Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 58-19 (2024) 658–663

Pure Data-Driven Machine Learning Challenges for *p*FMEA: A Case Study

Mahdi Mokhtarzadeh * **Jorge Rodríguez-Echeverría** * *** **Zafer Zeren** ** **Johan Van Noten** ** **Sidharta Gautama** *

* *Department of Industrial Systems Engineering and Product Design, Ghent University, 9052 Ghent, Belgium; FlandersMake@UGent – corelab ISyE, Lommel, Belgium (corresponding author email: Sidharta.Gautama@UGent.be)*

** *CodesignS, Flanders Make, Oude Diestersebaan, 133, Lommel, 3920, Belgium*

*** *ESPOL Polytechnic University, Escuela Superior Politécnica del Litoral, ESPOL, Facultad de Ingeniería en Electricidad y Computación, Campus Gustavo Galindo, Km 30.5 Vía Perimetral, P.O. Box 09-01-5863, EC090112 Guayaquil, Ecuador*

Abstract: Manufacturing processes are susceptible to quality defects, resulting in overall equipment effectiveness reduction. Proactive and reactive methods, such as process failure mode and effects analysis, and root cause analysis, have been developed to eliminate potential causes of failure modes. In this study, data from an assembly case is evaluated using supervised machine learning methods to analyze the challenges of purely data-driven failure mode detection. Assembly step execution times, as indicators, and end-of-the-line quality checklists, as the failure modes, are used to gain insights into failure mode detection. Challenges for data-driven methods are discussed and possible future research streams are proposed.

Keywords: FMEA, RCA, Failure analysis, Explainable AI, Hybrid intelligence systems

1. INTRODUCTION

Manufacturing processes are prone to quality defects, leading to costs, line stoppage, and wasted time. Proactive and reactive methods, such as process failure mode and effects analysis (*p*FMEA) and root cause analysis (RCA), have been developed to eliminate potential quality failure modes (QFMs) and identify underlying failure causes (Wu et al., 2021; Oliveira et al., 2023). *p*FMEA aims at minimizing the impact of potential quality defects by evaluating QFMs and their effects on the process. However, in a real industrial environment, it is impossible to preemptively eliminate all potential QFMs and causes. *p*FMEA is usually performed in the early stage of process development when limited data from the process is available as an expert-based procedure (Mokhtarzadeh et al., 2024). RCA holds importance in manufacturing for identifying and addressing the underlying causes of an occurred QFM. Such a reactive approach incurs costs of mitigation actions and a significant amount of expert analysis.

During a *p*FMEA procedure, each tuple of QFM, cause, and effect is evaluated using occurrence, detection, and severity criteria. Real-time monitoring, as enabled by Industry 4.0, connected devices, and IoT, offers a proactive approach to quality management. This data availability may contribute to answering three problems.

- Data-driven detection: to determine whether a production process is completed with no QFM or completed with one or more QFMs, which we focus on.
- Data-driven occurrence: to determine the probability of a failure occurrence under a specific situation.
- Data-driven RCA: to relate QFMs to underlying 4M causes, man, machine, material, and method.

Data from the early stage phase or operational phase of the production may help in determining a QFM occurrence under different situations, e.g., when a machine pressure is set to a specific value and specific raw materials are used. Indicators and production parameters may help to detect potential QFMs to mitigate the adverse impacts of QFMs, contributing to enhancing *p*FMEA.

In this study, *p*FMEA state-of-the-practice and data-driven failure analysis literature are discussed in Section 2. The data-driven QFM detection problem is defined in Section 3. A laboratory manual assembly case is described in Section 4. Execution times of each assembly step obtained from sensors are explored to identify possible correlations indicative of QFMs. The performance of five classification algorithms is evaluated to assess the feasibility of data-driven QFM detection in Section 5. Based on our results, we discuss the challenges of data-driven methods for detection rate in Section 6. In Section 7, we conclude the paper with future avenues of knowledge-driven approaches, counterfactual reasoning, and causal Bayesian networks.

⋆ This research is funded by Flanders Make under the project AQUME_SBO, project number 2022-0151.

10.1016/j.ifacol.2024.09.225

2. *P*FMEA STATE-OF-THE-PRACTICE/ART

During the industry needs analysis for the funding project of this paper, several industrial companies were interviewed for the requirements of *p*FMEA. *p*FMEA construction begins during the design phase of production as a brainstorming activity (Sharma and Srivastava, 2018). Potential QFMs, their causes, and their effects are identified. The occurrences, detections, and severities of these failures are also quantified as a preliminary estimation. Then, during the pilot phase of the process, the *p*FMEA table is checked and updated, mainly in an ad-hoc manner without the potential help of measured process data. During the real production phase, the *p*FMEA table settles into a more or less stable state, varying slightly as long as no modifications to the production line are implemented. In the absence of such modifications, potential variations may arise from factors such as time-dependent tool wear, a new operator, or varying environmental conditions.

The construction of *p*FMEA provides substantial information, but it is a time-costly procedure, and the information contained is expert-based and subjective. Subjectivity brings on issues such as reliability, reusability, and impedes the practical applicability (Siva et al., 2016; Spreafico et al., 2017; Wu et al., 2021). Also, *p*FMEA deals with one component and one QFM, neglecting the combination of different causes contributing to a single QFM when they both occur. These issues push companies to perform it as occasionally as possible. It further relegates the proactive aspect of *p*FMEA, leaving it more as an authority-obligated activity. Nevertheless, its proactive aspect alone is very encouraging since the early detection of QFMs has fundamental benefits for manufacturing processes. Therefore, the issues common today in *p*FMEA should be dealt with with robust techniques/methods to promote its use more effectively in manufacturing.

Different approaches have been explored to deal with *p*FMEA limitations such as subjectivity, reliability, and reusability. Recently, data-driven *p*FMEA and QFM detection have garnered attention (Mokhtarzadeh et al., 2024; Siva et al., 2016). For example, the identification of real-time success of robotic assembly in snap-fit assemblies and quality prediction of an assembly in the production process of electrical engines based on the machine data are studied by Doltsinis et al. (2019). Carvajal Soto et al. (2019) used machine learning for failure detection during the production of surface mount devices.

Schuh et al. (2021) analyzed the correlation between the processing time of assembly steps for automobile engines and the occurrence of end-of-the-line quality issues. They employed various classification methods and demonstrated an average prediction accuracy score of 0.75 for detecting a potential issue. Muhr et al. (2020) analyzed data partitioning based on domain knowledge in unsupervised machine learning to improve faulty internal combustion engine detection in the production line. Elanangai and Vasanth (2023) used neural networks for early defect detection in a steel assembly.

Krauß et al. (2020) developed a machine learning framework for integration, preparation, modeling, and model deployment to reduce the reliance on experts and expedite predictive quality in production. A comparison to the manual machine learning implementation showed the superiority of their framework. Filz et al. (2020) developed a data-driven framework to analyze product state propagation for quality-related cause-effect relationships in the electronic production industry. Ma et al. (2019) used digital twin for quality control improvement. More information is in Tercan and Meisen (2022); Kang et al. (2020) reviews.

3. PROBLEM DEFINITION AND METHODOLOGY

A QFM is defined as the specific manner or way by which a failure occurs. The detection rate is calculated based on the means or method by which a failure is detected and the time it takes. The detection rate for each QFM is usually determined by an expert (Sharma and Srivastava, 2018).

The detection of each QFM is determined by assessing the likelihood of detecting it using the available data. To determine the detection rank of QFMs, a binary classification problem is formulated. A set of input variables, such as the sensor data collected from the production line during various executions, is given. Each execution is classified as either faulty or not faulty based on the consideration of each QFM. In this study, execution times of 13 assembly steps (ETAS) are employed as homogeneous input features. The output comprises 11 binary variables predicting the occurrence of each of the 11 QFMs under a certain string of the 13 ETAS.

The performance of a classifier is quantified as the detection rate. If the classifier performance is high, indicating ease in classifying an execution as either faulty or not for a specific QFM, the detection rate for that QFM is also high. Five supervised classifiers in Python —namely, Convolutional Neural Network (CNN), Multilayer Perceptron Neural Network (MLP), Naïve Bayes (NB), Random Forest (RF), and Support Vector Machine (SVM)—were employed to explore the possibility of detecting QFMs. *Balanced accuracy* metric is used to measure the performance of the classifiers as our data is imbalanced. *Balanced accuracy* measures the average accuracy from both minority and majority classes in an imbalanced dataset.

The subsequent step involves analyzing the data to determine the root causes of the QFM categorized as 4M causes, i.e., data-driven RCA. Not all QFMs are easily detected by data. Thus, if the performance is low, signifying difficulty in detecting the QFM using the available data, the detection rate is correspondingly low. In this case, the data is not qualified for an objective assessment of QFM detection. Therefore, the next step is to integrate expert knowledge with the data to potentially improve the performance of the system, as detailed in Section 7.

4. CASE-STUDY

An experimental assembly case is developed in our laboratory to validate various technologies and strategies for manual assembly activities. This case involves three main models, the assembly workstation, open trailer, and garbage container, all integrated onto a truck. Our focus is on a sub-assembly of the open trailer, Figure 1. This model comprises two sub-models and involves 31 steps, requiring a total of 3 hours and 20 minutes of manual

Fig. 1. The final assembly vehicle

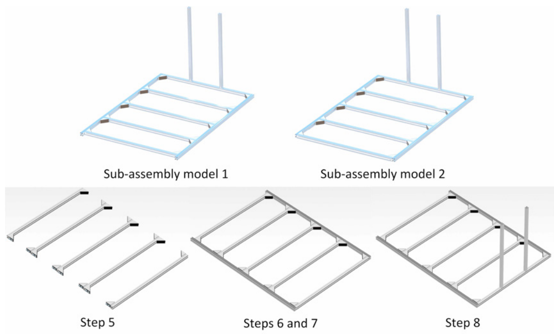

Fig. 2. The sub-assmebly models and steps

effort. The workstation has a working area of 4x4 square meters, a rack zone for storing raw materials, an assembly table on which the product is assembled, and several IoT tools including a microphone, five cameras that record the movements of the operator, sensitive floor, and input buttons to capture the operator's input during assembly.

Despite advancements in technology (Sader et al., 2022), data from manual assembly lines is primarily limited to the ETAS, imposing constraints on defect analysis (Schuh et al., 2021). To analyze possible QFM detection, a sub-assembly consisting of 14 steps—namely, 5.1, 5.2, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 7.1, 7.2, 7.3, 8.1, and 8.2—has been utilized. This sub-assembly requires 50 minutes of manual effort and comprises six different materials. Figure 2 shows that a total of 16 angle brackets are screwed onto five 1400mm profiles in Step 5. In Steps 6 and 7, these five profiles are assembled together using two 2000mm profiles from each side. In Step 8, two 920mm profiles are placed on the top of the sub-assembly. This sub-assembly comprises two distinct models. The distinguishing factor between Models 1 and 2 is the distance of the last assembled 1400mm profile from the end of the 2000mm profiles.

A total of 80 experiments were conducted at a single assembly station, involving 40 operators. Each operator conducted two experiments, assembling Model 1 and immediately proceeding to assemble Model 2. The ETAS were gathered through the automatic analysis of camera records. Quality checks were anonymously performed by a supervisor at the end of the line for 11 QFMs as follows.

Q1. Visually check. Are the holes facing upwards and lining up with the middle of the beams? Q2. Do you hear if anything is loose when shaking the assembly? Q3. Visually check. Is the assembly correctly assembled? Q4. Are all the correct parts used? Q5. Are all the correct fasteners being used? Q6. Are all the parts correctly oriented? Especially the 2000 beams? Q7. Are all the angle brackets correctly aligned? Q8. Is the correct (ID-dependent) alignment used for the last 1400 beam? Q9. Are the 2000 beams correctly aligned with the end beam? Q10. Are the parts fixated? Q11. Are the fasteners fixated?

24 experiments were excluded from the dataset during the data cleaning stage due to inconsistencies and incomplete experiments. Step 8.2 execution time was removed from the analysis because it lacked data for most experiments.

5. DATA-DRIVEN RESULTS SUMMARY

Results from the case study are now summarized. The data were split into train and test sets following the 80-20 rule. Each classifier was executed 20 times, to assess its stability, under five architectures. The average and best-balanced accuracies for each classifier are in Table 1.

For the first four architectures, a classification problem is solved for each QFM. The first (A1) represents a setting where the experience of the operators is disregarded, utilizing data from both Models 1 and 2. There is no standout method with the best-balanced accuracy; the average performances are around 50%. Q4 could not be assessed as its occurrence rate is very low.

Table 1. Summary of classifiers' performance

	Average Balanced Accuracy					Best Balanced Accuracy					
	CNN	MLP	NB	RF	SVM	QFM	CNN	MLP	NB	RF	SVM
A1	51%	50%	52%	48%	50%						
A2	51%	51%	53%	52%	53%	Q9	71%	59%	62%	63%	61%
A3	51%	49%	52%	50%	51%	Q6	61%	48%	65%	63%	61%
						Q2	51%	53%	60%	45%	60%

The assembly of the two models represents the experience level and the start-up/operational phases of production. In the first assembly, Model 1, operators have no experience with the product. In the second assembly, Model 2, they have some experience with the product. As such, the second (A2) represents a setting where operators are entirely inexperienced with the product, using only data from Model 1. Q5 could no longer be assessed as its occurrence rate is very low. The best-balanced accuracy belongs to Q9. We observe that the balanced accuracy for some of the QFMs has increased. The third (A3) represents a scenario where operators have some experience with the product, utilizing only data from Model 2. Q4 and Q8 could no longer be assessed as their occurrence rate is very low. The best-balanced accuracy belongs to Q6 and Q2.

The enhanced performance of the second and third architectures for some of the QFMs stems from the clustering of inexperienced and experienced executions. This results in a more uniform volume of uncertainty, noise, and operator speed across various executions for Model 1 or Model 2. The algorithms can adeptly eliminate noise from the data, enabling a clearer distinction between executions with and without a QFM. As shown in Figure 3, there is a clear difference between the execution time of Models 1 and 2. During the start-up execution, i.e., Model 1, the operators deal with learning the procedure, while during the second execution, i.e., Model 2, they already know how to assemble the product, resulting in a faster execution with less uncertainty and noise in execution times. Hence, it is advisable to cluster data before training a classifier.

The fourth represents the merging of similar QFMs into one QFM using a max operator; that is, if at least one of the QFMs is observed. Q1 and Q6, Q2 and A10, and Q2,

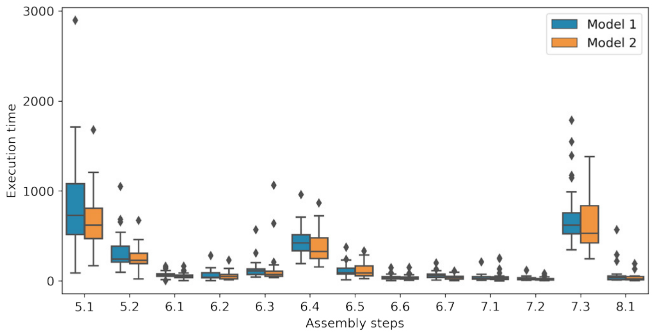

Fig. 3. Assembly steps execution times Boxplot

Q10, and Q11 were merged, resulting in approximately 50% balanced accuracy for both Model 1 and Model 2, with no improvement. The fifth considered the interdependencies between QFMs. We used CNN, considering all QFMs together, and NB with chain in Python. For CNN, results showed no improvement compared to disregarding the interdependencies for both Model 1 and 2. NB shows improvement for Q3 (63%), Q4 (70%), Q7 (62%), and Q9 (62%) in terms of balanced accuracy for Model 1. For Model 2, the results were the same as before.

Classification algorithms can assist in predicting some of the QFMs within an assembly line. The efficacy of these algorithms varies across different QFMs. Some QFMs, e.g., Q9, show a strong correlation with execution time, resulting in a high detection rate. For other QFMs, a low classifier performance suggests a low detection rate, meaning they are not easily detectable using the current datasets. To enhance the performance of a data-driven pFMEA, we discuss possible challenges in Section 6. However, the installation of high-precision IoT devices enabling the detection of all QFMs is necessary but comes at a high cost. The severity criterion in a pFMEA analysis would help, expressing the potential consequences of a QFM, such as the degree of impact on the system or customers.

6. DATA DRIVEN *P*FMEA CHALLENGES

While our study has limitations such as a fully manual process, a small number of cases, and a laboratory environment, tackling more complex cases with additional data wouldn't necessarily solve the problem (Mangal and Kumar, 2016; Schuh et al., 2021). We now elucidate the challenges gleaned from the experiments. Failure to address these challenges before employing a data-driven pFMEA/RCA may lead to unsatisfactory outcomes.

6.1 Small and Imbalance Data

The data available for analysis is usually limited, particularly for pFMEA, as it is usually conducted during the early stage of production. This scarcity of data poses challenges for training robust machine learning models, which often require large datasets to generalize well. For example, we only had access to 56 experimental trials.

More importantly, failure events are usually rare compared to normal operations. We may have a fair number of failed data for a pFMEA; however, this number will drastically decrease because of the learning effect, and the imbalanced

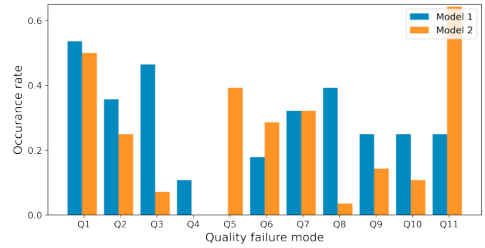

Fig. 4. Quality failure mode occurrence rates

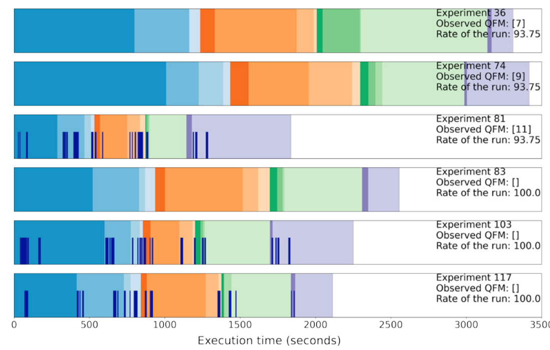

Fig. 5. Assembly steps execution times for some instances

data will be more observable for RCA. Classifiers lean towards predicting the majority class (normal operation) and overlook the minority class (failure events). Figure 4 shows the QFM occurrence rates for Models 1 and 2. This indicates that there is little or no failed data for Q4 and Q5 for Model 1 and Q3, Q4, Q8, and Q10 for Model 2. Comparing the accuracy of classifiers with the prior QFMs distributions reveals that the classifiers are learning from prior QFMs distributions instead of execution times in most cases. Mangal and Kumar (2016) observed that the failure rate for the Bosch case study was less than 1%.

6.2 Noise in Data

Noise in Inputs Noise refers to unwanted or random variations, in execution times for our case. Two reasons behind these noises are related to operators, such as the operator experience level and characteristics, and processes, such as the sequence of steps. Figure 5 presents an assembly summary of eight experiments with the least QFMs. The first part displays the execution time of each step, with each color representing a step, revealing a significant variation in the execution times. There are three reasons for our case. *Arrangement of the storage rack*: In Step 5.1, we observed that it was difficult for the operator to locate profiles on the rack *Unclear instructions*: For example, the use of the angle bracket fixture is not clear and is optional in the manual instructions. The operator did not use the angle bracket fixture. *Messy Table*: Sometimes operators (often in the first run) collect a lot of parts, spread them over the table, and then try to assemble them.

Noise in Outputs It is crucial to recognize the presence of noise in the outputs when assessing QFMs. This can be attributed to the subjectivity of a QFM assessment by a human or incorrect instructions given to an automatic quality control device. We observed such noise in for Q3 in our case study, where subjectivity is high.

Precision of Sensors Precision denotes the capacity of the sensor to deliver consistent and reproducible measurements under similar conditions. Capturing human behaviors, and delineating boundaries between different activities poses a challenge for sensors. Take, for instance, a camera sensor's inability to meticulously track steps in our case study, resulting in variations in the execution times. We observed execution times of less than 5 seconds for steps 6.1, 6.6, 6.7, 7.1, and 8.1 for multiple instances. Noise may manifest in other steps, e.g., in Step 5.1, the impact is less conspicuous due to the considerably larger standard execution time of that step compared to the noise.

6.3 Hidden and Confounder Variables

Hidden variables are factors not immediately obvious but can significantly impact the process. For instance, in our data, we observe learning effects, causing variations in execution times, as illustrated in Figure 3. Operators themselves introduce hidden variables due to differences in background, skills, experience level, and pace. In the first execution of Step 5.1, some operators initially struggled but learned over attempts, resulting in a faster second execution. Workforce changes, like new operators or turnover, can bring about learning curves and forgetting effects. Another hidden variable is optional tasks in assembly steps. Operators may encounter these tasks, adding variability to execution times that sensors may not easily detect.

Confounder variables, like the sequence of steps, can be overlooked, leading to potential misinterpretation of variable relationships. For example, the order in which operators perform tasks influences process efficiency; misalignment in the prescribed sequence may lead to an increase or decrease in execution times or other indicators. Operator support is another confounder introducing variation in execution times. Another factor is the inline fixing of errors, which might result in extended execution times.

6.4 Process and QFM Definitions

Process definition should be aligned with data-driven methods. However, other than the similarity between tasks in a step, there is no clear justification for defining a step in our case. Step 5.1 involves preparing 4 profiles of 1400mm, each with 4 angle brackets. Step 5.2 involves preparing 2 profiles of 1400mm, each with 2 angle brackets. Step 6.4 is about sliding and assembling 1 profile of 1400mm on a 2000mm profile. Step 6.5 is about sliding and assembling 3 similar profiles of 1400mm on a 2000mm profile.

Also, as sensors collect data from the production line, it is crucial to plan assembly processes based on the capabilities of the sensors to minimize noise in data gathering. Finally, QFMs should be well-defined, impartial, and objective, i.e., not influenced by personal opinions or feelings. However, the QFMs for the ProoVit case exhibit two flaws. Overlaps exist, e.g., Q2, Q10, and Q11 overlap. Some QFMs are ambiguous and subjective such as Q2 and Q3.

6.5 Other Considerations

Interdependencies between Failures The correlation analysis, illustrated in Figure 6, uncovers intricate relationships between ETAS and QFMs. Changes in processing

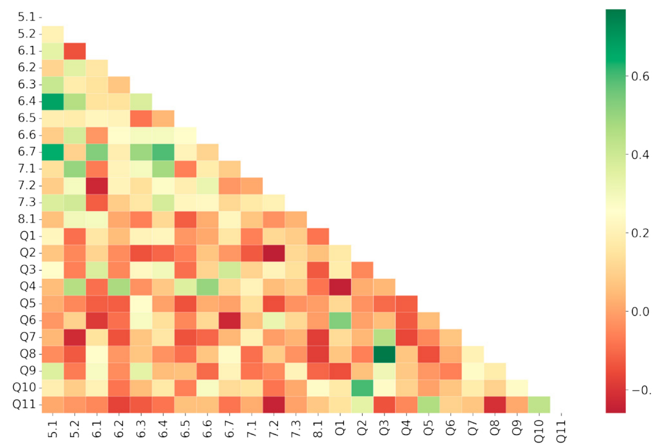

Fig. 6. Correlation between execution times and QFMs

time are seldom isolated to a single quality issue; instances marked by Q3 reveal additional underlying quality concerns. This underscores the necessity for a more robust analytical tool to comprehensively interpret the data. Considering the link between these issues and specific assembly steps, a comprehensive understanding requires exploring broader factors, including human elements and process intricacies, to inform effective root cause analysis and quality improvement strategies.

Outliers Figure 3 displays several significant outliers in the ETAS. Due to the nature of pFMEA, it is possible for a step to be performed with high variances, resulting in substantial outliers. A suitable strategy should be considered to address outliers.

7. CONCLUSION AND FUTURE RESEARCH

We focused on analyzing the execution times of assembly steps and identifying patterns related to QFMs. The QFM detection problem was formulated as a classification problem. Some QFMs showed more correlation with execution time, while others showed no correlation. Several challenges have been identified that hinder pure data-driven approaches. Although enriching the data with sources other than execution time through smart tools and/or data augmentation may improve the performance of a data-driven approach, these challenges should be carefully considered. To address these challenges, a major resource that may help is expert knowledge (Correia et al., 2023).

A future research direction is to develop a Bayesian network (BN). BNs can handle mixed data types, both categorical and continuous variables, while classification algorithms usually work with homogeneous data. Causal BNs explicitly model the cause-and-effect relationships between different variables, providing a deeper understanding of the underlying mechanisms in a system. Causal BNs offer a comprehensive view of how changes in one variable may influence others. This facilitates the identification of factors driving certain outcomes.

BNs present a possibility for the integration of data derived from sensors, human inputs, and other resources. Renu et al. (2016) highlights that pFMEA knowledge is typically expert-based and lacks organization. To address this, employing knowledge structuring techniques with ontology proves beneficial for reusing unorganized knowledge. Wehner et al. (2023) proposes the data-driven and knowledge-driven approaches because a data-driven approach could potentially have problems scaling real-world manufacturing processes. The primary advantage lies in the ontology's capacity to query concealed knowledge within the data. This integrated approach will involve the harmonization of data-driven and knowledge-driven methodologies to address challenges. Moreover, Razouk and Kern (2022) propose the use of the *p*FMEA to extract the causal relation knowledge to enhance the downstream methods in quality analysis such as RCA.

A method to explain hidden and confounder variables may improve QFM detection. Counterfactual reasoning has been developed for such situations where hidden variables influence the inputs and outputs. Counterfactual reasoning helps to evaluate what might have occurred if certain conditions or actions had been different. Counterfactual scenarios help to explore potential alternative outcomes of a given process or system. By considering hypothetical changes in variables or actions, analysts can identify critical points where the process might have deviated from the desired outcome, leading to QFMs. This method allows for the understanding of failure pathways, enabling robust preventive measures and improving the overall reliability.

REFERENCES

Carvajal Soto, J., Tavakolizadeh, F., and Gyulai, D. (2019). An online machine learning framework for early detection of product failures in an industry 4.0 context. *International Journal of Computer Integrated Manufacturing*, 32(4-5), 452–465.

Correia, A., Grover, A., Schneider, D., Pimentel, A.P., Chaves, R., De Almeida, M.A., and Fonseca, B. (2023). Designing for hybrid intelligence: A taxonomy and survey of crowd-machine interaction. *Applied Sciences*, 13(4), 2198.

Doltsinis, S., Krestenitis, M., and Doulgeri, Z. (2019). A machine learning framework for real-time identification of successful snap-fit assemblies. *IEEE Transactions on Automation Science and Engineering*, 17(1), 513–523.

Elanangai, V. and Vasanth, K. (2023). Defect detection and recognition of ss plate using deep autoencoder neural network. In *2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF)*, 1–8. IEEE.

Filz, M.A., Gellrich, S., Herrmann, C., and Thiede, S. (2020). Data-driven analysis of product state propagation in manufacturing systems using visual analytics and machine learning. *Procedia CIRP*, 93, 449–454.

Kang, Z., Catal, C., and Tekinerdogan, B. (2020). Machine learning applications in production lines: A systematic literature review. *Computers & Industrial Engineering*, 149, 106773.

Krauß, J., Pacheco, B.M., Zang, H.M., and Schmitt, R.H. (2020). Automated machine learning for predictive quality in production. *Procedia CIRP*, 93, 443–448.

Ma, Y., Zhou, H., He, H., Jiao, G., and Wei, S. (2019). A digital twin-based approach for quality control and optimization of complex product assembly. In *2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)*, 762–767. IEEE.

Mangal, A. and Kumar, N. (2016). Using big data to enhance the bosch production line performance: A Kaggle challenge. In *2016 IEEE International Conference on Big Data*, 2029–2035. IEEE.

Mokhtarzadeh, M., Rodríguez-Echeverría, J., Semanjski, I., and Gautama, S. (2024). Hybrid intelligence failure analysis for industry 4.0: a literature review and future prospective. *Journal of Intelligent Manufacturing*, In Press.

Muhr, D., Tripathi, S., and Jodlbauer, H. (2020). Divide and conquer anomaly detection: A case study predicting defective engines. *Procedia Manufacturing*, 42, 57–61.

Oliveira, E.E., Miguéis, V.L., and Borges, J.L. (2023). Automatic root cause analysis in manufacturing: an overview & conceptualization. *Journal of Intelligent Manufacturing*, 34(5), 2061–2078.

Razouk, H. and Kern, R. (2022). Improving the consistency of the failure mode effect analysis (FMEA) documents in semiconductor manufacturing. *Applied Sciences*, 12(4), 1840.

Renu, R., Visotsky, D., Knackstedt, S., Mocko, G., Summers, J.D., and Schulte, J. (2016). A knowledge based FMEA to support identification and management of vehicle flexible component issues. *Procedia CIRP*, 44, 157–162.

Sader, S., Husti, I., and Daroczi, M. (2022). A review of quality 4.0: Definitions, features, technologies, applications, and challenges. *Total Quality Management & Business Excellence*, 33(9-10), 1164–1182.

Schuh, G., Gützlaff, A., Thomas, K., and Welsing, M. (2021). Machine learning based defect detection in a low automated assembly environment. *Procedia CIRP*, 104, 265–270.

Sharma, K. and Srivastava, S. (2018). Failure mode and effect analysis (FMEA) implementation: A literature review. *Journal of Advance Research in Aeronautics and Space Science*, 5(1&2), 1–17.

Siva, V., Gremyr, I., Bergquist, B., Garvare, R., Zobel, T., and Isaksson, R. (2016). The support of quality management to sustainable development: A literature review. *Journal of Cleaner Production*, 138(2), 148–157.

Spreafico, C., Russo, D., and Rizzi, C. (2017). A state-of-the-art review of FMEA/FMECA including patents. *Computer Science Review*, 25, 19–28.

Tercan, H. and Meisen, T. (2022). Machine learning and deep learning based predictive quality in manufacturing: a systematic review. *Journal of Intelligent Manufacturing*, 33(7), 1879–1905.

Wehner, C., Kertel, M., and Wewerka, J. (2023). Interactive and intelligent root cause analysis in manufacturing with causal Bayesian networks and knowledge graphs. In *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, 1–7. IEEE.

Wu, Z., Liu, W., and Nie, W. (2021). Literature review and prospect of the development and application of FMEA in manufacturing industry. *The International Journal of Advanced Manufacturing Technology*, 112, 1409–1436.