Jolanta B. Krolczyk

ISSN 1333-1124
eISSN 1849-1391

THE USE OF THE CLUSTER ANALYSIS METHOD TO DESCRIBE THE MIXING PROCESS OF THE MULTI-ELEMENT GRANULAR MIXTURE

UDC 519.8:62-492.3

Summary

The paper presents a new idea of and approach to using the cluster analysis to describe a complex problem concerning the mixing process of granular mixture in industrial conditions. The phenomenon of mixing is still not well understood particularly in the process of mixing multi-component nonhomogeneous mixtures. Most real granular mixtures which are present in the industrial practice, such as feed mixing, are multi-component heterogeneous systems. The results and conclusions obtained from the modelling process in laboratory conditions cannot be directly translated into industrial conditions due to the change in the scale of the device. Real situations are usually more complex and bring with them new problems to be solved and the need to develop other methods of describing the process of mixing. The article describes the process of mixing compound feeds consisting of various cereal grains of different properties, namely of different size of grains and bulk density, carried out in a two-tone industrial mixer.

Key words: *mixture, mixer, granular mixture, multi-element mixture, cluster analysis, feed*

1. Introduction

The mixing process is a particular unit operation which occurs in a great number of practical applications and in every processing industry (mining, pharmaceuticals, energy, food and agriculture). The effect of mixing determines the quality of the products. Mixing of granular material has been defined as a process of dispersing a few components by chaotic, random movement of grains [1]. Complete mixing could be also defined as the case in which all the sub-mixtures are found to contain the components in the same proportion as the original mixture. Practical mixing issues (concerning fluids or solids) are rarely solved analytically, because the mechanism of mixing is too complicated. The mixing process is a result of mechanisms like diffusion, convection and shear [2, 3, 4, 5], but still the phenomenon of this process is not well understood particularly in the process of mixing multi-component nonhomogeneous mixtures. Most real granular mixtures which are present in the industrial practice, such as feed mixing, are multi-component heterogeneous systems. An extremely important aspect in the mixing of feed is to obtain a good quality mixture in the shortest time possible. Manufacturer of feed expects that every, even the smallest portion of the dosed feed, will always contain the specified composition of various components. By precise balancing of nutrients and their availability for animals both the excretion of unfavourable nutrients, as well as the environmental burden resulting from livestock production can be decreased [6]. Controlling the manufacturing process is a very important issue in every industrial process and many different statistical methods can be applied to monitor the production and management [7, 8].

None of the mixing mechanism models is perfect and each of them has disadvantages and advantages. We know also that the results and conclusions obtained from the modelling process in laboratory conditions cannot be directly translated into industrial conditions due to the change in the scale of the device and other complexities associated with real world conditions. Therefore, there is a need to develop other methods of describing the state of the mixture. Most of the granular mixtures used in the industry, such as feed mixing are multicomponent heterogeneous systems. Particularly in this field the mixing processes are not well understood.

Modelling of both mixing and segregation processes require a confluence of several tools, including discrete and continuum description, like Monte Carlo simulations or cellular automata computations [9, 10]. Even small differences either in size or density lead to the flow-induced segregation [11]. The segregation problem is unavoidable in industrial conditions, especially when industrial formulations are multicomponent mixtures. The behaviour of granular materials is of great technological interest [12, 13] and its investigation has a history of more than two hundred years. Nevertheless, the basic physical understanding of granular media is far from being complete [14]. However, a significant part of research has been related to models and studies carried out in laboratory conditions. Many of these studies concern the phenomenon of segregation in powder processes [12, 15, 16, 17]. Not much research concerns the feed industry. Practical issues concerning mixing are rarely solved analytically due to the fact that the mechanism is too complicated. In these cases, the problem is solved in an experimental model of the process.

The author of the article has made an attempt to describe the course of the mixing process that was taking place in an industrial feed mixer with the use of the cluster analysis, which is an innovative method in solid mixing processes in agricultural science. The aim of the research was to classify the components that are included in the ten element granular mixture, to find similarities in the behaviour of grains during mixing and to evaluate the impact of participation of particular elements on the course of the mixing process.

2. Materials and methods

The study was conducted on the mixing line in an industrial feed plant (Fig. 1) in the following units: a feed mixer, an intake hopper and a bucket elevator (Fig. 2, Fig. 3). The dimensions of the mixer have been presented in the author's previous paper [9].

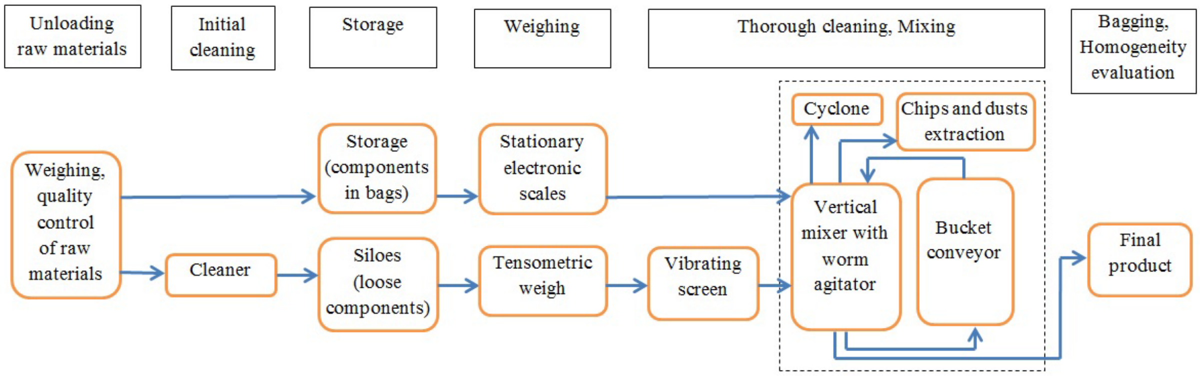

Fig. 1 Diagram of the feed production unit

Fig. 2 Industrial feed mixer

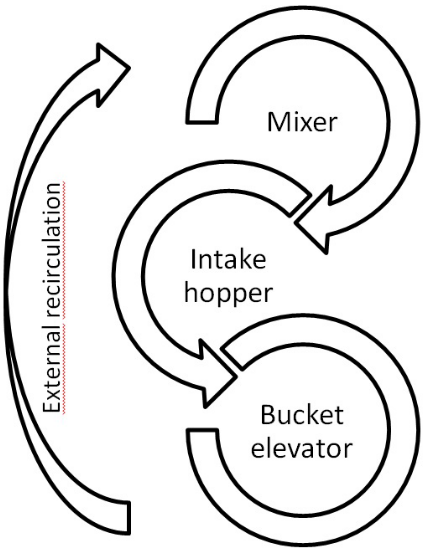

Fig. 3 Schematic diagram of mixing

The process was characterized by the recirculation of the components performed by the bucket elevator. The poured granular material was emptied in the place of discharge, and then directed to the bucket elevator from which it went back to the mixer (Fig. 3). Thus, the mixing took place owing to the movements of the mixing screw and the recirculation through the bucket elevator. The reason of recirculating the mixture was to clean the feed from small impurities (husks, shells, dust). The full mixing cycle including recirculation lasted 30 minutes. The samples to be examined were taken during the process of mixing with recirculation every 30 seconds at the discharge end of the mixer (60 samples in every experiment). The mass fractions of the samples were converted into percentage fractions. The experiment was repeated three times (three series). The granulate mixture used was poultry feed mixture with a total weight poured into the mixer of 2,200 kg. The composition of the granulate mixture, mass fractions and the volume (percentage) fractions are presented in Fig 4.

Fig. 4 Composition of the examined compound feed

3. Statistical assessment

The cluster analysis is one of the most important elements in a multi-dimensional, statistical data analysis and has been applied to a wide variety of research problems. In contemporary nomenclature, the cluster analysis should be classified under taxonomy [11]. Cluster analysis or clustering refers to the grouping of a set of objects in such a way that objects in the same group (cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). Cluster analysis is a set of methods serving to isolate homogeneous subpopulations among the items originating from heterogeneous population [18]. The notion of a 'cluster' cannot be precisely defined, which is one of the reasons why there are so many clustering algorithms [19]. There is, however, a common denominator: a group of data objects. Still, different researchers employ different cluster models, and for each of these cluster models different algorithms can be given. The notion of a cluster varies between algorithms and the clusters found by different algorithms vary significantly in their properties. Understanding these 'cluster models' is the key to understanding the differences between the various algorithms [19]. By applying the cluster analysis, the grouping of mixture components was made and similarities in the behaviour of grains during mixing were identified. Then, the obtained results were compared to the average size of the grain particles and the bulk density. Moreover, the influence of the participation of individual components on the course of the mixing process was assessed. In Fig. 5 the algorithm of the method is presented.

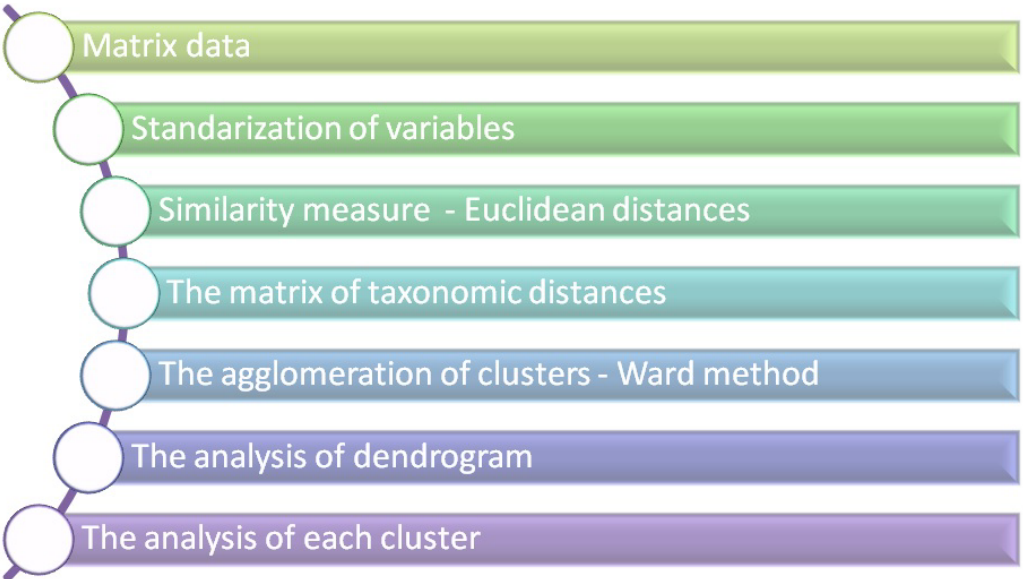

Fig. 5 Algorithm of the statistical assessment

The output data matrix is the percent portion of particular elements depending on the time of mixing (Table 1). After the output data standardization had been performed, the Euclidean measure of the distance between items was applied. The matrix of taxonomic distances is described in Table 2. The agglomeration of clusters was performed using the Ward method. The characteristic of this method is the assurance of the minimum variance within the cluster variance. This method provides homogeneity inside the clusters and heterogeneity between them (in the sense of minimization and maximization of the variances) [20]. The plan of the course of agglomeration obtained for the examined mixture for the selected research series is presented in Table 3. The obtained dendrograms are presented in Figs. 6 - 8. The number of homogenous clusters (subpopulations) can be decided by the researcher or supported by using one method suggested by other researchers. In this study, a method suggested by Hellwig was used [21]. In this method, two subsets are considered to be fundamentally different if the distance between a pair of points belonging to two different subsets is greater than a certain critical value (W_k). In order to evaluate the critical value, a minimum value in particular rows of the distance matrix has to be found. Then, for the variables so formed the arithmetic mean x and the standard deviation σ are calculated. The critical value is calculated on the basis of the following formula [22]:

$$W_k = x + 2\sigma \tag{1}$$

where:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}} \tag{2}$$

x_i - the following minimum values in the rows of the Euclidean distances matrix

$\overline{x}$ - the mean of the minimum values

n – the number of elements in the matrix of the Euclidean distances

The final step of the statistical assessment was the analysis of individual clusters to find out due to what characteristics a particular cluster was formed. In order to make the analysis, the method of arithmetic means proposed by Runge [23] was applied. In this method, for the output data matrix as well as for particular clusters, the arithmetic means of subsequent characteristics Xn – the subsequent minutes of mixing were calculated. Another step was to calculate a group averages X, which is average for a particular characteristic (minutes of mixing). The structure indicators of each cluster are quotients Xn/X. The quotient greater than one proves the advantage of a particular characteristic in the cluster. A calculation of the number of dominant clusters (cases) for 60 examined cases was made.

4. Results

The results of experimental studies are presented in Table 1.

Table 1 Percentage share of particular elements of the examined mixture depending on the time of mixing for selected research series

Item number	Components of the mixture	Percentage share during consecutive mixing time /s								
		30	60	90	120	150	...	1740	1770	1800
1	Maize	88.65	79.88	58.9	48.7	46.5	...	52.13	50.06	44.84
2	Wheat	10.12	10.02	20.42	42.83	73.56	...	47.33	45.59	52.88
3	Field pea	8.17	13.74	13.71	13.66	18.2	...	15.89	12.75	13.35
4	Green pea	1.59	0	0	0	0	...	1.52	1.58	2.76
...		...	...	...	...	...	...	...	...	...
9	Sunflower	0	0	0.1	0.3	0	...	1.63	2.08	2.59
10	Dari	0	0	0.06	0	0	...	5.25	3.91	3.5

Average masses of samples were as follows: 1st research series 138.98 g (standard deviation 6.05); 2nd research series 140.03 (standard deviation 11.06), 3rd research series 141.42 (standard deviation 6.43). The samples were separated manually into individual components. The matrix of taxonomic distances which describes the resemblance between the items is presented in Table 2. The values of taxonomic distances of neighbour joining for separated clusters are presented in Table 3.

Table 2 Matrix of Euclidean distances obtained for elements of 10 element mixture for selected research series

	Maize	Wheat	Field pea	Green pea	Yellow pea	Sorghum	Yellow millet	Brown vetch	Sunflower	Dari	Minimum values
Maize	0.0000	11.3743	16.4430	19.8869	17.6879	18.3866	19.0738	20.4711	19.9941	19.8318	11.3743
Wheat	11.3743	0.0000	13.9103	17.3563	15.2172	15.2562	16.9992	17.7018	17.2817	17.1787	11.3743
Field pea	16.4430	13.9103	0.0000	4.0057	1.7872	3.0614	3.5074	4.3057	3.9434	3.7270	1.7872
Green pea	19.8869	17.3563	4.0057	0.0000	2.5435	3.9897	2.2989	1.2298	0.6691	2.0070	0.6691
Yellow pea	17.6879	15.2172	1.7872	2.5435	0.0000	2.9586	2.4397	2.9242	2.4837	2.6039	1.7872
Sorghum	18.3866	15.2562	3.0614	3.9897	2.9586	0.0000	3.2376	3.7890	3.7567	2.6297	2.6297
Yellow millet	19.0738	16.9992	3.5074	2.2989	2.4397	3.2376	0.0000	2.1888	2.1879	1.7298	1.7298
Brown vetch	20.4711	17.7018	4.3057	1.2298	2.9242	3.7890	2.1888	0.0000	0.7788	1.4503	0.7788
Sunflower	19.9941	17.2817	3.9434	0.6691	2.4837	3.7567	2.1879	0.7788	0.0000	1.6877	0.6691
Dari	19.8318	17.1787	3.7270	2.0070	2.6039	2.6297	1.7298	1.4503	1.6877	0.0000	1.4503
										Mean	3.4250
										Standard deviation σ	4.2341
										Critical value W_k	11.8931

Table 3 Plan of the course of agglomeration obtained for the examined mixture for selected research series

Agglomerate distance	Object No.									
	1	2	3	4	5	6	7	8	9	10
0.6691	Green pea	Sunflower								
1.1160	Green pea	Sunflower	Brown vetch							
1.7298	Yellow millet	Dari								
1.7872	Field pea	Yellow pea								
2.9763	Green pea	Sunflower	Brown vetch	Yellow millet	Dari					
3.4176	Field pea	Yellow pea	Sorghum							
6.7845	Field pea	Yellow pea	Sorghum	Green pea	Sunflower	Brown vetch	Yellow millet	Dari		
11.3743	Maize	Wheat								
43.7397	Mazie	Wheat	Field pea	Yellow pea	Sorghum	Green pea	Sunflower	Brown vetch	Yellow millet	Dari

Another step in the analysis was to make agglomerations of the clusters in order to obtain a dendrogram (Figs. 6-8). A dendrogram is a tree diagram frequently used to illustrate the arrangement of clusters produced by hierarchical clustering. Smaller distances of taxonomic bonding means more similarity. The obtained dendrogram that illustrates the cluster hierarchy of particular elements was analysed. In order to determine the number of clusters and particular elements of the mixture that belong to that mixture, the critical value Wk was calculated. The critical value, after being replaced by the formula (1), was tantamount to Wk=10.76 (for the first research series), Wk =11.89 (for the second research series) and Wk=14.10 (for the third research series). The value is marked on the graph as full line. The critical value divided the dendrogram into two clusters. The composition of the particular clusters is presented in Figs. 6-8 and Table 4.

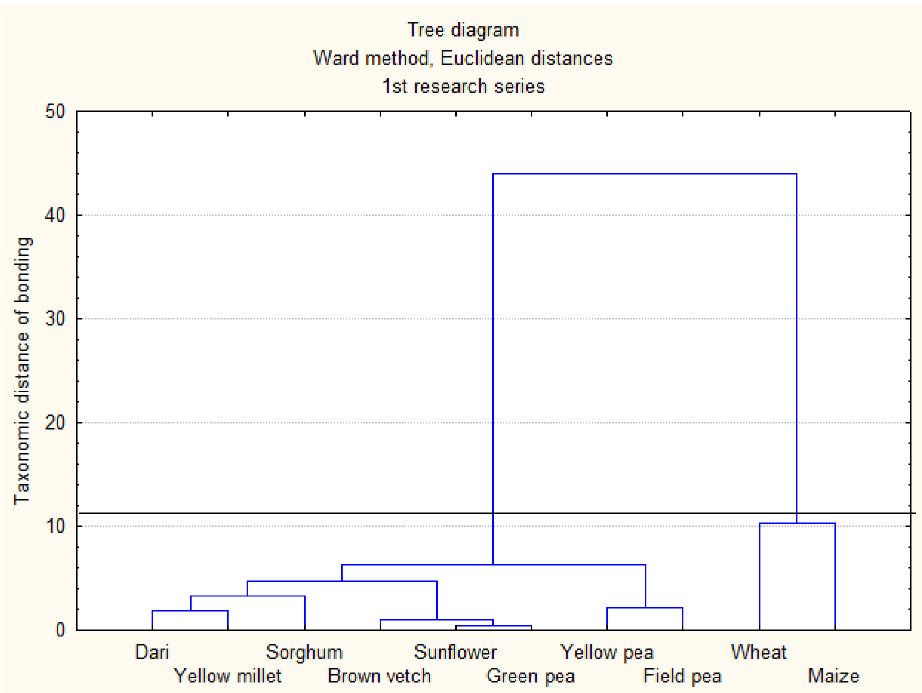

Fig. 6 Tree diagram of taxonomic distances of bonding for the first research series

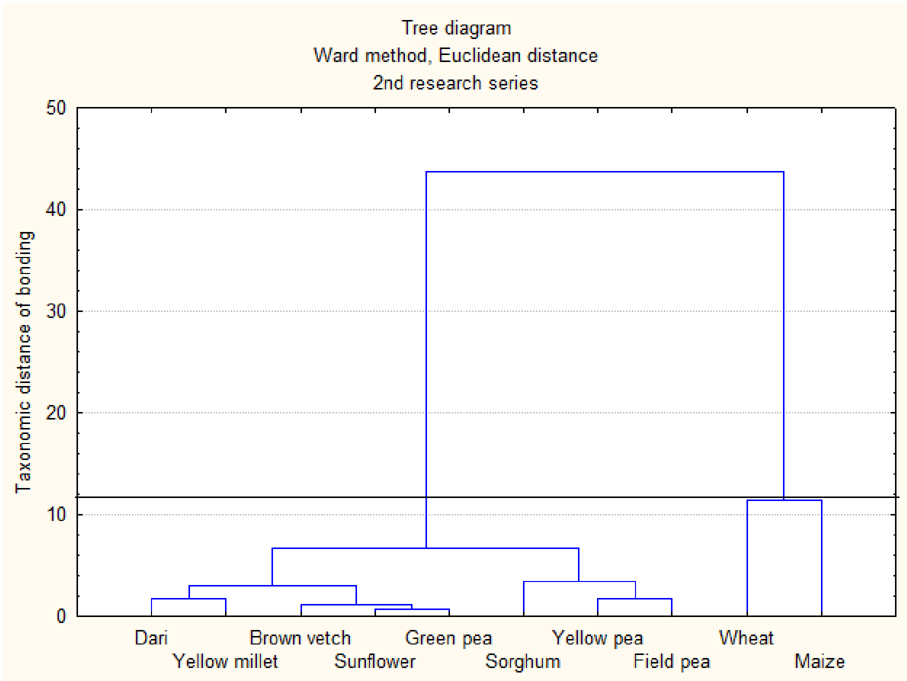

Fig. 7 Tree diagram of taxonomic distances of bonding for the second research series

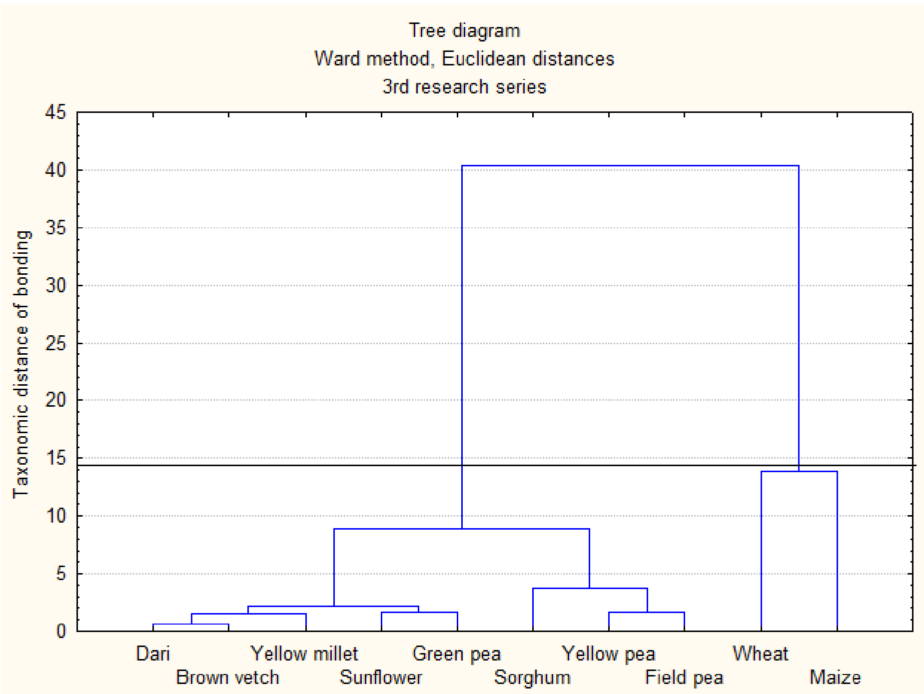

Fig. 8 Tree diagram of taxonomic distances of bonding for the third research series

Table 4 Division of mixture components into clusters

Number of the cluster	Components of the mixture	Portion of the component in the mixture, %	Bulk density /kg·m^{-3}	Average size of particles /mm	
1	Dari	0.91	723	3.38	
	Yellow millet	6.82	732	2.26	
	Sorghum	6.82	697	3.94	
	Brown vetch	1.14	800	4.62	
	Black sunflower	1.82	430	4.90	
	Green pea	2.27	767	7.24	
	Yellow pea	6.14	771	6.95	
	Field pea	8.64	793	6.86	
2	Wheat	29.09	718	3.74	
	Maize	36.35	726	8.16	

The last step was to make an analysis of each of the separated clusters. By calculating quotients Xn/X (arithmetic mean of subsequent clusters / group average) the advantage of a particular characteristic (minutes of mixing) in a cluster was determined (Table 5). Graphical interpretation of the calculated quotients for particular minutes of mixing is presented in Fig. 9.

Table 5 Assessment of particular cluster similarities according to the properties (min) of Xn/X quotient values for the third research series

Parametres	Mixing time /s				
	30	**60**	**…**	**1170**	**1800**
	Cluster 1				
Dari	0.00	0.00	…	2.41	1.15
Yellow millet	0.05	0.07	…	3.70	3.25
Sorghum	0.08	0.05	…	15.43	16.57
Brown vetch	0.00	0.00	…	1.28	1.04
Sunflower	0.21	0.07	…	1.91	2.09
Green pea	21.91	7.16	…	0.69	2.64
Yellow pea	22.54	12.69	…	10.96	7.98
Field pea	3.30	3.49	…	13.05	12.40
Mean X1	6.01	2.94	…	6.18	5.89
Mean X	14.26	15.00	…	14.71	13.73
X1/X	0.42	0.20	…	0.42	0.43
	Cluster 2				
Wheat	12.16	8.99	…	34.33	41.88
Maize	82.37	117.49	…	63.37	48.27
Mean X2	47.27	63.24	…	48.85	45.08
Mean X	14.26	15.00	…	14.71	13.73
X2/X	3.31	4.22	…	3.32	3.28

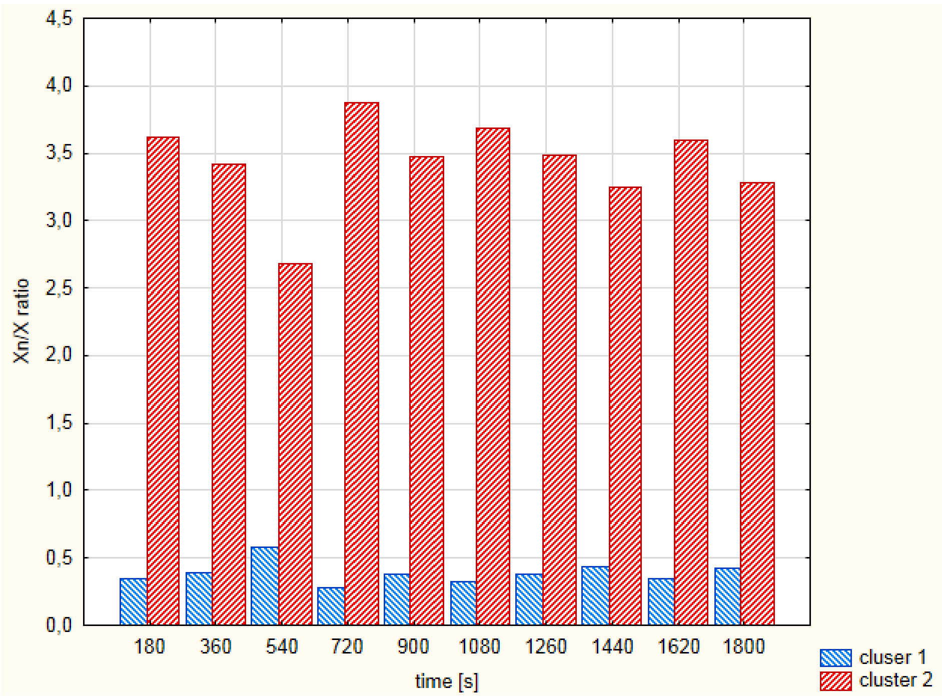

Fig. 9 Graphical interpretation of quotients Xn / X for selected minutes of mixing for selected research series

5. Discussion

The research objective was the classification of the components included in a ten element granulate mixture and finding a similarity in the behaviour of grains during mixing, as well as the assessment of the participation of individual components in the course of the mixing process. Based on the results the following conclusions may be made. The mixture components were divided into two groups, thus two clusters, differing in features from each other. Only two components with the largest percentage fractions, i.e. maize and wheat, exerted dominant influence on the course of the process. The components behaved in a similar way during the mixing process in the device with a mixing screw and recirculation. Fig. 9 shows that the values of quotients Xn/X of cluster 1, namely ofdari, yellow pea, brown vetch, sorghum, sunflower, yellow millet, green pea, yellow pea, and field pea, did not exceed the value of 1 in any of the examined cases. The components behaved similarly during the mixing with recirculation in the industrial feed mixer.

6. Conclusions

1. The cluster analysis is a new and very useful instrument that allows a description of mechanisms of mixing multi-element heterogeneous granular structures and helps the understanding of principles that govern this process.
2. A classification of the mixture elements concerning their behaviour in the course of the process in an industrial mixer has been made. A similarity in behaviour of wheat and maize grains has been determined - these are the elements with similar bulk densities that were grouped into one cluster.
3. The remaining elements (field pea, yellow pea, green pea, vetch, dari, sorghum, sunflower and millet) of varied bulk densities and component size behaved similarly during mixing and have been classified into one cluster.
4. The greatest impact on the course of the mixing process has been exerted by wheat and maize; these grains had the largest portion in the mixture that was tantamount to 65.44%.

REFERENCES

[1] Boss, J. (1987). Solid mixing process (in Polish). PWN, Warszawa – Wrocław

[2] Barbosa-Cánovas, G.V., Ortega-Rivas, E., Juliano, P., Yan H. (2005). Food Powders Physical Properties. Processing and Functionality. Springer-Verlag, New York

[3] Choldrich, R. (2002): Fundamentals of particle technology. Midland Information Technology and Publishing. Loughborough

[4] Ottino, J.M., Khakhar, D.V. (2000). Mixing and Segregation of Granular Materials. Annual Review of Fluid Mechanics. 32, 55-91

[5] Tang, P., Puri, V. M (2004). Methods for Minimizing Segregation: a Review. Part. Sci. Technol. 22, 321-337

[6] Gillespie, J., Flanders F. (2010). Modern Livestock & Poultry Production. 8th Edition Cengage Learning

[7] Krolczyk, G., Gajek, M., Legutko, S. (2013). Predicting the tool life in the dry machining of duplex stainless steel. Eksploatacja i Niezawodnosc - Maintenance and Reliability 01 15(1), 62-65

[8] Oyama, T., Miwa M. (2006). Mathematical modeling analyses for obtaining an optimal railway track maintenance schedule. Japan J. Indust. Appl. Math. 23, 2, 207-224

[9] Krolczyk, J.B. (2014). An attempt to predict quality changes in a ten-component granular system. Tehnicky Vjesnik - Technical Gazette Tehnički vjesnik. 21, 2255-261

[10] Krolczyk J., Tukiendorf M. (2005). Using the methods of geostatic function and Monte Carlo in estimating the randomness of distribution of a two-component granular mixture during the flow mixing. Electronic Journal of Polish Agricultural Universities. 8 Issue 4

[11] Mezzich, J. E., Solomon, H. (1980). Taxonomy and Behavioral Science. Academic Press. London. UK

[12] Jaeger, Y., Nagel,Y., Behringer R.P. (1996). Phys. Today. 4, 32

[13] Krolczyk J.B., Rezwiakow A., Tukiendorf M. (2014). Mixing of biomass and coal in a static mixer as an example of technological solutions involving implementation of renewable energy sources. Ecol CHEM ENG S. 4 vol. 21, 685-696. doi: 10.1515/eces-2014-0050

[14] Dury, C. M., Ristow, G. H. (1997). Radial segregation in twodimensional rotating drum. J. Phys. I France. 7, 737-745

[15] Daumann, B., Nirschl, H. (2008). Assessment of the mixing efficiency of solid mixtures by means of image analysis. Powder Technology. 182, 415–423

[16] Duffy, S. P., PURI, V. M. (2003). Development and Validation of a Constitutive Model for Size-Segregation during Percolation. KONA 21, 151-162

[17] Hogg, R. (2003). Characterization of Relative Homogeneity in Particulate Solids. Int. J. Miner. Process. 72, 477-487

[18] Tryon, R.C. (1939). Cluster Analysis. Edwards Press. Ann Arbor. Ml. USA

[19] Estivill-Castro, V (2002). Why so many clustering algorithms — A Position Paper. ACM SIGKDD Explorations Newsletter 4 (1): 65–75. doi:10.1145/568574.568575

[20] Marek, T. (1989). Cluster Analysis in Empiric Research. SAHN methods (in Polish). PWN Press, Warsaw, Poland

[21] Hellwig, Z. (1968). Application of the taxonomic method in typological division of countries based on the level of their development and resources as well as skilled employees structure (in Polish). Przegląd Statystyczny. 4, 307-326

[22] Pluta, W (1977). Multi-dimensional Comparison Analysis in Economic Research. PWN Press. Warsaw, Poland

[23] Runge, J. (1992). Selected problems of spatial analysis in geographical research (in Polish). Skrypt Uniwersytetu Śląskiego 469

Acknowledgements: The Author thank Professor Thomas G. Mathia and Prof. Andrzej Tukiendorf for professional guidance and valuable support in preparing the article.

Submitted: 25.9.2014

Accepted: 10.12.2014

Jolanta B. Królczyk
Department of Biosystems Engineering
Faculty of Production Engineering and Logistics
76 Prószkowska Street, 45-758 Opole, Poland