Gaziosmanpaşa Üniversitesi Ziraat Fakültesi Dergisi
Journal of Agricultural Faculty of Gaziosmanpasa University
https://dergipark.org.tr/tr/pub/gopzfd

Araştırma Makalesi/*Research Article*

JAFAG
ISSN: 1300-2910
E-ISSN: 2147-8848
(2022) 39(3) 143-150
doi:10.55507/gopzfd.1175502

A Decision Tree Model to Determine Some Environmental Factors Affecting 305-Day Milk Yield in Simmental Cows

Aslı AKILLI[1]* Hülya ATIL[2] Çiğdem TAKMA[2]
[1]Department of Agricultural Economics, Faculty of Agriculture, Kırşehir Ahi Evran University Kırşehir
[2]Department of Animal Science, Faculty of Agriculture, Ege University Izmir
***Corresponding author's email:** asliakilli@ahievran.edu.tr

Alındığı tarih (Received): 15.09.2022 Kabul tarihi (Accepted): 26.10.2022

Abstract: In this study, some environmental factors thought to be effective on 305-day milk yield in Simmental cows, were examined according to the decision tree method with regression tree algorithm. For this purpose, the effect levels of calving interval, somatic cell count, calving age, and parity variables on the 305-day milk yield of 148 Simmental cows were determined. As a result of the decision tree application, the factors affecting 305-day milk yield were found as parity, calving age, somatic cell count, and calving interval, in order of importance. In addition, it was determined that the 305-day milk yield of the cows with the calving age above 5 was high, and the cows with the somatic cell count greater than 104.500 were found to be the lowest. There is a need to use the decision tree approach in order to examine the effects of other environmental factors that are thought to be effective on milk yield or other economic characteristics in dairy farming and to provide appropriate conditions by correcting the relevant factors accordingly.

Keywords: Decision Tree, 305-day milk yield, Simmental cow

Simmental İneklerde 305 Günlük Süt Verimini Etkileyen Bazı Çevresel Faktörlerin Belirlenmesine Yönelik Bir Karar Ağacı Modeli

Öz: Bu çalışmada Simmental (Sarı Alaca) ineklerde 305 günlük süt verimi üzerinde etkili olduğu düşünülen bazı çevresel faktörler regresyon ağacı algoritması ile karar ağacı yöntemine göre incelenmiştir. Bu amaçla buzağılama aralığı, somatik hücre sayısı, buzağılama yaşı ve laktasyon sırası değişkenlerinin 148 adet Simmental ırkı ineğin 305 günlük süt verimine etki düzeyleri belirlenmiştir. Karar ağacı uygulaması sonucunda 305 günlük süt verimini etkileyen faktörler önem sırasına göre laktasyon sırası, buzağılama yaşı, somatik hücre sayısı ve buzağılama aralığı olarak bulunmuştur. Ayrıca buzağılama yaşı 5 yıl yaşın üzerinde olan ineklerin 305 günlük süt veriminin yüksek olduğu, somatik hücre sayısı 104,500'ün üzerinde olan ineklerin ise en düşük olduğu belirlenmiştir. Süt hayvancılıkta süt verimi veya diğer ekonomik özellikler üzerine etkili olduğu düşünülen başka çevresel etmenlerin etkilerini incelemek ve buna göre ilgili faktörlerin düzeltilerek uygun şartların sağlanması için karar ağacı yaklaşımının kullanılmasına ihtiyaç vardır.

Anahtar Kelimeler: Karar ağacı, 305 Günlük süt verimi, Simmental inek

1. Introduction

Milk contains protein, fat, carbohydrate, mineral substances, vitamins, etc. It is an important nutrient and animal food for humans due to its elements. In fact, approximately 26.5% of the animal protein consumed in the world and 268% of the energy is provided from milk. These values are respectively 49.8% and 54.5% in Turkey. The amount of milk produced from dairy cattle has been reported as 80.97% in the world and 90.57% in Turkey as of 2020 (Anonymous, 2020).

In order to increase animal production, optimizing the product obtained per animal instead of increasing the number of animals is a suitable approach in terms of a more accurate evaluation of feed resources. In addition to the use of qualified breeding animals, it is possible to increase animal production by improving environmental conditions by providing better care, feeding, and preventing diseases. Environmental conditions include external factors that affect the animal's productivity and biological functions at different rates.

The milk yield of an animal is a physiological condition shaped by genotype and environmental effects. Milk yield in cows increases immediately after birth and reaches its highest level in the 45-60th days of lactation. It stays at this level for a while and gradually decreases. The daily or annual amount of milk obtained from cows and the ratio of milk components vary depending on many factors such as genotype, lactation period, temperature, and a number of milking.

Simmental cattle are among the breeds with high milk yield in the world. Simmental cattle, breeding for combined efficiency with meat and milk characteristics, are one of the culture breed materials such as breeding bulls, pregnant heifers, and semen (Koç, 2016). There are various scientific studies on the research of milk yield in Simmental cattle (Koçak et al., 2008) compared the yield characteristics of Holstein, Brown, and Simmental cattle in terms of different variables. Another study in which Brown and Simmental cattle were handled comparatively in terms of milk yield was carried out by (Aksoy, 1995). Moreover, Akbulut (1998) comparatively evaluated the results obtained from studies on the yield characteristics of Simmental cattle in Turkey. In terms of genetic and phenotypic trends, 305-day milk yield in Simmentals was discussed by (Ulutaş et al., 2010). Furthermore Şekerden (1999) investigated the effects of calving season and lactation order on milk yield and components in Simmental cattle. The suitability of one of the time series methods (ARMA) for modeling and estimating test-day milk production data in Simmental cattle was investigated (Macciotta et al., 2002). Also, the milk production characteristics of selected animals of Polish Black and White cows with up to 50% Holstein-Friesian genes imported from the Czech Republic and reared under the same environmental conditions as Simmental cows born in Poland were comparatively evaluated (Sablik et al., 2019). Various scientific studies dealing with the milk yield characteristics of Simmental cattle in the light of environmental and genetic factors are available in the literature (Çilek and Tekin, 2006; Çilek et al., 2008; Ulutaş and Sezer, 2009). Today, apart from classical statistical approaches such as regression that examines the yield factor relationship; different models are developed for the analysis of large and complex data sets. Decision trees algorithm is one of these new approaches within the scope of data mining. The use of decision trees is becoming increasingly common in examining various factors affecting milk yield. Yordanova et al., (2015) examined the conformational properties effective on 305-day milk yield in Holstein cows and found that the most effective conformational properties on milk yield were udder width, locomotion, stature, and chest width. Genç and Mendeş (2021) determined the factors affecting the 305-day milk yield of dairy cattle by using Regression Tree (CRT) analysis. They studied eight different cattle breeds grown in Türkiye. Çak et al., (2013) used the decision tree method to examine the lactation milk yield of Brown cattle, depending on productivity and environmental factors. Lopez-Suarez (2018) used decision trees to extract patterns for dairy culling management. Sitkowska et al., (2017) aimed to investigate of detection of high levels of somatic cells in milk on farms equipped with an automatic milking system by the decision-tree technique. Slob et al., (2021) examined and interpreted the articles discussing machine learning applications, including the decision-tree method, in terms of dairy farm management. Kliś et al., (2021) aimed to determine the possibility of using automatic milking system data for periparturient cows to predict their lactation milk yield. Piwczyński et al., (2020) aimed to utilize the decision trees technique to determine the factors responsible for high monthly milk yield in Polish Holstein-Friesian cows from 27 herds equipped with milking robots. O'Leary and Lynch (2022) presented a review of machine learning approaches applied to dairy-specific data and meteorological signals to predict milk volumes in Ireland at a national level. One of the investigated methods is the Decision-Tree method. Aerts et al., (2022a) aimed to determine the best combination of factors and their levels suggested by the decision tree for high daily milk yield per automatic milking systems in dairy cattle herds. In their study, milk yield, number of cows, free robot time, milking speed, and cow treatment time were the significant determinants of total milk yield for all cows per milking robot and were most frequently used to construct the decision tree.

In this study, it was aimed to determine the effect levels of parity, calving age, somatic cell count, and calving interval variables affecting 305-day milk yield in Simmental cows by using the regression tree algorithm. At the same time, based on the results of the analysis, it is aimed to present some suggestions for the breeders on the correction of environmental factors.

2. Material and Methods

2.1. Material

The animal material of the study was provided from a dairy farm registered in Bursa province. 305-day milk yields of a total of 148 Simmental cows, raised from 2020-to 2021 and aged between 2-5 years were used in the analysis.

2.2. Method

A decision tree is a technique that allows the factors and their levels that affect the dependent variable to be determined graphically. According to the algorithm, the decision tree analysis determines whether the variable value is large or small according to a predetermined constant, and the tree is formed by a two-way split (Steensels et al., 2016). If all records at any node in the tree have the same classification, that part of the tree stops growing (Witten & Frank, 2005). There is a completely different way of partitioning the Decision-Trees. Successive binary partitions can achieve the partition based on the different predictors for both classification and regression. In the decision trees method, a partitioning algorithm is used that continues to divide the data set into smaller and smaller subsets until the stopping criteria are met (Irizarry, 2005). Decision trees start with root nodes. This beginning node was the most heterogeneous sub-group. Other homogenous sub-groups with subsets that are not split are called terminal nodes, and still, other sub-groups were child nodes (Çamdeviren et al., 2005). Accordingly, this technique uses different growing methods or algorithms to split the datasets, such as CHAID (Chi-square automatic interaction detection), exhaustive CHAID, CRT (Classification and Regression Trees, also called CART), and QUEST (Quick, Unbiased and Efficient Statistical Tree). The CHAID algorithm often splits the data into more than two categories and should be preferred when the outcome is categorical. On the other hand, QUEST and CRT tend to split the data into two categories. The CRT is useful regardless of whether the outcome is categorical or numerical.

The divisions in regression tree algorithms vary depending on the variable type. Depending on the split of the estimators, $(2^{I-1} - 1)$ splits are possible if the variable (X) is in a categorical structure. If the variable has continuous values (K different values), $K - 1$ separations can occur. A decision tree starts growing from root node X. The proceeding steps of the method are summarized in three steps for each node. In the first step, the best split of each estimator in the model is found. To decide the best split point for estimators ordered from a wide range to a narrow range, if $x \leq v$, v is called, the observation goes to the split to occur on the left. Otherwise, it goes to the right, thus performing the checking. There is one best split point, which splits to optimize the split criteria. For nominal estimators, possible categorical subsets are evaluated to find the best split. Here it is called A with $x \in A$; the observation first goes to the node to be born on the left; on the contrary, it advances to the node on the right. In the next step, the best split of the node is calculated. In the last step, when the stopping rules are observed to be insufficient, the node is divided using the best split calculated in the second step.

The purity criterion is important in a node. If the impurity criteria are defined for the node, the division criterion corresponding to the decrease in purity is defined with the mathematical expression in Equation 1. Here, "t" is the node that maximizes the division criterion, and the best division in this node is expressed by "s" ($\Delta_i(s,t)$).

$$\Delta\mathrm{I}(\mathrm{s},\mathrm{t}) = \mathrm{p}(\mathrm{t})\Delta_{\mathrm{i}}(\mathrm{s},\mathrm{t}) \quad (1)$$

If the variable is continuous, the division criterion is expressed as in Equation 2 using the least squares deviation. Mathematical representations of impurity are given in Equation 3-7 (Karadağ, 2014).

$$\Delta_{\mathrm{i}}(\mathrm{s},\mathrm{t}) = \mathrm{i}(\mathrm{t}) - \mathrm{p_L}\mathrm{i}(\mathrm{t_L}) - \mathrm{p_R}\mathrm{i}(\mathrm{t_R}) \quad (2)$$

$$\mathrm{i}(\mathrm{t}) = \frac{\sum_{\mathrm{n}\in\mathrm{h}(\mathrm{t})} \mathrm{W_n}\mathrm{f_n}(\mathrm{y_n} - \bar{\mathrm{y}}(\mathrm{t}))^2}{\sum_{\mathrm{n}\in\mathrm{h}(\mathrm{t})} \mathrm{W_n}\mathrm{f_n}} \quad (3)$$

$$\mathrm{p_L} = \frac{\mathrm{N_w}(\mathrm{t_L})}{\mathrm{N_W}(\mathrm{t})} \quad (4)$$

$$\mathrm{p_R} = \frac{\mathrm{N_w}(\mathrm{t_R})}{\mathrm{N_W}(\mathrm{t})} \quad (5)$$

$$\mathrm{N_W}(\mathrm{t}) = \textstyle\sum_{\mathrm{n}\in\mathrm{h}(\mathrm{t})} \mathrm{W_n}\,\mathrm{f_n} \quad (6)$$

$$\bar{\mathrm{y}}(\mathrm{t}) = \textstyle\sum_{\mathrm{n}\in\mathrm{h}(\mathrm{t})} \mathrm{W_n}\,\mathrm{f_n}/\mathrm{N_W}(\mathrm{t}) \quad (7)$$

In this study, the CRT growing method was used for the dependence of 305-day milk yield with respect to calving interval, somatic cell count, calving age, and parity independent factors. The maximum tree depth was 5, the minimum cases in Parent Node were 10, and the minimum cases in Child Node were 6. The CRT analysis was done using IBM SPSS Statistics Version 25.

3. Results and Discussion

The designed regression tree with the root (node 0), branch (nodes:2, 3, and 5), and leaf (nodes: 1, 4, 6, 7, and 8) nodes, in this investigation, was given in Figure 1.

As shown in the Figure 1, the average 305-day milk yield of Simmental Cows was found to be 8748,39±1211, 71 (Node 0). 305-day milk yield value was found to be higher than the values reported by (Erdem et al., 2015; Macciotta et al., 2002; Şekerden, 1999). In this study, the decision tree constructed for 305-day milk yields of Simmentals first were divided into two nodes according to parity 1 (Node 1: 8226,14±1067,47) and parity 2, 3 (Node 2: 8922,48±1211,14). Then parity 2, 3 splits into two nodes based on calving age, calving age is equal to or less than

5 years old (Node 3:8851,46±1192,25), and bigger than 5 years old (Node 4: 10165,33±863,05). Node 3 was grouped according to somatic cell count is equal to or less than 104,500 (Node 5: 8953,59±1181,91) and bigger than 104,500 (Node 6: 7978,73±925,23). Finally, Node 5 was divided for calving interval is equal to or less than 402,5 day (Node 7: 8707,93±1147,55) and bigger than 402,5 day (Node 8: 9315,61±1152,26).

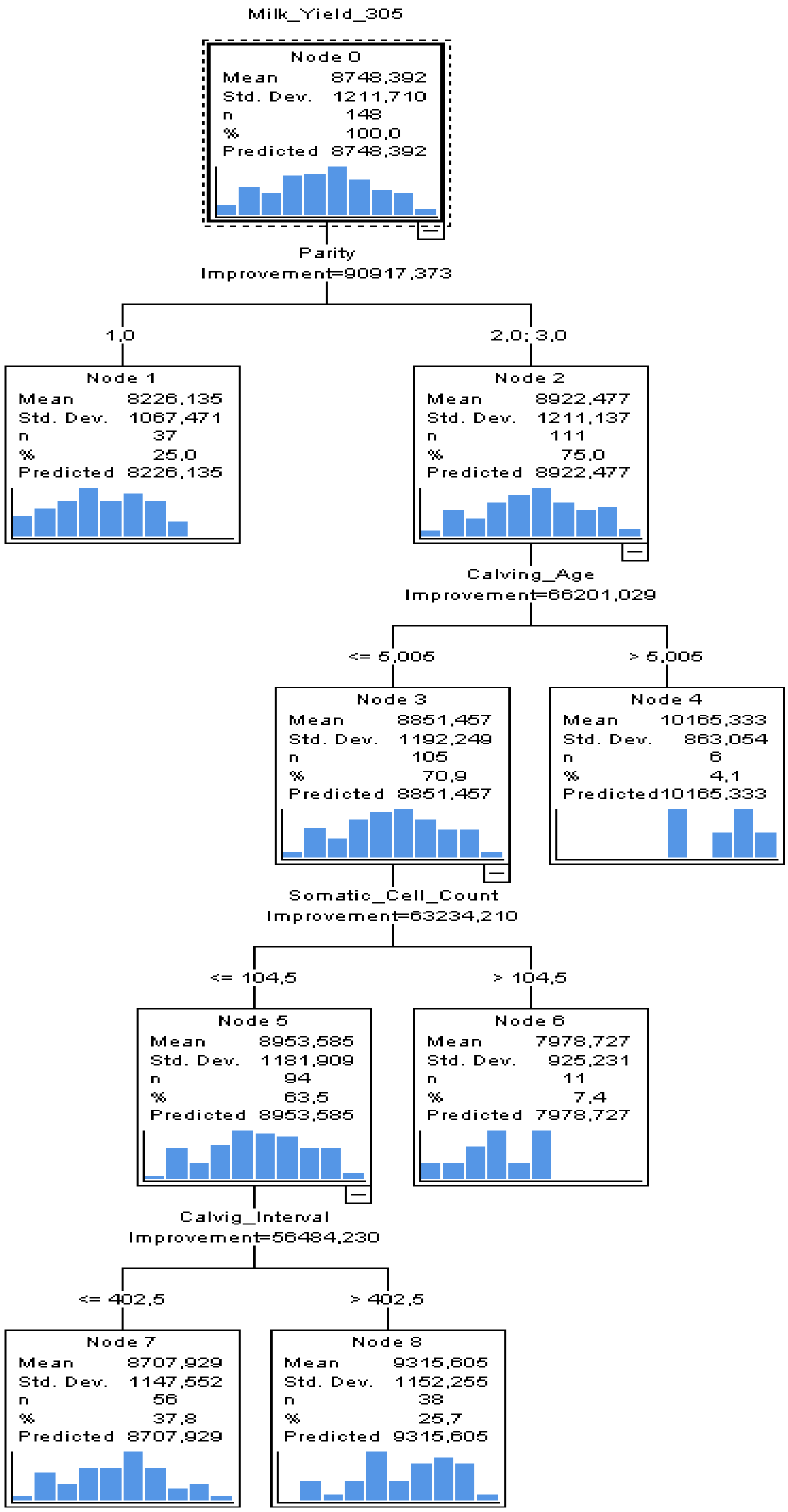

Figure 1. Regression Tree Diagram for 305 Day Milk Yield
Şekil 1. *305 günlük süt verimi için Karar Ağacı Diyagramı*

As a result of the decision tree analysis, the 305-day milk yield average of the cows with parity 2, 3 years old and calving age higher than 5 years old was found to be the highest (10165,33±863,05). Unlike our study, in the analyses performed by (Genç & Mendeş, 2021) with the decision-tree technique, calving age was found to have lower importance in the 305-day milk yield estimation. In another study on the decision-tree technique and milk yield in cattle, the significant effects of the calving age variable were observed, similar to our study (Aerts et al., 2022b). In their study, Bakir et al., (2010) were determined the effects of the dry period, parity, farm, calving season, and age on 305-days milk yield using the Regression Tree (CRT) method. Similar to our study, dry period, parity and calving season were determined to affect 305-day milk yield at first-, second- and third-degree factors, respectively. Similarly, according to the decision tree results, it was determined that the calving age of the cows should be greater than 5 for high milk yield (Mostert et al., 2001).

On the other hand, 305-day milk yield was found to be low in cows with lactation parity of 2- 3 years, calving age older than 5 years, and somatic cell count greater than 104,5 (7978,73±925,23). Moreover, in this research, some environmental factors that effect on 305-day milk yields were ranked according to their importance levels with the regression tree analysis method. The results of the importance of independent variables based on classification and regression trees (CRT) are presented in Table 1 and Figure 2.

Table 1. Independent Variable Importance
Çizelge 1. *Bağımsız Değişkenlerin Önemleri*

Variable	Importance	Normalized Importance
Calving Age	154460.26	100.0%
Calving Interval	147401.60	95.4%
Somatic Cell Count	133257.97	86.3%
Parity	90917.37	58.9%

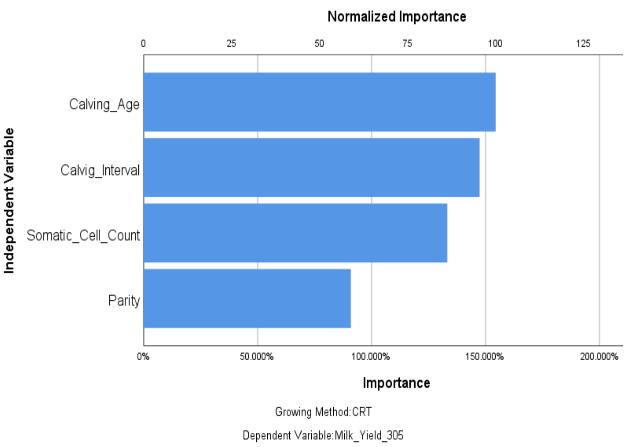

Figure 2. Normalized importance levels of the factors
Şekil 2. *Faktörlerin yüzde önem seviyeleri*

As seen in Table 1 and Figure 2 variables on 305-day milk yield of Simmentals were found as calving age, calving interval, somatic cell count, and parity in order of effect, respectively. Therefore, it was determined that the most effective factor on the 305-day milk yield was estimated as the calving age and the second important factor was the calving interval. Moreover, the least effective variable on 305-day milk yield was the parity.

As a matter of fact, it is known that the milk yield of cows increases until the lactation 4^{th} or 5^{th} in equal conditions. M'hamdi et al. (2012) reported that the effects of parity on dry period and lactation duration were statistically significant. The lactation duration was shortest in lactation 5^{th} and longest in lactation 4^{th}. The dry period was longest in 5^{th} lactation and shortest in 3^{rd} lactation. Since milk yields up to the 3^{rd} parity were analyzed in this study, the effect of parity was not sufficiently explained and found to be low.

The calving interval is a very important trait and environmental factor in dairy cattle breeding in terms of revealing the productivity of the herd and evaluating the management applied in the herd (Şahin & Ulutaş, 2011). In this study, the Simmental calving interval was divided as equal to or less than 402,5 days and bigger than 402.5 days. The cut-off value of the calving interval of division, 402,5 days, was found to be higher than 379,26 days reported by Kara et al., (2021) and approximately similar to 404.6 days reported by Gültekin (2019).

Cziszter et al., (2016), in their study evaluating the production and reproductive performance of Simmental cows with different temperaments, determined that the calving interval, which is among the characteristics they examined, was at least 405.3, similar to our study. A higher calving interval (446.6) was found in the animal group defined as nervous in the aforementioned study compared to the value obtained in our study. In the study by Bujko et al., (2018), the relationship between milk yield characteristics and calving interval in breeding herds of Slovak Simmental dairy cows was evaluated. Researchers reported the calving interval of cows in the second and third lactation to be 406.3 days, which is consistent with our study. Ulutaş and Sezer (2009) studied Simmental cattle to evaluate phenotypic and genetic parameters for 305- day milk yield, lactation length, dry period, calving interval and service periods in their study. It was observed that the calving intervals reported in the research results were below the average values determined in our study.

In addition, it has been reported that there is a negative correlation between somatic cell count and milk yield (Aytekin and Boztepe, 2011; Bartlett et al., 1990). There are many factors that directly and indirectly affect the number of somatic cells in milk in dairy cattle. In addition to direct factors such as cow's age, lactation period and feeding, factors such as milking place and shape indirectly cause changes in the number of somatic cells.

It is seen that the results of Franzoi et al (2020) examining the effects of somatic cell number on milk yield and related traits in Simmental and different cattle breeds are compatible with our study. Another study on the number of somatic cells in Simmental cattle was carried out by Kalinska et al., (2019). The results show that somatic cell number is an essential variable in milk yield and quality, similar to our study. In our study, it was determined that animals with high somatic cell counts had lower milk yield than other cattle in the herd, which is in line with the literature. These results obtained in our study are consistent with the research conducted by Barlowska et al (2009). The researchers aimed to determine the relationship between the number of somatic cells in milk and daily yield, chemical content, and technological usefulness of milk in different dairy cattle breeds. An increase in somatic cell count has been associated with a decrease in daily milk yield in high-yielding dairy cows.

In this study, it was determined that the highest milk yield average belonged to 6 cows during 2 and 3 parity and calving interval older than 5 years. It is thought that these results will be insufficient and misleading since they represent only 4% of the total cows evaluated. For this reason, it can be said that 305-day milk yield will be high (9315.61±1152.26) in conditions where the parity is 2 and above, the calving age is less than 5 years, the number of somatic cell counts are less than 104.000 and the calving interval is more than 402.5 days.

In the literature, there are successful applications of the decision tree method to analyze animal data, similar to the application carried out in our study (Bakır et al., 2010; Eyduran et al., 2013; Gocheva-Ilieva et al., 2022; Rodrigez et al., 2019; Topal et al., 2010).

4. Conclusion

Decision-tree method is a beneficial technique in terms of estimating the effects on the dependent variable in the analysis processes of large data sets containing multidimensional and complex variables. One of the essential advantages of the decision-tree technique is that it is unnecessary to choose which variable to add to the model and which not. This way, the research variables can be approached with a more flexible perspective. Even if the variables in the model are similar, all available data can be added, and the algorithm selects them according to their importance.

The environmental factors affecting the 305-day milk yield of Simmental cows were estimated as calving age, calving interval, somatic cell count and parity, respectively. However, there are other environmental factors that affect the number of calves to be taken from the animals during their lifetime and the total amount of milk. On the other hand, there are few studies using decision trees approach for yield traits in Simmental cows. Therefore, there is a need for more studies on the use of this method in determining the effect of different environmental factors.

References

Aerts, J., Sitkowska, B., Piwczyński, D., Kolenda, M., & Önder, H. (2022a). The optimal level of factors for high daily milk yield in automatic milking system. *Livestock Science,* 264, 105035. https://doi.org/10.1016/j.livsci.2022.105035.

Aerts, J., Kolenda, M., Piwczyński, D., Sitkowska, B., & Önder, H. (2022b). Forecasting Milking Efficiency of Dairy Cows Milked in an Automatic Milking System Using the Decision Tree Technique. *Animals*, 12(8), 1040. https://doi.org/10.3390/ani12081040.

Akbulut, Ö. (1998). Simental sığırların Türkiye'de verim performansı üzerine bir değerlendirme. *Atatürk Üniversitesi Ziraat Fakültesi Dergisi*, 29(1), 43-49.

Aksoy, A. R. (1995). Milk yield of Swiss Brown and Simmental cows in kars goose raising Station. *Yuzuncu Yil University, The Journal of Agricultural Faculty* (Turkey), 6, 55-57.

Anonymous, (2020). Food and Agriculture Organization of the United Nations. https://www.fao.org/faostat/en/#data/QCL.

Aytekin, İ., & Boztepe, S. (2014). Süt sığırlarında somatik hücre sayısı, önemi ve etki eden faktörler, *Türk Tarım Gıda Bilim ve Teknoloji Dergisi*, (2): 112-121.

Bakır, G., Keskin, S., & Mirtagioğlu, H. (2010). Determination of the effective factors for 305 days milk yield by Regression Tree (RT) method. *Journal of Animal and Veterinary Advances*, 9(1), 55-59.

Barłowska, J., Litwińczuk, Z., Wolanciuk, A., & Brodziak, A. (2009). Relationship of somatic cell count to daily yield and technological usefulness of milk from. *Polish Journal of Veterinary Sciences*, 12(1), 75-79.

Bartlett, P. C., Miller, G.Y., Anderson, C.R. & Kirk, J.H. (1990). Milk production and somatic cell count in Michigan dairy herds. *Journal of Dairy Science*, 73, 2794-2800. https://doi.org/10.3168/jds.S0022-0302(90)78966-7.

Bujko, J., Candrák, J., Strapák, P., Žitný, J., & Hrnčár, C. (2018). Evaluation relation between traits of milk production and calving interval in breeding herds of Slovak Simmental dairy cows. *Albanian Journal of Agricultural Sciences*, 17(1), 31-36.

Cziszter, L. T., Gavojdian, D., Neamt, R., Neciu, F., Kusza, S., & Ilie, D. E. (2016). Effects of temperament on production and reproductive performances in Simmental dual-purpose cows. *Journal of Veterinary Behavior*, 15, 50-55. https://doi.org/10.1016/j.jveb.2016.08.070.

Çak, B., Keskin, S., & Yılmaz, O. (2013). Regression tree analysis for determining of affecting factors to Lactation Milk Yield in Brown Swiss cattle. *Asian Journal of Animal and Veterinary Advances,* 8(4), 677-682. Doi: 10.3923/ajava.2013.677.682.

Çamdeviren, H., Mendeş, M., Özkan, M. M., Toros, F., Şasmaz, T., & Öner, S. (2005). Determination of depression risk factors in children and adolescents by regression tree methodology. *Acta Medica Okayama*, 59(1), 19-26. 10.18926/AMO/31985.

Çilek, S., & Tekin, M. E. (2006). Calculation of adjustment factors for standardizing lactations to mature age and 305-day and estimation of heritability and repeatability of standardized milk yield of Simmental cattle reared on Kazova state farm. *Turkish Journal of Veterinary & Animal Sciences*, 30(3), 283-289.

Çilek, S., Orhan, H., Kaygısız, A., & Şahin, E. H. (2008). Estimation of breeding values of Anatolian population of Simmental cows using monthly test day milk yields. *Archiva Zootechnica*, 11, 79-85.

Erdem, H., Atasever, S., & Kul, E. (2015). Relations of body condition score with milk yield and reproduction traits in Simmental cows. *Large Animal Review*, 21: 231-234. 2015.

Eyduran, E., Yılmaz, İ., Tariq, M. M., & Kaygısız, A. (2013). Estimation of 305-d milk yield using regression tree method in Brown Swiss cattle. *The Journal of Animal & Plant Sciences*, 23: 731-735.

Franzoi, M., Manuelian, C. L., Penasa, M., & De Marchi, M. (2020). Effects of somatic cell score on milk yield and mid-infrared predicted composition and technological traits of Brown Swiss, Holstein Friesian, and Simmental cattle breeds. *Journal of Dairy Science*, 103(1), 791-804. Doi: 10.3168/jds.2019-16916.

Genç, S., & Mendes, M. (2021). Determining the Factors Affecting 305-Day Milk Yield of Dairy Cows with Regression Tree. *Turkish Journal of Agriculture-Food Science and Technology*, 9(6), 1154-1158. Doi: https://doi.org/10.24925/turjaf.v9i6.1154-1158.4384.

Gocheva-Ilieva, S., Yordanova, A., & Kulina, H. (2022). Predicting the 305-Day Milk Yield of Holstein-Friesian Cows Depending on the Conformation Traits and Farm Using Simplified Selective Ensembles. *Mathematics*, 10(8), 1254.

Gültekin, İ. (2019). *Afyonkarahisar süt sığırı işletmelerinde bazı döl verimi parametrelerine etkili çevresel faktörler ve ekonomik kayıplar* (Master's thesis, Afyon Kocatepe Üniversitesi, Sağlık Bilimleri Enstitüsü).

Irizarry, R. A. (2006). *Statistical learning: Algorithmic and nonparametric approaches*. Course Notes. Department of Biostatistics Johns Hopkins University, Fourth Term 2005-2006.

Kalinska, A., Slosarz, J., Golebiewski, M., Wojcik, A. G. A. T. A., Przysucha, T. O. M. A. S. Z., & Kruzinska, B. (2019). Influence of cattle breed and country of origin on milk yield and milk composition in dairy cows. Annals of Warsaw University of Life Sciences-SGGW. *Animal Science*, 58(1), 21-28. Doi: 10.22630/AAS.2019.58.1.3.

Kara, N. K., Galiç, A., & Çevik, S. (2021). Simental ırkı ineklerde bazı verim özellikleri ve sağlık sorunları arasındaki ilişkinin belirlenmesi. *Akademik Ziraat Dergisi*, 10(2), 411-418. https://doi.org/10.29278/azd.881922.

Karadağ, M. (2014). Karar ağaçları ile lojistik regresyon analizinin performanslarının simülasyon çalışması ile karşılaştırılması (Master's thesis, Trakya Üniversitesi, Sağlık Bilimleri Enstitüsü).

Kliś, P., Piwczyński, D., Sawa, A., & Sitkowska, B. (2021). Prediction of lactational milk yield of cows based on data recorded by AMS during the periparturient period. *Animals*, 11(2), 383. https://doi.org/10.3390/ani11020383.

Koç, A. (2016). A review on Simmental raising: 1. Simmental raising in the World and in Turkey. *Journal of Adnan Menderes University Agricultural Faculty*, 13(2), 97-102.

Koçak, S., Tekerli, M., Özbeyaz, C., & Demirhan, I. (2008). Some production traits of Holstein, brown-swiss, and Simmental cattle reared in Lalahan livestock research institute. *Journal of Lalahan Livestock Research Institute* (Turkey), 48(2), 51-57.

Lopez-Suarez, M., Armengol, E., Calsamiglia, S., & Castillejos, L. (2018). Using decision trees to extract patterns for dairy culling management. In *IFIP international conference on artificial intelligence applications and innovations*. Springer New York LLC, pp. 231–239.

M'hamdi, N., Bouallegue, M., Frouja, S., Ressaissi, Y., Brar, S. K., & Hamouda, M. B. (2012). Effects of environmental factors on milk yield, lactation length and dry period in Tunisian Holstein cows. In Milk *Production-An Up-to-Date Overview of Animal Nutrition, Management and Health. IntechOpen*.

Macciotta, N. P. P., Vicario, D., Pulina, G., & Cappio-Borlino, A. (2002). Test day and lactation yield predictions in Italian Simmental cows by ARMA methods. *Journal of Dairy Science*, 85(11), 3107-3114. https://doi.org/10.3168/jds.S0022-0302(02)74398-1

Mostert, B. E., Theron, H. E., & Kanfer, F. H. J. (2001). The effect of calving season and age at calving on production traits of South African dairy cattle. *South African Journal of Animal Science*, 31(3), 205-214.

O'Leary, C., & Lynch, C. (2022, June). An Evaluation of Machine Learning Approaches for Milk Volume Prediction in Ireland. *In 2022 33rd Irish Signals and Systems Conference (ISSC)* (pp. 1-8). IEEE. Doi: 10.1109/ISSC55427.2022.9826160.

Piwczyński, D., Sitkowska, B., Kolenda, M., Brzozowski, M., Aerts, J., & Schork, P. M. (2020). Forecasting the milk yield of cows on farms equipped with automatic milking system with the use of decision trees. *Animal Science Journal*, 91(1), e13414. https://doi.org/10.1111/asj.13414.

Rodríguez, E., Waissman, J., Mahadevan, P., Villa, C., Flores, B. L., & Villa, R. (2019). Genome-wide classification of dairy cows using decision trees and artificial neural network algorithms. *Genetics and Molecular Research*, 18(4), 1-13.

Sablik, P., Szewczuk, M., Januś, E., & Skrzypiec, A. (2019). Comparison of selected milk production traits of Simmental and Polish Black-and-White cows raised in the buffer zone of Ujście Warty National Park. *Landbauforschung*, 68(3-4), 45-51. Doi:10.3220/LBF1538729088000.

Sitkowska, B., Piwczynski, D., Aerts, J., Kolenda, M., & ÖZKAYA, S. (2017). Detection of high levels of somatic cells in milk on farms equipped with an automatic milking system by decision trees technique. *Turkish Journal of Veterinary & Animal Sciences*, 41(4), 532-540. Doi: 10.3906/vet-1607-78.

Slob, N., Catal, C., & Kassahun, A. (2021). Application of machine learning to improve dairy farm management: A systematic literature review. *Preventive Veterinary Medicine*, 187, 105237. https://doi.org/10.1016/j.prevetmed.2020.105237.

Steensels, M., Antler, A., Bahr, C., Berckmans, D., Maltz, E., & Halachmi, I. (2016). A decision-tree model to detect post-calving diseases based on rumination, activity, milk yield, BW and voluntary visits to the milking robot. *Animal*, 10(9), 1493-1500. https://doi.org/10.1017/S1751731116000744.

Şahin, A., & Ulutaş, Z. (2011). Tahirova tarım işletmesi de yetiştirile siyah alaca i ekleri süt ve döl verim özellikleri i etkileye bazı çevresel faktörler. *Anadolu Tarım Bilimleri Dergisi*, 26(2), 156-168.

Şekerden, Ö. (1999). Effects of calving season and lactation order on milk yield and milk components in Simmental cows. Turkish *Journal of Veterinary & Animal Sciences*, 23(7), 79-86.

Topal, M., Aksakal, V., Bayram, B., & Yaganoglu, A. M. (2010). An analysis of the factors affecting birth weight and actual milk yield in Swedish red cattle using regression tree analysis. *Journal of Animal and Plant Sciences-Japs*, 20(2), 63-69.

Ulutaş, Z., & Sezer, M. (2009). Genetic study of milk production and reproduction traits of local born Simmental cattle in Turkey. Gaziosmanpaşa Üniversitesi Ziraat Fakültesi Dergisi, 26(1), 53-59.

Ulutaş, Z., Sezer, M., Saatçi, M., & Şahin, A. (2010). Estimation of genetic and phenotypic trends of 305-day milk yield for Simmentals reared in Kazova State Farm in Turkey, *Kafkas Üniversitesi Veteriner Fakultesi Dergisi,* 16(4), 533-536.

Doi:10.9775/kvfd.2009.1083.
Witten, I. H., & Frank, E. (2005). *Data mining; practical machine learning tools and techniques*, 2nd edition. Morgan Kaufmann, San Francisco, CA, USA.
Yordanova, A., Gocheva-Ilieva, S., Kulina, H., Yordanova, L., & Marinov, I. (2015). Classification and regression tree analysis in modeling the milk yield and conformation traits for Holstein cows in Bulgaria. *Agricultural Science and Technology*, 7(2), 208-213.