Annals of Emerging Technologies in Computing (AETiC)
Vol. 4, No. 3, 2020

Research Article

Failure Mode & Effect Analysis and another Methodology for Improving Data Veracity and Validity

Ana Elsa Hinojosa Herrera*, Chris Walshaw and Chris Bailey

School of Computing & Mathematical Sciences, University of Greenwich, UK
aehinojosa@ieee.org; C.Walshaw@greenwich.ac.uk; C.Bailey@greenwich.ac.uk
*Correspondence: aehinojosa@ieee.org

Received: 29th April 2020; Accepted: 1st June 2020; Published: 1st July 2020

Abstract: Failure Mode & Effect Analysis (FMEA) is a method that has been used to improve reliability of products, processes, designs, and software for different applications. In this paper we extend its usage for data veracity and validity improvement in the context of big data analysis and discuss its application in an electronics manufacturing test procedure which consists of a sequence of tests. Finally, we describe another methodology, developed as a result of the DVV-FMEA application which is aimed at improving the tests' repeatability and failure detection capabilities as well as monitoring their reliability.

Keywords: ***Big Data; Data Veracity; Data Validity; FMEA; Statistics; Electronics Manufacturing; Quality Assurance; Test Limits Optimisation***

1. Introduction

The market of data analytics was valued at USD 904.65 million in 2019 and is expected to reach USD 4.55 billion by 2025 [1]. Moreover, the use of data driven techniques is popular in smart manufacturing. Cost reduction can be achieved by mining data for predicting the quality of a batch, improving robustness of processes, or by reducing the process cycle time, for example.

With regards the definition of big data, the authors in [2] describe it using 1C for complexity and 11Vs for: Volume, Velocity, Variety, Volatility, Virtual, Visibility, Vendee, Vase, Value, Veracity, and Validity. In this paper we cover the last 2 Vs of the list.

Failure Mode and Effect Analysis (FMEA) is a method that has been used to improve reliability, testability and safety of hardware designs, processes, products, and software, for example [3-6]. In electronics, hardware (HW) FMEA has been used to improve electronics reliability [4], and in [7] software (SW) FMEA was used to validate embedded real time systems.

In this paper we extend the usage of the FMEA method to improve data veracity and validity. The proposed extension (DVV-FMEA) is illustrated with an electronics manufacturing application for quality assurance. From using DVV-FMEA in this application a novel methodology was motivated for evaluating, improving and monitoring the definition of production tests.

This article is organized as follows. Section 2 introduces the data veracity and validity concepts and main causes that commonly affect data quality. Section 3 discusses the usage of FMEA for data improvement and its application in production testing data. Sections 4 and 5 present the methodology for test definition evaluation, improvement, and monitoring, in addition to its application in a production test dataset, respectively. And finally, Section 6 concludes the article and states future work.

Ana Elsa Hinojosa Herrera, Chris Walshaw and Chris Bailey, "Failure Mode & Effect Analysis and another Methodology for Improving Data Veracity and Validity", Annals of Emerging Technologies in Computing (AETiC), Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 9-16, Vol. 4, No. 3, 1st July 2020, Published by International Association of Educators and Researchers (IAER), DOI: 10.33166/AETiC.2020.03.002, Available: http://aetic.theiaer.org/archive/v4/v4n3/p2.html.

2. Data Veracity and Validity

Poor data veracity and validity improvement is relevant for big data applications, because low quality data could generate inaccurate models and unreliable information, resulting in incorrect data-driven decision taking. In this section we discuss the characteristics of data veracity and validity.

2.1. Data Veracity

Data veracity is the ability to understand the data and the analytical process applied to a dataset. It covers aspects related to confidence in the dataset or data source, for example data integrity, availability, completeness, consistency, and accuracy and in addition, transparency and clarity in the processes used to generate, improve and analyse the dataset [2, 8, 9]. Authors in [10] discuss a general list of causes that frequently affect data veracity:

- Measurement system limits: For example, equipment calibration, human errors, and non-standard measurement processes.
- Limits of features extraction: This could be evaluated by measuring the precision of correctness and completeness.
- Data integration limits: In real applications it is useful to gather and combine information from different sources, but sometimes it is challenging due to the diversity of data sources or formats.
- Data ambiguity and uncertainty: In addition to the uncertainty due to data integration there are other sources of data ambiguity, for example ambiguities of natural language, uncertainty related to the information source and low relevance of the information with respect to other available information [11].
- Data falsification and source collusion: In [12] authors model data falsification attack as a constrained optimization problem with two parameters: efficacy and covertness of the attack. The first parameter is related to the degradation in the detection performance, and the second one is the probability that the attacker will not be detected. In the formulation, the attacker would maximize the attack efficacy while controlling its exposure to the defence mechanism.

2.2. Data Validity

Data validity refers to data worthiness, which may change over time and during the process under study. For example, data generated before relevant changes in the process is not valid to generate models of the current state [2].

The authors in [13] discussed data staleness for information systems where data is frequently updated. This data freshness characteristic is relevant, for example, in data streaming applications where information quickly becomes obsolete.

3. Data Veracity and Validity Failure Mode and Effect Analysis

In Section 2 we discussed the importance of veracity and validity. In addition, we noted its impact on data-based decision-making success. In this section we are going to present the DVV-FMEA steps to follow for improving these two elements of the big data definition, and the results of its usage in an electronics manufacturing quality assurance application.

3.1. Steps of DVV-FMEA

The DVV-FMEA is like HW FMEA, although with differences in System Identification, List of Failure Mode, Causes Identification, and Effect Analysis steps. The details as follows:

Step 1. System Identification: In data-driven analysis, it is common that the modules identified in the process before using datasets for analysis consist of data generation, data storage, data gathering, and data pre-processing. Nevertheless, in some applications where data is streaming the storage module could be different.

As in SW FMEA, the variables or features in the dataset must be listed for its evaluation. When working on big datasets which comprise a big quantity of variables, it seems sensible to group them based on engineering feature or data processes similarities.

Step 2. List of Failure Modes Generation: It make sense to split the meeting time into the different modules and generate a failure modes list for each of these. The brain-storming meeting(s) should include team members with know-how and expertise in the data process and application.

Step 3. Causes Identification: List the causes of failure modes and score them by its occurrence. We recommend including causes related to measurement system limits, features extraction limits, data integration limits, data ambiguity and uncertainty, data falsification and source collusion, data staleness. Ishikawa diagram is a useful tool which could be used as a guidance for causes identification. In Fig. 1 is the version we propose for causes identification in DVV-FMEA. It could be used for each failure mode identified in Step 2.

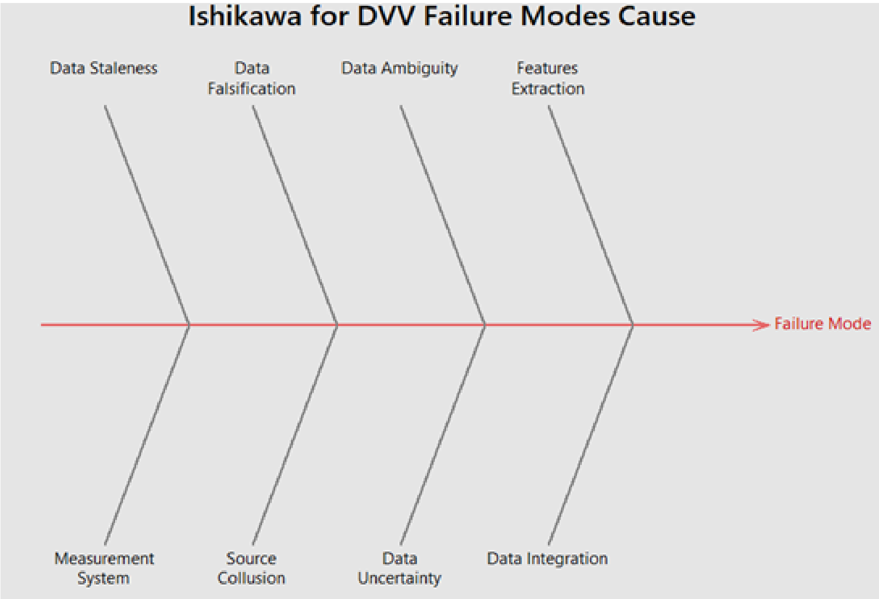

Figure 1. Ishikawa Diagram for DVV Failure Modes Causes

Step 4. Effect Analysis: In this step the effects of the failures are listed, and each of the effects is scored by its severity. It makes sense to include impacts to confidence in the dataset or data source, data integrity, data availability, data completeness, data consistency, data model, or analysis accuracy, execution time or efficiency, ability to replicate results or analysis, and data worthiness.

As a guidance during the meeting, the DVV-FMEA leader could ask if and how each of the impacts listed above impacts the failure mode and fill it in the DVV-FMEA table.

The following steps are the same as in HW FMEA.

Step 5. Detection mechanism identification: A list with the available mechanisms that helps detecting the failure modes is generated. Each failure mode should have a score of its detectability.

Step 6. Failure mode prioritization: In order to improve the efficiency of this method, the list of failure modes should be filtered based on the Risk Priority Number (RPN), which is calculated as in:

Equation 1. Risk Priority Number

$$RPN = Severity \ \times Occurrence \ \times Detection$$

Step 7. Process or Product Improvement: Based on the prioritization and resources available, the next step is to generate and execute an improvement plan, which contains actions to improve the data veracity and validity. These changes should reduce the score of severity, occurrence, or detection. It seems likely that severity score is less frequently reduced.

3.2. Severity, Occurrence, and Detection Scales

For the scaling it makes sense to use simple scales for severity, occurrence, and detection scores. For example, a 5 levels measure such as the Likert scale, which is easy to use. In Table 1 is detailed the ranking scale we recommend. Whenever historical data or a previous DVV-FMEA is available, it could be used to quantify the severity, likelihood, or detectability rates.

Table 1. Occurrence, Severity, and Detection Ranking Scale

Ranking	Occurrence	Severity	Detection
1	No known failures	Very low or none	Almost certain detection
3	Isolated failures	Low or minor	Remote chance of detection
5	Occasional failures	Moderate or significant	Moderate chance of detection
7	High rate of failure	High	High chance of detection
10	Failure is almost inevitable	Very high or catastrophic	Cannot be detected

3.3. DVV-FMEA Application in Production Testing

In this subsection we include DVV-FMEA usage to establish the pre-processing step of the data analysis of an electronics manufacturing application. Experts in the manufacturing and data processes were part of the team that generated the DVV-FMEA table.

In this application the input variables are the result of individual tests in a sequence that runs in a stop-on-fail scenario. For some tests in the sequence, a feature is measured and then compared to upper, lower or both limits to classify faulty devices. More details of the application and intermediate steps of the DVV-FMEA can be found in [14].

As a result of using the DVV-FMEA, and based on the RPN, the list of +60 failure modes related to data validity and veracity was reduced to 14. Some of them are included in Table 2. Most of the improvements comprise R scripts that pre-process data before its usage for analysis. The scripts detect incorrect data and eliminate it, correct formats, and standardize data pre-processing steps to ensure repeatability, consistency, efficiency, and confidence.

Table 2. DVV-FMEA for an Electronic Manufacturing Application

System Module	Input	Failure Mode	RPN
Data Generation	Overall result	The overall result is not consistent	490
Data Generation	Text File	The file format is not correct	100
Data Generation	Test: 90, 480	The test was unsuccessful to detect faulty devices	150
Data Generation	Test type	Different to test sequence ‘p’	50
Data Generation	Dataset	Data does not represent the current process conditions	250
Data Pre-processing	Data order	The data is not ordered by date-time	70
Data Pre-processing	Clean dataset	No clarity on how the data was processed before using it for analysis	49
Data Pre-processing	Test/Training datasets	The sampling is not repeatable	70

The failure mode that has the highest priority is that the overall test result is not consistent, impacting the effectiveness of the test but also its efficiency because extra analysis is performed to ensure the good quality of the devices. The definition of the limits is relevant not only to the accuracy of the tests and the overall result, but also to its efficiency, because in the application one faulty characteristic of the device could be detected by more than one test in the sequence, but the earlier the fault is detected, the shorter the length of the test procedure. In Section 4 we present a methodology proposed to improve the definition of the tests. It was automated using a Python script implemented in a Jupiter notebook.

Another failure mode with high priority is to avoid using out-of-date data for data analysis because the model would not be useful for the current state. This failure mode is relevant because in real applications it is very common that the processes change over time, for instance using new raw materials, updates to the design, or improvements to the manufacturing procedures. The methodology in Section 4 includes a monitoring phase which could be used for data analytics reliability as well.

4. Test Limits Evaluation, Improvement and Monitoring Methodology

The tests limits evaluation and improvement process we propose consists of four main phases: Test Efficiency Evaluation, Test Utility to Improve another Test Evaluation, Re-Define Test Limits, and Limits Monitoring.

4.1. Phase 1: Test Efficiency Evaluation

In this phase the aim is to evaluate each test in the sequence, comparing the data distribution versus test limits for FS-PTx, PS, and FTx samples.

Step 1. Select a Test_x in the Sequence: The earlier in the sequence the better because potentially there is more improvement when finding a fail early in the sequence.

Step 2. Split the Dataset into FS-PTx, PS, FTx: Here FS-PTx contains data of assets that failed the test sequence but in another test different to Test_x, PS contains the data of assets that passed the test sequence, and FTx is the data of assets that fail Test_x.

Step 3. Plot Histograms for FS-PTx, PS, FTx: In the histograms can be visualised how each of these datasets performs versus the Test_x limits, if there is a partition between the three datasets, and if the datasets correspond to the same distribution.

Step 4. Calculate Statistics for FS-PTx, PS, FTx: Descriptive statistics are useful for understanding the datasets. It makes sense to include mean, standard deviation, quartiles, maximum and minimum.

Step 5. Partition Evaluation: Quantify the distance between PS and FTx populations. We propose using the following formulas:

Equation 2. Partition Evaluation around Lower Limit

$$\max(\mathrm{FTx}_{below\ ll}) + 2 * \mathrm{np.std}(\mathrm{PS}_{between\ 0.15\ and\ 0.85\ quartiles}) < \ \mathrm{Tx\ lower\ limit}$$

Equation 3. Partition Evaluation around Upper Limit

$$\min(\mathrm{FTx}_{above\ ul}) - 2 * \mathrm{np.std}(\mathrm{PS}_{between\ 0.15\ and\ 0.85\ quartiles}) < \ \mathrm{Tx\ upper\ limit}$$

Where $\mathrm{FTx}_{\mathrm{below\ ll}}$ = {y in FTx ∣ y < Tx lower limit}, $\mathrm{FTx}_{\mathrm{above\ ul}}$ = {y in FTx ∣ y > Tx upper limit}, and $\mathrm{PS}_{\mathrm{between\ 0.15\ and\ 0.85\ quartiles}}$ = {y in PS ∣ y > PS quartile 15% & y < PS quartile 85%}.

Step 6. Is there a Partition Between PS and FS-PTx? Using results of Steps 3 to 5 of this phase, when the answer is positive, the recommendation is to add or update the limits for Test_x.

Step 7. Are PS & FTx Clearly Separated? Using results of Steps 3 to 5 of this phase, when the answer is negative, the recommendation is to reconsider the limits for Test_x.

Step 8. Is FTx Empty? If the data of FS-PTx, PS, FTx are a representative sample, it can be inferred that it is highly probable that Test_x is passed, as a result could be eliminated from the sequence, or reduced the frequency of its execution.

4.2. Phase 2: Test Utility to Improve another Test Evaluation

In this phase the aim is to identify relationships between tests and whether one test could be used to calculate the result of another one. The steps are as follows:

Step 1. Select Test_y in the sequence: Here Test_y is another test in the sequence which is executed after Test_x.

Step 2. Are both continuous variables? If Test_x and Test_y measurements are continuous values, calculate Pearson Correlation Coefficient to quantify its association. If the coefficient is > 0.9 or < -0.9 the conclusion is that both tests are highly associated.

Step 3. Are both discrete variables? If Test_x and Test_y measurements are discrete values, execute a Chi-Square Test to quantify their association. If the p-value is < 0.05 the conclusion is that both tests are highly associated. When the test sequence is run on stop-to-fail scenario, this test cannot be performed, since the dataset contains "pass" and "fail" data for Test_y but only "pass" for Test_x.

When associated Tests are found in Steps 2 and 3, sometimes the association between them could be used to estimate the value of Test_y instead of performing the reading. As a result, the test sequence potentially could be reduced.

4.3. Phase 3: Re-Define a Test Limit

In this phase, the results of previous phases are summarised and joined after solving possible conflicts, followed by the implementation and documentation of changes. The details as follows:

Step 1. Improvements Summary: Summarise the recommendations from Phase 1 and 2.

Step 2. Feasibility Evaluation: Evaluate if the new test limits are correct from customer and engineering point of view.

Step 3. Conflict Evaluation: Also evaluate if the recommendations are not in conflict, otherwise evaluate which is the recommendation that generates more improvement.

Step 4. Update Test Limits Definition: The automated test sequence should be updated with the new test limits definition. It is likely that this motivates a new software version, which may need to be certified as part of software quality processes.

Step 5. Document Changes: We recommend that these changes and verifications to be documented on the DVV-FMEA to have all information related to data quality improvement in a single document.

4.4. Phase 4: Limits Monitoring

The objective of this phase is to continuously evaluate whether the new limits are valid, or a re-definition is needed.

Step 1. Metrics Definition: It is relevant to select the most representative metrics to monitor, and it makes sense to choose only a few and to prefer the ones which are easy to measure.

Step 2. Continuous Monitoring: We recommend using statistical process control charts to monitor the key metrics. To keep the manufacturing process as simple as possible, it makes sense to have a small list of key elements to monitor, and also to automate this step, and consider automated flags or warnings when the key elements are not in control.

Step 3. Maintenance: Whenever any of the key monitored parameters are not in control it is time to revisit Phases 1 to 5 of this methodology.

5. Test_80 Evaluation and Improvement

In this subsection the methodology we proposed in previous section is illustrated using the Test_80, which is part of the test sequence analysed in the DVV-FMEA we included in Section 3.

In **Figure** *2* the histograms of assets that passed the test and in **Figure** *3* the histogram of assets that failed the test. In both figures, the upper and lower limits of Test_80 are indicated in vertical lines.

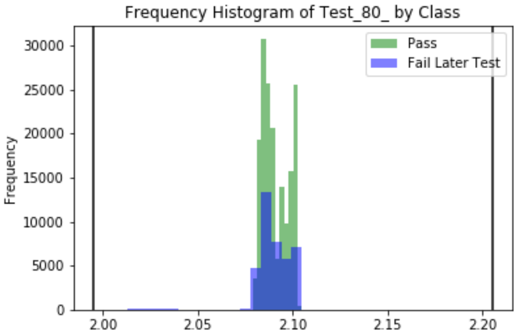

Figure 2. Histograms of Assets that Passed Test_80

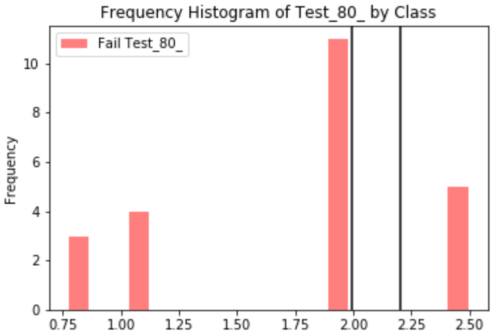

Figure 3. Histogram of Assets that Failed Test_80

Table 3. Statistics of Test_80 Samples

Statistics	PS	FS-PT80	FT80
Count	171131	39846	368
Mean	2.090	2.089	1.694
Std	0.006	0.010	0.432
Min	2.057	1.996	-0.140
25%	2.085	2.085	1.470
50%	2.088	2.089	1.473
75%	2.097	2.096	1.949
Max	2.104	2.104	2.697

From the histograms we can note that FS-PT80, PS and FT80 populations are not clearly separated. They are close around Test_80's lower limit. In addition, most of the assets, which failed

Test_80, are near its lower limit. The statistics in Table 3 are in line with this conclusion. Furthermore, the results of the partition evaluation recommend re-defining the Test_80 lower limit.

Following with the methodology, every test in the sequence was evaluated as stated in Phase 2. We found that there is a linear relation between Test_80 and Test_220. Furthermore, all are faulty assets when Test_80 < 2.05 & Test_220 > 2.05. Also, when Test_220 < 1.95 (Fig. 4).

Based on previous results, we recommend changing the lower limit of Test_220 to 2, and the lower limit of Test_80 to 2.05. After the company has implemented these changes in their software, we could continue with the monitoring stage.

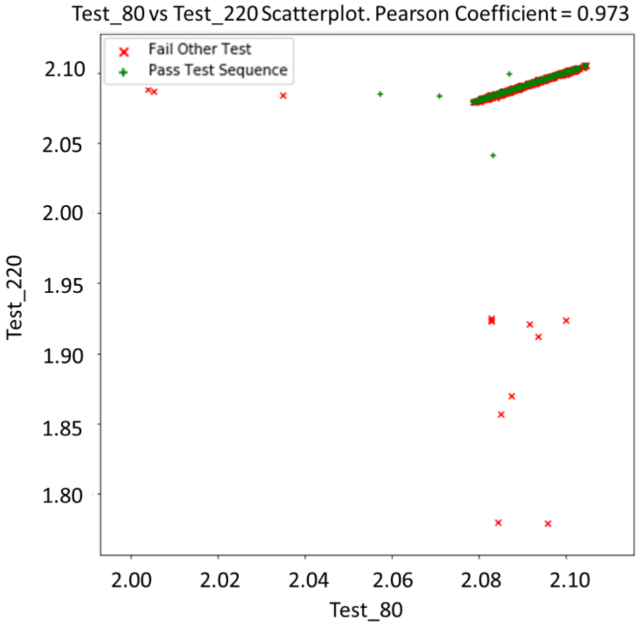

Figure 4. Test_80 and Test_220 Correlation

6. Conclusion and Future Work

In this paper an extension of the FMEA method was proposed to upgrade data veracity and data validity, two relevant characteristics in the big data paradigm. As discussed, an early identification and mitigation of potential failures in data processes can impact the accuracy of the data-driven models and analysis. Another benefit of using DVV-FMEA in the early stages of a data analysis project is that, as the method is applied, experts can transfer know-how, data understanding, and business priorities, which are relevant elements of big data and key elements for the success of further analysis.

The DVV-FMEA is presented as a complementary method to improve data veracity and validity, with improvements driven by feature engineering and experts' know-how rather than on purely data statistics analysis.

The proposed DVV-FMEA method applied to a dataset from production testing of electronical devices satisfactory improved the data in terms of veracity and validity. For this application, DVV-FMEA was able to obtain and document know-how of the experts in the electronical manufacturing and data generation processes, which would be useful for further analysis. In addition, a methodology for improving the definition of the tests was motivated, described and illustrated in this paper.

From applying the methodology proposed to improve tests limits, we were able to do a better definition of some tests in the sequence that is used as part of the electronics manufacturing quality assurance procedure of the electronics application discussed.

Future studies should aim to replicate results in other tests in the sequence, as well to implement the monitoring phase of the methodology proposed.

Note

This paper is an extended version of a presentation given at London Metropolitan University, London, UK in August 2019, as part of the IEEE International Conference on Computing, Electronics & Communications Engineering 2019 (iCCECE ’19) conference.

Acknowledgment

The authors would like to acknowledge the support from Microsemi Corporation (A Microchip company) for providing the dataset, collaborating with the FMEA elaboration and sharing their domain knowledge with us.

References

[1] Mordor Intelligence. (2019). Big data analytics in manufacturing industry market - growth, trends, and forecast (2020 - 2025). Mondor Intelligence. Available: https://www.mordorintelligence.com/industry-reports/big-data-analytics-in-retail-marketing-industry.

[2] Patgiri R. and Ahmed A. (2016). Big Data: The V's of the Game Changer Paradigm. 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS), Sydney, NSW, pp. 17-24.

[3] Huang B., Zhang H. and Lu M. (2009). Software FMEA approach based on failure modes database. 2009 8th International Conference on Reliability, Maintainability and Safety, Chengdu, pp. 749-753.

[4] Cui J., Ren Y., Yang D. and Zeng S. (2015). Model based FMEA for electronic products. 2015 First International Conference on Reliability Systems Engineering (ICRSE), Beijing, pp. 1-6.

[5] Yuan G., Shi G., Hou N. and Kuang X. (2014). The application of FMEA technology in the unstable failure analysis. 2014 10th International Conference on Reliability, Maintainability and Safety (ICRMS), Guangzhou, pp. 663-665.

[6] Zhang H., Wang T., Shao J., Lu F. and Yan H. (2015). Advantage analysis of FMEA technique based on EDA simulation, First International Conference on Reliability Systems Engineering (ICRSE), Beijing, pp. 1-5.

[7] Goddard P. L. (1993). Validating the safety of embedded real-time control systems using FMEA, Annual Reliability and Maintainability Symposium 1993 Proceedings, Atlanta, GA, USA, pp. 227-230.

[8] Moyne, J., Iskandar, J. (2017). Big Data Analytics for Smart Manufacturing: Case Studies in Semiconductor Manufacturing. Processes, 5, 39. Available online at: https://doi.org/10.3390/pr5030039.

[9] Batista A. F. M., da Silva D. L. and Correa P. L. P. (2017). Enabling Data Legitimacy in Data-Driven Projects, 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC), Mysore, pp. 50-54.

[10] Laure B., Javier B. (2015). Veracity of Data: From Truth Discovery Computation Algorithms to Models of Misinformation Dynamics, Morgan & Claypool.

[11] Dragos V. (2014). Assesment of uncertainty in soft data: A case study. 17th International Conference on Information Fusion (FUSION), Salamanca, pp. 1-8.

[12] Kailkhura B., Han Y. S., Brahma S. and Varshney P. K. (2013). On covert data falsification attacks on distributed detection systems, 2013 13th International Symposium on Communications and Information Technologies (ISCIT), Surat Thani, pp. 412-417.

[13] Chayka O., Palpanas T. and Bouquet P. (2012). Defining and Measuring Data-Driven Quality Dimension of Staleness.

[14] Hinojosa A., Walshaw C. and Bailey C. (2019). Failure mode & effect analysis for improving data veracity and validity, Proceedings of IEEE International Conference on Computing, Electronics & Communications Engineering, London, United Kingdom, pp. 100-105.