COMPUTER
METHODS
IN MATERIALS
SCIENCE

2023, vol. 23, no. 2, 29–38
https://doi.org/10.7494/cmms.2023.2.0812

Comparing deterministic and statistical approaches for predicting "short can" defects in aluminium beverage can production

Wojciech Baran[1], Krzysztof Regulski[2], Sławomir Kąc[2], Andrij Milenin[2]

[1] CanPack S.A., Business Support Service, Starowiejska 28 Street, 32-800 Brzesko, Poland.
[2] AGH University of Krakow, Faculty of Metals Engineering and Industrial Computer Science, al. A. Mickiewicza 30, 30-059 Krakow, Poland.

Abstract

In the production of beverage cans, "short can" defects in the form of material discontinuities can occur during the deep drawing of cylindrical thin-walled aluminium products. These defects have a significant impact on production efficiency and scrap generation, and their occurrence is influenced by material and process properties. To determine the main influence of material on defect occurrence, two approaches were used: deterministic analysis of mechanical properties and microstructure, as well as statistical processing of production data using decision tree models. The latter approach was found to be more efficient, and a numerical tool was developed based on this approach to predict and reduce defect occurrence in the production process.

Keywords: short can, deterministic, analysis, statistical methods, predict, defect, reduce, decision trees

1. Introduction

Aluminium cans have been a cornerstone of the beverage industry for over half a century. Since the introduction of the first aluminium can by the Coors Brewing Company in 1959, major beverage companies like PepsiCo and Coca-Cola have also adopted aluminium cans. Over time, advancements in aluminium beverage can manufacturing technology have resulted in an efficient and cost-effective process. The lightweight nature and recyclability of aluminium cans make them an environmentally friendly choice (Stewart et al., 2018). Additionally, aluminium cans help preserve the freshness of beverages and prevent spoilage, making them particularly suitable for carbonated drinks.

Manufacturing companies in the can production industry strive to optimize their processes, including minimizing the production of defective products. Even a one percent decrease in spoilage can save millions of products annually for a typical beverage can line. As a result, identifying the significant factors or combinations of factors that lead to the highest number of defective products is crucial.

The production process of aluminium cans involves several stages, including rolling aluminium sheets, deep drawing to form can bodies, separately forming can tops and bottoms, and attaching them to the can body through seaming (Folle et al., 2008). Despite technological advancements, minimizing defects remains a challenge in the manufacturing of aluminium beverage cans.

Previous studies have used deterministic approaches to investigate the impact of various process parameters on workability, force, and defects in aluminium

* Corresponding author: wojciech.baran@canpack.com
ORCID ID's: 0000-0001-7368-1255 (W. Baran), 0000-0001-8080-2254 (K. Regulski), 0000-0001-9556-2345 (S. Kąc), 0000-0002-6266-2881 (A. Milenin)

beverage can production. For example, Chang & Wang (1997) examined the influence of tool angles, thickness reduction, and friction on deep drawing, while Folle et al. (2008) investigated the impact of friction coefficient, ironing angles, material hardening, and punch-ironing clearance on pressing force. Other studies have analysed the effect of temperature, lubrication, and material parameters on the production process (Gao et al., 2009; Rękas et al., 2014a, 2014b; Schünemann et al., 1996; Simões et al., 2013; Venkateswarlu et al., 2010; Wędrychowicz et al., 2021).

However, practical analyses of the effect of workpiece parameters on formation defects in the real technological process are currently lacking in the literature. This study aims to address this gap by examining two different approaches, deterministic and statistical, for minimizing defects in aluminium beverage can production.

Beverage can production lines have the capacity to produce a significant number of cans per minute, resulting in millions of cans per day and billions per year. Even a small reduction in spoilage can yield substantial savings when considering the global scale of production lines. However, identifying the significant factors or combinations of factors that contribute to defect formation is a challenging task due to the numerous variables involved.

Defects can occur at various stages of the production process, including cup operation, redrawing and drawing operations, and during the formation of the bottom of the can. One of the most common defects during the horizontal press operation is the loss of material continuity known as "ironings." This defect can lead to the rupture of the can wall, resulting in a "short can" with a reduced height compared to a full-value product (see Fig. 1).

Fig. 1. Normal can (left) and the analysed defect – a "short can" (right)

This study takes a different approach by comparing deterministic and statistical methods for minimizing defects in aluminium beverage can production. The deterministic approach involves analysing the plasticity, microstructure, and material properties of aluminium sheets through tensile tests. The statistical approach involves processing data from the production process using decision tree models. By comparing these two approaches, the authors aim to determine which method is more effective in predicting and reducing defects.

The use of statistical methods in industrial processes has gained popularity in recent years. Rodríguez et al. (2016) employed decision tree models to predict roughness in face milling and found them to be more accurate than traditional regression analysis methods. Similarly, Baran et al. (2022) explored the use of a decision tree model to predict defects in aluminium can production. However, their study lacked a comparison between the decision tree model and deterministic methods, making it difficult to determine the optimal approach.

In recent years, research has focused on both deterministic and statistical approaches to minimize defects in aluminium can production.

Deterministic methods involve analysing material properties, process parameters, and microstructure of aluminium sheets, while statistical approaches involve processing production process data using decision tree models.

The goal of this study is to compare the effectiveness of deterministic and statistical approaches (decision tree model) in minimizing the "short can" defect in aluminium beverage can production. By understanding the complex interplay between material and process properties, the study aims to develop new techniques for defect reduction and improve the overall quality of aluminium beverage cans.

In summary, aluminium beverage cans play a significant role in the beverage industry, and their manufacturing process has undergone continuous advancements. Minimizing defects in production is crucial for cost savings and maintaining product quality. While previous studies have explored various factors and methods for defect reduction, practical analyses of workpiece parameters in the real technological process are lacking. This study bridges this gap by comparing deterministic and statistical approaches, aiming to determine the most effective method for minimizing defects in aluminium beverage can production. By achieving this objective, the study aims to contribute to developing techniques that enhance aluminium beverage cans' quality and reliability.

2. Material and methods

2.1. Material

To analyse the relationship between material parameters and jam occurrence in beverage can production, it is crucial to consider the type of material used. In this industry, the predominant material is aluminium alloy from the 3xxx series, specifically the 3104 alloy in the H19 temper. This temper is defined as strain-hardened extra hard material according to the PN-EN 515:1996 standard. Aluminium coils typically have a thickness range of 0.260–0.235 mm and are supplied as rolled coils. All coils must meet the specifications outlined in the Draw and Wall Ironing (DWI) can technology norms. Table 1 shows the nominal chemical composition of the 3104 aluminium alloy in the H19 temper.

Table 1. The nominal chemical composition of aluminium alloy 3104 in the H19 temper, according to EN 573-3:2009 [%]

Si	Fe	Cu	Mn	Mg	Zn	Ti
0.6	0.8	0.05–0.25	0.8–1.4	0.8–1.3	0.25	0.1

2.2. Aluminium beverage cans production technology

Several authors, including Wędrychowicz et al. (2021) and Baran et al. (2022), have described the production stages of aluminium beverage cans. The process begins by pulling a coil from an uncoiler, lubricating it, and feeding it to a cupper press. Next, a disc is cut from the coil, drawn into a cup using draw pad tools, and then further drawn into a can on the bodymaker machine. The can is then washed, coated with various chemicals and dried. The decorator machine is used to cover the can with lithography, lacquer, and dry it in a pin oven. The internal surface of the can is then sprayed and dried. The can is necked, flanged, and checked for quality, and then packed and prepared for transportation to the brewery. Meanwhile, on a separate production line, the lid is prepared, which is later closed with the can after cleaning and filling at the brewery plant.

This article specifically focuses on the early stages of the process, specifically the bodymaker machine, where the tool pack is responsible for reducing the diameter and side wall thickness of the cup while increasing the height of the can.

2.3. Tensile test for deterministic analysis of ductility

To perform a deterministic analysis of the possibility of "short can" defects, tensile studies were conducted on samples while observing the samples fracture surfaces through a scanning microscope (SEM). The methodology for conducting such experiments is elaborated in the work of Milenin et al. (2011). The samples used for this study are depicted in Figure 2.

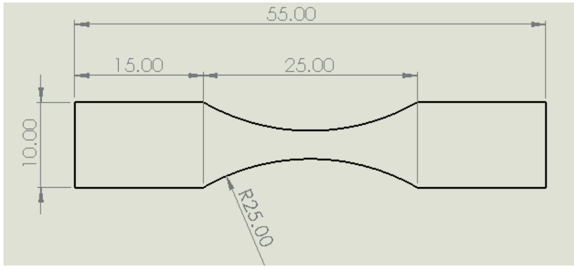

Fig. 2. Specimen for fracture analyses

Details of the experimental technique was described by Milenin et al. (2011).

2.4. Material for deterministic studies

For deterministic studies of mechanical properties, samples taken from three different coils were used. The difference between these coils was the number of "short can" defects per million produced cans when using the material of these coils as a blank. In the first case, this number was 357.9 pcs/mln; this variant of the material will be denoted by the toponym "Bad." The use of second and third coil material resulted in 127.3 pcs/mln and 79.1 pcs/mln defects. These materials in the article are designated as "Medium" and "Good," respectively.

2.5. Tensile tests

For testing the material in tension (before failure), samples were used, made in accordance with the PN-EN ISO 6892-1 standard. The MTS Exceed E43.504E machine was used for this study. 15 samples were tested from each of the studied materials. Sampling was carried out from the left, centre and right side of the coil (5 samples each).

2.6. Statistical-base approach (decision tree model)

Over a three-month production period, data were collected on material parameters and the number of "short cans" for each batch of cans produced. The material parameters were obtained from certificates that contained information about the chemical composition and mechanical properties of the coils used for production. The production monitoring system was used to collect information on the supplier name, coil number, date and time of production, the number of cans produced, the number of "short cans," and an indicator for "short cans" per million produced cans. Additionally, the results of measurements of chemical composition and strength parameters, such as yield strength (YS), ultimate tensile strength (UTS), ears and elongation, were entered for each coil separately. All the collected information was merged into a database to analyse the correlation between them using the STATISTICA program.

Initially, the correlation between the mechanical properties and chemical composition of the material and the number of jams was analysed to determine if there was any connection between them. The statistical significance of these connections was then determined. Next, a regression (see Fig. 3) and classification tree (see Figs. 4 and 5) were used to determine the range of values for each parameter that corresponded to a specific number of jams.

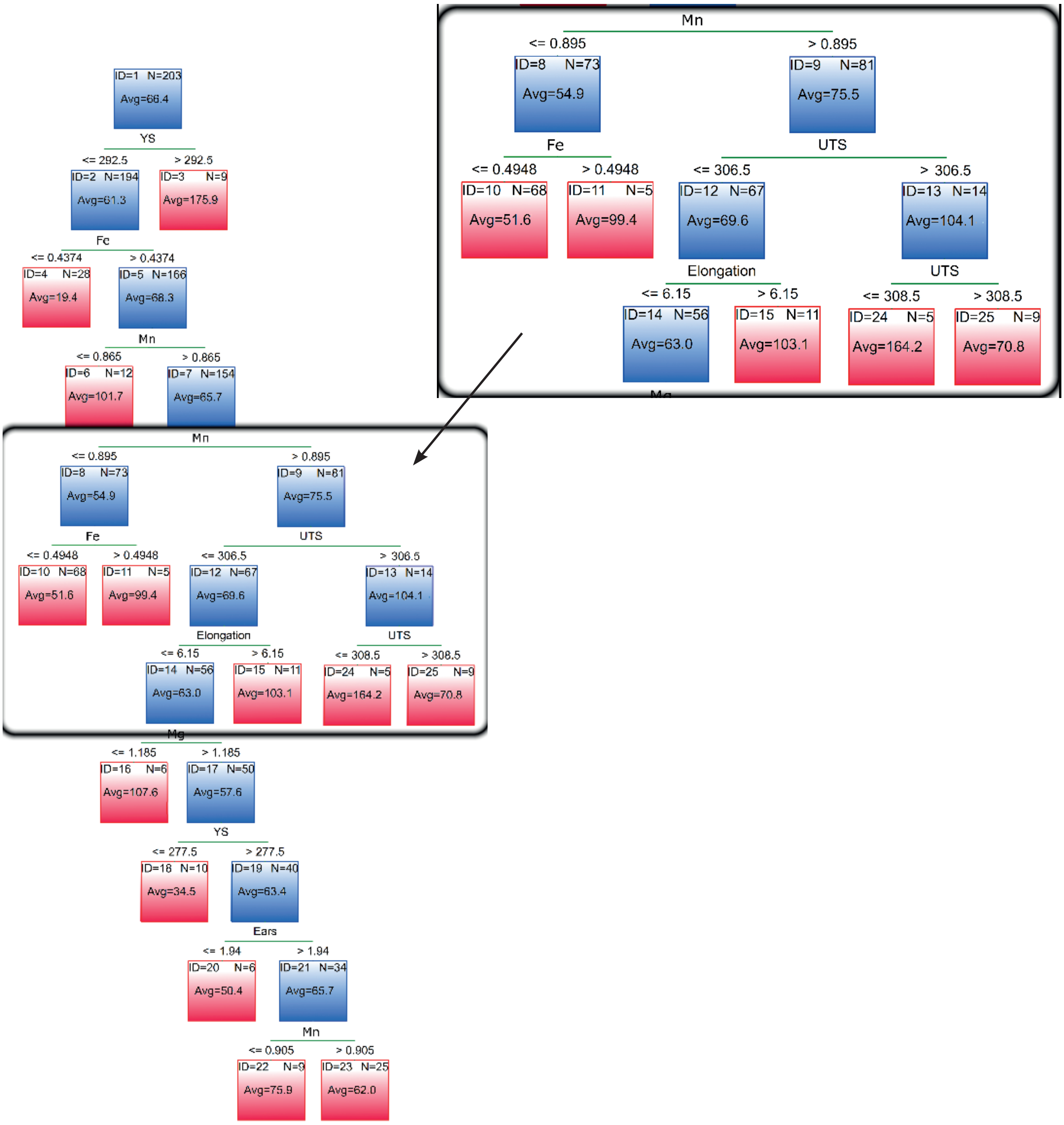

Fig. 3. Interactive regression tree for among of "short cans," Model C&RT

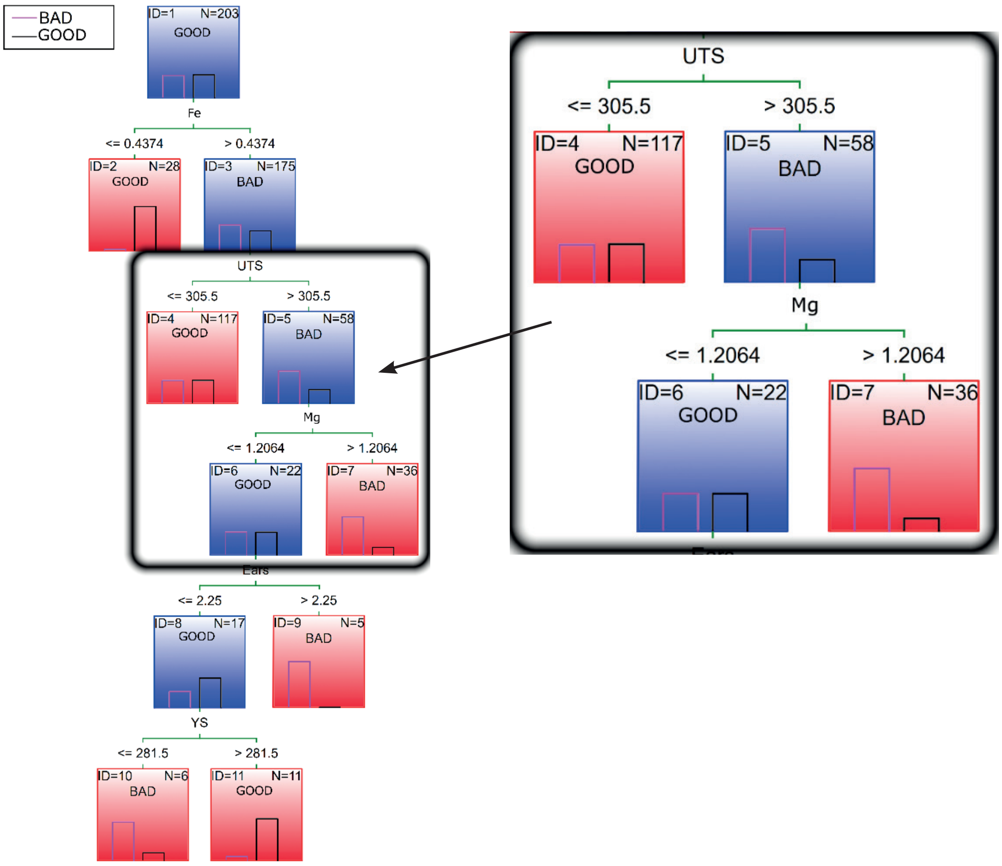

Fig. 4. Interactive classification tree for "short cans" (2Q), Model C&RT

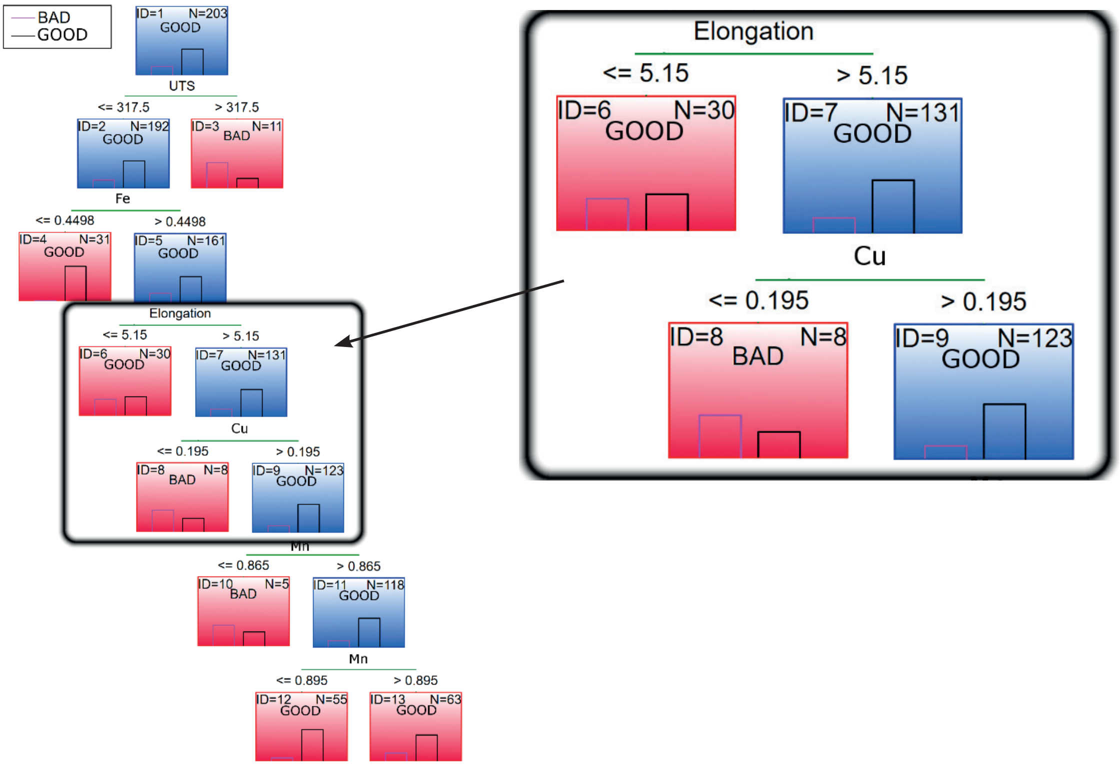

Fig. 5. Interactive classification tree for "short cans" (3Q), Model C&RT

The interactive regression tree (Fig. 3) was used to predict the number of damaged cans based on the content of chemical elements and mechanical properties. Mutual correlations of all parameters were taken into account simultaneously. The division was performed automatically by the "Statistica" program, classifying the groups based on the importance of the impact of each parameter on the others. Red blocks represent last leaf of the tree and base on which rules IF-THEN was created. The most important variables that had the greatest impact on the result were selected for the second model, which was developed with an interactive classification tree. This model classified the sheets into "Good" and "Bad" based on the number of damaged cans, using the labels added earlier.

The generated trees (Figs. 4 and 5) were used to create a set of rules that could assign a coil with entered parameters to a specific range of expected damaged products. The tree in (Fig. 4) represents the division of aluminium coils into two equal groups, where two quartiles (2Q) of sheets with the lowest "short can" index were classified as good. The division (3Q) (Fig. 5) means that three quartiles of coils with the lowest "short can" factor have been classified as good. This division was intended to isolate the group of sheets that pose the greatest risk during production. Software was developed using these rules to help operators make decisions when accepting an aluminium coil for production. This software in the form of subroutines in an Excel environment, would assist operators in determining the quality of the coils based on their parameters and expected damage levels.

3. Results and discussion

3.1. Deterministic analysis

The results of standard tensile tests are shown in Figure 6.

The standard elongation of samples from the three studied materials did not allow us to find any relationship between the elongation value and the number of defects "short cans" (Fig. 6a). Moreover, the standard deviation calculated for the obtained experimental data (shown by error bars in Figure 3a) suggests that the difference in elongation between the samples of the studied materials is not statistically significant. Analysis of strength characteristics (UTS, YS) suggests that fewer defects occur in less strong material (lower values of UTC and YS for "Good" coils). However, these changes are on the border of statistical significance.

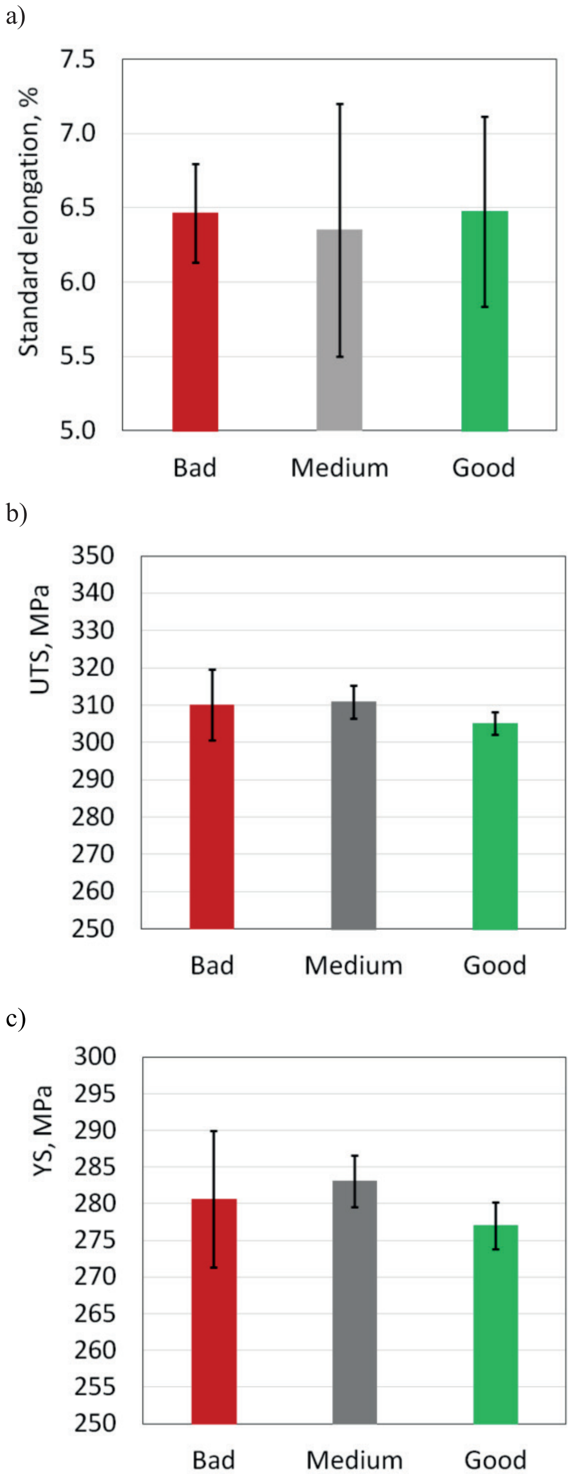

Fig. 6. Mechanical parameters, measured for "Bad", "Medium" and "Good" coils: a) standard elongation; b) UTS; c) YS

Since the standard tensile tests failed to find a statistically significant relationship between mechanical properties and the number of defective cans, we consider the results of tests performed for material from "Good" and "Bad" coils. The results of measuring the tensile force are shown in Figure 7.

a) b)

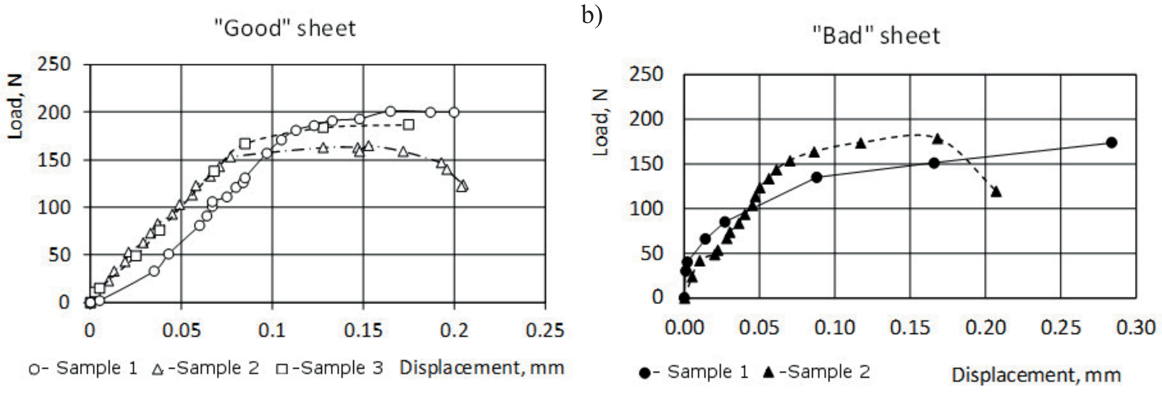

Fig. 7. Load – displacement curves during tests for "Good" (a) and "Bad" (b) coils

Three samples of "Good" sheet coils and two samples of "Bad" sheet coils were taken for testing. These curves, however, also do not correlate statistically significantly with the number of defective cans.

On the SEM images of the fracture surfaces of both samples, a characteristic plastic deformation of the materials is observed. The formation of a cup and cone fracture is visible. The differences are not significant, but in the case of the "Good" sample (Fig. 8a, b) the plastic deformation of the material is greater and relatively evenly distributed over the fracture surface. This causes the material to deform evenly throughout its volume.

In the case of the "Bad" sample (Fig. 8c, d), a smaller plastic deformation of the material around the cups is observed on the surface of the fractures, which indicates less plastic deformation of the material. In addition, the fracture surface is more heterogeneous. There are visible areas of plastic deformation (cups and cone fracture), but there are also numerous areas of large volume in the form of faults and fractures. The presence of such objects causes uneven deformation of the material during plastic processing and faster tearing of the material.

Analysis of the fault image of the samples (Fig. 8) obtained using SEM shows that there is some difference in the nature of the faults, but it is very difficult to interpret.

a) b) c) d)

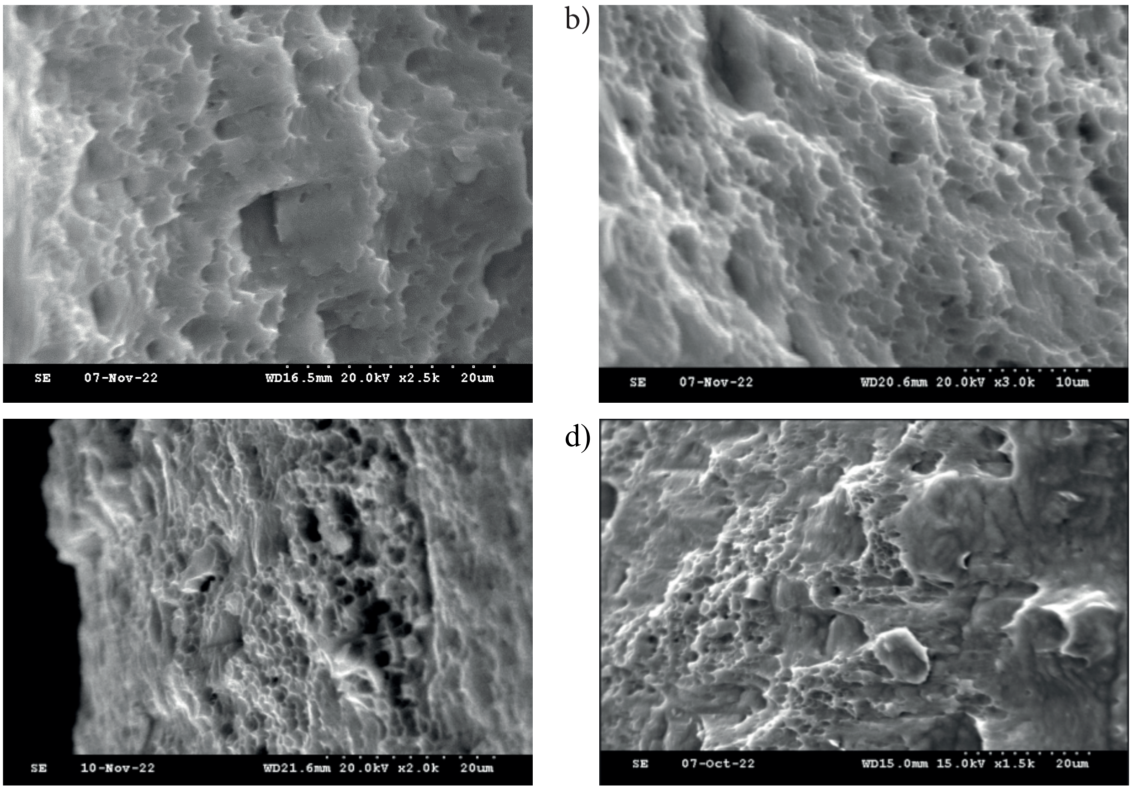

Fig. 8. Examples of microstructures of fracture surfaces for samples from the "Good" (a, b) and "Bad" (c, d) coils

Thus, all deterministic methods used to detect the correlation between mechanical properties and the number of defects were generally ineffective. The only conclusion that can be drawn is that a proportional relationship has been found between the strength of the initial material and the number of defects "short can" (Fig. 6).

3.2. Statistical analysis

The product defect prediction software developed uses statistical analysis through decision tree models. In the current case of the "short can" defect, where the material continuity breaks during the can's passage through smaller dies to thin the side wall, both classification and regression trees were employed. These models are intuitive and transparent, making them an interesting tool for a thorough analysis of dependencies, even in cases of weak correlations (where explanatory variables have little or insignificant effects on the dependent variable's variability, known as weak learners). Trees do not require additional assumptions regarding variable distribution or dependency type. The output of the trees consists of IF-THEN rules that allow for the creation of models using discretization at different levels, leading to a single model output incorporating all the acquired knowledge.

Based on these assumptions, software was created that provides three answers based on the input variables: chemical and strength parameters of the tested material. The response with priority 1 utilizes a regression tree model and returns the predicted number of defects along with the method error. The responses with priority 2 and 3 return the probability of defect occurrence within a given range, taking into account the rule's support, i.e., how strong the rule is.

To test the effectiveness of the program, 20 coils available for production were randomly selected and put on the production line for standard production. Information about the number of actual defects was collected using sensors that counted the number of jams. Subsequently, information about the chemical composition and strength parameters for material from each of the coils was entered into the program (Fig. 9) to predict the number of defects.

Based on the algorithms implemented, the program returned three answers for each sheet. The collected results are summarized in Table 2.

The table shows 20 randomly selected aluminium coils with information on the actual number of can damages that have been recorded on the production line and answers from three decision trees. The answers have been prioritized according to the importance of the answers. Priority 1 is the most important. The regression tree was chosen for this priority because it gives us a specific number of predicted faults. Priority 2 and 3 are answers from classification trees. These trees return an answer as to whether the test sample is above or below a predetermined threshold for the number of short cans. The last column counts the number of correct answers of decision trees consistent with the actual number of can damages.

The program correctly predicted three answers for 9 out of 20 coils, meaning that the actual number of defects fell within limits given by the program for all statistical tools used. Additionally, nine coils were accurately classified by two out of the three methods, while the remaining two coils were correctly identified by the same method. No misclassifications were observed. All trees predict correctly, but not all return answers. By creating a hybrid inference model, we can replace the lack of a response from one tree with a response from the next tree with a priority one lower.

From the presented results, we can conclude that the program is capable of predicting well, although its accuracy is not particularly high. The results show a relatively large standard deviation, which may be due, among other things, to insufficient training data. To improve the accuracy of the results, the statistical analysis should be re-run, and new rules created using a much larger amount of material and production data, which is also planned for future study.

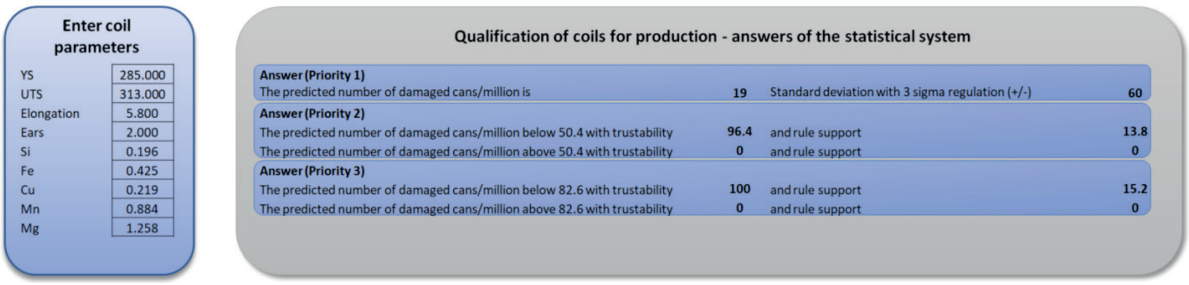

Fig. 9. Program for predicting the number of short cans

Table 2. Comparison of the predicted number of damaged cans with the actual value

Coil no.	Real demaged cans [pcs/mln]	Predicted damaged cans [pcs/mln]										Number of trees compatible with production reports
		priority 1		priority 2				priority 3				
		predicted number of demaged [cans/mln]	standard deviation	below 50.4 with trustability	with rule support	above 50.4 with trustability	with rule support	below 82.6 with trustability	with rule support	above 82.6 with trustability	with rule support	
1	357.9	164	475	0	0	83.3	17.7	0	0	0	0	2
2	305.9	175	579	0	0	83.3	17.7	0	0	72.7	5.4	3
3	217.9	175	579	0	0	83.3	17.7	0	0	72.7	5.4	3
4	176.9	175	579	0	0	100.0	2.4	0	0	72.7	5.4	3
5	165.3	101	402	0	0	0	0	0	0	0	0	1
6	152.0	164	475	0	0	83.3	17.7	0	0	0	0	2
7	123.7	70	192	0	0	83.3	17.7	0	0	0	0	2
8	122.7	0	0	0	0	0	0	0	0	62.5	3.9	1
9	114.5	70	192	0	0	100.0	2.4	0	0	0	0	2
10	111.9	101	402	0	0	83.3	17.7	0	0	0	0	2
11	107.6	70	192	0	0	83.3	17.7	0	0	0	0	2
12	105.5	175	579	0	0	83.3	17.7	0	0	72.7	5.4	3
13	105.1	175	579	0	0	100.0	2.4	0	0	72.7	5.4	3
14	103.5	0	0	0	0	100.0	2.4	0	0	72.7	5.4	2
15	97.9	175	579	0	0	100.0	2.4	0	0	72.7	5.4	3
16	56.5	0	0	0	0	83.3	3.0	90.9	27.0	0	0	2
17	42.6	0	0	90.9	5.4	0	0	90.9	27.0	0	0	2
18	18.1	19	60	96.4	13.8	0	0	100.0	15.2	0	0	3
19	13.6	19	60	96.4	13.8	0	0	100.0	15.2	0	0	3
20	13.3	19	60	96.4	13.8	0	0	100.0	15.2	0	0	3

4. Conclusion

Based on the results of the study, several conclusions can be drawn. The deterministic methods used to detect the correlation between mechanical properties and the number of defects were generally ineffective, as no statistically significant relationship was found. When using deterministic methods to measure strength parameters, we cannot be sure that the measurement point is the place where the continuity of the material was broken during the drawing of the can. The material is not uniform over the entire surface, and by examining its properties in a random place, we can obtain results that differ from the actual results in the place of the defect. Based on a large number of damages, using statistical methods, it is possible to find relationships that are not visible in individual measurements. However, a proportional relationship was found between the strength of the initial material and the number of defects "short can." The statistical analysis, using decision tree models, was successful in predicting the number of defects. The program developed correctly predicted the number of defects for 9 out of 20 coils, with no misclassifications observed. However, the accuracy of the program is not particularly high, and further studies with a larger amount of material and production data are planned to improve its accuracy. Overall, the results suggest that the use of statistical analysis and decision tree models can be an effective tool for defect prediction in industrial production processes.

References

Baran, W., Regulski, K., & Milenin, A. (2022). Influence of materials parameters of the coil sheet on the formation of defects during the manufacture of deep-drawn cups. *Processes*, *10*(3), 578. https://doi.org/10.3390/pr10030578.

Chang, D.-F., & Wang, J.E. (1997). Influence of Process Parameters on the Ironing of Deep-Drawn Cups. *Journal of Manufacturing Science and Engineering*, *119*(4B), 699–705. https://doi.org/10.1115/1.2836812.

Folle, L.F., Silveira Netto, S.E., & Schaeffer, L. (2008). Analysis of the manufacturing process of beverage cans using aluminum alloy. *Journal of Materials Processing Technology*, *205*(1–3), 347–352. https://doi.org/10.1016/j.jmatprotec.2007.11.249.

Gao, E.-z., Li, H.-w., Kou, H.-ch., Chang, H., Li, J.-s., & Zhou, L. (2009). Influences of material parameters on deep drawing of thin-walled hemispheric surface part. *Transactions of Nonferrous Metals Society of China*, *19*(2), 433–437. https://doi.org/10.1016/S1003-6326(08)60291-5.

Lewicki, P., & Hill, E. (2006). *Statistics. Methods and Applications*. StatSoft.

Milenin, A., Byrska, D.J., & Grydin, O. (2011). The multi-scale physical and numerical modeling of fracture phenomena in the MgCa0.8 alloy. *Computers & Structures*, *89*(11–12), 1038–1049. https://doi.org/10.1016/j.compstruc.2011.01.003.

Rękas, A., Latos, T., Budzyn, R., Fijałkowski, M., & Brodawka, Ł. (2014a). The analysis of influence of sheet properties on the ironing process of thin-walled cylindrical shell products from aluminum alloys. *Key Engineering Materials*, *641*, 232–245. https://doi.org/10.4028/www.scientific.net/KEM.641.232.

Rękas, A., Latos, T., Budzyn, R., Furman, A., & Siedlik, M. (2014b). The analysis of mechanical properties of 3XXX series aluminum sheets used for beverage can production. *Key Engineering Materials*, *641*, 246–256. https://doi.org/10.4028/www.scientific.net/KEM.641.246.

Rodríguez, J.J., Quintana, G., Bustillo, A., & Ciurana, J. (2016). A decision-making tool based on decision trees for roughness prediction in face milling. *International Journal of Computer Integrated Manufacturing*, *30*(9), 943–957. https://doi.org/10.1080/0951192X.2016.1247991.

Schünemann, M., Ahmetoglu, M.A., & Altan, T. (1996). Prediction of process conditions in drawing and ironing of cans. *Journal of Materials Processing Technology*, 59(1–2), 1–9. https://doi.org/10.1016/0924-0136(96)02280-7.

Simões, V.M., Coër, J., Laurent, H., Oliveira, M.C., Alves, J.L., Manach, P.Y., & Menezes, L.F. (2013). Sensitivity analysis of process parameters in the drawing and ironing processes. *Key Engineering Materials*, *554–557*, 2256–2265. https://doi.org/10.4028/www.scientific.net/KEM.554-557.2256.

Stewart, R., Niero, M., Murdock, K., & Olsen, S.I. (2018). Exploring the implementation of a circular economy strategy: the case of a closed-loop supply of aluminum beverage cans. *Procedia CIRP*, *69*, 810–815. https://doi.org/10.1016/j.procir.2017.11.006.

Venkateswarlu, G., Davidson, M.J., & Tagore, G.R.N. (2010). Influence of process parameters on the cup drawing of aluminium 7075 sheet. *International Journal of Engineering, Science and Technology*, *2*(11), 41–49. https://doi.org/10.4314/IJEST.V2I11.64553.

Wędrychowicz, P., Kustra, P., Paćko, M., & Milenin, A. (2021). A flow stress model of the AA3104-H19 alloy for the FEM simulation of the beverage can manufacturing process under large plastic deformations. *Materials*, *14*(21), 6408. https://doi.org/10.3390/MA14216408.