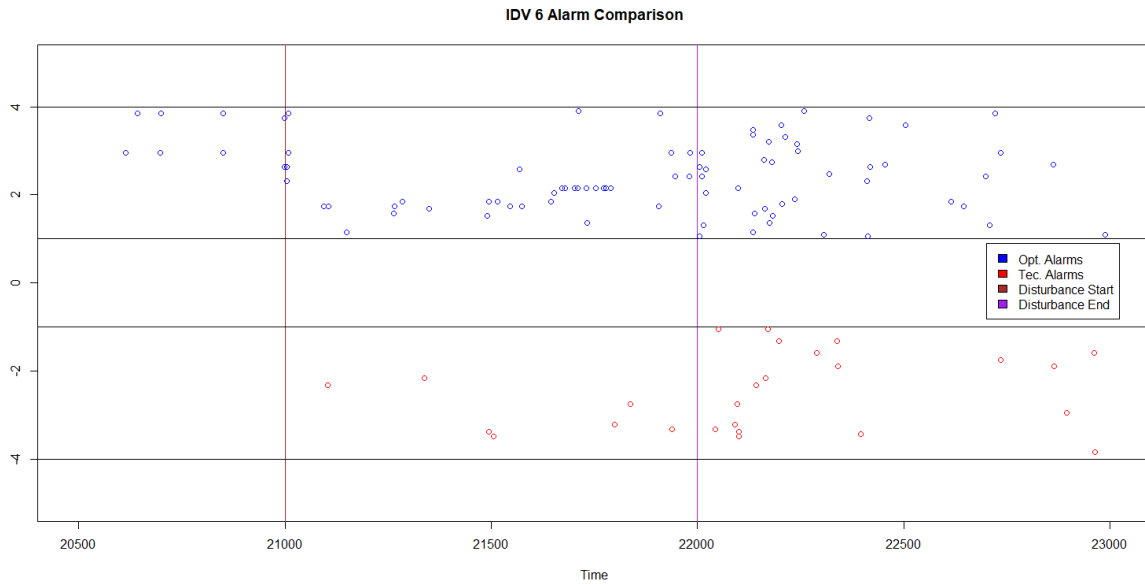


Figure 23: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV6

The 6th disturbance results differ from the previous results. Alarms generated through the proposed work arrive immediately after the moment of disturbance, similar to the previous ones, but alarms generated through the reference parameter set arrive quite late (1.02

hours). There is also a high-intensity alarm region for both results after disturbance. Similar to the first two disturbances, we can say that the process values after the disturbance did not reach stable conditions.

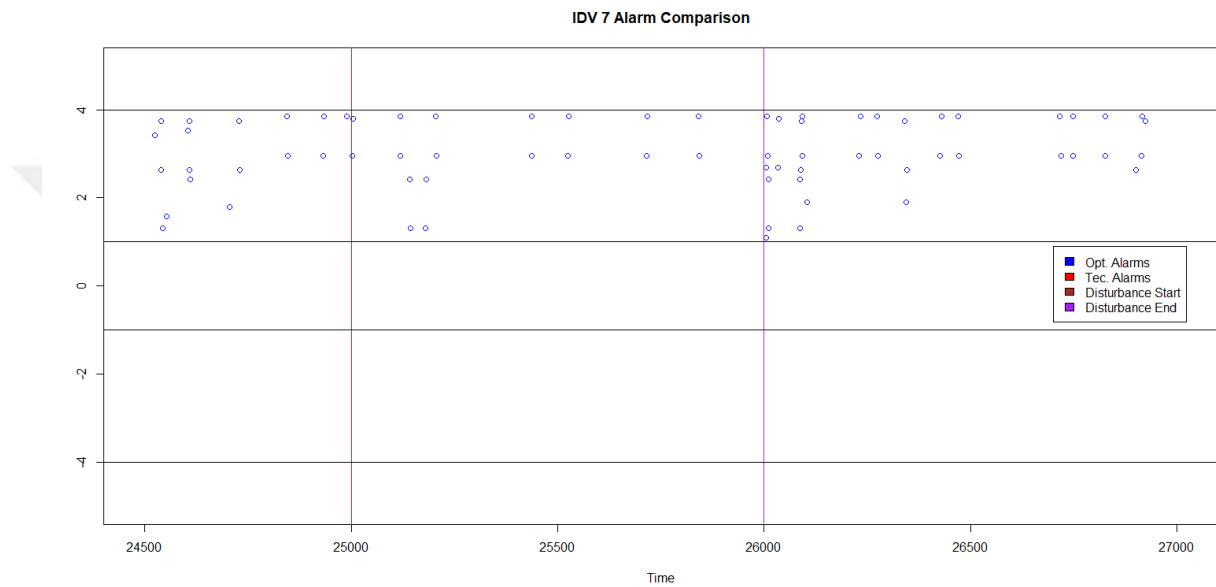


Figure 24: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV7

When we look at the seventh disturbance results, we see that no alarms were generated through the reference parameter set. When we look at the alarms generated through the proposed work, we see that the alarm intensity does not change during the disturbance, but the alarm types change. This change can be used to detect the moment of disturbance even on a rule-based basis. Apart from this, we see that the end of this disturbance also affects the process values and creates alarm intensity.

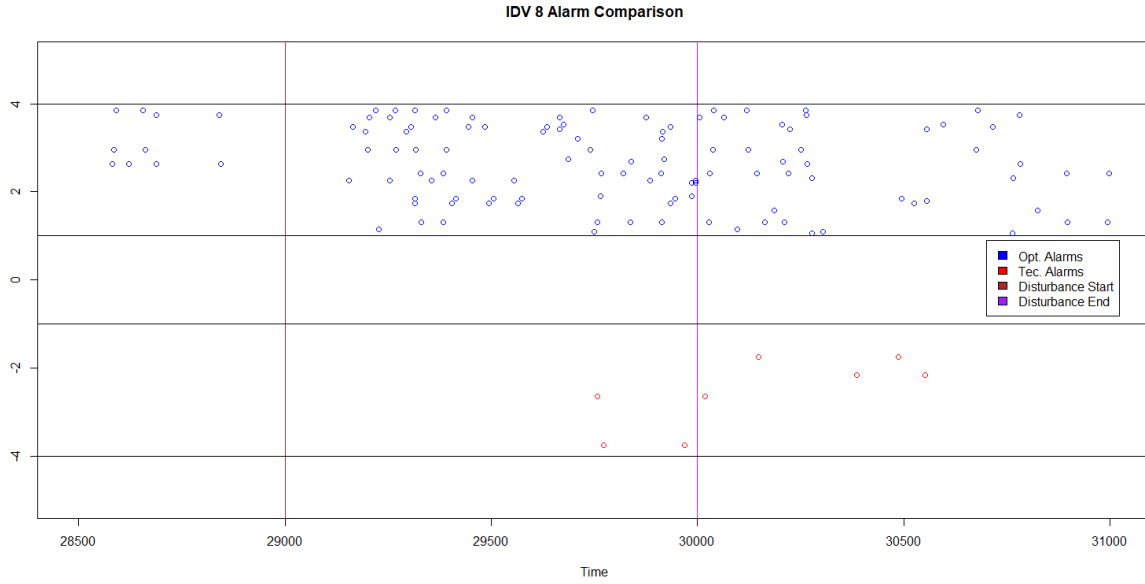


Figure 25: Proposed work (Optimized Results) versus reference parameters alarm comparison – IDV8

Finally, at the 8th disturbance, we see that both parameter sets produce alarms with a fairly high delay. However, when compared, we can say that the proposed work generated alarm much earlier.

The optimization method focuses on reducing the number of alarms while ensuring timely responses. By doing so, it effectively decreases the occurrence of false alarms, as fewer alarms mean fewer chances for false positives. Additionally, by minimizing the reaction delay, the method helps in reducing missed alarms, ensuring that critical alarms are addressed promptly. This dual approach enhances the overall reliability and effectiveness of the alarm system, addressing both false and missed alarms simultaneously.

CHAPTER 6

CONCLUSION AND FUTURE WORK

Alarm systems are critical in complex industrial plants as a safety barrier. Most of the accidents can be prevented with a healthy alarm system. Having a healthy alarm system requires optimized alarm parameters, which filter unnecessary alarms and show critical ones as soon as possible. This work proposes a data-driven optimization method to optimize alarm parameters.

Optimization is performed considering 3 KPIs: number of alarms, alarm-on times, and alarm reaction delay. The number of alarms KPI is selected to filter unnecessary alarms. The alarm-on times KPI is selected to minimize the time interval that an alarm is presented on a DCS screen. The alarm reaction delay KPI is selected to minimize the time interval between a disturbance and an alarm log.

A dataset from TEP is obtained to test the proposed work. To have a valid comparison, alarm parameters are taken from Manca (2020). At the same time, optimized parameters are obtained with the proposed work. As explained in the first scenario given in the results and comparisons section, when the alarm reaction delay KPI is prioritized, we get a very good result in this KPI, while we get relatively bad results in number of alarms and alarm on time KPI's. However, in the second scenario, where we evaluate all KPIs with equal weight, we get very close results in the number of alarms, while we get better results in the alarm reaction delay and alarm-on time values.

It can be said that there is a trade-off between timely alarm arrival and alarm filtering. An important point is, timely reactions require alarms closest to disturbances. Alarms can be filtered in further studies by detecting alarm patterns. However, the minimum value of the early detection time that can be achieved with further studies cannot be lower than the alarm reaction delay. As a result, the prerequisite for further studies to yield better results is the low alarm reaction delay.

6.1. Validity Threats

In any research study, it is crucial to acknowledge potential threats to the validity of the findings. This section discusses the internal and external validity threats that may have impacted the results of this study on data-driven alarm parameter optimization.

6.1.1 Internal Validity Threats

1. **Selection Bias:** The study utilized data from the TEP simulation, which may not fully represent all possible industrial scenarios. The selection of specific disturbances and process variables could introduce bias, affecting the generalizability of the results.
2. **Instrumentation:** The accuracy of the simulation tools and the parameters used in the TEP model could influence the outcomes. Any inaccuracies in the simulation model or measurement tools could lead to erroneous conclusions.
3. **Testing Effects:** Repeated testing and optimization iterations might have led to learning effects, where the system becomes increasingly fine-tuned to the specific disturbances used in the study, potentially reducing the applicability of the findings to other scenarios.
4. **Maturation:** Over the course of the study, changes in the simulation environment or updates to the software tools could have impacted the results. Ensuring consistency in the simulation setup throughout the study was essential to mitigate this threat.

6.1.2 External Validity Threats

1. **Population Validity:** The findings from the TEP simulation may not be directly applicable to other industrial processes or settings. The specific characteristics of the TEP may limit the generalizability of the results to other contexts.
2. **Ecological Validity:** The controlled environment of the TEP simulation may not accurately reflect the complexities and variabilities of real-world industrial settings. Factors such as operator behavior, environmental conditions, and unexpected disturbances in actual plants could affect the applicability of the results.
3. **Temporal Validity:** The study's findings are based on the current state of the TEP simulation and the specific disturbances applied. Changes in industrial practices, technological advancements, or new types of disturbances could impact the relevance of the results over time.

6.1.3. Mitigation Strategies

To address these validity threats, several strategies were employed:

1. **Randomization:** Where possible, randomization techniques were used to select disturbances and process variables to minimize selection bias.
2. **Consistent Instrumentation:** The same simulation tools and parameters were used throughout the study to ensure consistency and reduce instrumentation threats.
3. **Control Groups:** Control scenarios without disturbances were included to provide a baseline for comparison and to account for testing effects.
4. **Documentation:** Detailed documentation of the simulation setup, parameters, and procedures was maintained to ensure reproducibility and to mitigate maturation threats.

By acknowledging and addressing these validity threats, the study aims to provide a robust and reliable framework for data-driven alarm parameter optimization, while also highlighting areas for future research to further validate and extend the findings

6.2. Answers to Research Questions

6.2.1. *Research Question 1: How is the performance of an alarm management system or an alarm measured?*

The performance of an alarm management system or an alarm is measured using several key performance indicators (KPIs). In this study, three primary KPIs were selected to evaluate the performance: the number of alarms, alarm-on times, and alarm reaction delay. The number of alarms KPI helps in filtering unnecessary alarms and detecting alarm floods. Alarm-on times KPI minimizes the duration that an alarm is displayed on a Distributed Control System (DCS) screen, reducing operator distraction. Alarm reaction delay KPI measures the time interval between a disturbance and the corresponding alarm log, ensuring timely detection and response to disturbances.

6.2.2. *Research Question 2: Which key performance indicators should be utilized to measure the performance of an alarm management system?*

The key performance indicators utilized to measure the performance of an alarm management system in this study are:

1. **Number of Alarms:** This KPI helps in identifying and reducing alarm floods and chattering alarms, ensuring that only critical alarms are presented to operators
2. **Alarm-On Times:** This KPI measures the duration an alarm remains active, helping to minimize long-standing alarms that can clutter the DCS screen and distract operators
3. **Alarm Reaction Delay:** This novel KPI measures the time interval between the occurrence of a disturbance and the activation of the corresponding alarm. It ensures that alarms are generated promptly, allowing operators to take timely actions to mitigate potential issues

6.2.3. Research Question 3: How should alarm parameters be optimized for better performance? What is required for the optimization of alarm parameters?

Alarm parameters should be optimized using a data-driven approach that considers the selected KPIs. The optimization process involves the following steps:

1. **Data Collection:** Collect process data under stable conditions and during disturbances. This data is used to generate alarms based on predetermined alarm parameters
2. **Alarm-Disturbance Association:** Develop a method to associate disturbances with alarms, enabling the calculation of alarm reaction delay
3. **Simulation Platform:** Create an alarm simulation platform using the Tennessee Eastman Process (TEP) to validate the optimization method
4. **Optimization Method:** Implement a novel data-driven optimization method that adjusts alarm parameters to minimize the number of alarms and alarm-on times while ensuring minimal alarm reaction delay.

The optimization process requires a comprehensive understanding of the process dynamics, the ability to simulate disturbances, and the capability to adjust alarm parameters based on the selected KPIs. The proposed method in this study demonstrates that optimizing alarm parameters can significantly improve the performance of the alarm management system by balancing the trade-off between timely alarm generation and alarm filtering.

6.3. Future Work

In this work, a single process state is considered, which means that steady state values exist for each process value. As a result, alarm parameters can be kept constant, and generating these constant values is easier compared to a multi-state problem. However, multi-states exist in many processes. In future work, multi-state optimization will be studied. That study can also be considered as the beginning of our dynamic parameter optimization studies.

Another issue that can be addressed in future studies is unpostulated alarms. Since the moments of disturbances were known in advance in the proposed study, optimization was carried out for previously observed disturbances. However, the fact that the optimization is designed this way does not mean that the study will not work in cases of disturbances that have not been observed before. The main requirement for operability is that at least one disturbance is assigned to all process values in the disturbance and process value matching. In this way, alarm parameters can be generated for alarms that should occur when stable conditions are exceeded. However, at this point, the minimum delay may not be achieved for the new type of disturbance. A dynamic optimization structure can be evaluated to ensure high performance even in such disturbances.

By addressing these areas, future studies can build on the foundation laid by this work, providing more robust and adaptive alarm management solutions for complex industrial processes.





REFERENCES

- Ahmed, K., Izadi, I., Chen, T., Joe, D., & Burton, T. (2013). Similarity Analysis of Industrial Alarm Flood Data. *IEEE Transactions on Automation Science and Engineering*, 10(2), 452–457. <https://doi.org/10.1109/TASE.2012.2230627>
- Bathelt, A., Ricker, N. L., & Jelali, M. (2015). Revision of the Tennessee Eastman Process Model. *IFAC-PapersOnLine*, 48(8), 309–314. <https://doi.org/10.1016/j.ifacol.2015.08.199>
- Capaci, F., Vanhatalo, E., Kulahci, M., & Bergquist, B. (2018). The Revised Tennessee Eastman Process Simulator as Testbed for SPC and DoE Methods. *Quality Engineering*, 31, 1–57. <https://doi.org/10.1080/08982112.2018.1461905>
- Charbonnier, S., Bouchair, N., & Gayet, P. (2015). A weighted dissimilarity index to isolate faults during alarm floods. *Control Engineering Practice*, 45, 110–122. <https://doi.org/10.1016/j.conengprac.2015.09.004>
- Chenaru, O., Popescu, D., & Enache, D. (2016). Practical fault management using real-time decision tree analysis. 2016 24th Mediterranean Conference on Control and Automation (MED), 384–389. <https://doi.org/10.1109/MED.2016.7535916>
- Cheng, Y., Izadi, I., & Chen, T. (2013). Pattern matching of alarm flood sequences by a modified Smith–Waterman algorithm. *Chemical Engineering Research and Design*, 91(6), 1085–1094. <https://doi.org/10.1016/j.cherd.2012.11.001>
- Dong, Y., & Qin, S. J. (2018). A novel dynamic PCA algorithm for dynamic data modeling and process monitoring. *Journal of Process Control*, 67, 1–11. <https://doi.org/10.1016/j.jprocont.2017.05.002>
- Dorgo, G., & Abonyi, J. (2018). Sequence Mining Based Alarm Suppression. *IEEE Access*, 6, 15365–15379. <https://doi.org/10.1109/ACCESS.2018.2797247>

Downs, J. J., & Vogel, E. F. (1993). A plant-wide industrial process control problem. *Computers & Chemical Engineering*, 17(3), 245–255. [https://doi.org/https://doi.org/10.1016/0098-1354\(93\)80018-I](https://doi.org/https://doi.org/10.1016/0098-1354(93)80018-I)

Gao, H., Xu, Y., Gu, X., Lin, X., & Zhu, Q. (2015). Systematic rationalization approach for multivariate correlated alarms based on interpretive structural modeling and Likert scale. *Chinese Journal of Chemical Engineering*, 23(12), 1987–1996. <https://doi.org/https://doi.org/10.1016/j.cjche.2015.11.009>

Gyasi, P., & Wang, J. (2023). Design of alarm thresholds and delay timers for non-IID process variables based on alarm durations. *Process Safety and Environmental Protection*, 170, 1173–1187. <https://doi.org/https://doi.org/10.1016/j.psep.2022.12.069>

Gyasi, P., Wang, J., Yang, F., & Izadi, I. (2023). An adaptive method to update alarm deadbands for non-stationary process variables. *Process Safety and Environmental Protection*, 179, 493–502. <https://doi.org/https://doi.org/10.1016/j.psep.2023.08.097>

He, R., Chen, G., Dong, C., Sun, S., & Shen, X. (2019). Data-driven digital twin technology for optimized control in process systems. *ISA Transactions*, 95, 221–234. <https://doi.org/https://doi.org/10.1016/j.isatra.2019.05.011>

HSE Books. (1995). *The Explosion and Fires at the Texaco Refinery, Milford Haven, 24 July 1994*. Sudbury, U.K.: HSE Books.

Hopkins, A. (2008). *Failure to Learn: The BP Texas City Refinery Disaster*. CCH Australia Limited

Hu, W., Al-Dabbagh, A. W., Chen, T., & Shah, S. L. (2016). Process Discovery of Operator Actions in Response to Univariate Alarms**This work was supported by the Natural Sciences and Engineering Research Council of Canada via the CRD program. *IFAC-PapersOnLine*, 49(7), 1026–1031. <https://doi.org/https://doi.org/10.1016/j.ifacol.2016.07.337>

Hu, W., Chen, T., & Shah, S. (2018a). Detection of Frequent Alarm Patterns in Industrial Alarm Floods Using Itemset Mining Methods. *IEEE Transactions on Industrial Electronics*, PP, 1. <https://doi.org/10.1109/TIE.2018.2795573>

Hu, W., Chen, T., & Shah, S. L. (2018b). Discovering Association Rules of Mode-Dependent Alarms From Alarm and Event Logs. *IEEE Transactions on Control Systems Technology*, 26(3), 971–983. <https://doi.org/10.1109/TCST.2017.2695169>

Hu, W., Wang, J., & Chen, T. (2015). A new method to detect and quantify correlated alarms with occurrence delays. *Computers & Chemical Engineering*, 80, 189–198. <https://doi.org/https://doi.org/10.1016/j.compchemeng.2015.05.028>

International Society of Automation. (2016). ISA 18.2 Management of Alarm Systems for the Process Industries. ANSI/ISA.

Kirchhübel, D. (2020). Combining functional modeling and reasoning with on-line event analytics. Technical University of Denmark.

Kondaveeti, S. R., Shah, S. L., & Izadi, I. (2009). Application of Multivariate Statistics for Efficient Alarm Generation. IFAC Proceedings Volumes, 42(8), 657–662. <https://doi.org/https://doi.org/10.3182/20090630-4-ES-2003.00109>

Kondo, R., Rocha Loures, E., Santos, E., & Polito Braga, C. (2014). Alarm Rationalization Based on Process Mining Techniques. Advanced Materials Research, Volumes 1061-1062, 1258–1265. <https://doi.org/10.4028/www.scientific.net/AMR.1061-1062.1258>

Kumari, P., Lee, D., Wang, Q., Karim, M., & Kwon, J. (2020). Root Cause Analysis of Key Process Variable Deviation for Rare Events in the Chemical Process Industry. Industrial & Engineering Chemistry Research, XXXX. <https://doi.org/10.1021/acs.iecr.0c00624>

Lai, S., & Chen, T. (2017). A method for pattern mining in multiple alarm flood sequences. Chemical Engineering Research and Design, 117, 831–839. <https://doi.org/https://doi.org/10.1016/j.cherd.2015.06.019>

Ma, J., Liu, W., & Miller, P. (2010). Event Modelling and Reasoning with Uncertain Information for Distributed Sensor Networks. 6379, 236–249. https://doi.org/10.1007/978-3-642-15951-0_24

Ma, X., & Li, D. (2017). A hybrid fault diagnosis method based on fuzzy signed directed graph and neighborhood rough set. 2017 6th Data Driven Control and Learning Systems (DDCLS), 253–258. <https://doi.org/10.1109/DDCLS.2017.8068078>

Manca, G. (2020). “Tennessee-Eastman-Process” Alarm Management Dataset. IEEE Dataport. <https://doi.org/10.21227/326k-qr90>

Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press. <https://doi.org/10.1017/cbo9780511809071>

Martin-Villalba, C., Urquia, A., & Shao, G. (2018). Implementations of the Tennessee Eastman Process in Modelica. IFAC-PapersOnLine, 51(2), 619–624. <https://doi.org/https://doi.org/10.1016/j.ifacol.2018.03.105>

Müller, M. (2007). Information Retrieval for Music and Motion. Springer-Verlag.

Park, J. K., Choi, S. S., Hong, J. H., & Chang, S. H. (1997). Development of the effectiveness measure for an advanced alarm system using signal detection theory. IEEE Transactions on Nuclear Science, 44(2), 163–172. <https://doi.org/10.1109/23.568798>

Parsa, K., Hassall, M., & Naderpour, M. (2021). Process Alarm Modeling Using Graph Theory: Alarm Design Review and Rationalization. *IEEE Systems Journal*, 15(2), 2257–2268. <https://doi.org/10.1109/JSYST.2020.3019041>

Parsa, K., Hassall, M., & Naderpour, M. (2022). Enhancing Alarm Prioritization in the Alarm Management Lifecycle. *IEEE Access*, 10, 99–111. <https://doi.org/10.1109/ACCESS.2021.3137865>

Peng, X., Tang, Y., Du, W., & Qian, F. (2017). Multimode Process Monitoring and Fault Detection: A Sparse Modeling and Dictionary Learning Method. *IEEE Transactions on Industrial Electronics*, 64(6), 4866–4875. <https://doi.org/10.1109/TIE.2017.2668987>

Perdisci, R., Giacinto, G., & Roli, F. (2006). Alarm clustering for intrusion detection systems in computer networks. *Engineering Applications of Artificial Intelligence*, 19(4), 429–438. <https://doi.org/https://doi.org/10.1016/j.engappai.2006.01.003>

Ragab, A., El-Koujok, M., Poulin, B., Amazouz, M., & Yacout, S. (2018). Fault diagnosis in industrial chemical processes using interpretable patterns based on Logical Analysis of Data. *Expert Systems with Applications*, 95, 368–383. <https://doi.org/https://doi.org/10.1016/j.eswa.2017.11.045>

Shang, J., & Chen, T. (2019). Early Classification of Alarm Floods via Exponentially Attenuated Component Analysis. *IEEE Transactions on Industrial Electronics*, PP, 1. <https://doi.org/10.1109/TIE.2019.2949542>

Shirshahi, A., & Aliyari-Shoorehdeli, M. (2024). Diagnosing root causes of faults based on alarm flood classification using transfer entropy and multi-sensor fusion approaches. *Process Safety and Environmental Protection*, 181, 469–479. <https://doi.org/https://doi.org/10.1016/j.psep.2023.11.036>

Si, Y., Wang, Y., & Zhou, D. (2021). Key-Performance-Indicator-Related Process Monitoring Based on Improved Kernel Partial Least Squares. *IEEE Transactions on Industrial Electronics*, 68(3), 2626–2636. <https://doi.org/10.1109/TIE.2020.2972472>

Smith, T. F., & Waterman, M. S. (1981). Identification of common molecular subsequences. *Journal of Molecular Biology*, 147(1), 195–197. [https://doi.org/https://doi.org/10.1016/0022-2836\(81\)90087-5](https://doi.org/https://doi.org/10.1016/0022-2836(81)90087-5)

Stauffer, T., & Clarke, P. (2016). Using alarms as a layer of protection. *Process Safety Progress*, 35(1), 76–83. <https://doi.org/https://doi.org/10.1002/prs.11739>

Stephanou, H. E. (1995). Advanced automation in manufacturing and service industries. *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, 3, 3166 vols.3-. <https://doi.org/10.1109/ROBOT.1995.525739>

- Sudarshan, V., Seider, W. D., Patel, A. J., & Arbogast, J. E. (2021). Understanding rare safety and reliability events using forward-flux sampling. *Computers & Chemical Engineering*, 153, 107387. <https://doi.org/https://doi.org/10.1016/j.compchemeng.2021.107387>
- Sudarshan, V., Seider, W. D., Patel, A. J., Oktem, U. G., & Arbogast, J. E. (2024a). Alarm rationalization and dynamic risk analyses for rare abnormal events. *Computers & Chemical Engineering*, 184, 108633. <https://doi.org/https://doi.org/10.1016/j.compchemeng.2024.108633>
- Sudarshan, V., Seider, W. D., Patel, A. J., Oktem, U. G., & Arbogast, J. E. (2024b). Path-Sampling and Machine Learning for Rare Abnormal Events: Application to Polymerization CSTRs. *Chemical Engineering Science*, 300, 120513. <https://doi.org/https://doi.org/10.1016/j.ces.2024.120513>
- Sudarshan, V., Seider, W., Patel, A., Oktem, U., & Arbogast, J. (2023). Multivariate alarm systems to recognize rare unpostulated abnormal events. *AIChE Journal*, 70. <https://doi.org/10.1002/aic.18284>
- Tseng, F.-C. (2013). Mining frequent itemsets in large databases: The hierarchical partitioning approach. *Expert Systems with Applications*, 40(5), 1654–1661. <https://doi.org/https://doi.org/10.1016/j.eswa.2012.09.005>
- Vogel-Heuser, B., Schütz, D., & Folmer, J. (2015). Criteria-based alarm flood pattern recognition using historical data from automated production systems (aPS). *Mechatronics*, 31, 89–100. <https://doi.org/https://doi.org/10.1016/j.mechatronics.2015.02.004>
- Walker, J. S. (2004). *Three Mile Island: A Nuclear Crisis in Historical Perspective*. University of California Press
- Wang, H., Khan, F., Ahmed, S., & Imtiaz, S. (2015). Risk-based warning system design methodology for multimode processes. *IFAC-PapersOnLine*, 48(8), 663–668. <https://doi.org/https://doi.org/10.1016/j.ifacol.2015.09.044>
- Wang, J., & Chen, T. (2014). An online method to remove chattering and repeating alarms based on alarm durations and intervals. *Computers & Chemical Engineering*, 67, 43–52. <https://doi.org/https://doi.org/10.1016/j.compchemeng.2014.03.018>
- Wang, J., & Chen, T. (2016). Main causes of long-standing alarms and their removal by dynamic state-based alarm systems. *Journal of Loss Prevention in the Process Industries*, 43, 106–119. <https://doi.org/https://doi.org/10.1016/j.jlp.2016.05.006>
- Wang, J., Zhao, Y., & Bi, Z. (2018). Criteria and Algorithms for Online and Offline Detections of Industrial Alarm Floods. *IEEE Transactions on Control Systems Technology*, 26(5), 1722–1731. <https://doi.org/10.1109/TCST.2017.2723578>

Wang, Z., Bai, X., Wang, J., & Yang, Z. (2019). Indexing and Designing Deadbands for Industrial Alarm Signals. *IEEE Transactions on Industrial Electronics*, 66(10), 8093–8103. <https://doi.org/10.1109/TIE.2018.2885718>

Xu, Y., Wang, J., & Yu, Y. (2020). Alarm Event Prediction From Historical Alarm Flood Sequences Based on Bayesian Estimators. *IEEE Transactions on Automation Science and Engineering*, 17(2), 1070–1075. <https://doi.org/10.1109/TASE.2019.2935629>

Zang, H., Yang, F., & Huang, D. (2015). Design and Analysis of Improved Alarm Delay-Timers. *IFAC-PapersOnLine*, 48(8), 669–674. <https://doi.org/10.1016/j.ifacol.2015.09.045>

Zhao, H., Sun, S., & Jin, B. (2018). Sequential Fault Diagnosis Based on LSTM Neural Network. *IEEE Access*, 6, 12929–12939. <https://doi.org/10.1109/ACCESS.2018.2794765>

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Eylen, Tayfun

Nationality: Turkish

Marital Status: Married

EDUCATION

Degree	Institution	Year of Graduation
MS	METU Electrical – Electronics Engineering	2015
BS	METU Electrical – Electronics Engineering	2012
High School	Samsun Science High School	2007

WORK EXPERIENCE

Year	Place	Enrollment
2018-Present	TÜPRAŞ	Data Scientist – Head of data science
2017-2018	METEKSAN SAVUNMA	Embedded Software Engineer
2012-2017	METU EE	Research Assistant

FOREIGN LANGUAGES

Native Turkish, Advanced English

PUBLICATIONS

Eylen, T. & Bazlamaçcı C. F. (2015). One-Way Active Delay Measurement With Error Bounds. IEEE Transactions on Instrumentation and Measurement, vol. 64, no. 12, pp. 3476-3489. doi: 10.1109/TIM.2015.2469431

Eylen, T., Eren, P. E., & Koçyiğit, A. (2025). Data-driven alarm parameter optimization. Computers & Chemical Engineering, 196, 109041. <https://doi.org/https://doi.org/10.1016/j.compchemeng.2025.109041>

