

Analysis of Relationships Between Defects and Variables in Wire Mesh Production Using Association Rules and Simple Regression Analysis

Aree Wiriyaiphongsanon¹ and Rungchat Chompu-inwai²⁺

¹ Master's Degree Program in Industrial Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand

² Department of Industrial Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand

Abstract. Many categories of defects frequently occur in the production process of wire mesh used for construction. This research has the main objective of finding variables or factors in production which are related with the amounts of defects occurring, to find ways of reducing them in future. The procedure of this research started from studying the wire mesh production process in the case study company and selecting a product for the study. For the next step, as there were many categories of defect, these categories were prioritized to select those which required prior improvement using the Pareto principle and the principle of Association Rules. The latter were used to analyze relationships between defects such as whether any kinds of defects occurred together, to help reduce the categories of defects requiring study. The research found that there were two important categories of defect which should be addressed as priority, type S1 defect and type S2 defect. From there analysis was conducted to identify the causes of each kind of defect with a Cause and Effect Diagram. A brainstorming session with the relevant personnel found that machinery settings were the factors causing opportunities for the greatest amounts of defects to occur. The researchers then collected data of detailed machine settings and defects which had occurred in the past. Simple Regression Analysis was then used to identify the machine setting factors related with the amounts of defects. The research found that the type S2 defect was related with only the X_8 factor, while the type S1 defect was related with the factors X_3 , X_4 and X_{10} .

Keywords: defects, association rules, simple regression analysis.

1. Introduction

The wire mesh industry is growing and generating increased income from consumer demand. It produces both for domestic consumption and export. The case study company manufactures wire mesh which is further processed into many other products such as wire conveyor belt, sieves, wire mesh for construction and filtration products. Wire mesh for construction products are the highest volume wire mesh product category. The case study company was therefore interested in finding ways to improve the production process for this product group.

The wire mesh production process gives rise to many kinds of defects in large amounts on the sheet as shown in Fig. 1. The type S1 defect occurs from the weft thread twisting. The type S2 defect is irregular mesh holes from unequal spacing of the weft threads, occurring in series. The type S3 defect is when a hole is created by the warp thread breaking. Defects mean that either the sheet has to be repaired or removed from production to prevent defective product being shipped to customers. If there are so many defects, repairs are not possible. The company therefore seeks ways to maximally reduce defects to develop the product to be better in line with customer demand. This will reduce expenses associated with defective product and

⁺ Corresponding author. Tel.: + 66 53 944126; fax: +66 53 944185
E-mail address: rungchatc@hotmail.com

increase profits by cutting production costs. Therefore, this research has the principal objective to find variables or factors in production which are related with the amounts of defects occurring in the wire mesh sheet, to find ways to reduce defects in the future.

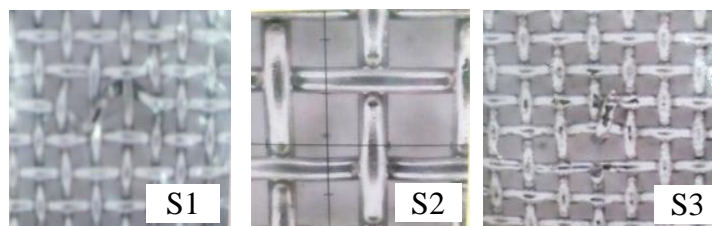


Fig. 1: Characteristic wire mesh defects.

Association Rules data mining was applied to identify relationships of two or more datasets. One example of Association Rules currently used is Market Basket Analysis to find relationships between products in a customer's shopping basket, for use in arranging sales promotions. There are many procedures and methods to find Association Rules, but a widely known and used method is Apriori. For example, research [1] using Association Rules which applied Apriori Algorithm to find a profile of shoppers' behavior, to improve the efficiency of sales processes in the textile industry. Study [2] sought Association Rules between parameters of yarn and fabric properties for use in designing fabric properties in line with the demands of consumers in the textile industry.

For this research, as there are many kinds of defect in the wire mesh, it was necessary to prioritize defects to select those needing prior attention. Association Rules were used initially to analyze relationships between defects, as to whether any kinds of defects occurred together, to reduce the categories of defects under study. This was done in conjunction with the use of a Pareto Diagram evaluating the losses from each category of defect.

In the analysis to find the relationships between factors from historical data, previous studies found that there were many applicable techniques such as Data Mining using Decision Trees or Correlation-Based Feature Selection. These methods are suitable for qualitative data. There are also techniques of Correlation Analysis which can be analyzed to find the relationships between factors and responses [3]. This technique is focussed on analysis to find a conclusion about how and to what extent the two factors are related. This is different from the Simple Regression Analysis, a method to find the linear relationship between an independent variable (X) and a dependent variable (Y). The objective of this is to use the resulting relationships obtained to make predictions of factors according to the dependent variable in a Simple Linear Regression Model.

Past studies such as [4] used a Simple Regression Analysis to find correlations between labour productivity and the value added. This analysis found that the value added of industry in Thailand had a direct effect on the labor factor. There is also research [5] dedicated to creating a prediction model for forecasting rice production which was able to reduce the factors of rice planting in the equation to 11 factors with a Simple Regression Analysis. In addition, there is research [6] which studied the relationships of factors in producing motorcycle front fork and wheels with manufacturing costs, with a Regression Analysis. This found that three factors in production were significantly related with costs with a confidence level of 95%. This research is thus using Simple Regression Analysis to find factors in machinery settings which have relationships with defects occurring in wire mesh products.

2. Research Methodology and Results

As stated above, this research has the principal objective to find variables or factors in production which are associated with the amounts of defects in wire mesh products, to find ways to reduce defects in future. The researchers then specified research methodology in 4 principal stages as shown in Fig. 2 as follows:

2.1. Studies of the Wire Mesh Production Process and Choosing Products for Study

The wire mesh production process starts from inspecting the quality of the wire and then winding the wire onto the beam, which is sent for weaving in the weaving room. When the weaving process is complete,

the workpieces are sent for quality inspection and then washed and cut to size according to the customer's need, and then packaged ready for shipment to the customer.

The wire mesh for construction products have the greatest production volumes compared to other wire mesh products. The case study company was thus interested in finding ways to improve the manufacturing process for this group of products.

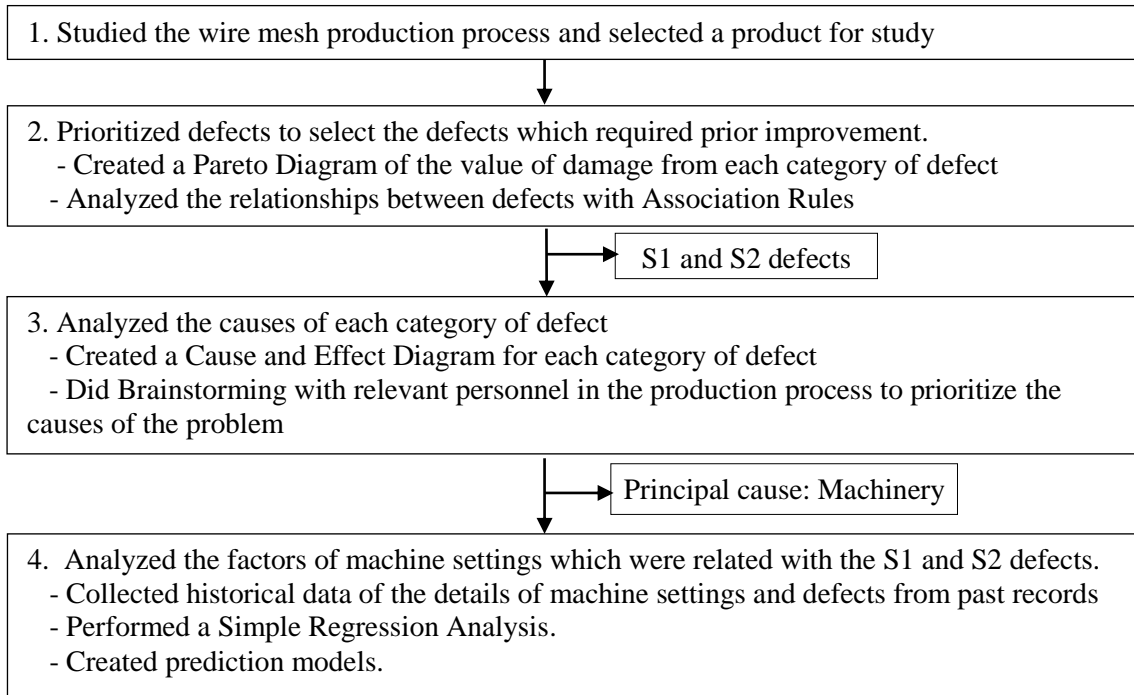


Fig. 2: Research methodology.

2.2. Prioritizing Defects to Select Defects for Prior Improvement

This stage is prioritizing defects to select the defects requiring prior improvement, using the Pareto Principle together with the principles of Association Rules. The latter are implemented to analyze the relationships between defects, as to whether any kinds of defects occur together, to reduce the categories of defects under study, with the following details:

2.2.1. Pareto Diagram

Fig. 3 shows the Pareto Diagram of the losses caused by each category of defect (unit: Thai Baht).

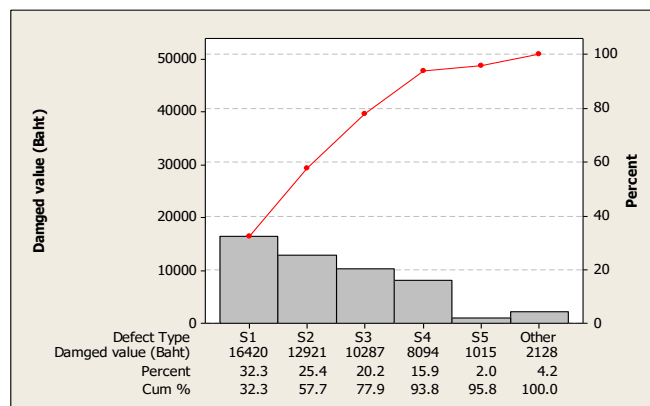


Fig. 3: Pareto Diagram showing the losses caused by each category of defect.

From the 80/20 Rule of Pareto, it was found that types S1, S2 and S3 defects were in the “Vital Few” group, i.e. few categories, but causing a great value of losses, while the many other remaining defects combined caused only minor losses [7].

2.2.2. Data collection and data table for association rule mining

This started from studying data about categories of defects occurring in the wire mesh workpieces, to ascertain the number of defects and which defects were occurring in wire mesh for construction products. Data were then gathered of defects for each production lot from year 2012 to 2015 for analysis to find the relationships between defects. Data were collected for 52 lots and 20 categories of defects.

Table 1 is the data table for applying Association Rules mining. The first column is the production lot which total 52. The other columns are the categories of defect arising, which total 20 categories of defect. If any category of defect occurs in each lot, it is given the value of 1 in the column of that defect category. A value of 0 is given in the column of a defect category, if it has not occurred in that lot.

Table 1: Data table for association rules mining

Lot	S1	S3	S4	S5	S6	...	S2
1	0	0	0	1	0	...	0
2	0	0	1	1	0	...	1
3	0	1	1	1	0	...	1
...
52	0	1	0	1	0	...	0

2.2.3. Association rules implementation and results

Association Rules Mining is searching for relationships hidden in the dataset and comprises two main stages as follows [8-10]:

The first stage is searching for itemsets which occur together frequently (frequent itemset generation). Many algorithms can be implemented at this stage such as FP-Growth, PrePost+, Hmine and Apriori. This research selected the Apriori Algorithm as has been popular in past research. It has the principle of mining relationships by making a wide search first, as follows:

- Creating candidate itemsets level-wise to seek frequent itemsets. All transactions are read to find the support value.
- Looking at the support values of that candidate itemset. If the support value is below the specified minimum, that candidate itemset is pruned. Candidate itemsets with equal or greater support value are considered candidate itemsets at the next level.
- Creating candidate itemsets from the previous frequent itemset to reduce the number of candidates.
- Apriori Algorithm creates and discovers whether an itemset is a frequent itemset by counting down transactions in this matter continuously until the data is exhausted and no more candidates which pass the minimum support value can be created.

Stage 2 is to create Association Rules from frequent itemsets discovered in the previous stage, which have support values and confidence values above the minimum (called “strong rules”).

This research used the program Rapid Miner Studio 7.3 for the analysis by defining a minimum support value of 45% and a minimum confidence value of 60%.

Table 2: Strong association rules derived from data mining analysis

No.	Association Rules	Support	Confidence
1	S3 ==> S2	0.462	0.632
2	S2 ==> S3	0.462	0.774

Table 2 found support values of 0.462 which means 46.2% of lots had type S2 and type S3 defects occurring together. Considering No.2, the Confidence level was found to be 0.774, showing that in 77.4% of the lot with type S2, type S3 defects also occurred. To reduce the amounts of defects to study, the researchers thus chose only type S2 defect for this study.

The reason for reducing the amounts of defects to study was that future research would use the results of a Simple Regression Analysis (to be mentioned under the next heading) to specify “potential design factors” in the analysis using Design of Experiments. If the number of design factors is too many, this requires much time and expense for the experiment, as one change of machine settings for test production takes 3 days and costs 31,500 Thai baht, while also requiring a shutdown of normal production for the test.

Therefore, by combining Association Rules and the Pareto Diagram, only type S1 and type S2 defects were selected for study at the next stage.

2.3. Analysis to Find the Causes of type S1 and type S2 Defects

The researchers brainstormed with the relevant parties from the production line to find the causes of type S1 and type S2 defects using Cause and Effect Diagrams. Fig. 4 shows the causes of type S1 defect from analysis of the 4M; Man, Machine, Method and Material (Cause and Effect diagram to find the causes of type S2 defect is not shown here but is done in the same manner).

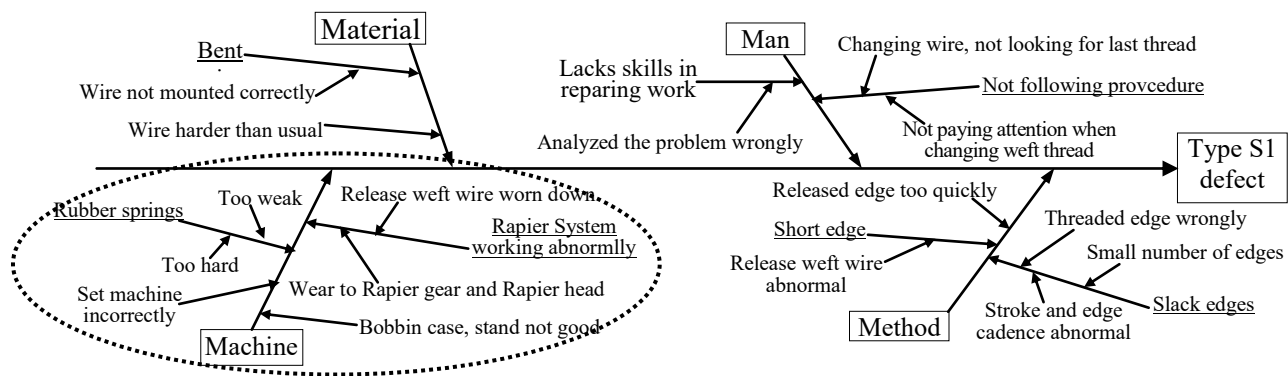


Fig. 4: Cause and effect diagram showing the causes of type S1 defect.

Then there was another brainstorming session with the relevant production people to prioritize the causes of defects. It was found that the Machine group had the most result in the chance of producing both categories of defect. This group was also feasible to study and improve further.

2.4. Analysis of Factors in Making Machine Settings which are Related with the Occurrence of type S1 and type S2 defects

When it was known that Machine causes had the most effect on the chance of both kinds of defect to occur, at this stage the researchers collected historical data from year 2012 to 2015 from the following two departments:

(1) Details of the machine settings for each lot. This data revealed that there were wide ranges for the machine settings in the manual. However, each worker set the machine differently, depending on the experience and expertise of each worker to set the machine.

(2) The amount of type S1 and type S2 defects occurring in each lot from the Quality Control Department.

Then was performed a Simple Regression Analysis to find factors in setting the machine which were related with each category of defect. Simple Regression Analysis is estimating the value of a dependent variable (here being the amounts of defects occurring) by relying on data obtained from independent variables (here being the factors of machine settings), and then doing Hypothesis Testing using the t-Test. The statistical hypotheses are specified as follows:

H_0 : Amount of type S1 (or S2) defects in the lot t (y_t) not related with factors in setting the machine (X_i)

H_1 : Amount of type S1 (or S2) defects in the lot t (y_t) related with factors in setting the machine (X_i)

where y_t is the amount of type S1 (or S2) defects occurring in lot t , where $t = 1, 2, 3, \dots, 52$

X_i are the factors in making settings on the machine, where $i = 1, 2, \dots, 12$

The results of the analysis are shown in Table 3. With a Confidence Level of 90%, it is found that the factor significantly related with type S2 defect is X_8 , while the factors related with type S1 defect are X_3 , X_4 and X_{10} . The prediction models are created as follows:

Prediction model of the amount of type S2 defects

$$\hat{y}_{S2} = 7.85 - 0.137X_8 \quad (1)$$

Prediction model of the amount of type S1 defects

$$\hat{y}_{S1} = 26.3 + 26.3X_3 + 0.98X_4 - 0.231X_{10} \quad (2)$$

When the variables or factors of machine settings which correlate to the occurrence of defects are known, this can be used to find ways to reduce defects in the future.

Table 3: Analysis of relationship between variables and defects by simple regression analysis

Factors	Type S2 Defect			Type S1 Defect		
	Correlation coefficient	t-test	P-Value	Correlation coefficient	t-test	P-Value
X_1	0.188	1.35	0.183	-0.181	-1.30	0.199
X_2	0.178	1.28	0.206	-0.215	-1.56	0.126
X_3	-0.095	-0.67	0.503	0.266	1.95	0.057*
X_4	-0.092	-0.65	0.516	0.236	1.72	0.092*
X_5	0.093	0.66	0.514	0.089	0.64	0.528
X_6	0.159	1.14	0.260	0.175	1.26	0.215
X_7	0.166	1.19	0.241	0.152	1.08	0.283
X_8	-0.258	-1.89	0.065*	-0.084	-0.60	0.552
X_9	0.192	1.39	0.170	-0.071	-0.52	0.603
X_{10}	-0.218	-1.58	0.121	-0.288	-2.13	0.038*
X_{11}	0.022	0.02	0.987	-0.108	-1.37	0.177
X_{12}	0.188	0.28	0.778	-0.181	-1.51	0.138

3. Conclusion and Discussion

This research had the principal objective of finding variables in production which are related to the number of defects occurring in the wire mesh workpiece. The results of this research found that the important defects to be addressed first were type S2 defect and type S1 defect. The type S2 defect was related to the factor of only the X_8 , while the type S1 defect was related to the factors X_3 , X_4 and X_{10} .

In future research, the researchers will bring the factors obtained from this research to study the level of each factor to cause the minimum amount of defects, to specify working standards for the staff. The technique of Design of Experiments will be used, in the format of 2^{k-p} Fractional Factorial Design. The four design factors mentioned above (X_8 , X_3 , X_4 and X_{10}) will be studied simultaneously at 2 levels per factor. Specifying the level of the factor in the experiment will refer to the original manual and the capability of the machine, alongside directly consulting the company's relevant staff in production. The response variables of the experiment will be the number of type S2 and type S1 defects. The experiment will be replicated twice, with three center point experiments.

As the time and expense required for each experiment is rather high, this research has attempted to prioritize defects to reduce the amounts of defects in the experiment. However, in the future, there may be studies of factors in setting the machinery which are directly related to the amounts of other defects.

4. Acknowledgments

The authors gratefully acknowledge the financial support provided for the research by the Research and Researcher for Industry (RRI) fund 2017, a part of the Thailand Research Fund (TRF). The researchers thank the Graduate School, Chiang Mai University for providing a fund for teaching and research assistants

(TA/RA) for 2016 - 2017 and the Graduate Research Scholarships 2017. The researchers would also like to thank the case study company and its staff for facilitating the research and giving their generous cooperation.

5. References

- [1] V. Istrat, and N. Lalic. Association Rules as a Decision Making Model in the Textile Industry. *Fibres & Textiles in Eastern Europe*. 2017, 25 (4): 8-14.
- [2] P. Yildirim, D. Birant, and T. Alpyildiz. Discovering the relationships between yarn and fabric properties using association rule mining. *Turkish Journal of Electrical Engineering & Computer Sciences*. 2017, 25: 4788-4804.
- [3] S. Sinsomboonthong. *Elementary statistics*. Jamjuree Product Press, 2009.
- [4] P. Thaithae. *A Measurement of Labour Productivity of Electronic Component Industry in Thailand*: Master's Thesis Industrial Engineering and Management. Silpakorn University. Thailand, 2011.
- [5] A. Kunlerd, and P. Seresangtakul. Rice Yield Prediction Model based on Artificial Neural Networks. *Proc. of The Tenth National Conference on Computing and Information Technology*. Phuket. 2014, pp. 7-12.
- [6] K. Sangthong. *Cost Reduction of Spring Wheels (Front Fork) of Motorcycle by Reduction Cost Activity in Samutprakarn Province*: Master's Thesis Business Administration. Stamford International University. Thailand, 2014.
- [7] K. Phiphatpanyanukun. *Statistical Quality Control*. Top Press, 2014.
- [8] W.S. Lasiritaworn, N. Singtorn, and P. Tapeng. Spare Parts Storage Improvement with Association Rules. *Proc. of the World Congress on Engineering and Computer Science*. San Francisco. 2016, pp. 1023-1026.
- [9] J. Han, and M. Kamber. *Data Mining Concepts and Techniques*. Morgan Kaufmann Press, 2006.
- [10] N. Chairporn. *Data mining techniques*. 2015, [online] <http://www.erp.mju.ac.th/articleDetail.aspx?qid=551>. Online Search dated: 30 May 2017.