Original scientific paper
Journal of Central European Agriculture, 2019, 20(3), p.770-779
DOI: /10.5513/JCEA01/20.3.2242

Grouping of hen lines according to some productive indicators through a combination of mathematical approaches

Групиране на линии кокошки според някои показатели за продуктивност чрез комбинация от математически подходи

Neli KERANOVA[1] (✉), Pavlina HRISTAKIEVA[2], Magdalena OBLAKOVA[2]

[1] Agricultural University-Plovdiv, Faculty of Economics, Department of Mathematics, Informatics and Physics, 12 "Mendeleev" Blvd., 4000, Plovdiv, Bulgaria

[2] Agrarian Academy, Agricultural Institute, 6000, Stara Zagora, Bulgaria

✉ Corresponding author: nelikeranova@abv.bg

ABSTRACT

The object of this study includes the following lines of hens: egg-type hens (line D, line B, line CZ80M, line CZ80B) and dual-purpose type hens (line NG, line E, line Ss, line StR, line ChS). After some measurements, biometric data were obtained, related to the following groups of indicators: reproductive traits of lines of hens (fertility percentage and hatchability percentage); body weight of 1-day old chickens: male, female (g); body weight of 5-month-old: male, female (kg); average egg mass (g), age of sexual maturity (days), body weight of 10-month-old hens (kg), egg production for 180 days (number). The main objectives in the present study are two: on the one hand to group the indicated lines of hens into clusters according to similarity in the relevant groups of indicators and on the other hand to determine which features have the greatest impact in the formation of the individual clusters. A combination of two mathematical-statistical methods was applied that provide objective and comprehensive information about the questions asked. A hierarchical cluster analysis was first used, followed by a factor analysis using the method of the main components. For the lines in the egg-type hens group, it was found that line D was farthest from the rest of the examined lines according to most of the analyzed indicators. The same is Line ChS from dual-purpose-type hens.

Keywords: cluster analysis, factor analysis, hens, productivity

РЕЗЮМЕ

Обект на изследване са следните линии кокошки: яйценосен тип (линиа Д, линиа Б, линиа ЦЗ-80М, линиа ЦЗ 80Б) и общоползувателен тип (линиа НГ, линиа Е, линиа Сс, линиа СтР, линиа СтР). След измервания са получени биометрични данни, свързани с групите показатели: репродуктивност (оплоденост на яйцата (%) и люпимост (%)); жива маса на еднодневно пиле: мъжки, женски (г); жива маса на петмесечна възраст: мъжки, женски (кг); жива маса на десетмесечна възраст на кокошките (кг), яйчна продуктивност за 180-дневна (брой), средна маса на яйцата (г), възраст на полова зрялост (дни). Основните цели в настоящото изследване са две. От една страна, да се групират посочените линии кокошки в кластери според сходство в съответните групи показатели, а от друга да се определи кои признаци оказват най-силно влияние при формирането на отделните клъстери. Приложени са два математически метода, даващи обективна и цялостна информация на поставените въпроси. Използван е йерархичен клъстерен анализ, последван от факторен анализ по метода на главните компоненти. За линиите от групата на яйценосните кокошки се установи, че линиа Д е най-отдалечена от останалите изследвани линии по по-голямата част от анализираните показатели. От общоползувателните линии такава е линиа ЧС.

Ключови думи: факторен анализ, кластерен анализ, кокошки, продуктивност

INTRODUCTION

A classic mathematical-statistical method for research in the field of genetics and selection is the cluster analysis (Vanhala et al., 1998; Rosenberg et al., 2001). The use of cluster analysis for the purpose of grouping the lines of hens according to certain indicators is a widespread study tool in poultry production (Rosenberg et al., 2001; De Marchi et al., 2006; Granevitze et al., 2007).

Some authors use a hierarchical cluster analysis to group broiler lines according to their bacterial characteristics in the gut (Yang et al., 2008; Corrigan et al., 2015). Savegnano et al. (2011) apply several mathematical approaches to study and analyze the relations between breeding values and egg production. The objectives of this paper are to identify the phenotypic egg-laying patterns in a White Leghorn line mainly selected for egg production, to estimate genetic parameters of traits related to egg production and to evaluate the genetic association between these by principal components analysis to identify trait(s) that could be used as selection criteria to improve egg production. Grouping is performed according to the genotype specificity through a non-hierarchical cluster analysis using the method of K-mean analysis.

Tadano et al. (2012) examine five lines of chickens and grouped them through a cluster analysis based on indicators related to their genotype features.

The genomes of Chinese native chicken populations are screened using microsatellites as molecular markers (Chen et al., 2004). By using fuzzy cluster, 12 Chinese native chicken populations are divided into three clusters. The first cluster comprised Taihe Silkies, Henan Game Chicken, Langshan Chicken, Dagu Chicken, Xiaoshan Chicken, Beijing Fatty Chicken and Luyuan Chicken. The second cluster included Chahua Chicken, Tibetan Chicken, Xianju Chicken and Baier Chicken. Gushi Chicken formed a separate cluster and demonstrated a long distance when comparing with other chicken populations. The cluster could be confirmed from three aspects which maybe geographical, bodily form and economical purposes.

Cluster analysis is used to investigate the genetic divergence among five lines of laying hens by Barbosa et al. (2005). There is genetic divergence between the evaluated laying hens' lines, so that two lines showed higher laying rate, egg weight and body weight. The trait that contributed mostly to the genetic divergence is body weight at 48 weeks of age.

The genetic diversity of the Turkish native chicken breeds Denizli and Gerze are evaluated with 10 microsatellite markers (Kaya and Yildiz, 2008). A phylogenetic tree is constructed using genetic distance and the neighbor-joining method. Its topology reflects the general pattern of genetic differentiation among the Denizli and Gerze breeds. Denizli and Gerze subpopulations have a rich genetic diversity. The information about Denizli and Gerze breeds estimated by microsatellite analysis may also be useful as an initial guide in defining objectives for designing future investigations of genetic variation and developing conservation strategies.

The aim of the present work is to examine the indicated lines of hens by grouping them in clusters through a hierarchical cluster analysis. Thus, on the one hand, establish similar lines on certain indicators and, and on the other – lines with proven differences.

MATERIALS AND METHODS

Nine lines of hens are subject of research in this work: egg-type hens (line D, line B, line CZ80M, line CZ80B) and dual-purpose type (line NG, line E, line Ss, line StR, line ChS), raised in the poultry farm at Stara Zagora - Agricultural Institute. The studies were developed in the section "Selection of population genetics and technologies of birds and rabbits". Their productivity is controlled through: reproductive traits of lines hens (fertility percentage and hatchability percentage), body weight of 1-day old chickens (male, female) (g), body weight of 5-month-old birds (male, female) (kg), body weight of 10-month-old hens (kg), egg production for 180 days (number), average egg mass (g), age of sexual maturity (days).

The hens were hatched in June, mated randomly (panmixia) in common breeding groups and sex ratio 1:10 on a deep permanent litter. A selection for exterior traits is made at 3 months of age. Birds are fed with commercial rations according to their production type and age (Lalev et al., 2012).

The indicators on the basis of which the cluster analysis was conducted are grouped as follows:

-first group - reproductive traits of breeding lines: egg fertility percentage, dead embryos at first and second inspection, hatchability percentage of eggs set, hatchability percentage of fertile eggs;

-second group - live body weight of 1-day-old chickens (g): female, male;

-third group - live body weight of 5-month-old laying birds (kg): female, male;

-fourth group - live body weight of 10-month-old laying hens (kg);

-fifth group - egg production for 180 days, average egg mass (g).

The mathematical processing in the current work is based on data for nine lines of hens grouped in two groups: egg-type (line D, line B, line CZ80M, line CZ80B) and dual-purpose type (line NG, line E, Line Ss, line StR, line ChS). For each group, biometric data were obtained on the following groups of indicators: reproductive traits of breeder lines: egg fertility percentage, dead embryos at first and second inspection, hatchability percentage of eggs set, hatchability percentage of fertile eggs; live body weight of 1-day-old chickens (g): female, male; egg production for 180 days, average egg mass (g); live body weight of 5-month-old laying hens (kg): female, male; live body weight of 10-month-old laying hens (kg).

To achieve the goals set at the beginning of the study, a hierarchical cluster analysis was used. This method was applied for each group of hens and for each group of indicators respectively. The notes, further included in the statement, apply to each clustering procedure pertaining to each of the groups considered (this will not be explicitly noted in order to avoid repetition). A dendrogram has been obtained that visualizes the results of the clustering procedure. Following a hierarchical cluster analysis using different clustering agglomerative methods, have chosen the inter-group binding method and the measure of similarity is the quadratic Euclidean distance.

To determine a suitable agglomerative clustering method giving optimal results, the one-dimensional distribution method was applied. Cross-tables were built and the contingency ratio was found to be greatest in the preferred method. This determines its choice in the current mathematical processing.

A factor analysis is used to determine the criteria by which some lines form a cluster and others are separated into another clusters. In order to avoid the impact of the different dimensions, the data is standardized. It is found that they have a distribution close to normal. The factor analysis is a statistical approach by which a plurality of n correlated variables is transformed into a plurality of m uncorrelated variables (m<n) which most influence the change of the initial data.

The mathematical data processing was performed using the statistical program product IBM Statistics SPSS 24 (Meyers et al., 2013; Abramowitz and Weinberg, 2016; Cronk, 2016).

RESULTS AND DISCUSSION

IBM Statistics SPSS 24 was used to establish the normal distribution of data (using the Kolmogorov-Smirnov test). All observations are independent. In order for a factor analysis to be applied, the data needs to be subjected to some tests. The determinant of the correlation matrix in all performed factor analyses in the present work is a positive number. For each individual procedure for applying a factor analysis, the Kaiser-Meyer-Olkin (KMO) test (>0.5) and the Bartlett's Sphericity test (Sig.<0.05) were performed. A principal component analysis (PCA) was applied. From the mathematical point of view, a new basis of variables, which are mutually orthogonal, is obtained through a factor analysis. For this purpose, the standard rotation method: Varimax was applied. The results of the factor analysis of the two groups of hens according to the five groups of indicators lead to the qualitative description of the formed clusters set out below. The factor analysis revealed that the factor Body weight of 5-month-old male exhausts 73.197% of the total dispersion in the second group of factors. In the first group, the factor: egg fertility percentage exhausts 66.468% and together with hatchability percentage of eggs set explains 98.229% of the total dispersion.

From the dendrogram in Figures 1 and 3 it can be seen that according to the indicators in the first and third groups the four hens' lines are grouped in two clusters in an identical manner. The first includes line CZ80M, line CZ80B and line B, which have the largest egg fertility percentage, the smallest hatchability percentage of eggs set, hatchability percentage of fertile egg and body of weight of 5-month-of-female and the second consists of line D whose values are in the opposite direction.

In the second group of indicators, the hens from egg-type form two independent clusters (Figure 2), which combine at a large Euclidean distance. The first one includes line D and line CZ80M (minimum body weight of 1-day-old chickens (g), male, female) and the second one – line B and line CZ80B, respectively maximum.

Figure 4 shows that line B, line CZ80M and line CZ80B form a common cluster, due to the similarity in body weight of 10-month-old laying hens (kg), and line D has proven difference from the other (an indicator with minimal values), which results in separating it in a cluster.

According to the indicators in the fifth group (Figure 5), the examined lines of hens are again divided into two clusters. Line B and line CZ80B form the first one, as they have maximum size (number or value) of average egg production for 180 days, while line D and line CZ80M have minimum size, statistically proven different from the rest, which is the reason for their separation in an independent cluster.

The dendrograms of figures 6-10 show the results of the cluster analysis according to the degree of similarity

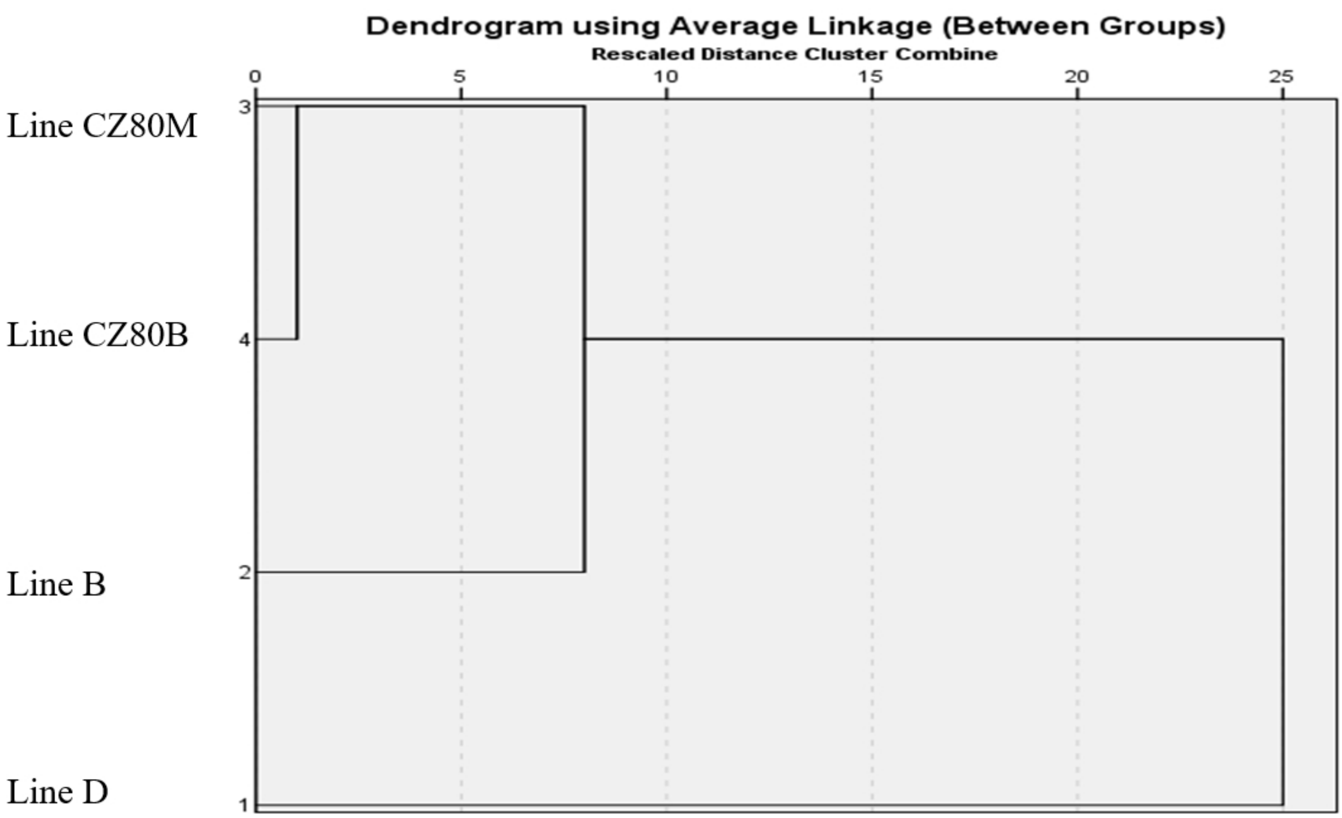

Figure 1. Dendrogram of egg-type-hens, first group of indicators

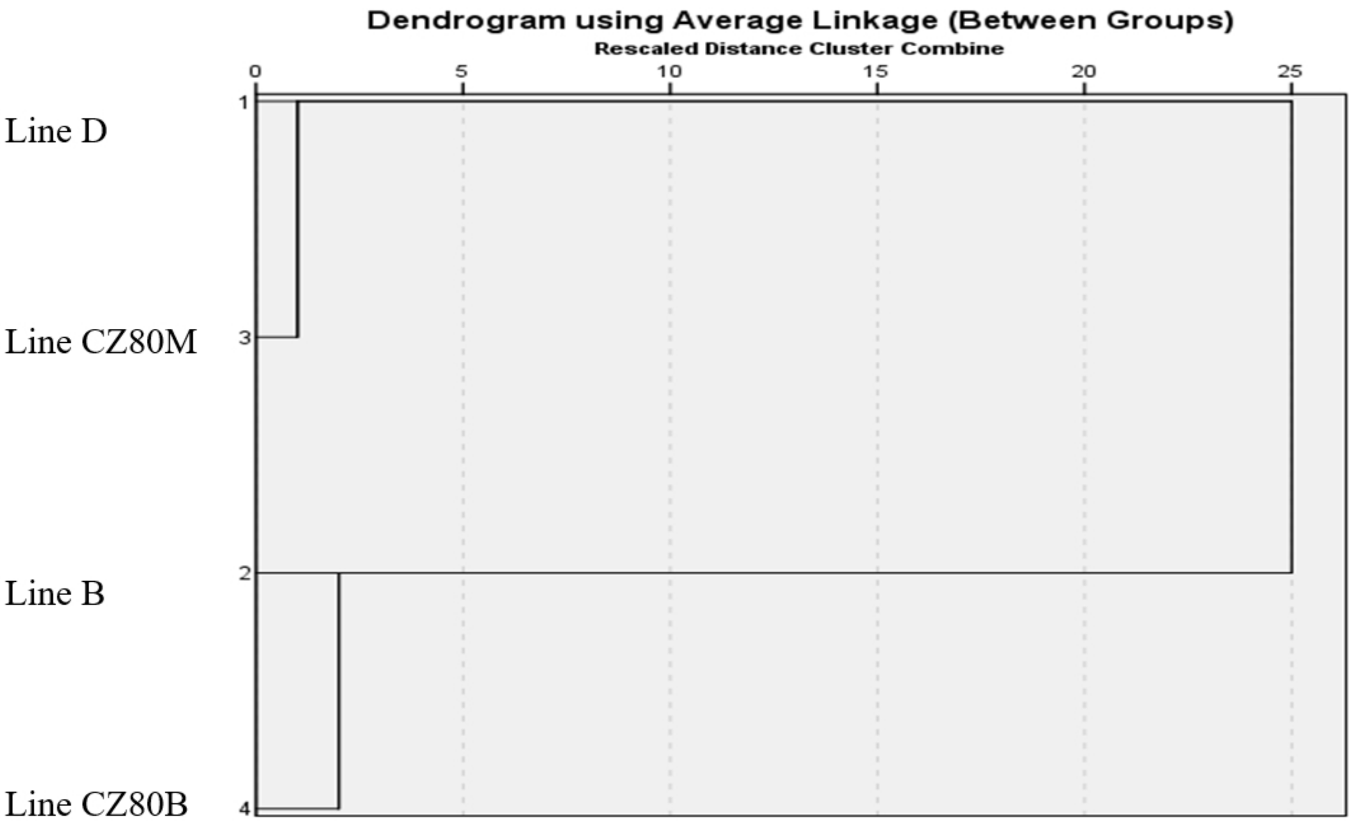

Figure 2. Dendrogram of egg-type-hens, second group of indicators

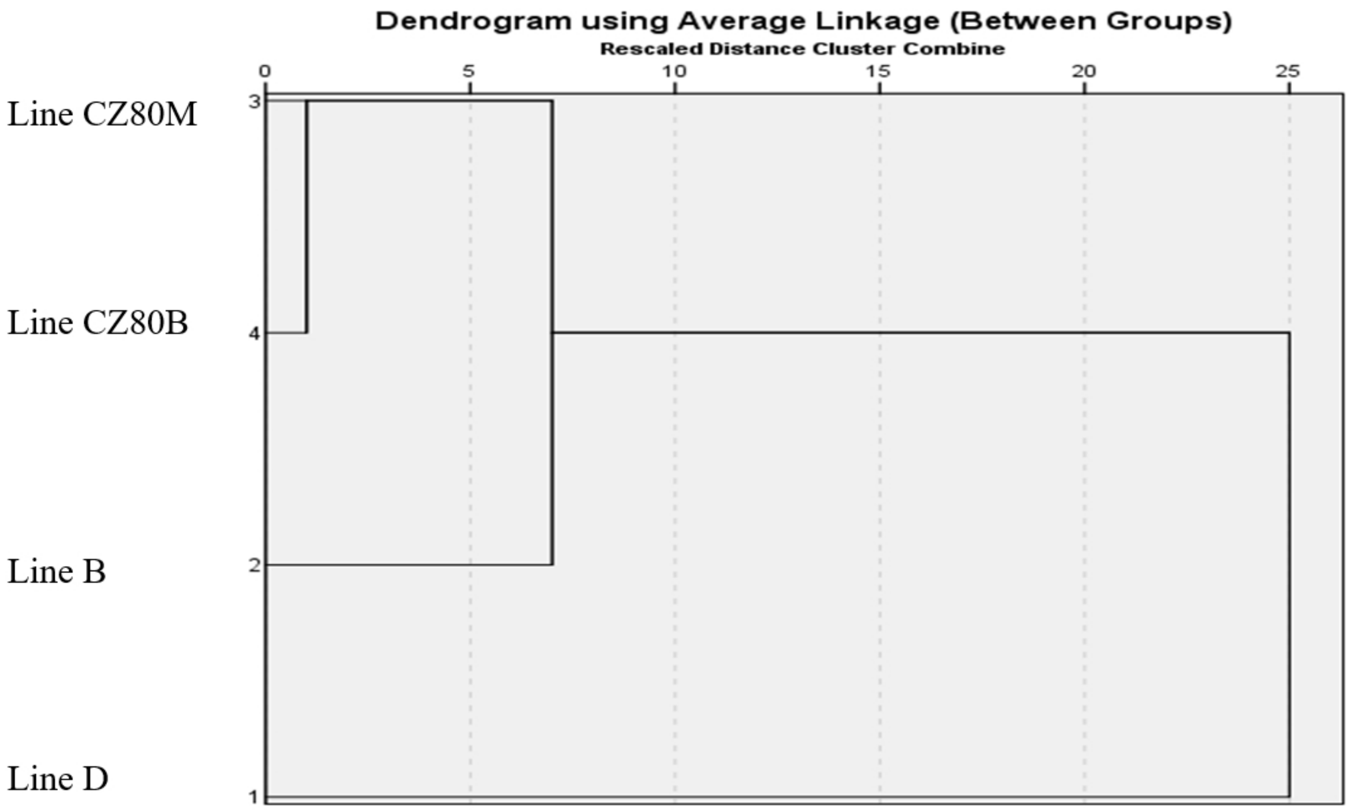

Figure 3. Dendrogram of egg-type-hens, third group of indicators

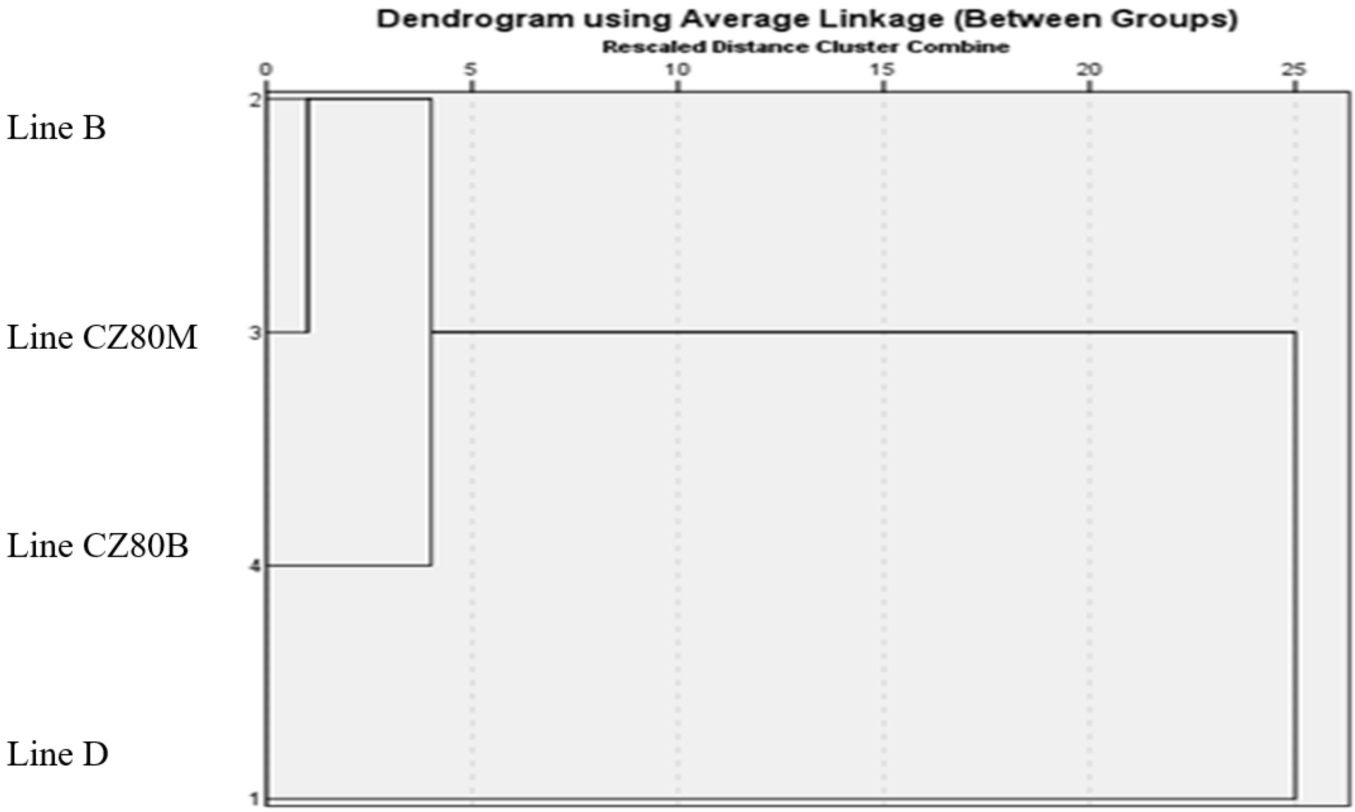

Figure 4. Dendrogram of egg-type-hens, fourth group of indicators

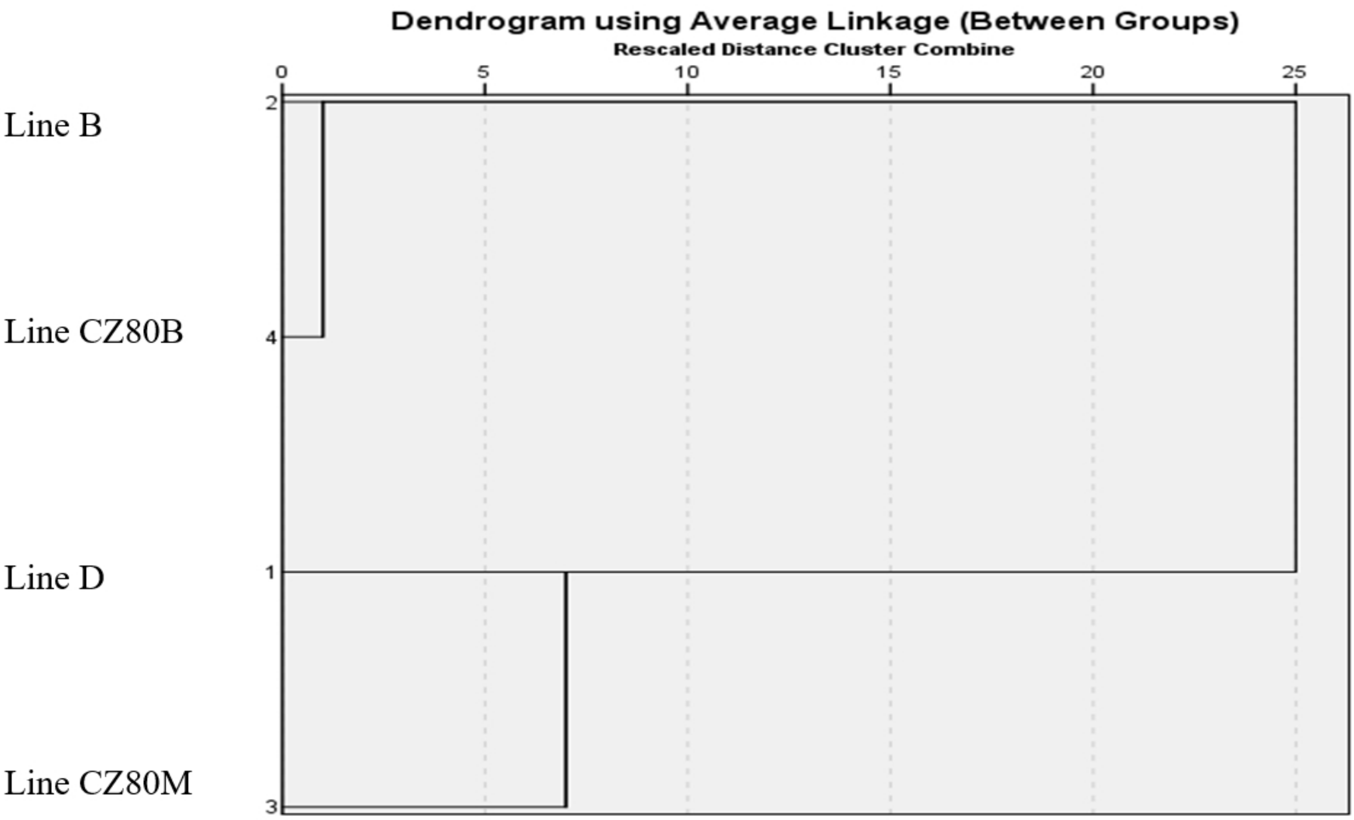

Figure 5. Dendrogram of egg-type-hens, fifth group of indicators

of the Dual-purpose-type hens for each of the five groups of features. According to the first group, line NG, line StR and line E form a cluster because they have maximum of eggs set (%), while line Ss and line ChS have minimal hatchability of eggs set (%) and hatchability of fertile eggs (%). A factor analysis revealed that the other indicators in this group do not affect the clustering (Figure 6).

On the dendrogram in Figure 7 two clusters are formed according to the degree of similarity of the five lines of hens according to the indicators of the second group. Line Ss and Line ChS form a separate cluster because they have minimum values according to both indicators in the group. The indicators of the other three lines have larger values, with a statistically proven similarity, which determines their separation into a new cluster.

According to the third group of factors, the studied five lines of hens were divided into three clusters (Figure 8). The first one includes: line NG, line StR and line Ss, the second one consists of: line E (with the highest values according to the two indicators in this group) and the third one consists of line ChS (with the highest values according to the two indicators from this group) which is farthest from the other two.

According to the indicator in the fourth group, the five lines of hens under consideration were grouped into two clusters (Figure 9). Line E forms a stand-alone cluster and joins at a distance of 25 Euclidean units to the cluster with the other lines, proving its remoteness according to the surveyed indicators. This is due to the large value of body weight of 10-month-old laying hens (kg), which is statistically proven different from the rest.

Three clusters are obtained according to the indicator in the fifth group (Figure 10). Line Ss and line ChS form the first cluster. These are the lines with minimal Average egg production for 180 days and Average egg mass (g). The second cluster consist of line E and line StR - the lines with maximal Average egg mass. The third cluster consists of line NG, which has a maximum size of Average egg production for 180 days and statistically different from other lines.

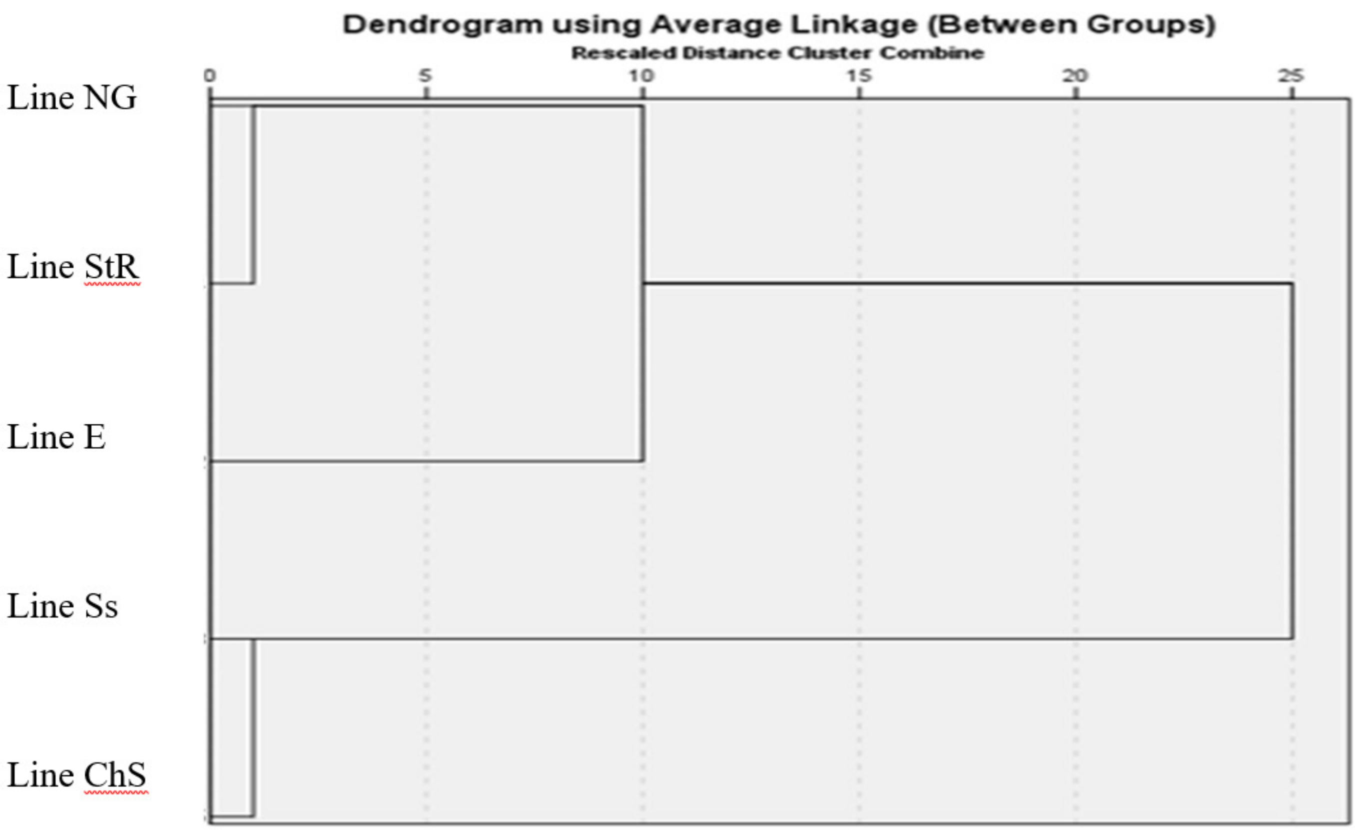

Figure 6. Dendrogram of dual-purpose-type hens, first group of indicators

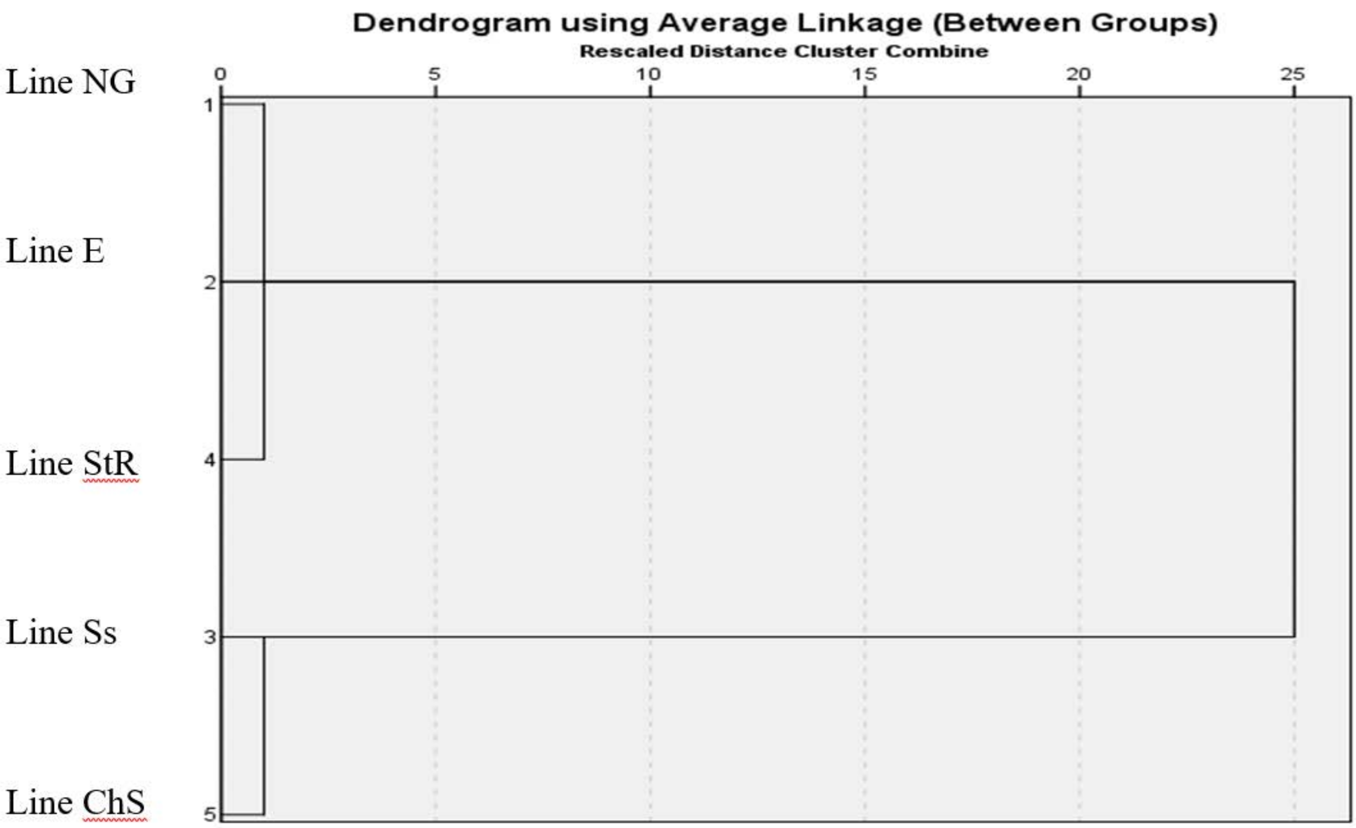

Figure 7. Dendrogram of Dual-purpose-type hens, second group of indicators

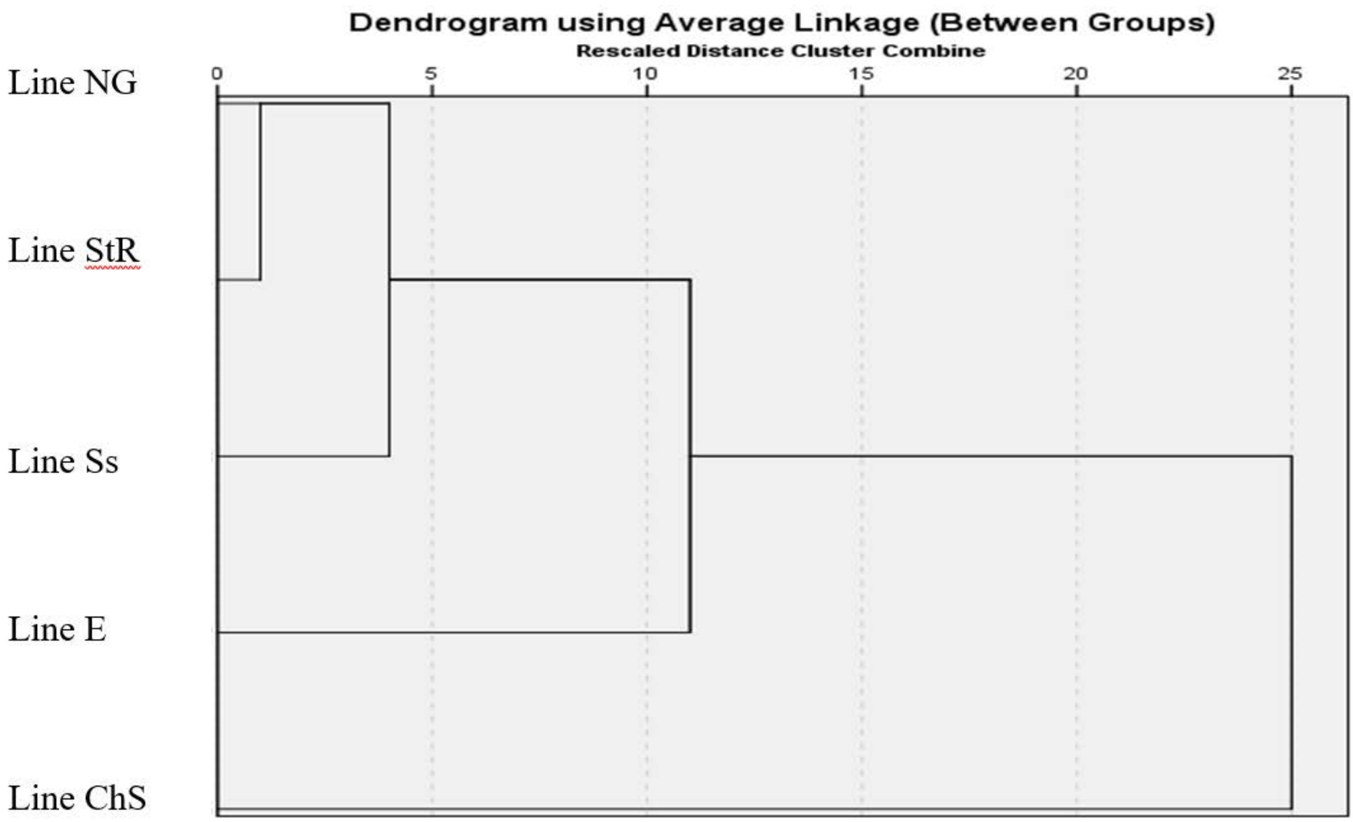

Figure 8. Dendrogram of Dual-purpose-type hens, third group of indicators

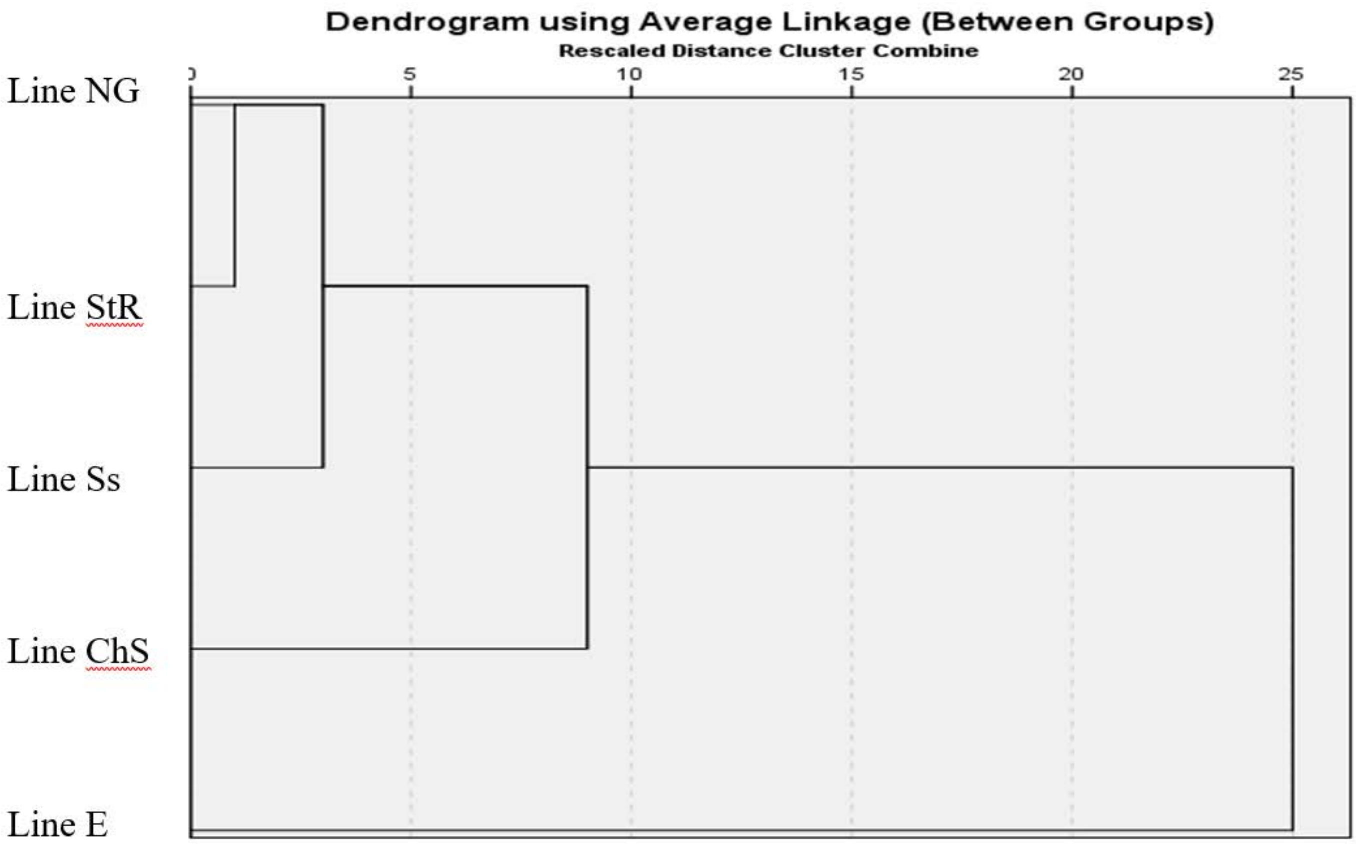

Figure 9. Dendrogram of Dual-purpose-type hens, fourth group of indicators

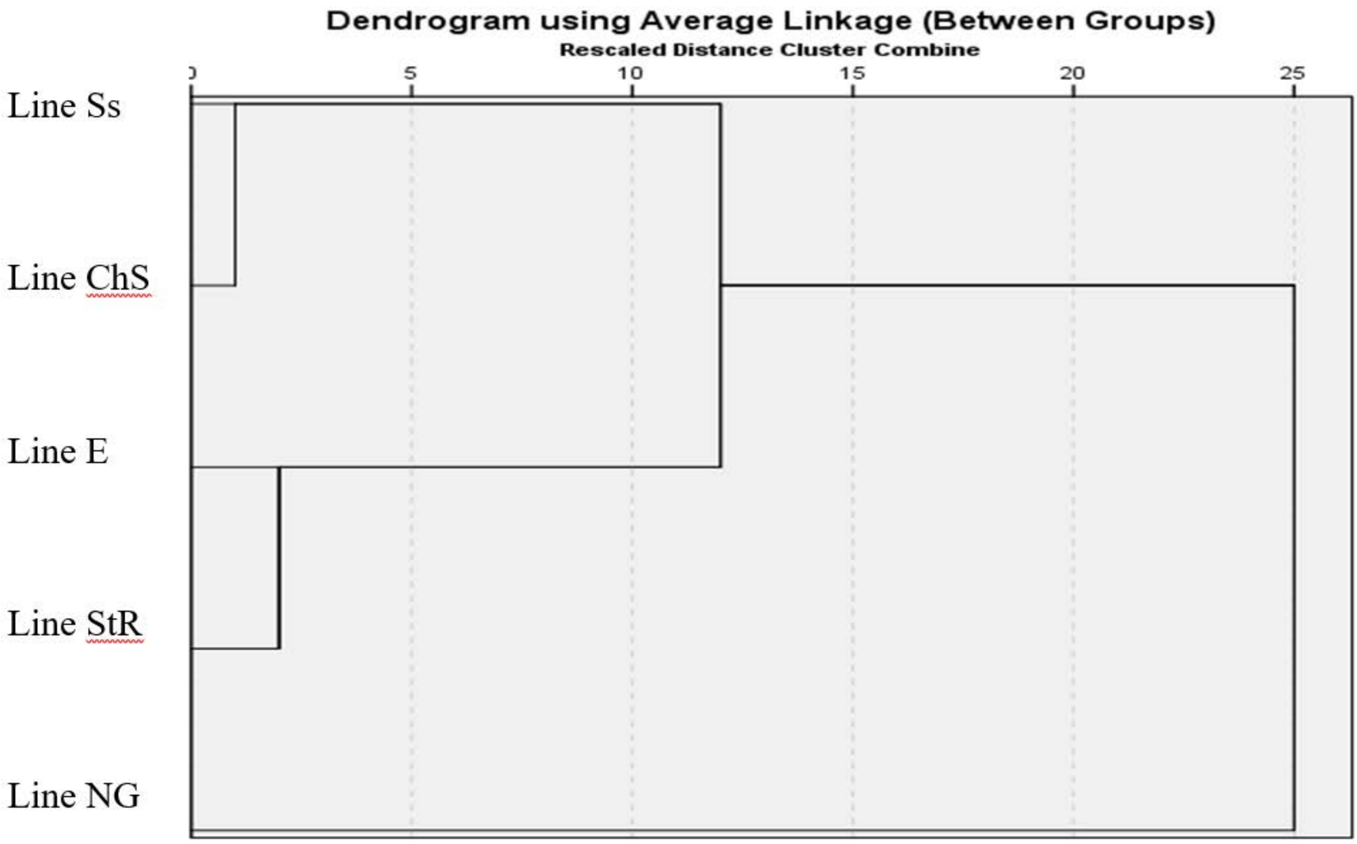

Figure 10. Dendrogram of Dual-purpose-type hens, fifth group of indicators

CONCLUSIONS

After applying a cluster analysis with respect to the lines of the Egg-type hens, it was found that according to the indicators from the first, third and fourth group, line D is the farthest away from the other lines studied. Among the lines of Dual-purpose-type hens it was found in most of the clustering, that line ChS forms independent clusters joining the rest at a large Euclidean distance, proving its differences with the other lines of that group. The results of this study could be taken into account in future scientific developments for the purpose of selecting individual lines of hens according to the relevant indicators.

REFERENCES

Abramowitz, S., Weinberg, S. (2016) Statistics using IBM SPSS, An integrative approach. Cambridge: Cambridge University Press.

Barbosa, L., Regazzi, A., Lopes, P., Breda, F., Sarmento, J., Torres, R., Torres, P. (2005) Evaluation of genetic divergence among lines of laying hens using cluster analysis. Revista Brasiliera de Ciencia Avicola, 7 (2). Available at: http://www.scielo.br/scielo.php?pid=S1516-635X2005000200003&script=sci_arttext&tlng=pt

Chen, G., Wu, X., Wang, D., Qin, J., Wu, S., Zhou, Q., Xie, F., Cheng, R., Xu, Q., Liu, B., Zhang, X., Olowofeso, O. (2004) Cluster analysis of 12 Chinese native chicken populations using microsatellite markers. Asian-Australasian Journal of Animal Sciences, 17 (8), 1047-1052. DOI: https://dx.doi.org/10.5713/ajas.2004.1047

Corrigan, A., de Leeuwb, M., Penaud-Frézetc, S., Dimova, D., Murphy, R.A. (2015) Phylogenetic and functional alterations in bacterial community compositions in broiler ceca as a result of mannan oligosaccharide supplementation. Applied and Environmental Microbiology, 81 (10), 3460-3470. DOI: https://dx.doi.org/10.1128/AEM.04194-14

Cronk, B. (2016) How to use SPSS: A step-by-step guide to analysis and interpretation. New York: Taylor & Francis Group.

De Marchi, M., Dalvit, C., Targhetta, C., Cassandro, M. (2006) Assessing genetic diversity in indigenous Veneto chicken breeds using AFLP markers. Animal genetics, 37 (2), 101-105. DOI: https://dx.doi.org/10.1111/j.1365-2052.2005.01390.x

Granevitze, Z., Hillel, J., Chen, G., Cuc, N., Feldman, M., Eding, H., Weigend, M. (2007) Genetic diversity within chicken populations from different continents and management histories. Animal genetics, 38 (6), 576-583. DOI: https://dx.doi.org/10.1111/j.1365-2052.2007.01650.x

Kaya, M., Yildiz, M. (2008) Genetic diversity among Turkish native chickens, Denizli and Gerze, estimated by microsatellite markers. Biochemical Genetics, 46 (7-8), 480-491. DOI: https://dx.doi.org/10.1007/s10528-008-9164-8

Lalev, M., Oblakova, M., Mincheva, N., Hristakieva, P., Ivanova, I. (2012) Evaluation of productive traits of chicken lines from the national gene pool. Trakia Journal of Sciences, 10 (10), 38-42.

Meyers, L., Gamst, G., Guarino, A. (2013) Performing data analysis using IBM SPSS. United States of America: John Wiley & Sons.

Rosenberg, N., Burke, T., Elo, K., Feldman, M., Freidlin, P., Groenen, M., Hillel, J., Mäki-Tanila, A., Tixier-Boichard, M., Vignal, A., Wimmers, K., Weigend, S. (2001) Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. Genetics, 159 (2), 699-713.

Savegnano, R., Caettano, S., Ramos, S., Nascimento, G., Schmidt, G., Ledur, M., Murani, D. (2011) Estimates of genetic parameters, and cluster and principal components analyses of breeding values related to egg production traits in a White Leghorn population. Poultry Science, 90 (10), 2174-2188. DOI: https://dx.doi.org/10.3382/ps.2011-01474

Tadano, R., Nakamura, A., Kino, K. (2012) Analysis of genetic divergence between closely related lines of chickens. Poultry Science, 91 (2), 327-333. DOI: https://dx.doi.org/10.3382/ps.2011-01879

Vanhala, T., Tuiskula-Haavisto, M., Elo, K., Vilkki, J., Maki-Tanila, A. (1998) Evaluation of genetic variability and genetic distances between eight chicken lines using microsatellite markers. Poultry Science, 77 (6), 783-790. DOI: https://dx.doi.org/10.1093/ps/77.6.783

Yang, K., Lin, C., Huang, H., Liou, J., Wu, C., Huang, C., Ou, B., Chen, C., Lee, Y., Lin, E., Tang, P., Ding, S., Cheng, W., Huang, M. (2008) Expressed transcripts associated with high rates of egg production in chicken ovarian follicles. Molecular and Cellular Probes, 22 (1), 47-54. DOI: https://dx.doi.org/10.1016/j.mcp.2007.06.001