Anna Szewczyk, Klaudia Górska, Adrian Wolski, Dariusz Fydrych

The Application of Cluster Analysis in the Assessment of the Weldability of Unalloyed Steels

Abstract: Non-alloy steels constitute a large group of steels characterised by diversified chemical composition, structural morphology and a wide range of mechanical properties (determining weldability). The paper presents results of multidimensional analyses (based on cluster analysis) of 110 selected unalloyed steel grades. Properties adopted as diagnostic features included the chemical composition, mechanical properties (yield point) and values of selected indicators concerning susceptibility to technological crack formation. The analyses (performed using Ward's and k-means methods) resulted in a division of the 110 steels into five steel groups (clusters). The comparison of results obtained using two clustering methods and involving various classification criteria revealed that multidimensional analyses constituted a prospective method making it possible to assess the weldability of steels. However, results of such multidimensional analyses should be subjected to thorough and substantive analyses.

Keywords: non-alloy steels, cluster analysis, weldability

Introduction

Because of the vast range of available steel grades varying not only in terms of their properties but also weldability, the selection of appropriate steel for a specific welded structure is a responsible and, often, laborious task. The notion of weldability is defined in many ways as there are numerous factors affecting the course of the joining process, technological and operational properties of joints as well as the complexity of joining processes. Factors defining weldability and contained in the PN-EN 1011-2 standard include the joint structure, hydrogen-induced cracking, the hardness and toughness of the heat affected zone (HAZ), solidification cracking, lamellar cracking and corrosion.

Presently, there is no uniform system of weldability assessment not only as regards groups of engineering materials but even in relation to significantly smaller groups including, e.g. iron alloys, aluminium alloys or copper alloys. When selecting materials for welded structures, engineers must rely on information concerning properties of base materials and their behaviour during welding and operation [1].

Most factors affecting the weldability of steel are taken into consideration during experimental weldability tests (e.g. Tekken, CTS, T-test, implant test, Varestraint test, etc.) [1–3]. However, the performance of such tests is often excessively time-consuming in terms of industrial applications and, as a result, is usually limited to special cases, i.e. concerning new materials and research. The simplest and fastest methods enabling the assessment of steel weldability were developed on the basis of experimental tests. Such methods are based on the chemical composition analysis-based forecasting of joint susceptibility to cracking. Various reference publications and related standards contain tens of formulas enabling the calculation of values indicating susceptibility to cold, hot, lamellar and annealing crack formation. Some of the

inż. Anna Szewczyk, inż. Klaudia Górska – Politechnika Gdańska, Wydział Inżynierii Mechanicznej i Okrętownictwa, Koło Naukowe Spawalników MMA (Gdańsk University of Technology; Faculty of Mechanical Engineering and Ship Technology, Research Society of MMA Welding Engineers)

mgr inż. Adrian Wolski, dr hab. inż. Dariusz Fydrych, prof. PG – Politechnika Gdańska, Wydział Inżynierii Mechanicznej i Okrętownictwa, Instytut Technologii Maszyn i Materiałów (Gdańsk University of Technology; Faculty of Mechanical Engineering and Ship Technology, Institute of Manufacturing and Materials Technology)

above-named correlations take into consideration the impact of additional factors such as thicknesses of elements subjected to welding, thermal conditions (cooling time t8/5) or the content of diffusive hydrogen in weld deposit. Presently, the aforesaid factors are treated as additional tools used in the initial assessment of the weldability of metals and their alloys (particularly various groups of steels) [2, 3].

The analysis of steel weldability is immensely important, particularly as regards high-strength steels, used in the fabrication of elements of crucial structures, e.g. cranes, bridges, wind turbines as well as hydro and ocean engineering structures [4–6]. The use of the above-named steels makes it possible to reduce the thickness of elements, which translates into the obtainment of lighter structures, the possibility of imposing greater loads on such structures as well as the reduction of welding time and that of welding consumables (resulting from smaller volumes of welds) [4].

Problems concerning many objects described simultaneously by many factors can be solved using data analysis methods, creating the so-called multidimensional analyses [7] and including traditional methods and regression analysis (multiple and logistic regression), data mining techniques (i.e. cluster analysis) as well as dimensional reduction methods (factor analysis and correspondence analysis). The area of cluster analysis applications include numerous areas of humanities and economics as well as social, natural, medical and technical sciences [8, 9]. The primary purpose of cluster analysis consists in the division of elements into homogenous clusters. The above-named methods are divided into hierarchical and non-hierarchical ones and can be treated as complementary. Particularly good results are obtained using Ward's method (hierarchical) and, next, the k-means method (non-hierarchical). Elements from various clusters should be maximally different from one another, whereas elements belonging to the same cluster should be selected according to maximum similarity [10, 11].

The issue of weldability has been thoroughly addressed in certain aspects, yet relatively rarely analysed in relation to the possible solution of multidimensional problems using statistical methods. The study discussed in the article aimed to verify the possibility of the objective assessment the weldability of unalloyed steels using multidimensional methods (cluster analysis).

Individual study

The scope of research work included the gathering of data subjected to analyses, the selection of diagnostic features, the preparation of data used in tests (verification of variability, analysis of correlations and normalisation), the performance of analyses and the substantive analyses of results.

Diagnostic features adopted in analyses included chemical compositions of unalloyed steels (alloying elements) and values of weldability indicators (hot cracking susceptibility, cold cracking susceptibility (Pcm) and lamellar cracking susceptibility (Pl)) as well as steel yield point (Re). Data of 110 steel grades were obtained from commonly available sources, i.e. standards, scientific articles and suppliers' catalogues [12–18]. In cases of steel grades provided with ranges of chemical element contents it was necessary to adopt their average values (Table 1).

The analyses were performed using the Statistica package. The method used as first was hierarchical (Ward's method with the Euclidean distance), whereas the subsequently applied method was non-hierarchical (the k-means method). The mode applied during the analysis could be summarised as "sort distances and observe at constant interval". The lack of initial information in reference publications was responsible for the fact that the authors planned three analysis, where diagnostic features only included contents of alloying elements (analysis I), contents of alloying elements and indicators of cracking susceptibility (analysis II) as well as contents of alloying elements, values of yield points and indicators of cracking susceptibility (analysis III). The analysis of data variability was performed in accordance with the following dependence:

$$\omega = \frac{s_j}{x_j} \qquad (1)$$

where
ω – coefficient of variation,
s_j – standard deviation,
x_j – arithmetic mean of a feature value.

Table 2 contains diagnostic features subjected to analysis along with coefficients of variation. In all cases the above-named value amounted to a minimum of 0.15, which led to the conclusion that the selected features were characterised by appropriate discriminant ability.

Table 1. Fragment of the set with raw input data used in statistical analyses

	S275NLH	S355J2	S420M	S460NLH	S500ML	S650MLH	S690QH	S700MH	S960MLH	S1100MC	S1300Q
max C %	0.2	0.2	0.16	0.22	0.16	0.16	0.25	0.16	0.2	0.18	0.12
max Si %	0.4	0.55	0.5	0.6	0.6	0.6	0.8	0.6	0.6	0.5	0.21
Mn %	0.95	1.6	1.7	1.4	1.7	2	1.7	2.1	2.2	1.3	0.9
max P %	0.03	0.025	0.03	0.03	0.025	0.02	0.025	0.02	0.02	0	0
max S %	0.025	0.025	0.025	0.025	0.02	0.012	0.015	0.015	0.01	0	0
max Nb %	0.05	0	0.5	0.05	0.5	0.09	0.06	0.09	0.09	0.09	0.02
max V %	0.08	0	0.12	0.2	0.12	0.2	0.16	0.2	0.2	0.2	0.02
min. Al %	0.02	0	0.02	0.02	0.02	0.015	0	0.015	0.015	0.015	0.038
max Ti %	0.03	0	0.05	0.03	0.05	0.22	0.05	0.22	0.25	0.01	0.01
max Cr %	0.3	0	0.03	0.3	0.03	0.3	1.5	0.3	1.6	1.5	0.48
max Ni %	0.3	0	0.8	0.8	0.8	0.8	2	0.8	0.8	2.5	1.23
max Mo %	0.1	0	0.2	0.1	0.2	0.5	0.7	0.5	1	0.8	0.4
max Cu %	0.35	0.55	0.55	0.7	0.55	0.55	0.5	0.55	0.55	0	0.01
max N %	0.015	0	0.025	0.025	0.025	0.025	0.02	0.025	0.025	0	0.06
max B %	0	0	0	0	0	0	0.005	0	0	0	0
max W %	0	0	0	0	0	0	1.5	0	0	0	0
max Zr %	0	0	0	0	0	0	0.15	0	0	0	0
HCS	3.31	2.08	1.62	2.53	1.33	0.74	1.36	0.77	0.64	0.01	0.003
Pcm	0.31	0.33	0.33	0.40	0.33	0.39	0.58	0.39	0.54	0.45	0.25
Pl	0.55	0.56	0.56	0.63	0.54	0.54	0.76	0.57	0.68	0.54	0.33
Re	275	355	420	460	500	650	690	700	960	1100	1300

Table 2. Analysis results concerning the variability of diagnostic features

Diagnostic feature	Coefficient of variation ω
C %	0.2
Si %	0.35
Mn %	0.23
P %	0.41
S %	0.40
Nb %	1.61
V %	0.82
Al %	0.95
Ti %	1.25
Cr %	1.09
Ni %	1.16
Mo %	1.38
Cu %	0.49
N %	1.17
Re	0.44
HCS	0.5
Pcm	0.27
Pl	0.15

The normalisation of features was performed assuming that all the features were stimulants in accordance with the following dependence:

$$x'_{ij} = \frac{x_{ij} - \min \{x_{ij}\}}{\max \{x_{ij}\} - \min \{x_{ij}\}} \qquad (2)$$

The normalised diagnostic features were additionally designated with the letter "n".

Analysis I

The Ward's method-based cluster analysis results are presented in the form of a dendrogram in Figure 1. By using the diagram of the distance of linkage in relation to the stage of linkage (Figure 2) it was possible to obtain five clusters. The clustering results are presented in Table 3. The determination of the number of clusters enabled the performance of analyses using the k-means method (Figure 3).

Analysis II

The Ward's dendrogram in relation to analysis including such diagnostic features as the chemical composition and indicators of cracking susceptibility is presented in Figure 4. Based on the diagram presented in Figure 5, the authors proposed the division of elements into five clusters. The clustering results are presented in Table 4. The graphic interpretation of analysis results obtained using the k-means method is presented in Figure 6.
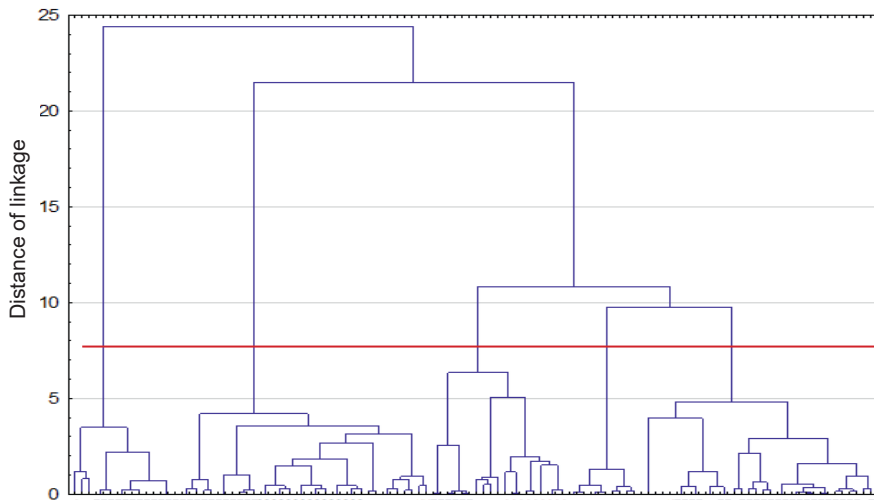
Fig. 1. Dendrogram of the classification of unalloyed steels based on the following features: Cn, Sin, Mnn, Pn, Sn, Nbn, Vn, Aln, Tin, Crn, Nin, Mon, Cun and Nn; Ward's method; Euclidean distance; the red line designates the level of division into clusters

Fig. 2. Diagram of the distance of linkage in relation to the stage of linkage

Fig. 3. Line graph of the k-means in relation to clusters of unalloyed steels on the basis of the following features: Cn, Sin, Mnn, Pn, Sn, Nbn, Vn, Aln, Tin, Crn, Nin, Mon, Cun and Nn

Table 3. K-means method-based results of clustering of unalloyed steels in relation to the following features: Cn, Sin, Mnn, Pn, Sn, Nbn, Vn, Aln, Tin, Crn, Nin, Mon, Cun and Nn

Cluster I	S275NL, S275M, S275ML, S275MH, S275MLH, S355N, S355NL, S355NH, S355NLH, S355M, S355ML, S355MLH, S420M, S420N, S420NL, S420ML, S420MH, S420MLH, S420NH, S420NLH, S460N, S460NL, S460JR, S460J0, S460J2, S460K2, S460M, S460ML, S460MH, S460MLH, S460NH, S460NLH, S500J0, S500M, S500ML, S500MH, S500MLH
Cluster II	S460QH, S460QLH, S460QL1H, S500QH, S500QLH, S500QL1H, S690QH, S690QLH, S960QH, S960QLH, S960QL1H, S1100MC, S1300QL
Cluster III	S650MH, S650MLH, S700MH, S700MLH, S960MC, S960MH, S960MLH
Cluster IV	S275N, S275JR, S275J0, S275J2, S275J0H, S275J2H, S275NH, S275NLH, S355JR, S355J0, S355J2, S355K2, S355J0WP, S355J2WP, S355J0W, S355J2W, S355K2W, S355J4W, S355J5W, S355J0WH, S355J2WH, S355J2H, S355K2WH, S355J0H, S355K2H, S420J0W, S420J2W, S420K2W, S420J4W, S420J5W, S420K2WH, S460J0W, S460J2W, S460K2W, S460J4W, S460J5W, S460K2WH, S500K2WH
Cluster V	S355MC, S355NC, S355J2, S420MC, S420NC, S460MC, S500MC, S650MC, S690Q, S690QL, S700MC, S1300Q

Fig. 4. Dendrogram of the classification of unalloyed steels based on the following features: Cn, Sin, Mnn, Pn, Sn, Nbn, Vn, Aln, Tin, Crn, Nin, Mon, Cun, Nn, HCSn, Pcmn and Pln; Ward's method; Euclidean distance; the red line designates the level of division into clusters
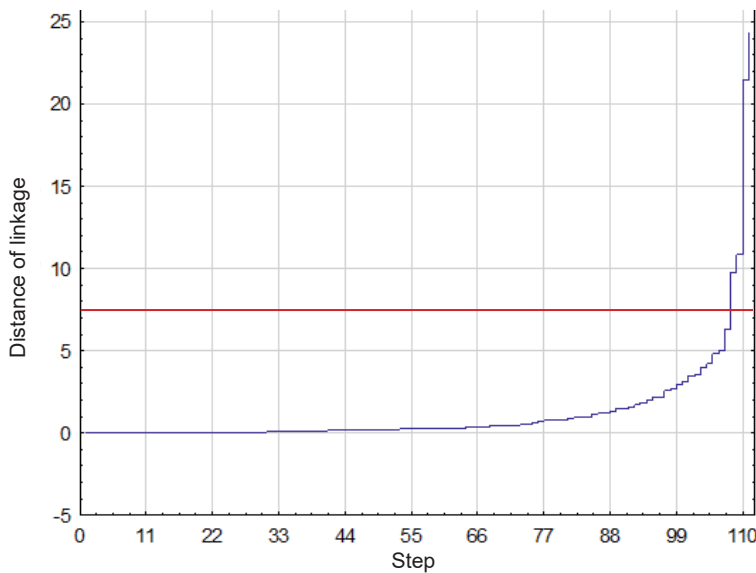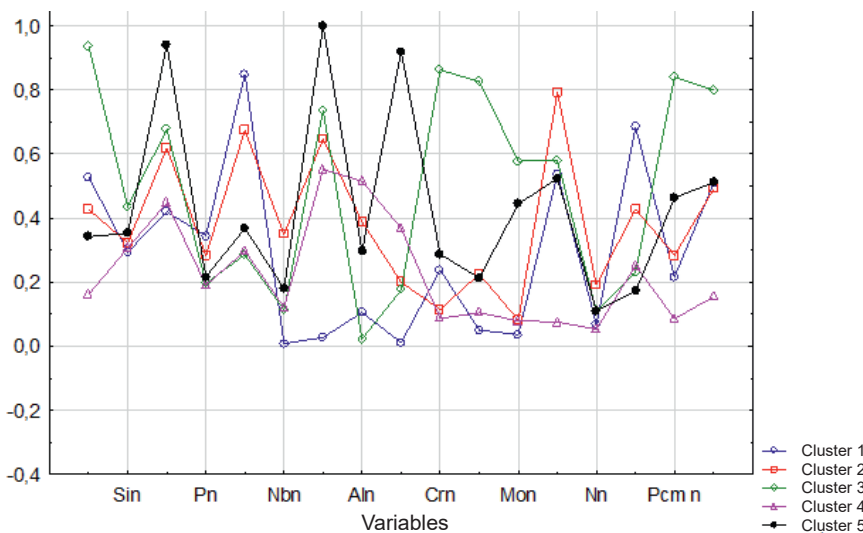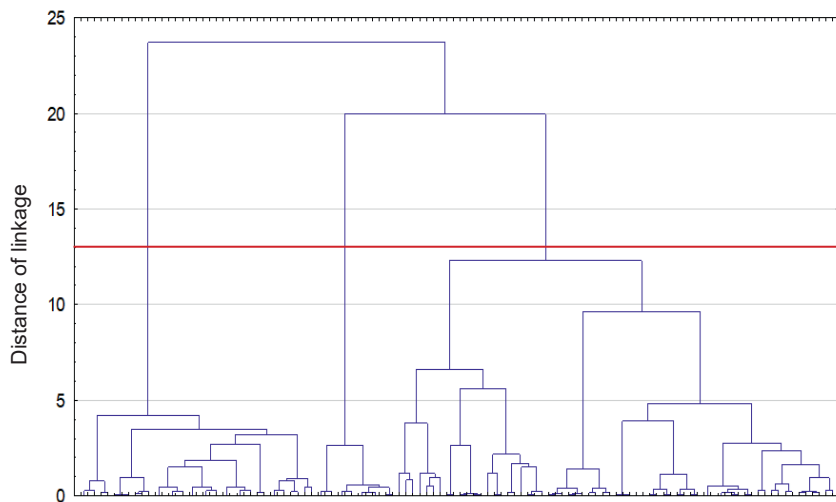

Fig. 5. Diagram of the distance of linkage in relation to the stage of linkage

Fig. 6. Line graph of the k-means in relation to clusters of unalloyed steels on the basis of the following features: Cn, Sin, Mnn, Pn, Sn, Nbn, Vn, Aln, Tin, Crn, Nin, Mon, Cun, Nn, HCSn, Pcmn and Pln

Table 4. K-means method-based results of clustering of unalloyed steels in relation to the following features: Cn, Sin, Mnn, Pn, Sn, Nbn, Vn, Aln, Tin, Crn, Nin, Mon, Cun, Nn, HCSn, Pcmn and Pln

Cluster I	S275N, S275JR, S275JO, S275J2, S275J0H, S275J2H, S275NH, S275NLH, S355JR, S355J0, S355J2, S355K2, S355J0WP, S355J0W, S355J2W, S355K2W, S355J4W, S355J5W, S355J0WH, S355J2WH, S355K2WH, S355J2H, S355K2H, S420J0W, S420J2W, S420K2W, S420J4W, S420J5W, S420K2WH, S460J0W, S460J2W, S460K2W, S460J4W, S460J5W, S460K2WH, S500K2WH
Cluster II	S275NL, S275M, S275ML, S275MH, S275MLH, S355N, S355NL, S355M, S355ML, S355MH, S355MLH, S355NH, S355NLH, S420N, S420NL, S420M, S420ML, S420MH, S420MLH, S420NH, S420NLH, S460N, S460NL, S460J0, S460J2, S460K2, S460M, S460ML, S460MH, S460NH, S460NLH, S500J0, S500M, S500ML, S500MH, S500MLH
Cluster III	S460QH, S460QLH, S460QL1H, S500QH, S500QLH, S690QH, S690QLH, S960QLH, S960QL1H, S1100QL, S1300QL
Cluster IV	S355MC, S355NC, S355J2, S420MC, S420NC, S460MC, S500MC, S690QL, S690Q, S1300Q
Cluster V	S650MC, S650MLH, S700MC, S700MLH, S960MC, S960MH, S960MLH

Fig. 7. Dendrogram of the classification of unalloyed steels based on the following features: Cn, Sin, Mnn, Pn, Sn, Nbn, Vn, Aln, Tin, Crn, Nin, Mon, Cun, Nn, Ren, HCSn, Pcmn and Pln; Ward's method; Euclidean distance; the red line designates the level of division into clusters
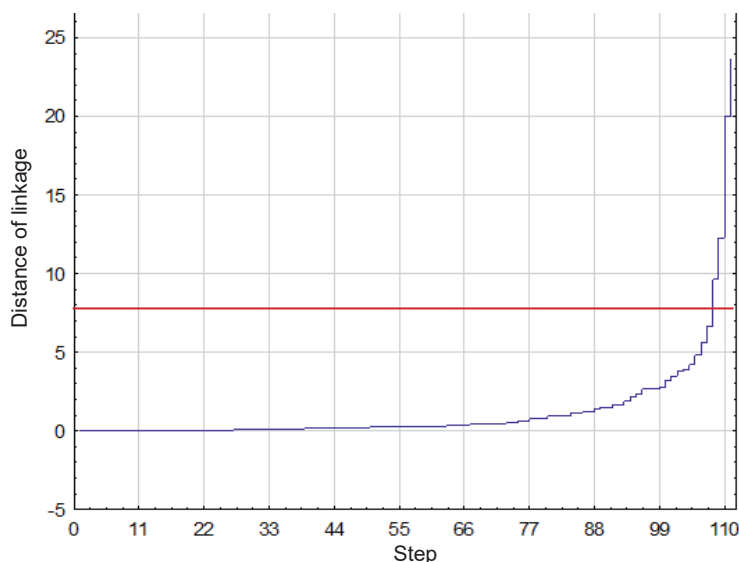

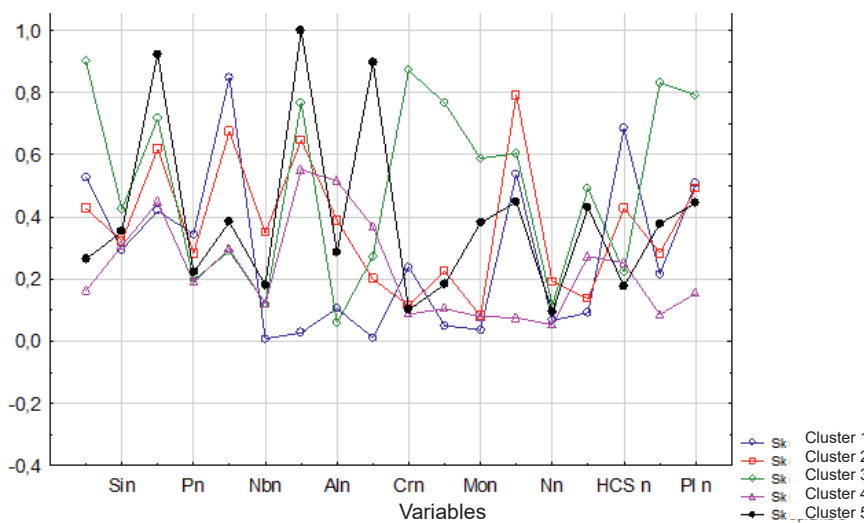Fig. 8. Diagram of the distance of linkage in relation to the stage of linkage

Fig. 9. Line graph of the k-means in relation to clusters of unalloyed steels on the basis of the following features: Cn, Sin, Mnn, Pn, Sn, Nbn, Vn, Aln, Tin, Crn, Nin, Mon, Cun, Nn, Ren, HCSn, Pcmn and Pln

Table 5. K-method-based results of clustering of unalloyed steels in relation to the following features: Cn, Sin, Mnn, Pn, Sn, Nbn, Vn, Aln, Tin, Crn, Nin, Mon, Cun, Nn, Ren, HCSn, Pcmn and Pln

Cluster I	S275N, S275JR, S275J0, S275J2, S275J0H, S275J2H, S275NH, S275NLH, S355JR, S355J0, S355J2, S355K2, S355J0W, S355J2W, S355J0WP, S355J2WP, S355K2W, S355J4W, S355J5W, S355J0WH, S355J2WH, S355K2WH, S355J2H, S355J0H, S355J2H, S355K2H, S420J0W, S420J2W, S420K2W, S420J4W, S420J5W, S420K2WH, S460J0W, S460J2W, S460K2W, S460J4W, S460J5W, S460K2WH, S500K2WH
Cluster II	S275NL, S275M, S275ML, S275MH, S275MLH, S355N, S355NL, S355M, S355ML, S355MLH, S355NH, S355NLH, S420N, S420NL, S420M, S420ML, S420MH, S420MLH, S420NH, S420NLH, S460N, S460NL, S460NH, S460ML, S460MLH, S460NLH, S500J0, S500M, S500ML, S500MH, S500MLH
Cluster III	S460QH, S460QLH, S460QL1H, S500QH, S460K2, S460J2, S460J0, S690QLH, S690QL1H, S960MH, S960MLH, S1100QL, S1100MC, S1300QL
Cluster IV	S355MC, S355NC, S355J2, S420MC, S420NC, S460MC, S500MC, S690Q, S690QL, S1300Q
Cluster V	S650MC, S650MLH, S650MLH, S700MC, S700MLH, S960MC

Analysis III

Results of the Ward's method-based cluster analysis results including such diagnostic features as the chemical composition, indicators of cracking susceptibility and values of yield points are presented in Figure 7. Similar to the previous analyses, the set under investigation was divided into five clusters (Figure 8). The clustering results are presented in Table 5. The results of analysis obtained using the k-means method are presented in Figure 9.

Summary

The article discusses the applicability of cluster analysis methods (Ward's and k-means) in the classification of unalloyed steels with respect to their weldability. To this end, it was necessary to perform the cluster analysis-based grouping of 110 selected steel grades by various criteria including the chemical composition (analysis I), the chemical composition and cracking susceptibility indicators (analysis II) as well as the chemical composition, mechanical properties represented by yield points and cracking susceptibility indicators (analysis III). All the three analyses resulted in the division of the above-named steel grades into 5 clusters. The comparison of results concerning steel classification based on Ward's and k-means methods revealed that each analysis led to the obtainment of similar clustering results. However, in none of the cases it was possible to obtain the state without changes in the assignment of steel grades to clusters. The direct comparison of the results obtained using the three analyses led to the conclusion that changes of diagnostic features (as input parameters) did not result in significant differences as regards clustering results. The most important changes were observed in relation to the different assignment of steel S650MC and steel S700MC after the addition of weldability indicators to diagnostic features as well as the change of the classification of steel S960MH and steel S960MLH in analysis III (including values of yield points). The most surprising results of the analyses was the non-intuitive assignment of two steel grades, i.e. S1300Q and S1300QL, which were not classified in the same cluster by the clustering algorithms applied in the study.

The analyses discussed in the article were performed assuming that weldability was a material-related feature (which was a significant simplification). It should also be emphasized that the applicability of cluster analysis in weldability assessment based on criteria used in analyses II and III was encumbered with an error resulting from the relatively low correlation between values of weldability coefficients and susceptibility to the formation of individual types of cracks. For this reason, the results of analyses and their practical applications should be preceded by a thorough and substantive analysis. The above-presented results led to the conclusion that the cluster analysis could be an effective and helpful tool when assessing the weldability of unalloyed steels, yet it can only constitute the initial stage of the decision-making process.

References

[1] Fydrych D., Rogalski G., Świerczyńska A.: Zastosowanie próby implant do oceny spawalności stali. Welding Technology Review, 2016, no. 88, pp. 28–32.

[2] Kozioł P., Organek P.: Ocena spawalności stali konstrukcyjnych. Przegląd Budowlany, 2015, vol. 86, no. 1, pp. 22–26.

[3] Grigorenko G.M., Kostin W.A.: Spawalność stali i kryteria jej oceny. Welding Technology Review, 2013, vol. 85, no. 7, pp. 11–17.

[4] Tuz L., Sulikowski K.: Assessment of the possibility of the laser welding of high strength steel having a yield point of 1100 MPa. Biuletyn Instytutu Spawalnictwa w Gliwicach, 2019, vol. 63, no. 6, pp. 25–31.

[5] Tuz L., Sulikowski K.: Ocena możliwości spawania stali wysokowytrzymałych ulepszanych cieplnie. Przegląd Spawalnictwa, 2018, vol. 90, no. 4, pp. 9–13.

[6] Szczucka-Lasota B., Węgrzyn T., Kalwas J., Jurek A., Wilczyński K.I.: Udarność spawanych podpór podestu ruchomego ze stali S690 QL. Aparatura Badawcza i Dydaktyczna, 2020, vol. 25, no. 1, pp. 31–34.

[7] Fydrych D., Świerczyńska A., Rogalski G., Łabanowski J.: Application of Multivariate Analysis Methods in Welding Engineering. Biuletyn Instytutu Spawalnictwa w Gliwicach, 2018, no. 5, pp. 137–145.

[8] Młodak A., Józefowski T., Wawrowski Ł.: Zastosowanie metod taksonomicznych w estymacji wskaźników ubóstwa. Wiadomości Statystyczne. The Polish Statistician, 2016, no. 2, pp. 1–24.

[9] Porębski A:. Analiza skupień zorganizowanych grup przestępczych jako przykład zastosowania metod statystycznych w kryminologii. Czasopismo Prawa Karnego i Nauk Penalnych, 2021, vol. 25, no. 3, pp. 113–137.

[10] Fydrych D., Rogalski G., Świerczyńska A., Łabanowski J.: Ocena przydatności komercyjnych elektrod otulonych do spawania mokrego pod wodą z wykorzystaniem analizy skupień. Przegląd Spawalnictwa, 2015, vol. 87, pp. 27–30.

[11] Stanisz A.: Przystępny kurs statystyki: z zastosowaniem STATISTICA PL na przykładach z medycyny. Analizy wielowymiarowe. StatSoft, Kraków 2007.

[12] PN-EN 10210-1:2006: Kształtowniki zamknięte wykonane na gorąco ze stali konstrukcyjnych niestopowych i drobnoziarnistych – Część 1: Warunki techniczne dostawy.

[13] PN-EN 10210-3:2020 (E): Kształtowniki zamknięte wykonane na gorąco ze stali konstrukcyjnych – Część 3: Warunki techniczne dostawy stali o wysokiej wytrzymałości i odporności na warunki atmosferyczne.

[14] PN-EN 10025-2:2019: Wyroby walcowane na gorąco ze stali konstrukcyjnych – Część 2: Warunki techniczne dostawy stali konstrukcyjnych niestopowych.

[15] PN-EN 10025-3:2019: Wyroby walcowane na gorąco ze stali konstrukcyjnych – Część 3: Warunki techniczne dostawy spawalnych stali konstrukcyjnych drobnoziarnistych po normalizowaniu/walcowaniu normalizującym.

[16] PN-EN 10025-4:2019 (E): Wyroby walcowane na gorąco ze stali konstrukcyjnych – Część 4: Warunki techniczne dostawy spawalnych stali konstrukcyjnych drobnoziarnistych po walcowaniu termomechanicznym.

[17] PN-EN 10025-5:2019 (E): Wyroby walcowane na gorąco ze stali konstrukcyjnych – Część 5: Warunki techniczne dostawy stali konstrukcyjnych o zwiększonej odporności na korozję atmosferyczną.

[18] PN-EN 10149-2:2013 (E): Wyroby płaskie walcowane na gorąco ze stali o podwyższonej granicy plastyczności do obróbki plastycznej na zimno – Część 2: Warunki techniczne dostawy wyrobów walcowanych termomechanicznie.

[19] PN-EN 10149-3:2013 (E): Wyroby płaskie walcowane na gorąco ze stali o podwyższonej granicy plastyczności do obróbki plastycznej na zimno – Część 3: Warunki techniczne dostawy wyrobów normalizowanych lub walcowanych normalizująco.

[20] EN 10219-3:2020 (E): Kształtowniki zamknięte ze szwem wykonane na zimno ze stali konstrukcyjnych – Część 3: Warunki techniczne dostawy stali o wysokiej wytrzymałości i odporności na warunki atmosferyczne.