Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 15(3), 765-777, 2025
Journal of the Institute of Science and Technology, 15(3), 765-777, 2025
ISSN: 2146-0574, eISSN: 2536-4618
Computer Engineering **DOI: 10.21597/jist.1594801**
Review Article
Received: 02.12.2024 **Accepted: 12.02.2025**

To Cite: Gurgen Erdogan, T. (2025). Process Mining for Root Cause Analysis: A Systematic Literature Review. *Journal of the Institute of Science and Technology*, 15(3), 765-777.

Process Mining for Root Cause Analysis: A Systematic Literature Review

Tugba GURGEN ERDOGAN[1*]

Highlights:
- Process Mining
- Root Cause Analysis
- Systematic Literature Review

Keywords:
- Process Mining
- Systematic Literature Review
- Root Cause Analysis

ABSTRACT:

Process mining techniques mostly focus on process discovery, conformance checking, and business process performance. However, relatively little research has been done on analyzing and identifying the root causes of compliance and performance issues in process mining. In this paper, we present a systematic literature review (SLR) that explores the use, stages, approaches, challenges of applying process mining to root cause analysis (RCA) in business processes from an empirical perspective. This SLR offers an overview of the state-of-the-art and applications of process mining in RCA, highlighting opportunities to address challenges and practical limitations in this domain. The findings aim to support the implement of novel RCA techniques and guide the selection of appropriate RCA methods with specific features tailored to various needs.

[1] Tugba GURGEN ERDOGAN (**Orcid ID:** 0000-0003-1491-8739), Hacettepe University, Department of Computer Engineering, Ankara, Türkiye

***Corresponding Author:** Tugba GURGEN ERDOGAN, e-mail: tugba@cs.hacettepe.edu.tr

INTRODUCTION

Process mining serves as bridge between model-based process analysis and data-oriented analysis techniques, encompassing a set of methods that support the examination of processes based on event logs (Qafari and Aalst, 2019). One of the main objectives of process mining is to identify the root causes of compliance or performance issues in processes (Hompes et al., 2016). Root cause analysis (RCA) is a systematic investigation aimed at uncovering the underlying reasons behind faults (Andersen and Fagerhaug, 2006). However, in the field of process mining, extracting root causes from event logs is a significant challenge and a complex task. To date, relatively limited research has focused on applying machine learning and statistical techniques for RCA in process mining applications.

The awareness and application of RCA techniques using process mining contribute to improving business processes. RCA seeks to explain risk incidents by identifying their root causes, ideally with a minimal number of contributing factors (Suriadi et al., 2012). RCA is a subset of causal analysis. RCA is more specific and action-oriented, while causal analysis is broader and more exploratory. These root causes may stem from various contexts, including case, process, social, and environmental context (Van Der Aalst and Dustdar, 2012). In particular, flexible, complex, and dynamic business processes—such as those in healthcare—often undergo continuous changes due to internal and external factors. Additionally, multiple resources and case-specific attributes may influence performance and compliance issues. In healthcare, process variability and compliance issues are critical due to the involvement of multiple stakeholders in patient care. In manufacturing, identifying the root causes of inefficiencies can drive cost savings and productivity gains. Despite the increasing interest in process mining, comprehensive studies systematically reviewing and classifying RCA techniques remain scarce, particularly regarding their applicability, effectiveness, and integration with advanced analytical methods such as machine learning.

In this study, we conduct a systematic literature review (SLR) of RCA techniques and process mining studies. Following systematic guidelines (Kitchenham et al., 2009), we searched six digital libraries to collect, categorize, and synthesize findings based on predefined research questions. Our review includes 40 process mining studies focusing on RCA and provides an overview of the state-of-the-art techniques and approaches. The primary contribution of this study is a comprehesive list of proposed RCA techniques and approaches, detailing their stages, along with a classification scheme encompassing contribution type, research type, study context, RCA stages, process mining-specific attributes, and supporting techniques employed in RCA.

The rest of this article is structured as follows: Section 2 gives an overview of the process mining background and relevant secondary studies, highlighting the research gap addressed by this work. Section 3 describes the research design used for the review. Section 4 presents and discusses the review findings in the context of the research questions and concludes with key insights and recommendations for future research.

PROCESS MINING AND RELATED WORK

Process mining is an process management technique situated at the intersection of process science and data science. It can be defined as the analyzing event data to improve operational processes (van der Aalst, 2022). Event data is typically structured into event logs, where each event corresponds to (1) a specific process instance (referred to as a case), (2) an activity, and (3) a timestamp (vander Aalst, 2019). Process mining generally encompasses four main types: process discovery, conformance checking, process enhancement, and operational support.

Random experiments leveraging various data science techniques and the causality theory provide two ways for identifying the causes of problems and predicting the effects of interventions on processes. Causality modeling methods in the literature are broadly categorized into data-based and knowledge-based approaches (Chiang and Braatz, 2003). RCA is often viewed as a process enhancement technique within the context of process mining applications. Numerous studies combine process mining and RCA techniques to uncover the root causes of process issues and implement process improvements. These studies typically examine the methods and prior work related to applying process mining for RCA in their respective related work sections.

While there are several review studies focused on specific process mining applications such as event abstraction (van Zelst et al., 2021), healthcare (Erdogan and Tarhan, 2018), industrial domains (Corallo et al., 2020; Akhramovich et al., 2024), educational contexts (Ghazal et al., 2017), and broader scopes (dos Santos Garcia et al., 2019), we found no secondary study that conducted a comprehensive literature review on the combined use of process mining and RCA techniques. Our study addresses this gap by investigating the integration of process mining and RCA to provide insights into their joint applications.

MATERIALS AND METHODS

We used SLR procedure proposed by Kitchenham et al. (2009) to identify and categorize process mining studies in RCA. SLR procedure begins with the defining research goal and questions, and requires an search protocol along with inclusion and exclusion criteria for screening and selection of relevant publications. It systematically investigates trends, challenges, and opportunities in the research area by developing a classification scheme and evaluating its findings. Figure 1 illustrates our research process and the steps followed in this review.

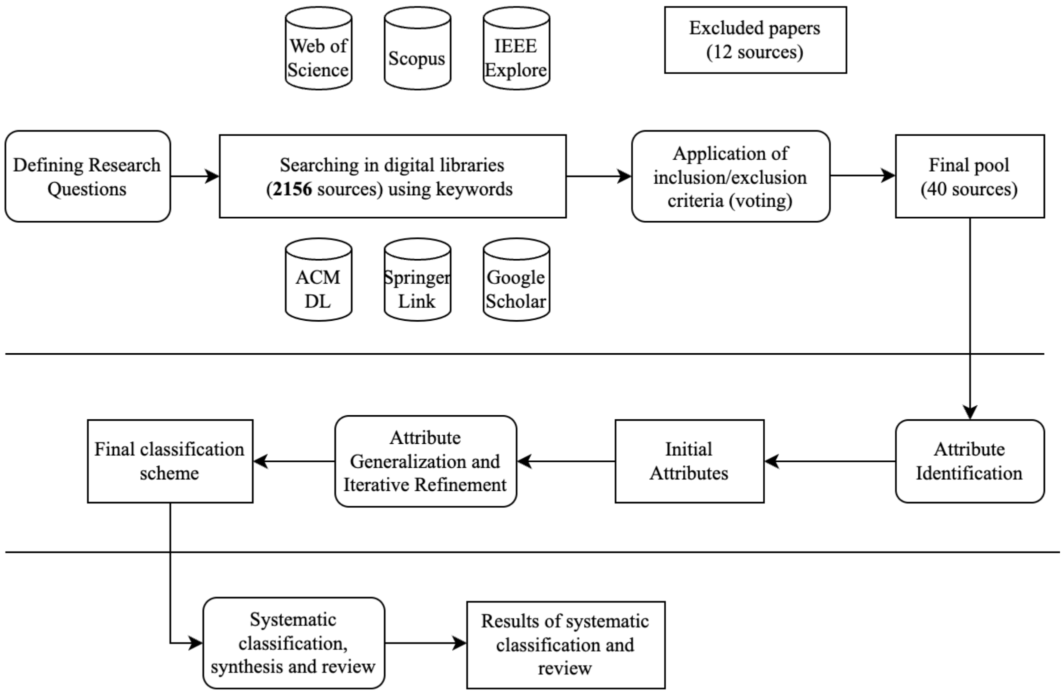


Figure 1. SLR Process Overview

The goal of this study is to understand the use, stages, approaches, and challenges of process mining in RCA of processes from an empirical perspective. Based on this goal, we derived the following research questions (RQs):

– RQ1: What are the research and contribution types?
– RQ2: What is the context of the study? (e.g., type of business process, number of staff involved
– RQ3: At what stage of RCA is process mining applied, and what stages of RCA are addressed?
– RQ4: What are the process mining-specific attributes used? (e.g., algorithms, techniques, tools)
– RQ5: What supporting techniques are employed in RCA (in addition to process mining), particularly with respect to the stages of RCA, if specified?

To achieve the study's goal, we searched seven different databases using the keywords ”process mining” and ”root cause analysis” in November, 2024. The number of sources initially retrieved and those selected by the search queries are presented in Table 1. From the initial pool of studies, primary studies were identified by screening abstracts and conducting high-level readings of the full texts. We have eliminated studies by identifying and applying inclusion criteria (IC) and exclusion criteria (EC). The inclusion criteria were as follows: (IC1) papers that explicitly focus on root cause analysis (RCA) with the context of process mining, (IC2) papers that propose, apply, or evaluate techniques, methods, or frameworks for RCA in process mining, (IC3) papers published in peer-reviewed venues, and (IC4) papers which are written in English. The exclusion criteria were as follows: (EC1) papers that discuss process mining or RCA in isolation without integrating the two concepts, (EC2) papers that are not accessible in full text. After applying these inclusion and exclusion criteria, the final pool comprised 40 studies. The classification scheme was developed iteratively based on the research questions. RCA stages in process mining applications were derived as a synthesis of SLR results. The final pool and the classification scheme can be accessed via the following link: https://tinyurl.com/SLR-PM-RCA.

Table 1. Number of sources initially retrieved and uniquely selected by search query

Digital Library	#Initially Retrieved	#Uniquely Selected
Web of Science	21	10
IEEE Explore	275	4
Scopus	263	7
ACM DL	47	1
Springer Link	178	6
Science Direct	69	2
Emerald	23	1
Google Scholar	1280	9

RESULTS AND DISCUSSION

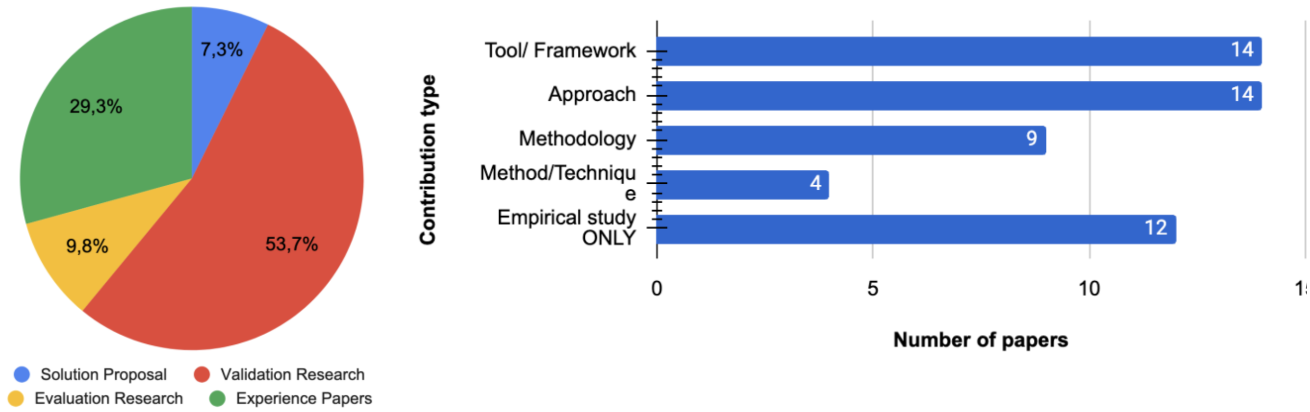
In this section, the results and findings of SLR study for each research question are presented.

RQ1. Research and contribution types

Research type reflects the general research approach used in the sources, independent of a specific focus area. We adopted the classification of research approaches proposed by Wieringa et al. (2006). Figure 2 (a) illustrates the distribution of research types across the 40 sources analyzed. The majority of sources (53.7%, n = 22) are categorized as validation research, followed by experience papers (29.3%, n = 12), evaluation research (9.8%, n = 4), and solution proposals (7.3%, n = 3).

In terms of research content, validation research papers dominate, reflecting the empirical maturity of the field. However, the limited presence of evaluation research and solution proposals suggests that the field may benefit from a greater emphasis on theoretical development and rigorous empirical testing. The increasing number of experience papers indicates a growing interest in documenting and understanding how RCA and process mining techniques are applied in practice. This trend highlights a demand for detailed insights into their practical implementation and effectiveness.

Contribution type refers to the type of contribution proposed in each source, categorized as one of the following: tool, model, metric, methodology/process, method/technique, or other (Petersen et al., 2008). Figure 2 (b) illustrates the distribution of sources by contribution type. The majority of sources (14 in total) propose new RCA approaches, while four studies introduce new RCA methods or techniques. Additionally, 14 tools/frameworks are presented, six of which involve the development of a ProM plugin to implement these newly proposed RCA approaches. Recently, the increasing focus on tool development has aligned with the rise in experience papers, as both trends reflect a growing emphasis on practical application and usability.



(a) (b)

Figure 2. (a) Distribution of sources by research type (b) Distribution of sources by contribution types

RQ2. Context of the study

We focus on the context of the studies and their application details, specifically distinguishing between real and synthetic data to examine the application areas that have received the most attention from researchers. Figure 3 shows the distribution of sources across different application contexts.

Process mining for RCA is most commonly applied to healthcare processes, with eight studies focusing on this domain. Applications related to loan application processes and software development processes follow. Publicly available datasets, such as the BPIC'12 dataset (Bautista et al., 2012), the Dutch hospital dataset (Mans et al., 2008), and research project datasets like the CoSeLoG project (Buijs, 2014), influence the context of many cases in this field. Additionally, synthetic data generated by CPN Tools (Ratzer et al., 2003) is used in six case studies alongside real-life cases.

Domain-specific applications are relatively rare, and publicly shared datasets or synthetic data are frequently used. This indicates that the field remains open to further research, particularly for cases that require domain knowledge.

RQ3. Stages of RCA for PM

We identified the main stages involved in applying RCA techniques with process mining across various domains. This process typically includes gathering process data from event logs, enriching the event logs with process mining results, transforming these data into various formats such as time series, and then applying machine learning, data mining, or statistical analysis techniques. All of these methods fall under the category of RCA techniques.

To highlight the key steps in RCA with process mining, activities in an RCA project are grouped into five categories, as outlined in Table 2.

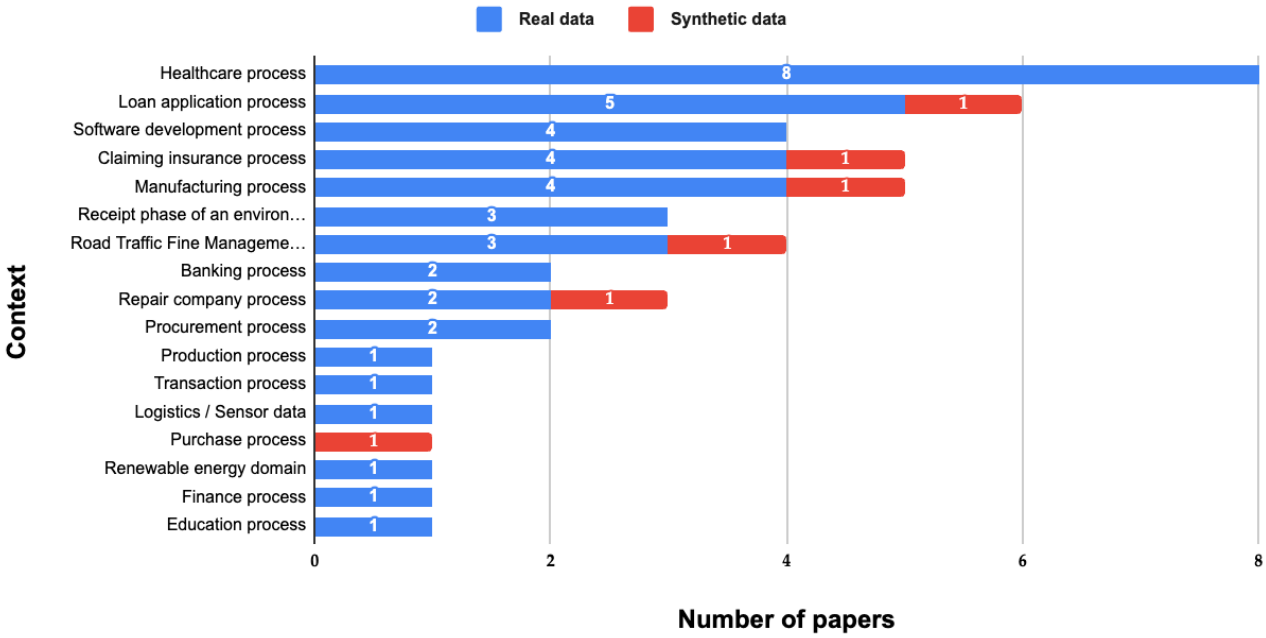


Figure 3. Distribution of sources by context

Table 2. RCA Stages with Process Mining

Process	Activity 1	Activity 2	Activity 3
1-(14)	Applying process mining	RCA Analysis	
2-(11)	Enriching event logs	Transforming event logs	RCA Analysis
3-(7)	Transforming event logs	RCA Analysis	
4-(4)	Applying process mining	Evaluating of PM results qualitatively	
5-(1)	RCA Analysis	Applying process mining	

In most studies (n = 14), after applying process mining techniques to event logs, various RCA analysis techniques are subsequently applied. The primary application of process mining is process model discovery. In six studies, conformance checking techniques are applied after the process models are discovered. Causality theory and process mining intersect in newly introduced techniques. For example, Bozorgi et al. (2020) proposed a causal rules discovery approach, while Leemans and Tax (2022) developed a causal process mining technique and presented it as a ProM plugin called *Visual Miner*, designed for process discovery and visualizing probabilistic dependencies, provided they are causal.

Discovered process models may serve as input for RCA techniques. As-is processes are discovered using the Inductive Visual Miner (Dogan and Areta Hiziroglu, 2024), Celonis DFG Graphs (Khakpour et al., 2024), or PM4PY abstraction techniques for large language models (LLMs) (Berti et al., 2023). In Li et al. (2023), bottlenecks are identified, and the scope is narrowed to handle large volumes of sensor data, which are then visualized with Performance Spectrum Miner. Domain experts collaborate in diagnosing root cause . Lastly, trace clustering techniques are utilized in Suriadi et al. (2013), Cai et al. (2019), and Goel et al. (2021).

The RCA process typically follows these key steps, as propose by Suriadi et al. (2012): 1) Enriching event logs, 2) Transforming event logs, and 3) RCA analysis. The first step, enriching event logs, may involve multi-perspective process mining results (De Leoni et al., 2016; Qafari and Aalst, 2019; Qafari and van der Aalst, 2021; Erdogan and Tarhan, 2022) or context data with newly derived features (Suriadi et al., 2012; Lehto et al., 2017; Van Houdt et al., 2022; Lehto et al., 2017). These features are sometimes referred as situation features (Qafari and van der Aalst, 2021). In four studies, enriched event logs are transformed into classification-ready logs, either at the event level or by aggregating the logs at the case level. Classification models are then applied. Qafari and Aalst create a situation feature table from enriched event logs and apply various RCA techniques, such as fair decision trees (Qafari and Aalst, 2019) or counterfactual reasoning (Qafari and van der Aalst, 2021).

In seven studies, event logs are transformed into different formats to apply RCA techniques, whereas four studies apply only process mining and evaluate its results qualitatively with domain experts for RCA analysis. These studies emphasize the importance of asking the right questions and incorporating domain knowledge into process mining applications. Interestingly, one recent study applies process mining over RCA analysis results by discovering process models to gain deeper insights into RCA outcomes.

RQ4. Process mining-specific attributes

We analyzed the distribution of the studies by process mining activities, as shown in Figure 4. We observe that process discovery is the most frequently used activity in process mining applications, as is common in the field and serves as the starting point for a process mining project. 25 studies in the pool employed process discovery techniques. The most commonly used process discovery techniques are Fuzzy Miner (Gunther and Van Der Aalst, 2007), Inductive Miner (Leemans et al., 2014), and Heuristic Miner (Weijters and Ribeiro, 2011). These results align with the frequently cited process discovery techniques in the literature.

Additionally, we found that process discovery and conformance-checking techniques were used together in 10 studies, although process enhancement was applied in only four studies. To further refine the discovered processes with different perspectives, researchers employed multi-perspective process mining techniques. These findings suggest that, to identify root causes beyond the problems in the process, researchers aim to extend their analyses by incorporating various process mining activities such as performance analysis and trace clustering.

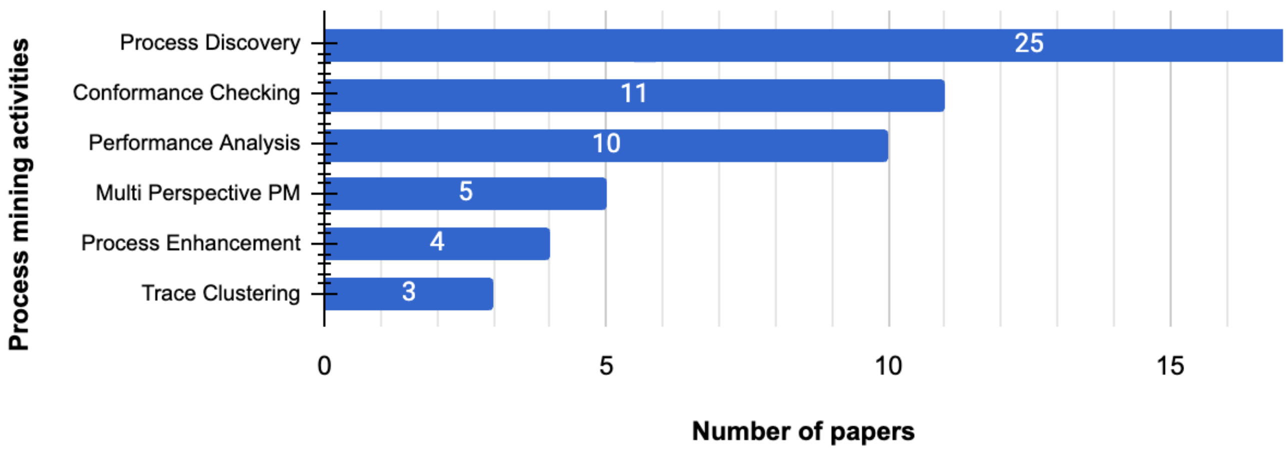


Figure 4. Distributions of sources by process mining activities

RQ5. Supporting techniques used (in RCA in addition to process mining)

Process mining techniques are often combined with various data science methods to enhance RCA. To gain a deeper understanding of RCA stages, we examined the supporting techniques used alongside process mining. These techniques are listed in Table 3. The most commonly used supporting technique is the decision tree, which is applied after process mining to analyze the behavior of the processes. Decision trees or regression trees are primarily used on enriched event logs to identify patterns and contributin factors. Additionally, process mining results are often evaluated by human expertise or domain specialists. This suggests that process mining techniques may uncover insights into processes to help identify root causes beyond those initially recognized by domain experts. However, this approach can be time-consuming and subjective. To mitigate these trade-offs, researchers increasingly integrate causal machine learning techniques with process mining to improve efficiency and objectivity.

Table 3. Used Supporting Techniques in RCA

Technique	#	Sources
Decision Tree/ Regression Tree	13	De Leoni et al., 2016; Qafari and van der Aalst, 2021; Nadim et al., 2022; Qafari and Aalst, 2019; Lehto et al., 2017; Erdogan and Tarhan, 2022; Jagadeesh Chandra Bose et al., 2015; Suriadi et al., 2012; Vasilyev et al., 2013; Ferreira andVasilyev, 2015; Dogan and Areta Hiziroglu, 2024; Southier et al., 2023; Khakpour et al., 2024
Human Expertise	10	Verboven and Martin, 2022; Nadim et al., 2022; Mahendrawathi et al., 2017; Qafari and van der Aalst, 2021; Erdogan and Tarhan, 2022; Stefanini et al., 2018,Qafari and van der Aalst, 2020; Southier et al., 2023; Van Houdt et al. 2023; Li et al., 2023
Statistics	6	Adams et al., 2021; Hompes et al., 2016; Senderovich et al., 2016; Hompes et al., 2017; Van Houdt et al., 2023; Dogan and Areta Hiziroglu, 2024
Causal Machine Learning	5	Qafari and van der Aalst, 2021; Verboven and Martin, 2022; Qafari and van der Aalst, 2021; Bozorgi et al., 2020; Qafari and van der Aalst, 2020
Probabilistic Temporal Logic	2	Van Houdt et al., 2022; Van Houdt et al., 2023
Anomaly Detection	1	Cai et al., 2019
Influence Analysis	1	Lehto et al., 2017
Value Stream Mapping	1	Knoll et al., 2019
Fault Tree analysis	1	Xu et al., 2014
LLMs	1	Berti et al., 2023

Additionally, statistics is another important area of research that can support root cause analysis with process mining. Other techniques, such as value stream mapping (Knoll et al., 2019), fault tree analysis (Xu et al., 2014), and graph-based anomaly detection (Cai et al., 2019), are also considered as supporting techniques. Knoll et al. (2019) integrated multi-perspective process mining with value stream mapping to identify root causes of inefficiencies in a manufacturing production line. The study revealed that long lead times were primarily caused by machine downtime and inefficient inventory management. By implementing lean production strategies based on these findings, the company achieved a 15% increase in production efficiency and a significant reduction in waste.

Recently, PM4PY abstraction techniques have been employed to leverage LLMs for RCA analysis on event logs. These findings suggest that additional supporting techniques beyond process mining are necessary to effectively detect root causes. Among the causal machine learning techniques, Structural Equation Models (SEM) (Qafari and van der Aalst, 2020; Qafari and van der Aalst, 2021), a neural network-based HTE estimator (Verboven and Martin, 2022), and Uplift Tree modeling (Bozorgi et al., 2020) have been used. Qafari and van der Aalst (2020, 2021) developed an RCA approach using SEMs and develop a ProM plugin for causality inference in process mining. In the context of software development, they identified that the root causes of delayed bug resolution were often related to insufficient testing and unclear task assignments. By addressing these issues, the development team was able to reduce the average bug resolution time by 30%. Verboven & Martin (2022) developed an RCA methodology for analyzing clinical and operational efficiency in healthcare processes. Their study focused on patient admission and discharge bottlenecks, utilizing HTE inference techniques. The findings led to specific recommendations for process improvements, such as reallocating staff during peak hours and streamlining documentation procedures, resulting in a 20% reduction in patient wait times.

Bozorgi et al. (2020) propose a causality discovery approach. The documentation and source code for conducting the experiments are available at https://github.com/zahradbozorgi/CausalRulesDiscovery. In their approach, classification-ready logs are created by identifying candidate treatments using action rule mining. Causal rules are then discovered and visualized with Uplift Tree modeling. Finally, the rules are ranked based on a cost-benefit model.

Various statistical techniques have also been used in RCA: Granger Causality (Adams et al., 2021; Hompes et al., 2017), statistical hypothesis testing (Hompes et al., 2016; Senderovich et al., 2016) and Statistical Process Charts (SPC) (Dogan & Hiziroglu, 2024). Granger causality determines the probability of correlation between two-time series given a time lag and can be seen as a form of predictive causality. Adams et al. (2021) and Hompes et al. (2017) transformed event logs into time series to apply Granger causality. Adams et al. (2021) applied this method logistics company event logs, identifying the delays in the supply chain due to inefficiencies in the procurement and inconsistent supplier performance. As a result, the company implemented supplier communication improvements and optimized inventory levels, leading to a 25% improvement in on-time deliveries. Hompes et al. (2017) analyzed performance indicators as time series, using performance functions such as case duration, activity duration, activity sojourn time, and activity waiting time. They examined cause-effect relationships at the process level to address performance issues like bottlenecks.

Dogan and Hiziroglu (2024) used SPC to monitor data attributes and identify out-of-control processes that deviated from specification limits. Adams et al. (2021) used the PELT algorithm, a change point detection technique for multivariate time series, and applied Granger causality for cause-effect analysis. Hompes et al. (2016) applied statistical hypothesis testing to understand whether a context label can explain performance differences. This technique enables the automatic analysis of the effect of various contexts on key performance indicators (KPI). Senderovich et al. (2016) constructed a queueing network to model both scheduled and actual process execution. They detected deviations in operations, applied statistical inference and similarity measures, identified root causes, and presented a technique for process improvement technique based on these root causes.

Knoll et al. (2019) integreate multi-perspective process mining, lean production, and value stream mapping. Waste analysis and strategy are employed to identify both part-specific (e.g., long lead times) and process specific (e.g., low trace fitness) root causes after multi-perspective process mining applications. Xu et al. (2014) use fault trees to visualize and structure data based on process mining results.

Lastly, the temporal trends of RCA techniques, based on the cited sources in Table 3, are analyzed. Decision Trees/Regression Trees have been consistently used since 2012, with a notable increase in recent years. This suggests their robustness and adaptability in RCA. Causal Machine Learning, while emerging in 2020, gained traction in 2021 and 2022, reflecting a shift toward advanced data-driven methods. Human Expertise has remained a stable and critical component across all years. Notably, LLMs were introduced in 2023, signaling a potential new direction in RCA research. These trends highlight the evolving landscape of RCA methodologies, with a growing emphasis on machine learning and AI-driven approaches.

CONCLUSION

This study provides a comprehensive analysis of the application of process mining techniques in RCA across various domains. By systematically reviewing the literature, we identified key trends, challenges, and opportunities within the field. The findings highlight that process discovery remains the most commonly employed process mining activity, with techniques like Fuzzy Miner, Inductive Miner, and Heuristic Miner frequently used to uncover process models. These techniques are favored for their ability to handle complex, real-world event logs and generate interpretable process models. However, they also have limitations: Fuzzy Miner, for instance, may struggle with highly unstructured processes (Gunther et al. 2007), while Inductive Miner, despite its robustness, can produce overly complex models that require simplification (Leemans et al., 2013). Heuristic Miner, though effective in handling noise, may miss subtle process behaviors due to its dependency on frequency thresholds (Weijters and Aalst, 2003). A critical analysis of these techniques reveals that their suitability depends heavily on the nature of the dataset and the specific requirements of the RCA task.

Additionally, we observed a growing interest in the integration of causality theory and process mining, as well as the application of advanced data science techniques, such as causal machine learning and statistical analysis, to enhance the RCA process. For example, decision trees, often used for their interpretability and ease of integration with process mining, may oversimplify complex causal relationships or struggle with high-dimensional data (Breiman et al, 1984). Similarly, causal machine learning methods, though powerful, require large volumes of high-quality data and are often computationally intensive (Pearl and Mackenzie, 2018). These limitations highlight the need for careful technique selection based on the context and goals of the RCA effort.

Our analysis reveals that while many studies rely on real-world or synthetic datasets, there is still a significant need for domain-specific research and the development of tailored approaches for different industries. Furthermore, the combination of process mining with supporting techniques, including decision trees, human expertise, and causal inference methods, underscores the importance of a multi-faceted approach to root cause analysis. However, the integration of these techniques is not always straightforward. For instance, while human expertise is invaluable for interpreting results and providing domain-specific insights, it introduces subjectivity and can be time-consuming (Aalst, 2018). Similarly, causal inference methods, though theoretically sound, often require strong assumptions that may not hold in practice.

The paper also discusses the importance of transforming event logs and enriching them with additional features to make them more suitable for RCA analysis.

Overall, our findings suggest that there is considerable potential for further research in the intersection of process mining and root cause analysis.

Conflict of Interest

The article author declares that there is no conflict of interest.

REFERENCES

Adams, J. N., van Zelst, S. J., Quack, L., Hausmann, K., van der Aalst, W. M., & Rose, T. (2021). A framework for explainable concept drift detection in process mining. International Conference on Business Process Management, 400–416.

Akhramovich, K., Serral, E., & Cetina, C. (2024). A systematic literature review on the application of process mining to industry 4.0. *Knowledge and Information Systems*, 66 (5), 2699–2746.

Andersen, B., & Fagerhaug, T. (2006). Root cause analysis: Simplified tools and techniques. Quality Press.

Bautista, A. D., Wangikar, L., & Akbar, S. M. K. (2012). Process mining-driven optimization of a consumer loan approvals process. BPI Challenge.

Berti, A., Schuster, D., & van der Aalst, W. M. (2023). Abstractions, scenarios, and prompt definitions for process mining with llms: A case study. International Conference on Business Process Management, 427–439.

Bozorgi, Z. D., Teinemaa, I., Dumas, M., La Rosa, M., & Polyvyanyy, A. (2020). Process mining meets causal machine learning: Discovering causal rules from event logs. 2020 2nd International Conference on Process Mining (ICPM), 129–136.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). Classification and Regression Trees. CRC Press.

Buijs, J. (2014). Receipt phase of an environmental permit application process ('wabo'), coselog project. Eindhoven University of Technology.

Cai, Z., Li, W., Zhu, W., Liu, L., & Yang, B. (2019). A real-time trace-level root- cause diagnosis system in alibaba datacenters. *IEEE Access*, 7, 142692–142702.

Chiang, L. H., & Braatz, R. D. (2003). Process monitoring using causal map and multivariate statistics: Fault detection and identification. *Chemometrics and intelligent laboratory systems*, 65 (2), 159–178.

Corallo, A., Lazoi, M., & Striani, F. (2020). Process mining and industrial applications: A systematic literature review. *Knowledge and Process Management*, 27 (3), 225–233.

De Leoni, M., van der Aalst, W. M., & Dees, M. (2016). A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs. *Information Systems*, 56, 235–257.

Dogan, O., & Areta Hiziroglu, O. (2024). Empowering manufacturing environments with process mining-based statistical process control. *Machines*, 12 (6), 411.

dos Santos Garcia, C., Meincheim, A., Junior, E. R. F., Dallagassa, M. R., Sato, D. M. V., Carvalho, D. R., Santos, E. A. P., & Scalabrin, E. E. (2019). Process mining techniques and applications–a systematic mapping study. *Expert Systems with Applications*, 133, 260–295.

Erdogan, T. G., & Tarhan, A. (2018). Systematic mapping of process mining studies in healthcare. *IEEE Access*, 6, 24543–24567.

Erdogan, T. G., & Tarhan, A. K. (2022). Multi-perspective process mining for emergency process. *Health Informatics Journal*, 28 (1), 14604582221077195.

Ferreira, D. R., & Vasilyev, E. (2015). Using logical decision trees to discover the cause of process delays from event logs. *Computers in Industry*, 70, 194–207.

Ghazal, M. A., Ibrahim, O., & Salama, M. A. (2017). Educational process mining: A systematic literature review. 2017 European Conference on Electrical Engineering and Computer Science (EECS), 198–203.

Goel, K., Leemans, S., Wynn, M. T., ter Hofstede, A., & Barnes, J. (2021). Improving phd student journeys with process mining: Insights from a higher education institution. Proceedings of the Industry Forum (BPMIF 2021) co-located with 19th International Conference on Business Process Management (BPM 2021), 39–49.

Gunther, C. W., & Van Der Aalst, W. M. (2007). Fuzzy mining–adaptive process simplification based on multi-perspective metrics. International conference on business process management, 328–343.

Hompes, B. F., Buijs, J. C., & van der Aalst, W. M. (2016). A generic framework for context-aware process performance analysis. OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", 300–317.

Hompes, B. F., Maaradji, A., La Rosa, M., Dumas, M., Buijs, J. C., & van der Aalst, W. M. (2017). Discovering causal factors explaining business process performance variation. International Conference on Advanced Information Systems Engineering, 177–192.

Jagadeesh Chandra Bose, R., Gupta, A., Chander, D., Ramanath, A., & Dasgupta, K. (2015). Opportunities for process improvement: A cross-clientele analysis of event data using process mining. International Conference on Service-Oriented Computing, 444–460.

Khakpour, R., Ebrahimi, A., & Seyed-Hosseini, S.-M. (2024). Lean process mining: Adopting process mining in lean manufacturing for dynamic process mapping and avoiding waste occurrence in real time. *International Journal of Lean Six Sigma*.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology*, 51 (1), 7–15.

Knoll, D., Reinhart, G., & Pr¨uglmeier, M. (2019). Enabling value stream mapping for internal logistics using multidimensional process mining. *Expert Systems with Applications*, 124, 130–142.

Leemans, S. J., Fahland, D., & Van Der Aalst, W. M. (2014). Process and deviation exploration with inductive visual miner. BPM (demos), 1295 (8).

Leemans, S. J., & Tax, N. (2022). Causal reasoning over control-flow decisions in process models. International Conference on Advanced Information Systems Engineering, 183–200.

Lehto, T., Hinkka, M., Hollm´en, J., et al. (2017). Focusing business process lead time improvements using influence analysis. SIMPDA, 54–67.

Li, C.-Y., Shinde, T., He, W., Lau, S. S. F., Hiew, M. X. B., Tam, N. T., Joshi, A., & van der Aalst, W. M. (2023). Unveiling bottlenecks in logistics: A case study on process mining for root cause identification and diagnostics in an air cargo terminal. International Conference on Service-Oriented Computing, 291–307.

Mahendrawathi, E., Zayin, S. O., & Pamungkas, F. J. (2017). Erp post implementation review with process mining: A case of procurement process. *Procedia Computer Science*, 124, 216–223.

Mans, R. S., Schonenberg, M., Song, M., van der Aalst, W. M., & Bakker, P. J. (2008). Application of process mining in healthcare–a case study in a dutch hospital. International joint conference on biomedical engineering systems and technologies, 425–438.

Nadim, K., Ragab, A., & Ouali, M.-S. (2022). Data-driven dynamic causality analysis of industrial systems using interpretable machine learning and process mining. *Journal of Intelligent Manufacturing*, 1–27.

Pearl, J., & Mackenzie, D. (2018). The Book of Why: The New Science of Cause and Effect. Basic Books.

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic mapping studies in software engineering. 12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12, 1–10.

Qafari, M. S., & Aalst, W. v. d. (2019). Fairness-aware process mining. OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", 182–192.

Qafari, M. S., & van der Aalst, W. (2020). Root cause analysis in process mining using structural equation models. International Conference on Business Process Management, 155–167.

Qafari, M. S., & van der Aalst, W. (2021). Feature recommendation for structural equation model discovery in process mining. arXiv preprint arXiv:2108.07795.

Qafari, M. S., & van der Aalst, W. M. (2021). Case level counterfactual reasoning in process mining. International Conference on Advanced Information Systems Engineering, 55–63.

Ratzer, A. V., Wells, L., Lassen, H. M., Laursen, M., Qvortrup, J. F., Stissing, M. S., Westergaard, M., Christensen, S., & Jensen, K. (2003). Cpn tools for editing, simulating, and analysing coloured petri nets. International conference on application and theory of petri nets, 450–462.

Senderovich, A., Weidlich, M., Yedidsion, L., Gal, A., Mandelbaum, A., Kadish, S., & Bunnell, C. A. (2016). Conformance checking and performance improvement in scheduled processes: A queueing-network perspective. *Information Systems*, 62, 185–206.

Southier, L. F. P., Krugger, E. M. R., Garcia, C. D. S., & Scalabrin, E. E. (2023). Process mining and root cause analysis for detecting inefficiencies in business processes: An applied case in a brazilian telecommunications provider. 2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 53–58.

Stefanini, A., Aloini, D., Benevento, E., Dulmin, R., & Mininno, V. (2018). Performance analysis in emergency departments: A data-driven approach. Measuring Business Excellence.

Suriadi, S., Ouyang, C., van der Aalst, W. M., & ter Hofstede, A. H. (2012). Root cause analysis with enriched process logs. International Conference on Business Process Management, 174–186.

Suriadi, S., Ouyang, C., van der Aalst, W. M., & ter Hofstede, A. H. (2013). Root cause analysis with enriched process logs. Business Process Management Workshops: BPM 2012 International Workshops, Tallinn, Estonia, September 3, 2012. Revised Papers 10, 174–186.

Suriadi, S., Wynn, M. T., Ouyang, C., ter Hofstede, A. H., & van Dijk, N. J. (2013). Understanding process behaviours in a large insurance company in australia: A case study. Advanced Information Systems Engineering: 25th International Conference, CAiSE 2013, Valencia, Spain, June 17-21, 2013. Proceedings 25, 449–464.

Van Der Aalst, W. M., & Dustdar, S. (2012). Process mining put into context. IEEE Internet Computing, 16 (1), 82–86. van der Aalst, W. M. (2019). Object-centric process mining: Dealing with divergence and convergence in event data. International Conference on Software Engineering and Formal Methods, 3–25.

van der Aalst, W. M. P. (2016). Process Mining: Data Science in Action. Springer.

Van der Aalst, W. M. (2022). Process mining: A 360 degree overview. In Process mining handbook (pp. 3–34). Springer.

Van Houdt, G., Depaire, B., & Martin, N. (2022). Root cause analysis in process mining with probabilistic temporal logic. International Conference on Process Mining, 73–84.

Van Houdt, G., Martin, N., & Depaire, B. (2023). Aitia-pm: Discovering the true causes of events in a process mining context. *Engineering Applications of Artificial Intelligence*, 126, 107145.

Van Zelst, S. J., Mannhardt, F., de Leoni, M., & Koschmider, A. (2021). Event abstraction in process mining: Literature review and taxonomy. *Granular Computing*, 6, 719–736.

Vasilyev, E., Ferreira, D. R., & Iijima, J. (2013). Using inductive reasoning to find the cause of process delays. 2013 IEEE 15th Conference on Business Informatics, 242–249.

Verboven, S., & Martin, N. (2022). Combining the clinical and operational perspectives in heterogeneous treatment effect inference in healthcare processes. International Conference on Process Mining, 327–339.

Weijters, A., & Ribeiro, J. T. S. (2011). Flexible heuristics miner (fhm). 2011 IEEE symposium on computational intelligence and data mining (CIDM), 310–317.

Wieringa, R., Maiden, N., Mead, N., & Rolland, C. (2006). Requirements engineering paper classification and evaluation criteria: A proposal and a discussion. *Requirements engineering*, 11 (1), 102–107.

Xu, X., Zhu, L., Weber, I., Bass, L., & Sun, D. (2014). Pod-diagnosis: Error diagnosis of sporadic operations on cloud applications. 2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 252–263.