J. Agr. Sci. Tech. (2019) Vol. 21(Suppl.): 1783-1792

Application of Classification Tree Method to Determine Factors Affecting Somatic Cell Count in Holstein Cows

H. Ghiasi[1], A. Sadeghi-Sefidmazgi[2], R. Taherkhani[1], M. Khaldari[3], D. Piwczyński[4], and M. Kolenda[4*]

ABSTRACT

In the current study, the effect of phenotypic factors on Somatic Cell Count (SCC) was evaluated by using classification tree technique. The current study used a total of 1,972,031 test day records of *SCC* in parity 1 to 4 collected from 1,281 Iranian Holstein-Friesian cows' herds through 2004–2013. The SCC records were converted to binary trait, defined as 1: If *SCC*≤ 200,000, and 2: Otherwise. The CART (Classification And Regression Tree) algorithm for classification trees, with *GINI* index and Entropy function as the division criteria, was used to develop the tree. Statistical analysis was performed using 'rpart' package in R software. The constructed tree had 12 leaves and it was 6 levels deep. The results of classification tree procedure for ranking of importance of the variables responsible for the variation in *SCC* were, respectively, parity, test-day milk production, year of calving, season of calving, and days in milk (stage of lactation). Based on the obtained classification tree, different combination of variables associated with *SCC* could be identified. According to the classification tree, the lowest amount of *SCC* was expected in the group of cows that were in the 1st or 2nd parity; their test-day milk production was > 30 kg; they were calved from year 2004 to 2013; and their calving season was autumn or winter.

Keywords: Entropy function, *GINI* index, Holstein-Friesian cows, Mastitis, Milk production.

INTRODUCTION

To have a hygienic milk production and to increase the profitability, it is important to increase the cows' udder health (Sargeant *et al*., 1998). One of the most common diseases that affect dairy cows' udder health is mastitis (Sordillo *et al*., 1997). Several economic losses that are associated with mastitis include reduction in the milk production, change in the milk compositions (Beck *et al*., 1992, Harmon, 1994), reduction in the profitability (Sadeghi-Sefidmazgi *et al*., 2011), and increase of the *SCC* (Sharma *et al*., 2011). Additionally, Østergaard *et al*. (2005) stated that elimination of mastitis in dairy cowherd would increase a net return of 146 Euro per cow per year.

Somatic cells (mainly blood cells) that are present in milk are produced by immune system to combat with infection in udder (Norman *et al*., 2011). The Somatic Cell Counts (SCC) in milk is an indicator of udder health, which is widely used to detect clinical and subclinical mastitis (Detilleux *et al*., 1997). The *SCC* is also an indicator for milk quality (Ma *et al*., 2000). As mentioned

[1] Department of Animal Science, Faculty of Agricultural Science, Payame Noor University, Tehran, Islamic Republic of Iran.
[2] Department of Animal Science, College of Agriculture, Isfahan University of Technology, Isfahan, Islamic Republic of Iran.
[3] Department of Animal Science, Faculty of Agricultural Science, Lorestan University, Khorramabad, 68159 Lorestan, Islamic Republic of Iran.
[4] Department of Biotechnology and Animal Genetics, Faculty of Biology and Animal Breeding, UTP University of Science and Technology in Bydgoszcz, Poland.
*Corresponding author: e-mail: kolenda@utp.edu.pl

by Norman *et al*., (2011), the legal maximum amount of *SCC* in milk varies across countries (e.g., 1,000,000 cells mL^{-1} in Brazil, 750,000 cells mL^{-1} in US; 500,000 cells mL^{-1} in Canada; 400,000 cells mL^{-1} in much of Europe, New Zealand and Australia). Through reducing somatic cell count, several benefits such as increased milk quality and decreased mastitis would be achieved. There is a correlation between *SCC* with milk quality (Ma *et al*., 2000) and mastitis (Ǿdegard *et al*., 2003; Vallimont *et al*., 2009). Decreased *SCC* in milk would lead to an increase in coagulating properties and cheese yield, and a decrease in loss of fat and casein in whey. It would also result in keeping the quality of milk, and increasing milk shelf-life (Ma *et al*., 2000). Schukken *et al*. (1992) concluded that reduction of *SCC* would increase the fat and lactose percentage. Nowadays, milk processors are implementing the premium quality payment programs for milk using low *SCC*.

From the herd management perspective, the *SCC* in milk can be reduced based on two methods. The first method, or short-term solution, is cows' culling and the second method, or long-term solution, is reducing mastitis in herd by genetic selection and hygiene (Looper, 2012). Since *SCC* has a high genetic correlation with mastitis i.e. ranging from 0.53 to 0.91, it has been utilized to increase udder health through genetic selection (Ǿdegard *et al*., 2004; Vallimont *et al*., 2009). Factors influencing hygienic practices are the other influential factors that can reduce *SCC*. Several other factors such as parity (Skrzypek *et al*., 2004; Cengiz *et al*., 2015), stage of lactation (Tančin, 2013; Koc and Kizilkaya, 2009), season of calving (Singh and Ludri, 2001; Green *et al*., 2006), and year of calving (Ødegard *et al*., 2003; Faraji-Arugh *et al*., 2012) have been reported to affect *SCC*. Complex interaction might be one of the factors influencing *SCC*. To help milk producers to produce milk with low *SCC*, it is crucial to know the combination of factors that affect *SCC*. This can also help producers to group total milk of herd according to *SCC*. One of the suitable statistical methods to identify the group of cows producing low and high level of *SCC* is classification tree. Piwczyński and Sitkowska (2012) used classification tree for statistical modelling of *SCC* in Polish Holstein cows. They concluded that complex interactions between several phenotypic factors affected *SCC*.

Our study aimed to use classification tree method to evaluate the impact of various phenotypic factors on *SCC* in Iranian Holstein cows to identify in which group of cows *SCC* level would be low or high.

MATERIALS AND METHODS

This study was conducted on 1,972,031 test day records of *SCC*, coming from 78,881 cows kept in 1,281 Iranian Holstein cow's herds. Data was collected by the Animal Breeding Center of Iran, between 2004 and 2013 throughout the entire country. The following rules were used to edit the data: (1) Cows with age at first calving before 18 and after 36 months were deleted from data file; (2) Only records gathered between days 5 and 400 after calving were used; (3) *SCC* lower than 5,000 cells mL^{-1} and higher than 6,000,000 cells mL^{-1} was omitted from data; and (4) The *SCC* records were defined as binary trait as '1': If *SCC*<= 200,000, and '2': Otherwise.

The following factors that could have an effect on *SCC* were considered in statistical analysis: Parity (1 to 4 levels), year of calving (2004 to 2013), season of calving (1: Spring, 2: Summer, 3: Autumn, 4: Winter), the amount of test-day milk production (1: <= 15 kg, 2: > 15 kg and <= 30 kg, 3: > 30 kg and <= 45 kg, 4: > 45 kg), days (d) in milk (stage of lactation) (1: <= 35 d, 2: > 35 d and <= 100 d, 3: > 100 d and <= 200 d, and 4: > 200 d). The effect of phenotypic factors on SCC was conducted by using classification tree analysis, also known as Classification And Regression Trees (CART). This method is a powerful and

[DOR: 20.1001.1.16807073.2019.21.7.3.4]

popular predictive machine learning technique that is used for both classification and regression. Therefore, in the current study, the CART algorithm with *GINI* index [1] and Entropy function (ENTROPY, [2]) as division criteria were used in creating the classification tree.

$$\text{GINI index} = 1 - \sum_{i=1}^{k} p_j^2 \quad (1)$$

$$\text{ENTROPY} = -\sum_{i=1}^{k} p_j log_2 (p_j) \quad (2)$$

Where, p_1, p_2,...p_k is probability vector of object assignment to classes and k= Number of class.

Statistical analysis and creating the classification tree was conducted using rpart and rpart plot packages in R software. Tree was built with the following constraints: it was assumed that the minimum size of the final node should not be less than 20 observations, the depth of tree (the length of the longest path from a root to a leaf) no higher than 6, and split must decrease the overall lack of fit of the model by a factor of cost-complexity parameter (cp). When the tree is constructed, the cross-validation techniques were used for further pruning it. In this method, data was divided to 10 subsets, 9 set for "learning samples" to create tree and 1 set for "test samples" to estimate cross validation error. The cross-validation error was used to prune the tree using the '1-SE' rule (Breiman *et al*., 1984) and then the corresponding *cp* value was used to optimally pruned tree.

RESULTS

The distribution of *SCC* according to phenotypic factors is shown in Table 1. All the investigated factors have a statistically significant effect on variation of *SCC* (P< 0.01). As illustrated in Table 1, most of the animals with *SCC*< 200,000 were observed in the group of cows that were in 1st parity; Test-day milk production> 45 kg; calved in 2010; calved in autumn, and days in milk < 36.

Table 1. The distribution of *SCC* according to tested factors.

Factor	Level	% of *SCC*<= 200000
Parity	1	58.6
	2	53.0
	3	47.4
	4	43.3
Test-day milk production (kg)	<=15	36.7
	> 15 and <= 30	49.5
	> 30 and <= 45	55.3
	> 45	56.5
Year of calving	2004	42.1
	2005	49.6
	2006	50.6
	2007	50.0
	2008	52.1
	2009	54.0
	2010	55.6
	2011	54.6
	2012	54.0
	2013	53.0
Season of calving	Spring	51.6
	Summer	53.4
	Autumn	54.4
	Winter	52.5
Days in milk (Days)	≤ 35	55.4
	> 35 and ≤ 100	52.7
	> 100and ≤ 200	52.7
	> 200	52.5

The rankings of variable importance are illustrated in Table 2. All the importance measures were scaled to maximum value of 100. Values shown in the “importance” column in Table 2 are the crucial factors in reducing node impurity. Accordingly, parity is the most important factor for predicting *SCC*, followed by test-day milk production, year of calving, season of calving, and days in milk.

The graphical model of constructed classification tree is depicted in Figure 1. To avoid over-fitting of the data, the tree was pruned to create an optimal classification tree. The cross-validation error and its standard error were 0.915 and 0.0007,

Table 2. Importance variable.

Variable	Importance
Parity	52
Test-day milk production	29
Year of calving	12
Season of calving	5
Days in milk	2

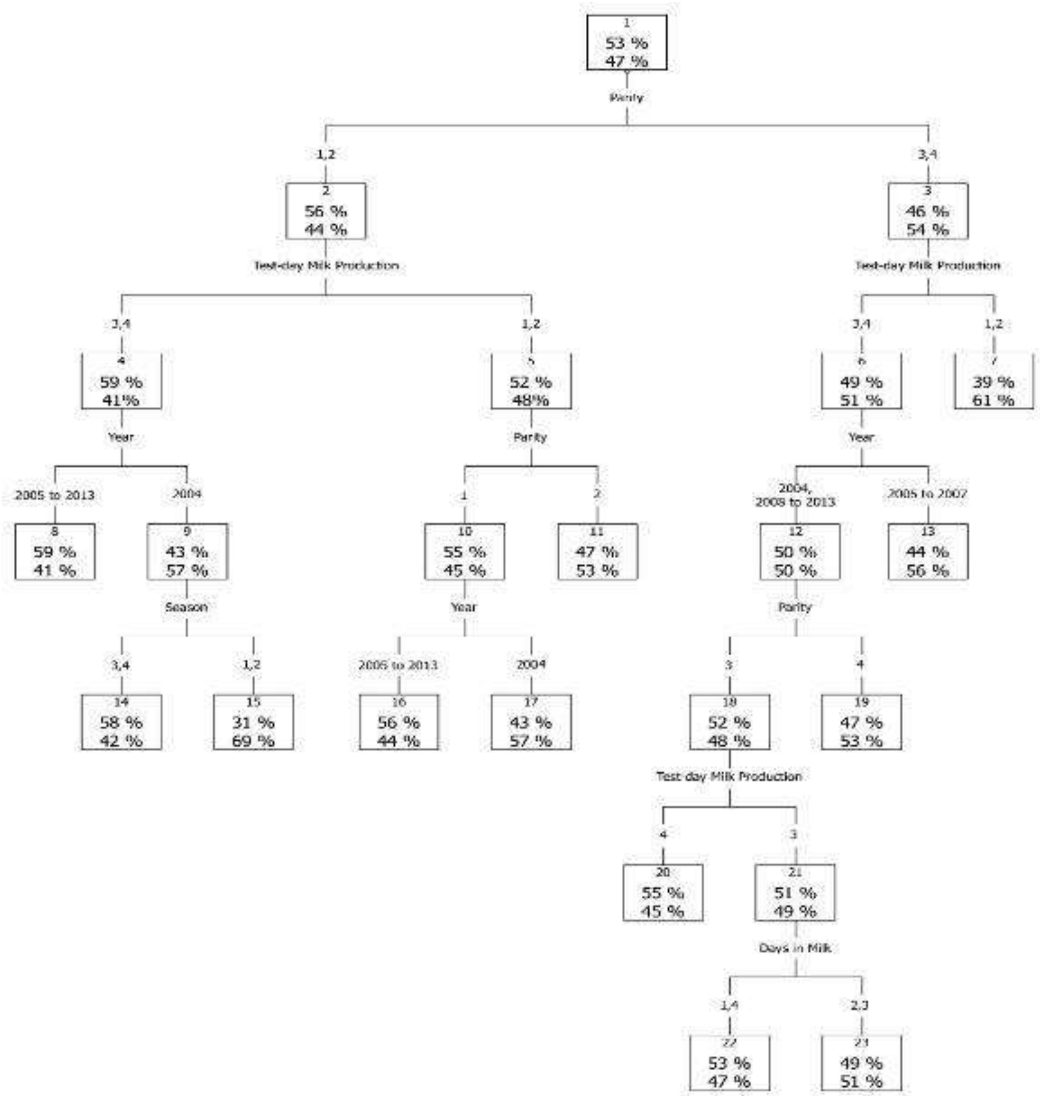

Figure 1. The classification tree, where: Season (season of calving) – 1: Spring, 2: Summer, 3: Autumn, 4: Winter; Test-day Milk Production (the amount of test-day milk production, kg – 1: <= 15, 2: > 15 and <= 30, 3: > 30 and <= 45, 4: > 45); Days in milk (stage of lactation) – 1: <= 35, 2: > 35 and <= 100, 3: > 100 and <= 200, 4 : > 200.

[DOR: 20.1001.1.16807073.2019.21.7.3.4]

respectively. According to the '1-SE' rule, the cross-validation error to identify the point to prune the tree was 0.915+(1×0.0007)= 0.9157. The corresponding *cp* (0.001) for 0.9157 was applied to prune the tree (Figure 3). The resulting tree had 12 leaves and it was six levels deep. To construct the tree, the greatest division was carried out based on parity, test-day milk production, and year of calving (3 divisions). The lowest division, however, was conducted based on the season of calving and days in milk (1 division). Each node and leaf presented in the classification tree, has the following information: (1) The ID of a node, (2) The percentage of *SCC*<= 200,000 (on the left), and the percentage of *SCC*>200,000 (on the right) (Figure 2).

The most important variable affecting *SCC* was parity (Table 2). In the classification tree, the first division occurred based on parity (Figure 1). Concerning the parity, *SCC* was divided in two branches: Parity< 3 (node 2) and Parity≥ 3 (node 3). In node 2, *SCC*<= 200,000 was 10 percent higher than node 3 (56 *vs* 46%).

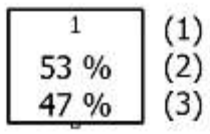

Figure 2. Description of node 1 (root node), where (1) The ID of a node, (2) The percentage of *SCC*<= 200,000 (on the left), and the percentage of *SCC*>200,000 (on the right).

Nodes 2 and 3 were split according to the test-day milk production. Division node 2 created two branches: Cows with test-day milk production>= 30 kg (node 4) and cows with test-day milk production lower than 30 kg (node 5). Percentage of *SCC*<= 200,000 in nodes 4 and 5 was 59 and 52%, respectively. Node 3 was divided in node 6 (Cows with test-day milk production>= 30 kg) and node 7 (Cows with test-day milk production< 30 kg). In node 6, *SCC*<= 200,000 was about 10 percent higher than node 7 (49 *vs* 39%).

Node 7 became the leaf. Node 6 was further branched into nodes 12 and 13 based on the year of calving. Cows calved in 2004, 2008 till 2013 were assigned to node 12 (Percentage of SCC<= 200,000= 50%). Moreover, cows calved from the year 2005 to 2007 were assigned to node 13 (Percentage of SCC<= 200,000= 44%). Node 13 became the leaf, while node 12 was further divided according to parity. Cows with Parity= 3 were assigned to node 18 (Percentage of SCC<= 200,000= 52%) and cows with Parity= 4 were assigned to node 19 (Percentage of SCC<= 200,000= 47%). Node 19 became the leaf. Node 18 created nodes 20 and 21 based on the test-day milk production. Cows with test-day milk production > 30 kg and <= 45 kg belonged to node 21 (Percentage of SCC<= 200,000= 51%), whereas cows with test-day milk production > 45 kg were assigned to node 20 (Percentage of SCC<= 200,000= 55%). Node 20 became the leaf, while node 21 was further divided based on the stage of lactation or days in milk. Cows with days in milk <=35 d and >200 d were assigned to node 22 (Percentage of SCC<= 200,000= 53%). Whereas cows with Days in milk>35 or <=200 were assigned to node 23 (Percentage of SCC<= 200,000= 49%). Both nodes 21 and 22 became the leaves.

Node 4 was divided based on the year of calving. Thus, nodes 8 and 9 were created. Node 8 represented cows' year of calving from 2005 to 2013. The percentage of *SCC*<= 200,000 was 59% in this group. Node 9 included cows calved in the year 2004. The percentage of *SCC*<= 200,000 in this group was 43%.

Node 8 became the leaf, whereas node 9 was branched further according to the season of calving. This resulted in nodes 14 (Percentage of SCC<= 200,000 in cows calved in autumn and winter was 58%) and node 15 (Percentage of SCC<= 200,000 in cows calved in spring and summer was 31%). Nodes 14 and 15 became the leaves.

[DOR: 20.1001.1.16807073.2019.21.7.3.4]

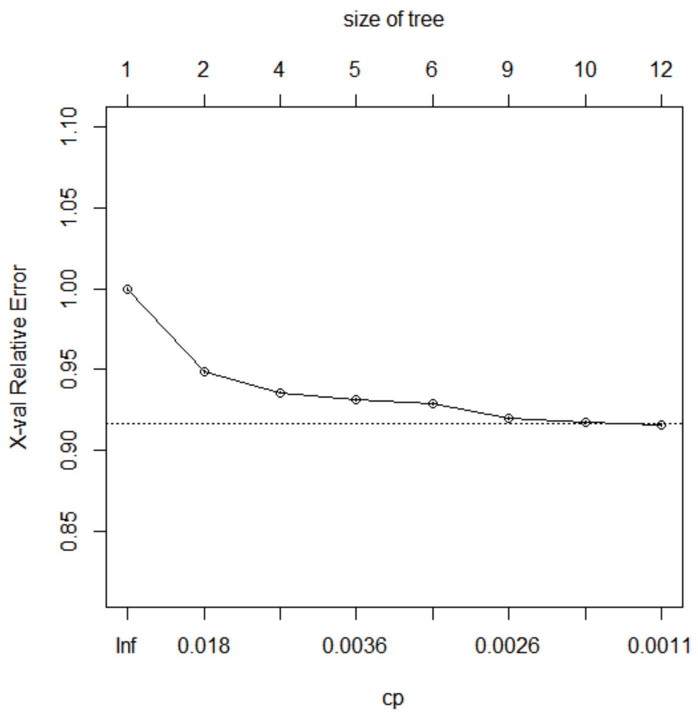

Figure 3. Cost-complexity parameter (cp), the cross-validation error (X-val Relative Error) and tree size i.e., number of nodes, where, *Inf* stands for Infinity.

Node 5 was divided further according to parity and created nodes 10 and 11. Cows in parity 1 were assigned to node 10 (Percentage of SCC<= 200,000 was 55%), while cows in parity 2 were allocated to node 11 (Percentage of SCC<= 200,000 was 47%). Node 11 became the leaf, whereas node 10 was further branched based on the year of calving. Therefore, it resulted in nodes 16 and 17. Cows calved in 2004 were assigned to node 17 (Percentage of SCC<= 200,000 was 43%) and cows calving in 2005 and later were assigned to node 16 (Percentage of SCC<= 200,000 was 56%).

DISCUSSION

The results of descriptive analysis in the current study were inconsistent with the results reported by Piwczyński and Sitkowska (2012) in Polish Holstein cows, except for season of calving. Piwczyński and Sitkowska (2012) showed that cows calved in summer had the lowest level of *SCC*. However, in the current study, the results indicated that animals calved in autumn had the lowest amount of *SCC*. In this study, percentage of records with *SCC* lower than 200,000 gradually decreased from early lactation to the end of lactation.

One of the features of classification tree is calculation of “variable importance”. The variables showed in the classification tree can be considered as deemed important. However, the variables that are not included in the classification tree are not necessarily unimportant since their effect might be masked by other correlated factors (Breiman *et al*., 1984). According to the results of current study, parity is the most important factor affecting SCC. The other important factors are test-day milk production, year of calving, season of calving, and days in milk.

The results revealed that *SCC* could be increased with the advancement of parity order, which is in line with the results obtained by Skrzypek *et al*. (2004) and

Cengiz *et al*. (2015). Nevertheless, some prior studies demonstrated that parity does not affect *SCC* (Singh and Ludri, 2001). Somatic cell count increases during the mastitis infection. According to Hagnestam *et al*. (2007), mastitis occurs more frequently in multiparous cows compared to primiparous cows.

The result of this study showed that the amount of *SCC* in high producing cows was lower than low producing cows. It can be due to genetic selection for increasing milk production and reducing *SCC* in Iranian Holstein population. The genetic correlation between *SCC* and milk production is unfavourable for the first lactation cows (Koivula *et al*., 2005; Boettcher *et al*., 1992; Banos and Shook, 1990). However, a favourable negative genetic correlation has been reported between milk production and somatic cell count for the second and later parities (Banos and Shook, 1990). Additionally, a negative or almost zero phonotypic correlation has been reported between daily somatic cell count and milk yield in the first and second lactation (Yamazaki *et al*., 2013). This indicated that selection for lower SCC might not affect milk production in later lactations and that lower level of SCC may be expected in milk of high producing cows.

The results of other studies indicated that *SCC* varies during the lactation stage. The lowest amount of SCC could be expected in early (Days in milk<= 35) and the subsequent stages of lactation (Days in milk> 35) which is in contrast with the results obtained by Tančin (2013). In particular, the mentioned study reported that there was a linear increase in *SCC* throughout lactation and the highest amount of *SCC* was observed in the later stages of lactation. Besides, Monardes *et al*. (1983) observed the highest level of *SCC* shortly after calving. It was then rapidly declining to a minimum level in cows with 25 to 45 days in milk. It was also shown that the amount of *SCC* decreased until the third month of lactation and then fluctuated until the end of lactation (Koc and Kizilkaya, 2009).

According to results of this study, after 2005, the amount of *SCC* in the investigated population decreased with time. The decline of *SCC* over the calving years might be due to genetic selection to decrease *SCC*. Favourable genetic and phonotypic trend for *SCC* have been reported in Iranian Holstein cows (Faraji-Arugh *et al*., 2012) as well as in Holstein cows in other countries (Ødegard *et al*., 2003; Pagnacco *et al*., 1994).

This study revealed that animals calved in autumn and winter had lower amount of *SCC* than the ones calved in spring and summer. This is in agreement with the results obtained by Singh and Ludri (2001) and Green *et al*. (2006). Accordingly, the high amount of *SCC* during spring and summer can be because of heat, stress, and high risk of clinical mastitis in these seasons (De Vliegher *et al*., 2004).

The classification tree constructed in this study reveals that somatic cell count is diversified by parity, test-day milk production, year of calving, season of calving, and days in milk. Study results showed that the low level of somatic cell count can be expected in the group of cows that were in Parity< 3, were calved in Year>= 2005, with Test-day milk production> 30 kg, and their calving season was in autumn or winter, and were in early (Days in milk<= 35) or subsequent stages of lactation (Days in milk> 35).

REFERENCES

1. Banos, G. and Shook, G. E. 1990. Genotype by Environment Interaction and Genetic Correlations among Parities for Somatic Cell Count and Milk Yield. *J. Dairy Sci.*, **73**: 2563–2573.
2. Beck, H. S., Wise, W. S. and Dodd F. H. 1992. Cost Benefit Analysis of Bovine Mastitis in the UK. *J. Dairy Res.*, **59:** 449–460.
3. Boettcher, P. J., Hansen, L. B. and Ernst, C. A. 1992. Genetic Evaluation of Holstein Bulls for Somatic Cells in Milk of Daughters. *J. Dairy Sci.*, **75:** 1127–1137.

4. Breiman, L., Friedman, J., Stone, C. J. and Olshen, R. A. 1984. *Classification and Regression Trees*. CRC Press, Chapman and Hall, New York, USA.
5. Cengiz, M., Kaynar, O., Cannazik, O., Ileriturk, M., Cengiz, S. and Hayirli, A. 2015. Sampling Factors Causing Variability in Milk Constituents in Early Lactation Cows. *Vet. Med-Czech*, **60:** 6–15.
6. De Vliegher, S., Barkema, H. W., Stryhn, H., Opsomer, G. and de Kruif, A. 2004. Impact of Early Lactation Somatic Cell Count in Heifers on Somatic Cell Counts over the First Lactation. *J. Dairy Sci.*, **87:** 3672–3682.
7. Detilleux, J., Leroy, P. and Volckaert, D. 1997. Alternative Use of Somatic Cell Counts in Genetic Selection for Mastitis Resistance, *Interbull Bull.,* **15:** 34–44.
8. Faraji-Arugh, H., Aslami-Nejhad, A. A. and Rokuyi, M. 2012. The Genetic and Phenotypic Trend and Environmental Factor Affecting Somatic Cell Score in Cattle Iranian Holstein, *Iran. J. Anim. Sci. Res.,* **3:** 459–454. (in Farsi)
9. Green, M. J., Bradley, A. J., Newton, H. and Browne, W. J. 2006. Seasonal Variation of Bulk Milk Somatic Cell Counts in UK Dairy Herds: Investigations of the Summer Rise. *Prev. Vet. Med.*, **74**: 293–308.
10. Hagnestam, C., Emanuelson, U. and Berglund, B. 2007. Yield Losses Associated with Clinical Mastitis Occurring in Different Weeks of Lactation. *J. Dairy Sci.*, **90:** 2260–2270.
11. Harmon, R. J. 1994. Physiology of Mastitis and Factors Affecting Somatic Cell Counts. *J. Dairy Sci.*, **77:** 2103–2112.
12. Koc, A. and Kizilkaya, K. 2009. Some Factors Influencing Milk Somatic Cell Count of Holstein Friesian and Brown Swiss Cows under the Mediterranean Climatic Conditions. *Arch. Tierz.*, **52:** 124–133.
13. Koivula, M., Mäntysaari, E. A., Negussie, E. and Serenius, T. 2005. Genetic and Phenotypic Relationships among Milk Yield and Somatic Cell Count before and after Clinical Mastitis. *J. Dairy Sci.*, **88:** 827–833.
14. Looper, M. 2012. *Reducing Somatic Cell Count in Dairy Cattle.* Division of Agriculture Research and Extension, University of Arkansas System, FSA4002, 2012, Available at: http://www.uaex.edu/publications /pdf /fsa-4002.pdf
15. Ma, Y., Ryan, C., Barbano, D. M., Galton, D. M., Rudan, M. A. and Boor, K. J. 2000. Effects of Somatic Cell Count on Quality and Shelf-Life of Pasteurized Fluid Milk. *J. Dairy Sci.*, **83:** 264–274.
16. Monardes, H. G., Kennedy, B. W. and Moxley, J. E. 1983. Heritabilities of Measures of Somatic Cell Count per Lactation. *J. Dairy Sci.*, **66:** 1707–1713.
17. Norman, H. D., Lombard, J. E., Wright, J. R., Kopral, C. A., Rodriguez, J. M. and Miller, R. H. 2011. Consequence of Alternative Standards for Bulk Tank Somatic Cell Count of Dairy Herds in the United States. *J. Dairy Sci.*, **94:** 6243–6256.
18. Ødegård, J., Heringstad, B. and Klemetsdal, G. 2004. Short Communication: Bivariate Genetic Analysis of Clinical Mastitis and Somatic Cell Count in Norwegian Dairy Cattle. *J. Dairy Sci.*, **87:** 3515–3517.
19. Ødegard, J., Klemetsdal, G. and Heringstad, B. 2003. Variance Components and Genetic Trend for Somatic Cell Count in Norwegian Cattle. *Livest. Prod. Sci.*, **79:** 135–144.
20. Østergaard, S., Chagunda, M. G. G., Friggens, N. C., Bennedsgaard, T. W. and Klaas, I. C. 2005. A Stochastic Model Simulating Pathogen-Specific Mastitis Control in a Dairy Herd. *J. Dairy Sci.*, **88:** 4243–4257.
21. Pagnacco, G., Miglior, F., Zhang, W. C., Dekkers, J. C. M. and Burnside E. B. 1994. Genetic Evaluation for Somatic Cell Count and Relationship with Inbreeding in Canadian Holstein. *In Proceedings of the World Congress on Genetics Applied to Livestock Production*, 7–12 August, Guelph, Canada, 93 PP.
22. Piwczyński, D. and Sitkowska, B. 2012. Statistical Modelling of Somatic Cell Counts Using the Classification Tree Technique, *Arch. Tierz.*, **55:** 332–345.
23. Sadeghi-Sefidmazgi, A., Moradi-Shahrbabak, M., Nejati-Javaremi, A., Miraei-Ashtiani, S. R. and Amer, P. R. 2011. Estimation of Economic Values and Financial Losses Associated with Clinical Mastitis and Somatic Cell Score in Holstein Dairy Cattle. *Animal*, **5:** 33–42.
24. Sargeant, J. M., Schukken, Y. H. and Leslie, K. E. 1998. Ontario Bulk Milk Somatic Cell Count Reduction Program: Progress and Outlook. *J. Dairy Sci.*, **81:** 1545–1554.
25. Schukken, Y. H., Leslie, K. E., Weersink, A. and Martin, S. W. 1992. Ontario Bulk Milk

Somatic Cell Count Reduction Program. 1. Impact on Somatic Cell Counts and Milk Quality. *J. Dairy Sci.*, **75:** 3352–3358.
26. Sharma, N., Singh, N. K. and Bhadwal, M. S. 2011. Relationship of Somatic Cell Count and Mastitis: An Overview. *Asian Austral. J Anim.*, **24:** 429–438.
27. Singh, M. and Ludri, R. S. 2001. Influence of Stages of Lactation, Parity and Season on Somatic Cell Counts in Cows. *Asian Austral. J Anim.*, **14:** 1775–1780.
28. Skrzypek, R., Wojtowski, J. and Fahr, R. D. 2004. Factors Affecting Somatic Cell Count in Cow Bulk Tank Milk: A Case Study from Poland. *J. Vet. Med. A*, **51:** 127–131.
29. Sordillo, L. M., Shafer-Weaver, K. and DeRosa, D. 1997. Immunobiology of the Mammary Dland. *J. Dairy Sci.*, **80:** 1851–1865.
30. Tančin, V. 2013. Somatic Cell Counts in Milk of Dairy Cows under Practical Conditions. *Slovak J. Anim. Sci.*, **46:** 31–34.
31. Vallimont, J. E., Dechow, C. D., Sattler, C. G. and Clay, J. S. 2009. Heritability Estimates Associated with Alternative Definitions of Mastitis and Correlations with Somatic Cell Score and Yield. *J. Dairy Sci.*, **92**: 3402–3410.
32. Yamazaki, T., Hagiya, K., Takeda, H., Sasaki, O., Yamaguchi, S., Sogabe, M., Saito, Y., Nakagawa, S., Togashi, K., Suzuki, K. and Nagamine, Y. 2013. Genetic Correlations between Milk Production Traits and Somatic Cell Scores on Test Day within and across First and Second Lactations in Holstein Cows. *Livest. Sci.*, **152:** 120–126.

کاربرد درخت تصمیم گیری جهت تعیین فاکتورها مؤثر بر شمار سلول های بدنی در گاوهای هلشتاین ایران

ح. قیاسی، ع. صادقی سفید مزگی، ر. طاهرخانی، م. خالداری، د. پیوزنسکی، و م. کولندا

چکیده

در این مطالعه اثر فاکتورهای محیطی بر شمار سلول های بدنی (SCC) با استفاده از تکنیک درخت تصمیم گیری مورد ارزیابی قرار گرفت. در این مطالعه ۱۹۷۲۰۳۱ رکورد روز آزمون SCC که در شکم های ۱ تا ۴ در ۱۲۸۱ گله پرورش دهنده گاو هلشتاین ایران که طی سالهای ۱۳۸۳ تا ۱۳۹۲ جمع آوری شده بود، مورد استفاده قرار گرفت. رکورد های SCC به صورت باینری به صورت زیر تبدیل گردیدند : اگر SCC کمتر از ۲۰۰۰۰۰ بود عدد ۱ در غیر این صورت عدد ۲ به آن اطلاق گردید. از الگوریتم CART(classification and regression tree) و شاخص GINI و تابع آنتروپی به عنوان معیار طبقه بندی و ایجاد درخت استفاده شد. آنالیزهای آماری و ایجاد درخت با استفاده از بسته نرم افزاری rpart در محیط نرم افزاری R انجام گرفت. درخت حاصل شده دارای ۶ سطح و ۱۲ برگ بود. بر اساس نتایج حاصل شده از درخت تصمیم گیری مهمترین فاکتورهای موثر بر صفت SCC از لحاظ اهمیت به ترتیب عبارتند از : شکم زایش، مقدار شیر تولیدی، سال زایش، فصل زایش و روز شیردهی (مرحله شیردهی). بر اساس درخت تصمیم گیری حاصل شده ترکیب های مختلفی از فاکتورهای فنوتیپی موثر بر SCC قابل شناسایی می باشد. درخت تصمیم گیری ایجاد شده در این

[DOR: 20.1001.1.16807073.2019.21.7.3.4]

مطالعه نشان می دهد که کمترین مقدار SCC در گاوهایی مشاهده شد که شکم زایش آنها اول یا دوم و مقدار تولید شیر روزانه آنها بیشتر از ۳۰ کیلوگرم و در فصل پاییز یا زمستان زایش و سال زایش آنها بین ۱۳۸۴ تا ۱۳۹۲ می باشد.

[DOR: 20.1001.1.16807073.2019.21.7.3.4]