

# Çoklu Doğrusal Regresyon ve Yapay Sinir Ağı Modellerinin Laktasyon Süt Verimlerine Uyum Yeteneklerinin Karşılaştırılması

Çiğdem TAKMA \*  Hülya ATIL \* Vecihi AKSAKAL \*\*

\* Ege Üniversitesi, Ziraat Fakültesi, Zootečni Bölümü, Biyometri ve Genetik Anabilim Dalı, TR-35100 Bornova/ İzmir - TÜRKİYE

\*\* Gümüşhane Üniversitesi, Aydın Doğan Meslek Yüksekokulu, TR-29100 Kelkit, Gümüşhane – TÜRKİYE

Makale Kodu (Article Code): KVFD-2012-6764

## Özet

Bu çalışmada Siyah Alaca ineklerin laktasyon süt verimleri üzerine laktasyon süresi (LS), buzağılama yılı (BY) ve servis periyodunun (SP) etkisi çoklu regresyon ve yapay sinir ağı (YSA) ile modellenmiş ve modellerin uyum yetenekleri karşılaştırılmıştır. Analizler 2006, 2007 ve 2008 tarihlerinde buzağılayan toplam 305 adet Siyah Alaca ineğin ilk beş laktasyonuna ait süt verimleri üzerinde uygulanmıştır. Yapay sinir ağları modelinde tekrarlanan denemeler sonucunda seçilen mimariye göre, gizli katman sayısı bir ve bu katmandaki gizli düğüm sayısı üç olarak alınmıştır. Ayrıca yakınsama kriteri  $1.10^{-6}$ , maksimum iterasyon sayısı 50 ve algoritmanın sonlanması her bir çalıştırma için 20 devir (epoch) alınmıştır. Yapay sinir ağları ile çoklu regresyon modelinin iyi uyumunun karşılaştırılmasında düzeltilmiş belirleme katsayısı ( $R^2$ ), hata kareler ortalamasının karekökü (Root Mean Square Error-RMSE), ortalama mutlak sapma (Mean Absolute Deviation-MAD) ve ortalama mutlak yüzde hata (Mean Absolute Percentage Error-MAPE) performans kriterleri kullanılmıştır. Analizler sonucunda yapay sinir ağları modeli ile ilk beş laktasyona ait  $R^2$  değerleri 0.62-0.85 arasında, RMSE değerleri 480.9-1682.8 arasında, MAD değerleri 325.2-1381.7 ve MAPE değerleri 6.1-20.2 arasında değişim göstermiştir. Çoklu doğrusal regresyonda ise bu değerler sırası ile  $R^2$  için 0.30 ile 0.75 arasında, RMSE için 1964.8-3008.7 arasında, MAD değeri 1576.6-2458.3 arasında ve MAPE değeri 24.7-35.6 arasında bulunmuştur. Söz konusu kriterlere göre bu modeller karşılaştırıldığında, bu çalışmada kullanılan süt verimlerinin tahminlenmesinde yapay sinir ağı modelinin çoklu doğrusal regresyon modelinden daha iyi uyum sağladığı gözlenmiştir. Dolayısıyla, yapay sinir ağlarının regresyon analizine alternatif bir metod olabileceği sonucuna ulaşılmıştır.

**Anahtar sözcükler:** *Yapay sinir ağları, Çoklu doğrusal regresyon, laktasyon süt verimi, Siyah Alaca, Çok katmanlı ağ*

## Comparison of Multiple Linear Regression and Artificial Neural Network Models Goodness of Fit to Lactation Milk Yields

### Summary

In this study, effects of lactation length, calving year and service period on lactation milk yield of Holstein Friesians were modeled with multiple regression and artificial neural networks (ANN) and compared goodness of fit of models. Analyses were carried on five lactations milk yields of 305 Holsteins calved at 2006, 2007 and 2008 years. After several experiments, hidden layer number was taken one and hidden nodes number were found three for the chosen architecture. Moreover, convergence criteria, maximum iteration number and epoch number were taken as  $1.10^{-6}$ , 50 and 20, respectively. Adjusted coefficient of determination ( $R^2$ ), root mean square error (RMSE), mean absolute deviation (MAD), mean absolute percentage error performance criteria (MAPE) were used for comparison of artificial neural network and multiple linear regression models goodness of fit. After analysis  $R^2$  values were found among 0.62-0.85 for the five lactations with neural networks model. RMSE, MAD and MAPE criteria also were found among 480.9-1682.8, 325.2-1381.7 and 6.1-20.2, respectively. These criteria were found for  $R^2$ , RMSE, MAD and MAPE among 0.30-0.75, 1964.8-3008.7, 1576.6-2458.3 and 24.7-35.6, respectively for multiple linear regression. When the models were compared, artificial neural networks model gave better fit than multiple linear regression models. Consequently, artificial neural networks was determined an alternative method to multiple regression analysis.

**Keywords:** *Artificial Neural Networks, Multiple Linear Regression, Lactation Milk Yield, Holstein Friesian, Multilayer perceptron*



**İletişim (Correspondence)**



+90 232 3112708



cigdem.takma@ege.edu.tr

## GİRİŞ

Laktasyon süresi (LS), buzağılama yılı (BY) ve servis periyodu (SP) süt verimini etkileyen ve sürekli varyasyon gösteren makro çevre faktörlerindedir<sup>1-3</sup>. Söz konusu faktörlerin süt verimi üzerine etkilerinin incelenmesinde çoklu doğrusal regresyon modelinin uygulanışı oldukça yaygın bir tekniktir. Atıl<sup>4</sup>, Siyah Alaca ineklerde yaptığı çalışmada buzağılama ayı ve yılı ile servis periyodu ve kuruda kalma süresinin 305 günlük süt verimi üzerine doğrusal ve kuadratik etkisini önemli bulunmuştur. Ayrıca, Duru ve Tuncel<sup>5</sup> Siyah Alaca ineklerde kuruda kalma süresi, servis periyodu ve ilkine buzağılama yaşının laktasyon süresi, 305 gün süt verimi ve laktasyon süt verimine etkisini belirlemek amacıyla bir araştırma yapmışlardır. Söz konusu çalışmada servis periyodunun incelenen süt verim özelliklerine etkisinin önemli olduğu belirlenmiştir. Ayrıca, kuruda kalma süresi bu üç özellik üzerine etkili bulunmadığı halde, ilkine buzağılama yaşı ise sadece laktasyon süresine etkisiz bulunmuştur. Akbulut ve Hausmann<sup>6</sup> ise, Alman Esmerinde yaptıkları çalışmada ilkine buzağılama yaşının ilk laktasyon süt verimi özelliklerine doğrusal ve kuadratik etkisini incelemiştir. Buna göre, 305 günlük süt verimi üzerine ilkine buzağılama yaşının doğrusal ve kuadratik etkisi önemli bulunmuştur.

Süt veriminin çoklu doğrusal regresyon modeline alternatif olarak yapay sinir ağları tekniği ile modellenmesi ise yeni bir yaklaşımdır. Yapay sinir ağları, insan beyninin fizyolojisinden yararlanılarak oluşturulan bilgi işleme modelleridir. Bu nedenle, yapay sinir ağları birçok basit işlemci elemandan oluşmaktadır. Bu elemanlar farklı formda ifade edilebilen nümerik verileri taşıyan ağırlıklar ile birbirine bağlı bulunmaktadır<sup>7</sup>. Yapay sinir ağları tekniği söz konusu bağlantılarla herhangi bir olay hakkında girdiler ve çıktılar arasındaki doğrusal veya doğrusal olmayan ilişkiyi öğrenmekte daha sonraki olay üzerinde gerekli çağrışımları yaparak çözümler üretebilmektedir.

Grzesiak ve ark.<sup>8</sup> yapay sinir ağları ve çoklu doğrusal regresyon modelleri kullanarak ilk dört denetim günü verimlerinden 305 günlük süt verimini tahminlemiştir. Yapılan çalışmada sürü, sağılan gün sayısı ve buzağılama ayı bağımsız değişkenler olarak tanımlanmıştır. Ayrıca, yapay sinir ağları modelinin çoklu doğrusal regresyon modeline alternatif olabileceği belirlenmiştir. Edriss ve ark.<sup>9</sup> ise çoklu doğrusal regresyon modeli ile ilk laktasyon bilgisini kullanarak ikinci laktasyondaki süt verimini tahminlemiş ve bu tahminleri yapay sinir ağları modeli ile elde edilen tahminler ile karşılaştırmıştır. Ülkemizde hayvancılık alanında yapay sinir ağlarının kullanımı henüz başlangıç aşamasındadır. Bu nedenle yapay sinir ağlarının performansını diğer modellerle karşılaştıran çalışmalar literatürde henüz mevcut değildir.

Bu çalışmada ilk beş laktasyon süt veriminin modellenmesinde laktasyon süresi, buzağılama yılı ve servis periyodu faktörlerine göre çoklu doğrusal regresyon ve yapay sinir ağları modellerinin uyum yeteneklerinin karşılaştırılması hedeflenmiştir.

## MATERYAL ve METOT

### Materyal

Çalışmanın materyalini 2006, 2007 ve 2008 tarihleri arasında buzağılayan Siyah Alaca inekler oluşturmaktadır. Toplam 305 adet ineğin ilk beş laktasyondaki laktasyon süt verimine laktasyon süresi, buzağılama yılı ve servis periyodu etkilerinin modellenmesi amacıyla çoklu doğrusal regresyon ve yapay sinir ağları analizleri uygulanmıştır.

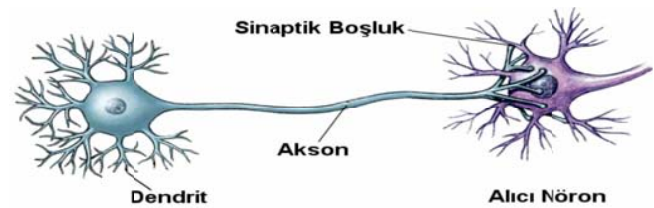
### Metot

Araştırmada kullanılan yapay sinir ağları, insan beyninin deneme yolu ile öğrenme ve genelleştirme esasına dayanarak geliştirilmiştir. Biyolojik sinir sistemi, merkezinde sürekli olarak bilgiyi alan, yorumlayan ve uygun bir karar üreten beynin bulunduğu üç katmanlı bir sistem olarak açıklanabilmektedir. Alıcı sinirler organizma içerisinden ya da dış ortamlardan algıladıkları uyarıları beyne bilgi ileten elektriksel sinyallere dönüştürmektedir. Tepki sinirleri ise, beynin ürettiği elektriksel darbeleri organizma çıktısı olarak uygun tepkilere dönüştürmektedir. *Şekil 1*'de bir sinir hücresinin (nöron) blok gösterimi verilmiştir.

Yapay sinir ağlarında ise girdi bilgileri, girdilere karşılık gelen çıktı bilgileri verilmekte ve ağırlıklı girdi-çıkıtı arasındaki ilişkiyi öğrenmesi sağlanmakta, böylece ağırlıklı eğitimi gerçekleştirilmektedir. Yapay sinir ağları, bu bilgi akışı biçimi bakımından ileri beslemeli (feedforward) ve geri beslemeli (feedback ya da recurrent) olarak gruplandırılabilir. İleri beslemeli ağ yapısında bilgi akışı ileri yönde olmaktadır. Diğer yandan, ileri beslemeli ağda bir girdi katmanı, genellikle bir veya iki gizli katman ve çıktı katmanları bulunmaktadır. Bu ağlara çok katmanlı ileri beslemeli ağlar da denilmektedir.

Ayrıca yapay sinir ağlarında danışmanlı, danışmansız ve takviyeli öğrenme kuralları ile öğrenme gerçekleştirilmektedir. Danışmanlı öğrenme kurallarından biri olan geri yayılım algoritmasında iki türlü bağlantı vardır. Birincisi ileri yönlü bağlantıdır ve ağırlıklar kullanılarak girişlere karşılık çıkışlar elde edilir. İkinci bağlantı ise geriye doğrudur, çıktı katmanında oluşan hatanın en aza indirilmesi için geriye doğru ağırlıkların düzenlenmesi işlemine dayanmaktadır. İleri besleme işlemi giriş katmanına gelen verilerin ilk gizli katmana giriş olarak aktarılması işlemidir. Tüm bu aşamalarda katmanlardaki işlemci eleman sayısı da önemli olmaktadır.

Yapay sinir ağlarında giriş katmanında herhangi bir işlem



Şekil 1. Nöronun şematik yapısı

Fig 1. Schema of neuron

yapılmadığından giriş ve çıkış katmanındaki işlemci elemanı sayısı tamamen uygulanan probleme bağlı olmaktadır. Gizli katman sayısı ve gizli katmanlardaki işlemci elemanı sayısı ise her ne kadar giriş ve çıkış katmanındaki sınır sayıları dikkate alınarak belirlense de kesin karara deneme yanılma yolu ile ulaşılmaktadır. Şekil 2'de 2-girişli 1-çıkışlı çok katmanlı ileri beslemeli bir ağ yapısı gösterilmektedir.

Bu çalışmada çok katmanlı ileri beslemeli geri yayımlı yapay sinir ağı kullanılarak laktasyon süt verimi tahminlenmiştir (Şekil 3). Söz konusu model ile tekrarlanan denemeler sonucunda yapay sinir ağı modelinde gizli katman sayısı bir ve bu katmandaki gizli düğüm sayısı üç olarak alınmıştır. Katmanlar arasında sigmoid aktivasyon fonksiyonu kullanılmıştır. Yakınsama kriteri  $1.10^{-6}$ , maksimum iterasyon sayısı 50 ve her bir çalıştırmada algoritmanın sonlandırılmasında 20 devir (epoch) alınmıştır. Analizlerde JMP <sup>10</sup> istatistik paket programı kullanılmıştır.

Çalışmada kullanılan çoklu doğrusal regresyon modelinin temelini oluşturan regresyon analizi, aralarında sebep-sonuç bağlantısı bulunan bağımlı ve bağımsız değişkenler arasındaki ilişkinin modelini saptamakta ve bu model kullanılarak gelecekle ilgili tahminler yapılmasına imkan tanımaktadır. Birden fazla bağımsız değişken kullanılarak yapılan regresyon analizine çoklu regresyon analizi (multiple regression analysis) adı verilmektedir.

Analizdeki çoklu doğrusal regresyon modeli  $Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \epsilon_i$  biçimindedir. Modelde Y bağımlı değişken;  $X_1$ , laktasyon süresi;  $X_2$ , buzağılama yılı ve  $X_3$ , servis periyodu olmak üzere  $b_0$ , regresyon eğrisinin y eksenini kesim noktası,  $b_1$  birinci bağımsız değişkenin regresyon katsayısı,  $b_2$  ikinci bağımsız değişkenin regresyon katsayısı,  $b_3$  üçüncü bağımsız değişkenin regresyon katsayısı,  $\epsilon_i$  ise ortalaması sıfır, varyansı  $\sigma^2$  olan normal dağılışa sahip şansa bağlı hata değişkenidir.

Yapay sinir ağları ve çoklu doğrusal regresyon modellerin karşılaştırılmasında ise düzeltilmiş belirleme katsayısı ( $R^2$ ), hata

kareler ortalamasının karekökü (Root Mean Square Error-RMSE), ortalama mutlak sapma (Mean Absolute Deviation-MAD) ve ortalama mutlak yüzde hata (Mean Absolute Percentage Error-MAPE) kriterleri kullanılmıştır. Bu kriterlere göre yüksek  $R^2$  ve düşük RMSE, MAD ve MAPE değerleri iyi uyum sağlayan modeli belirlemektedir. Söz konusu eşitlikler ve eşitliklerde yer alan terimler aşağıda verilmiştir.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad MAD = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

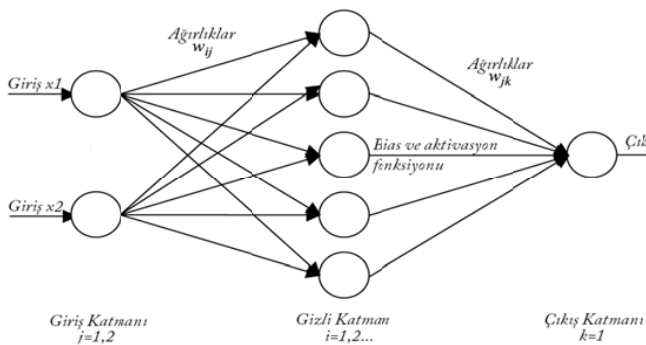
$$MAPE = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} \times 100, (y_i \neq 0)$$

$$\bar{R}^2 = 1 - (1 - R^2) \frac{(n - 1)}{(n - p - 1)}$$

Burada, n: Kayıt sayısı,  $y_i$ : Gözlenen değer,  $\hat{y}_i$ : Tahminlenen değer, p: Bağımsız değişken sayısını göstermektedir.

## BULGULAR

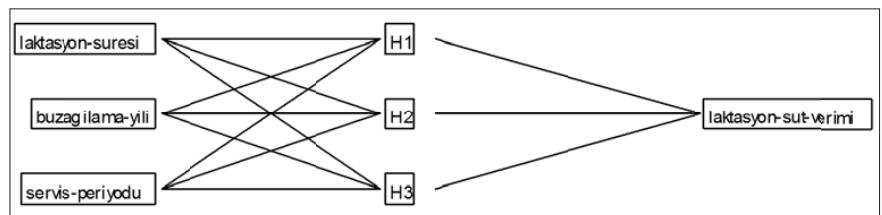
Çoklu doğrusal regresyon ve yapay sinir ağları modellerinin karşılaştırılmasında kullanılan belirleme katsayısı, hata kareler ortalamasının karekökü, ortalama mutlak sapma ve ortalama mutlak yüzde hata kriterlerinin değerleri Tablo 1'de verilmiştir. Sonuçlara göre yapay sinir ağları modeli kullanıldığında ilk beş laktasyonda  $R^2$  değerleri 0.62 ile 0.85 arasında elde edilmiştir. Çoklu doğrusal regresyon modeli ile bu değerler 0.30 ile 0.75 arasında değişmiştir. Buna göre yapay sinir ağları modeli ile çoklu doğrusal regresyon modelinden daha yüksek  $R^2$  değerleri bulunmuştur. Diğer yandan yapay sinir ağları ile RMSE değerleri 480.9-1682.8, MAD değerleri 325.2-1381.7 ve MAPE değerleri 6.1-20.2 sınırları arasında belirlenmiştir. Söz konusu kriterler çoklu doğrusal regresyon modeli



Şekil 2. Çok katmanlı ileri beslemeli ağ yapısı  
Fig 2. Multilayer feedforward neural network

Şekil 3. Laktasyon süresi, buzağılama yılı ve servis periyodu girişli, laktasyon süt verimi çıkışlı ve üç gizli düğümlü (H1, H2, H3) çok katmanlı ileri beslemeli yapay sinir ağı modeli

Fig 3. Lactation length, calving year and service period inputs, lactation milk yield output and three hidden nodes (H1, H2, H3), multilayer feedforward neural network model



**Tablo 1.** Yapay sinir ağları ve çoklu regresyon modellerine ilişkin performans kriterleri  
**Table 1.** Performance criteria for neural network and multiple linear regression models

Laktasyon	N	Yapay Sinir Ağları				Çoklu Regresyon			
		R <sup>2</sup>	RMSE	MAD	MAPE	R <sup>2</sup>	RMSE	MAD	MAPE
1	65	0.85	910.55	748.37	10.86	0.75	2781.26	2034.70	24.70
2	69	0.63	1682.78	1381.72	20.24	0.53	3008.69	2458.29	35.64
3	63	0.77	1399.142	1133.21	17.38	0.66	2984.69	2355.33	34.78
4	87	0.76	1447.38	1155.19	16.68	0.42	2213.02	1829.61	28.63
5	21	0.62	480.91	325.17	6.10	0.30	1964.84	1576.59	30.95

*N: İlgili laktasyondaki hayvan sayısı, R<sup>2</sup>: Düzeltilmiş belirlenme katsayısı, RMSE: Hata kareler ortalamasının karekökü, MAD: Ortalama mutlak sapma, MAPE: Ortalama mutlak yüzde hata*

kullanıldığında ise daha yüksek saptanmıştır. RMSE, MAD ve MAPE değerleri sırası ile 1964.8-3008.7, 1576.6-2458.3 ve 24.7-35.6 olarak bulunmuştur (Tablo 1).

## TARTIŞMA ve SONUÇ

Bu çalışmada Siyah Alaca ineklerin laktasyon süt verimi kayıtları üzerine çoklu doğrusal regresyon ve yapay sinir ağları modeli uygulanmıştır. Yapay sinir ağlarının pek çok modelleri arasından, yaygın ve başarılı sonuçlar veren bir model olması nedeniyle de ileri beslemeli geri yayımlı yapay sinir ağı kullanılmıştır. Yapılan ön çalışmalara göre, süt verimine etkili olduğu düşünülen tüm faktörlerden istatistiksel olarak önemli bulunan bağımsız değişkenler kümesi ayrıca çoklu doğrusal regresyon modeli ile de analiz edilmiş ve yapay sinir ağları modellerinin sonuçları ile karşılaştırılmıştır.

En iyi uyumu veren modelin belirlenmesinde belirleme katsayısı (R<sup>2</sup>) ve hata varyansını ifade eden kriterler (RMSE, MAD ve MAPE) dikkate alınmıştır. Buna göre, yüksek R<sup>2</sup> ve düşük RMSE, MAD ve MAPE değerli model iyi model olarak dikkate alınmıştır. Çalışmada elde edilen R<sup>2</sup> değerleri yapay sinir ağları için 0.62 ile 0.85 arasında, çoklu doğrusal regresyon kullanıldığında ise 0.30 ile 0.75 arasındadır. Bu değerler Grzesiak ve ark.<sup>8</sup> tarafından bildirilen değerlerden düşük; Edriss ve ark.<sup>9</sup> tarafından bildirilen değerlerden yüksek bulunmuştur. Öte yandan her iki yöntem için de ilk dört laktasyon süt verimi için Edriss ve ark.<sup>7</sup> ile Grzesiak ve ark.<sup>8</sup> bildirdiği değerlerden daha yüksek RMSE değerlerine ulaşılmıştır. İnceleme konusu olan çalışmalarda ulaşılan performans değerlerine göre yapay sinir ağları modelinin çoklu doğrusal regresyon analizinden daha etkin bir performans sergilediği görülmektedir. Bu sonuç Edriss ve ark.<sup>7</sup> ve Grzesiak ve ark.<sup>8</sup> ile uyumlu bulunmuştur. Böylece, yapay sinir ağlarının çoklu doğrusal regresyon analizine göre daha yüksek bir öngörüye sahip olduğu ve daha az hataya sahip sonuçlar verdiği görülmektedir. Bu durumda regresyon analizi varsayımlarının sağ-

lanamadığı ve analizin gerçekleştirilemediği durumlarda yapay sinir ağlarının rahatlıkla tercih edilebileceği açıktır.

Ülkemizde ise yapılan hayvancılık çalışmalarında yapay sinir ağlarının kullanımı bakımından henüz yok denecek sayıda araştırma mevcuttur. Bu çalışmada süt veriminin tahminlenmesinde yapay sinir ağlarının çoklu doğrusal regresyon modeli ile karşılaştırmalı analizi yapılmıştır. Bundan sonra hayvancılık alanında yapılacak çalışmalar ile yapay sinir ağlarının geleneksel yaklaşımlara üstünlüğünü araştıran daha fazla sayıda araştırmacının literatüre katkıda bulunacağı düşünülmektedir.

## KAYNAKLAR

- Düzgüneş O, Akman N:** Varyasyon Kaynakları. Ankara Üniv. Ziraat Fak. Ders Kitabı: Yayın No: 1408, s. 146, 1995.
- Tuncel, E:** Hayvan Islahı. Uludağ Üniv. Ziraat Fak, Ders Notları, No: 46, Bursa s. 217, 1994.
- Çilek S, Bakır G:** Malya Tarım İşletmesi'nde yetiştirilen Esmer ırkı ineklerin süt verim özellikleri ve bunlar üzerine bazı çevre faktörlerinin etkileri. *Kafkas Univ Vet Fak Derg*, 16 (2): 347-350, 2010.
- Atıl H:** Ratio and regression factors for predicting 305 days production from part lactation milk records in a herd of Holstein Friesian cattle. *J Biol Sci*, 2 (1): 31-37, 1999.
- Duru S, Tuncel E:** Siyah Alaca sığırlarda kuruda kalma süresi, servis periyodu ve ilkine buzağılama yaşı ile bazı süt verimi özellikleri arasındaki ilişkiler. *Uludağ Üniv Ziraat Fak Derg*, 18, 69-79, 2004.
- Akbulut Ö, Haussman H:** Buzağılama aralığının süt verim özelliklerine etkisi. *Atatürk Üniv Ziraat Fak Derg*, 25, 1-13, 1994.
- Duman N:** Yapay sinir ağları ve bir uygulama. *Yüksek Lisans Tezi*, Cumhuriyet Üniv. Sosyal Bil. Enst., 2006.
- Grzesiak W, Lacroix R, Wójcik J, Blaszczyk P:** A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records *Can J Anim Sci*, 83, 307-310, 2003.
- Edriss MA, Hosseinnia P, Edrisi M, Rahmani HR, Nilforoos MA:** Prediction of second parity milk performance of dairy cows from first parity information using artificial neural network and multiple linear regression methods. *Asian J Anim Vet Adv*, 3 (4): 222-229, 2008.
- JMP:** JMP Users Guide. Version 5.0.1.2, SAS Institute Inc., Cary, NC, 2003.