



RİDGE REGRESYON ANALİZİ VE ZOOTEKNİ ALANINDA KULLANIMI

Metehan Sarp ESKİTOROS

Yüksek Lisans Tezi

Zootekni Anabilim Dalı

Danışman: Prof. Dr. Eser Kemal GÜRCAN

2020

T.C.
TEKİRDAĞ NAMIK KEMAL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

RİDGE REGRESYON ANALİZİ VE ZOOTEKNİ
ALANINDA KULLANIMI

Metehan Sarp ESKİTOROS

ZOOTEKNİ ANABİLİM DALI

DANIŞMAN: Prof. Dr. Eser Kemal GÜRCAN

TEKİRDAĞ-2020

Her hakkı saklıdır.



Bu tezde görsel, işitsel ve yazılı biçimde sunulan tüm bilgi ve sonuçların akademik ve etik kurallara uyularak tarafımdan elde edildiğini, tez içinde yer alan ancak bu çalışmaya özgü olmayan tüm sonuç ve bilgileri tezde eksiksiz biçimde kaynak göstererek belirttiğimi beyan ederim.

Metehan Sarp ESKİTOROS

Prof. Dr. Eser Kemal GÜRCAN danışmanlığında, Metehan Sarp ESKİTOROS tarafından hazırlanan "Ridge Regresyon Analizi ve Zootekni Alanında Kullanımı" isimli bu çalışma aşağıdaki jüri tarafından Zootekni Anabilim Dalında Yüksek Lisans tezi olarak oy birliği ile kabul edilmiştir.

Jüri Başkanı : Prof. Dr. Eser Kemal GÜRCAN

İmza:

Üye : Doç. Dr. Aylin AĞMA OKUR

İmza:

Üye : Doç. Dr. Süleyman KÖK

İmza:

Fen Bilimleri Enstitüsü Yönetim Kurulu Adına

Doç. Dr. Bahar UYMAZ

Enstitü Müdürü

ÖZET

Yüksek Lisans Tezi

RİDGE REGRESYON ANALİZİ VE ZOOTEKNİ ALANINDA KULLANIMI

Metehan Sarp ESKİTOROS

Tekirdağ Namık Kemal Üniversitesi

Fen Bilimleri Enstitüsü

Zootekni Anabilim Dalı

Danışman: Prof. Dr. Eser Kemal GÜRCAN

Çalışmada çoklu doğrusal regresyon analizinde kullanılan Ridge Regresyon (RR) yöntemi açıklanmaya çalışılmıştır. Çalışmada RR yönteminin; bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olduğu zamanlarda En Küçük Kareler (EKK) yöntemine göre tercih edilerek kullanılmasının önemi özellikle Zootekni alanında yapılmış çeşitli çalışmalar ile gösterilmiştir. Çalışmada ayrıca çoklu doğrusal bağlantı sorunu, çoklu doğrusal bağlantı sorununun belirlenme yolları, RR yönteminde kullanılan yanlılık sabiti olan (k) değerinin belirlenme yöntemleri belirtilerek bu yöntemin nasıl uygulandığı ve bulunan değerlerin nasıl yorumlanacağı açıklanmıştır. Çalışmanın kolay açıklanabilmesi için Zootekni alanında verilen basit bir örnekle örneklendirilmiştir. Bu örnek için Esmer sığırların karkas ağırlıkları beş farklı vücut ölçüsü kullanılarak modellenmiştir. Çalışmada çoklu doğrusal bağlantı sorunu varlığında RR analizine göre elde edilen model ile aynı koşulda EKK yöntemi ile elde edilen iki model karşılaştırılmıştır. Çalışmada her iki modelinde belirleme katsayısı düşük olup ($R^2 < 0,25$), bir bağımlı ve beş farklı bağımsız değişkenin beraber kullanıldığı model için ele alınan özellik için istatistik olarak çok anlamlı bulunmamıştır. Buna rağmen RR yönteminin kullanılması beklenildiği gibi varyans artırıcı faktör (VIF) değerlerini azalttığı, çoklu doğrusal bağlantı sorunundan kaynaklanan bağımsız değişkenlere ait regresyon katsayılarının standart hata değerlerini EKK yöntemine göre düşürdüğü görülmüştür. Bu şekilde kullanılacak yöntem seçiminin isabetin artırılması için özellikle çoklu doğrusal bağlantı sorununun varlığında, doğru yöntemin kullanılmasının ne derece önemli olduğu ve bu konuda araştırmacıların daha dikkatli ve seçici davranmalarının gerekliliği bu çalışma ile vurgulanmıştır.

Anahtar kelimeler: Çoklu Regresyon, Ridge Regresyon, En Küçük Kareler Analizi, Sığır Karkas

2020, 53 sayfa

ABSTRACT

MSc. Thesis

RIDGE REGRESSION ANALYSIS AND ITS USE IN ANIMAL SCIENCE

Metehan Sarp ESKİTOROS

Tekirdag Namık Kemal University
Graduate School of Natural and Applied Sciences
Department of Animal Science

Supervisor: Prof. Dr. Eser Kemal GÜRCAN

In the study, the Ridge Regression (RR) method used in multiple linear regression analyze was tried out to be explained. The significance of RR method usage preferred according to the method of Least Square method (LS) in case having multicollinearity between multiple variables, was presented especially with several researches in the area of animal science. On the other hand, in the study, the multicollinearity problem, determination of multicollinearity problem, determination methods of bias stability (k) value used in RR method and application of the method and the way of interpretation of the presented values. So as to be easily explained, it was exemplified via a simple sample in the area of animal science. It was modeled by using five different body measurements of Brown Swiss cattle for the sample. In the study, in case of a multicollinearity, the determination coefficient was found low for both models obtained with LS method under the same environmental condition of the model obtained with RR analyze ($R^2 < 0,25$) and the approached feature for the model which was used with a dependent and five independent variables and was not found statistically significant. Nevertheless, the usage of RR method decreased the variance inflation factor (VIF) as it was expected and decreased the standard error values belonging to coefficient of regression for independent variables sourced from multicollinearity problems to LS method. For the method choice, it was emphasized as significant that the usage of correct method in case of multicollinearity for increasing of accuracy and the necessity of attentive and selective behaviors of researchers.

Key words: Multiple Regression, Ridge Regression, Least Square Analysis, Cattle Carcass

2020, 53 pages

İÇİNDEKİLER

	Sayfa
ÖZET	i
ABSTRACT	ii
İÇİNDEKİLER	iii
ÇİZELGE DİZİNİ	iv
ŞEKİL DİZİNİ	v
SİMGELER ve KISALTMALAR	vi
ÖNSÖZ	vii
1. GİRİŞ	1
2. LİTERATÜR TARAMASI	3
3. MATERYAL ve YÖNTEM	25
3.1. Araştırma Materyali	25
3.2. Araştırma Yöntemi	25
4. ARAŞTIRMA BULGULARI ve TARTIŞMA	27
5. SONUÇ ve ÖNERİLER	41
6. KAYNAKLAR	43
ÖZGEÇMİŞ	46

ÇİZELGE DİZİNİ

Çizelge 4.1. Kullanılan örnek veri setini oluşturan hayvanların karkas ağırlığı ve çeşitli vücut ölçülerine ait tanımlayıcı istatistikleri	27
Çizelge 4.2. Vücut ölçüleri arasındaki korelasyon katsayıları	28
Çizelge 4.3. EKK analizi sonucu bağımsız değişkenlere ait varyans arttırıcı (VIF) değerleri	29
Çizelge 4.4. Korelasyon katsayılarının özdeğerleri	30
Çizelge 4.5. Çeşitli k-değerleri için elde edilen standardize edilmiş Ridge regresyon (RR) katsayıları	31
Çizelge 4.6. Ridge Regresyon katsayıları, standart hataları ve VIF değerleri	33
Çizelge 4.7. Ridge regresyon (RR) ve En Küçük Kareler (EKK) yönteminin karşılaştırılması	33

ŞEKİL DİZİNİ

Sayfa

Şekil 4.1. Ridge (trace) iz grafiği yardımı ile k-sabitinin belirlenmesi 31

Şekil 4.2. Varyans Arttırıcı Faktör (VIF) grafiği yardımı ile k-sabitinin belirlenmesi 32



SİMGELER ve KISALTMALAR

β_0	: Modelde yer alan kesme katsayısı (İntercept)
β_i	: Modelde yer alan parametrelere ait kısmi regresyon katsayısı
R^2	: Belirleme katsayısı
R^2_d	: Düzeltilmiş belirleme katsayısı
ϵ	: Modele ait hata
N	: Kullanılan gözlem sayısı
S	: Hata kareler ortalamasının standart sapması
k	: Yanlılık sabiti
r	: Korelasyon katsayısı
X_n	: Bağımsız değişken
Y	: Bağımlı değişken
CI	: Koşul indeksi
CN	: Koşul sayısı
CY	: Cidago yüksekliği
EKK	: En küçük kareler
GÇ	: Göğüs çevresi
KA	: Karkas ağırlığı
KSY	: Kuyruk sokumu yüksekliği
PCR	: Temel bileşenler regresyonu
RR	: Ridge regresyonu
SY	: Sağrı yüksekliği
TV	: Tolerans değeri
VIF	: Varyans arttırıcı faktör
VU	: Vücut uzunluğu
P	: Modeldeki parametre sayısı

ÖNSÖZ

Bu çalışmada bir bağımlı değişken ve bu bağımlı değişkeni açıkladığı düşünülen çok sayıda bağımsız değişkeni kullanıldığı çoklu doğrusal regresyon analizinde bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olduğunda kullanılması önerilen Ridge regresyon analizinin kullanılmasının önemi vurgulanmıştır. Çalışmada Ridge regresyon analiz yönteminin nasıl uygulanacağı basit bir somut örnek kullanılarak açıklanmaya çalışılmıştır. Ayrıca bu konuyla ilgili Zootekni alanında yapılan çalışmalar ve sonuçları bir bütün olarak sunulmaya çalışılmıştır. Bu şekilde çoklu doğrusal regresyon analizi ile yapılacak çalışmalarda kullanılacak yöntem seçiminin ne derece önemli olduğu bu konuda araştırmacıların daha dikkatli ve seçici davranmalarının gerekliliğine dikkat çekilmiştir. Bu yönü ile yapılan çalışmasının teorik ve aynı zamanda araştırmacılara pratik anlamda katkı sunacak bir çalışma olması sevindiricidir. Bu nedenle çalışmanın yapılmasında bana yardımcı olan başta danışmanım Prof. Dr. Eser Kemal GÜRCAN'a ve yetişmeme katkısı bulunan tüm hocalarım ve beni sabırla destekleyen aileme teşekkürü borç bilirim.

Mayıs 2020

Metehan Sarp ESKİTOROS

1. GİRİŞ

Dünya nüfusun hızla artmasına karşın tarım alanlarının hızla azalması, bitkisel ve hayvansal ürünlerin bu artış hızını karşılayacak seviyede olmaması nedeniyle insanoğlu yetersiz beslenme sorunu ile karşı karşıyadır. Bu sorunun çözümü ise tarımsal kaynakların efektif kullanılarak topluca elde edilen verim seviyelerinin artmasına bağlıdır. Özellikle hayvanlardan elde edilen verim seviyesini artırmak için genetik yapının ve çevre şartlarının uygun bileşime getirilmesi gerekmektedir. Bu amaçla yapılan çalışmalarda verim özellikleri ile bu verime etkili olduğu düşünülen faktörleri kullanarak çeşitli modellemelerin yapılması ve en doğru modele ait tahminler ile verim özelliklerinin belirlenmesi istenen hedefe ulaşmak için önemli bir adım olacaktır. En doğru modelin belirlenmesi ve bu modeli kullanılarak yapılacak tahminler de ki isabet en doğru istatistik analiz tekniklerinin kullanılmasına bağlıdır.

Yanlış veya uygun olmayan istatistik analiz teknikleri kullanarak doğru sonuçlar elde etme olasılığı bulunmamaktadır. Bu konuda gösterilecek hassasiyet tüm çalışmaların doğruluğu ve sağlanacak ilerlemenin hızına direkt etkide bulunacaktır.

Hatta bu konuda Dr. Israeli isimli araştırmacı “istatistik metotlar kullanarak nasıl yalan söylenir” isimli bir kitap yazmıştır. Araştırmacı; kitabında istatistik metotlarının yalan söylemediği, fakat bilerek veya bilmeyerek araştırmacıların istatistik metotlarını yanlış kullandıkları için yanlış sonuçlara ulaştığını ifade etmiştir (aktaran Soysal 2012).

Bu tez araştırmasında kullanılan istatistiksel analiz yöntemi ise tek bir bağımlı değişken ve bu bağımlı değişkeni açıkladığı düşünülen çok sayıda bağımsız değişken yardımı ile bağımlı değişkenin modellenmesi işlemidir. Bu amaçla çoklu doğrusal regresyon analiz yöntemi sıklıkla kullanılmaktadır. Çoklu doğrusal regresyon analizi için bazı varsayımların gerçekleştiği varsayılarak en küçük kareler analiz yöntemi kullanılır. Eğer bu varsayımlar gerçekleşmediği ve bağımsız değişkenler arasında çoklu doğrusal bağlantı var ise yanlış bir tahmin metotlarından biri olan Ridge Regresyon (RR) yönteminin kullanılmasının önemi vurgulanmıştır. Çalışmada öncelikle çoklu doğrusal regresyon içinde yer alan En Küçük Kareler (EKK) yöntemi, çoklu doğrusal bağlantı sorunu, çoklu doğrusal bağlantı sorununun belirlenme yolları, çoklu doğrusal bağlantı sorununun çözümü, bu durumda kullanılan yöntemlerden biri olan RR yönteminin ve bu yöntemde kullanılan yanlışlık sabiti olan (k) değerinin belirlenme yöntemleri sırasıyla açıklanmıştır.

Çalışmada bu yöntemin nasıl uygulandığı ve bulunan değerlerin nasıl yorumlanacağı hakkında bilgiler verilmiştir. Bununla birlikte çoklu doğrusal regresyon analiz yöntemi Zootekni alanı içinde çeşitli çalışmalarda kullanılmıştır. Yapılan bazı çalışmalarda ise EKK ile RR yöntemi karşılaştırılmıştır. Bu çalışmaların da bir arada değerlendirilerek elde edilen sonuçların toplu bir sonuç olarak tartışılması araştırmacılara ayrı bir fayda sağlayacaktır. Bu şekilde bağımlı değişkenlerin çok sayıda bağımsız değişken tarafından modellenmesi yapılacağı çalışmalar için kullanılacak yöntem seçiminin önemi sebebiyle araştırmacıların daha dikkatli ve seçici davranmalarının gerekliliği vurgulanmış olacaktır.



2. LİTERATÜR TARAMASI

Araştırmacılar yıllardır olaylar arasında sebep-sonuç ilişkisi kurarken çoklu doğrusal regresyon yöntemlerini kullanmıştır. Bu yöntemler kullanılırken bazı şartların mevcut olması gerekir. Bu karşılaşılan şartlara göre araştırmacı kullanacağı yöntemi belirlemelidir. Araştırmacı bir bağımlı değişken ve bu değişkeni açıkladığı düşündüğü çok sayıda bağımsız değişkeni ele aldığı anda bağımsız değişkeni en iyi açıklayan çoklu doğrusal regresyon denklemini kullanmak ister. Bu nedenle çoklu doğrusal regresyonun en küçük kareler (EKK) yönteminin şartlarını yerine getirmediği veya çoklu doğrusal bağlantı sorunu varlığında EKK yöntemi yerine sapmalı tahmin yöntemlerinde biri olan Ridge Regresyon (RR) yöntemi kullanılabilir (Büyüküysal, 2016; Hoerl ve Kennard, 1970a, 1970b; Orhunbilge, 1996, 2017).

Çoklu doğrusal regresyonda bir adet bağımlı değişken (cevap değişkeni) ve bu değişkeni açıklamak için çok sayıda bağımsız değişkenler kullanılarak bir modelleme yapılır. Burada bağımlı değişkendeki değişkenliğin bağımsız değişkenler tarafından oluşturulan kısmının belirlenmesi, bağımlı değişken üzerinde etkili olan ve olmayan bağımsız değişkenlerin saptanması ve bağımsız değişkenleri kullanarak bağımlı değişkeni tahmin edilmesi amaçlanır. Bu amaçla çoklu doğrusal regresyon denkleminde yer alan parametre tahminleri yapılır. Çoklu doğrusal regresyon analizinde parametre tahmini genelde bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olarak bilinen (multicollinearity) olmadığında EKK yöntemine göre yapılır. EKK yönteminin temel kuramı hata kareler toplamını minimum yapan parametre tahmincilerinin belirlenmesidir.

Çoklu doğrusal regresyon analizinde bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu var ise EKK yöntemi yerine yanlı tahmin yöntemlerinden olan RR veya Temel Bileşenler Regresyonu (PCR) teknikleri kullanılması önerilmektedir. Çoklu doğrusal bağlantı durumunda, EKK tekniği ile yapılan parametre tahminleri yansız fakat parametrelere ait standart hatalar büyük olduğundan populasyon ölçeğinde parametreler istatistik; olarak anlamlı bulunmamaktadır ($P > 0,05$). Bunun yerine bu tahminlere küçük bir yanlılık sabiti (k) ekleyerek parametrelerin varyansları dolayısıyla standart hata değerleri küçülmekte ve çok daha anlamlı modeller elde edilebilmektedir ($P < 0,05$) (Büyüküysal, 2016; Hoerl ve Kennard, 1970a, 1970b; Orhunbilge, 1996,).

Çoklu doğrusal regresyon analizinde modelde kalması gereken veya çıkması gereken bağımsız değişkenlerin belirlenmesinde modellerin veya regresyon denklemlerinin çoklu

belirleme katsayısı olan R^2 değerlerine bakarak veya modelin istatistik olarak önem seviyesi değerlendirilerek karar verilir. Genel kanı olarak bağımlı değişken ile yüksek ilişkili olan bağımsız değişkenlerin modelde yer alması modelin anlamlılığını artıracaktır.

Çoklu doğrusal regresyon analizi genellikle EKK yöntemine göre yapılır. Bu yöntemi kullanmak için bazı varsayımların gerçekleşmiş olması gerekir. Bu varsayımlar kısaca hataların beklenen değer ortalaması 0 olması, tahmin hataları (e) arasında bağımlılık (otokorelasyon) olmaması, tahmin hata varyanslarının eşit olması, hatalar ile bağımlı değişken arasında korelasyon olmaması, hatalar ile bağımsız değişkenler birbirinden bağımsız olması, bağımsız değişkenler arasında anlamlı bir ilişki olmaması yani çoklu doğrusal bağlantının olmaması (multicollinearity), bağımsız değişkenleri varyansının sıfırdan büyük olması ve hataların dağılımının normal dağılım göstermesidir (Albayrak, 2005; Orhunbilge, 1996, 2017). Araştırmacı uygun yöntemi kullanmadan önce elindeki verilerde söz edilen varsayımların durumunu araştırmalı ve ona uygun yöntemi belirlemelidir. Aksi halde kullandığı yöntemin yanlışlığından dolayı yapmış olduğu modelleme yanlış olacaktır.

Regresyon analizinde hatalar birbirine bağımlı (otokorelasyon) olmamalıdır. Bunun testi için ($n \geq 15$) olması durumunda Durbin-Watson testi ile karar verilir. Bu teste ait eşitlik 2.1'de verilmiştir. Bu değer (1,5-2,5) arasında bulunursa otokorelasyon olmadığına hükmedilir (Karagöz, 2016).

$$d = \left[\frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \right] \quad (2.1)$$

Hata varyanslarının eşit olmasına eşit varyanslılık (homoscedasticity) denir ve çoklu regresyon analizinde şartlardan biridir. Hata varyanslarının farklı olmasına ise (heteroscedasticity) denir. Bu durum ise regresyon analizlerinde istenmez. Bu koşula uyum sağlanıp sağlanmadığı ise x ekseninde tahminle bulunan \hat{Y} ve y ekseninde ise standardize edilmiş hata değerleri (e / \hat{S}_{yx}) gösterilerek elde edilen serpilme grafiğine bakarak karar verilebilir. Standardize hatalar 0 etrafında simetrik dağılım gösterirse eşit varyanslılık vardır denir. Simetrik olmayan bir dağılım var ise yani standardize hataların 0 etrafında azalan veya artan şeklinde bir dağılım var ise bu sefer farklı varyanslılık olduğuna karar verilir (Orhunbilge, 2017).

Çoklu doğrusal bağlantı sorunu (multicollinearity) bu sorun bir den fazla bağımsız değişken kullanarak yapılan modellemede bağımsız değişkenler arasında basit korelasyon katsayılarının ($r \geq 0,75$) büyük olması yani bağımsız değişkenlerin istenmeyecek şekilde birbirine bağımlı olması durumunda karşılaşılır. Bu durum çoklu doğrusal regresyon analizinde istenmez. Bu sorunun varlığında tek bir değer bile değiştirilirse kısmi regresyon katsayılarında büyük farklılıklar olur, tahmin edilen regresyon katsayıları belirsiz ve standart hataları sonsuz olabilir. Ayrıca, modelin çoklu belirleme katsayısı R^2 yüksek ama bağımsız değişkenlerin çok azı yapılan t testinde istatistik olarak önemli ($P < 0,05$) çıkar, bağımsız değişken ile bağımlı değişken arasındaki ilişki beklenmedik şekilde ters (mantık dışı) bulunabilir (Albayrak, 2005; Orhunbilge, 1996, 2017). Bu durumda bu sorunu çözmek için bazen araştırmacı EKK yöntemi kullanmadan vazgeçmemek için toplanmış verilerdeki birim (n) sayılarını artırmak veya modelde yer alan ve birbiri ile ilişkili bulunduğu bazı bağımsız değişkenleri modelden çıkarmak veya birleştirmek yoluna başvurulabilir. Buna alternatif olarak böyle bir durumda yansız tahmin yapan EKK yöntemi yerine çalışmadaki birim sayısını ve değişkeni azaltmadan yanlı tahmin yapan RR yöntemini kullanmayı tercih edebilir. Bu ifadede tahmin edilen istatistiğin beklenen değeri populasyon parametresine eşitse buna yansız tahmin aksi durumda yanlı tahmin denir.

Ele alınan veri setinde çoklu doğrusal bağlantı sorunu tespit edilmesi için çeşitli ölçütlere bakılabilir. Bu ölçütler birbiri ile paralel sonuçlar vermekle birlikte birkaç tanesinin görülmesi çoklu doğrusal bağlantının varlığı hakkında karar vermek için yeterli görülebilir.

Bu durumu belirleme yolları, öncelikle bağımsız değişkenler arasındaki ilişkiyi göstermek için basit korelasyon katsayılarına bakılabilir. Bu korelasyon katsayısı ($r \geq 0,75$) büyük ise çoklu doğrusal bağlantıdan kuşkulunıp diğer ölçütler ele alınmalıdır. Kullanılan modele yeni bir bağımsız değişken eklendiğinde belirleme katsayısında önemli bir artış olmaması, ikili korelasyonların istatistik olarak önemli bulunmuşken çoklu veya kısmi korelasyon katsayılarının istatistik olarak anlamsız bulunması çoklu doğrusal bağlantıda durumundan araştırmacıyı uyaran ölçütlerdir. Bu ölçütlerin yanında bu konuda en etkin ve çok sık kullanılan başka ölçütlerde vardır. Bunlar varyans arttırıcı faktör ($VIF = \text{Variance Inflation Factor}$), bağımsız değişkenler için tolerans değeri ($TV = \text{Tolerance Value}$), koşul sayısı ($CN = \text{Condition Number}$) veya koşul indeksi ($CI = \text{Condition Index}$) değerleridir (Albayrak, 2005).

Varyans Arttırıcı Faktör (VIF) değerinin belirlenmesi, bu değer hesaplanması çoklu doğrusal bağlantının varlığının gösterilmesi için kabul edilen etkili yöntemlerden biridir. Bu değer tahmini regresyon katsayıları varyanslarının, bağımsız değişkenlerin doğrusal olarak birbirine bağımlı olmadığı duruma göre ne kadar artırdığını ortaya çıkarır. Bu değer bulunması eşitlik 2.2’de gösterilmiştir. Bu eşitlikte R_k^2 değeri, k bağımsız değişkeninin diğer bağımsız değişkenler ile arasındaki çoklu belirleme katsayısıdır. Bu değer 1’e yakın olması çoklu belirleme katsayısının 0’a yakın yani bağımsız değişkenler ile doğrusal bir ilişkinin olmadığını gösterir. Bu değer pratikte 10 ve üzerinden büyük olması ($R_k^2 \geq 0,90$) çoklu doğrusal bağlantının önemli bir göstergesi kabul edilir (Orhunbilge, 2017).

$$(VIF)_k = (1 - R_k^2)^{-1} \quad (2.2)$$

Çoklu doğrusal bağlantı göstergesi olarak kullanılan diğer ölçütler ise tolerans değeridir (TV) bu değer bulunması eşitlik 2.3’de gösterilmiştir. Bu değer çoklu belirleme katsayısına bağlı olup çoklu belirleme katsayısı büyüdükçe TV değeri küçülür buna karşın VIF değeri büyür.

$$TV = (1 - R_k^2) \quad (2.3)$$

Çoklu doğrusal bağlantı sorunu gösteren diğer ölçütler ise koşul sayısı (CN) veya koşul indeksi (CI) olarak bilinen kriterlerdir. Bunların hesaplanması eşitlik 2.4 ve 2.5’de gösterilmiştir. Koşul sayısı korelasyon katsayıları için bulunan en büyük özdeğerin en küçük özdeğere oranlanması ve koşul indeksi ise bu oranın karekökünün alınması şeklinde bulunur. Koşul sayısı 100-1000 arasında ve koşul indeksinin 10-30 arasında bulunması çoklu bağlantının işaretidir (Albayrak, 2005; Orhunbilge, 1996, 2017).

$$CN = \left(\frac{\lambda_{max}}{\lambda_{min}} \right) \quad (2.4)$$

$$CI = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} \quad (2.5)$$

Çok değişkenli doğrusal regresyon modelinde bir bağımlı değişken (Y) ve bu bağımlı değişkeni açıkladığı düşünülen iki yada daha fazla bağımsız değişken (X_1, X_2, \dots, X_n) bulunur. Bu modelin regresyon denklemi eşitlik 2.6’da verilmiştir.

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_n X_{ni} + \varepsilon \quad (2.6)$$

Burada Y bağımlı (cevap) değişkeni, X_2 ile X_3 ise bağımsız (açıklayıcı) değişkenler, ε hata terimi ve i indisi ise i 'inci gözlemdir. Modelde yer alan β_1 sabit terimi $X_{2i}, X_{3i}, \dots, X_{ni}$ değerleri sıfır iken Y'nin ortalama değeri veya modele alınmayan tüm değişkenlerin bağımlı değişkendeki ortalama etkisini verir. Modelde yer alan $\beta_2, \beta_3, \dots, \beta_n$ katsayılarına ise kısmi regresyon katsayıları denir. Kısmi regresyon katsayılarının anlamı ise β_2 X_2 hariç tüm bağımsız değişkenler sabit iken X_2 deki bir birimlik değişime karşılık bağımlı değişkendeki (Y) değişmeyi ölçer ayrıca β_2 değerini X_2 hariç tüm bağımsız değişkenler sabit iken Y'nin X_2 'ye göre eğimi şeklinde ifade edilebilir. Benzer şekilde diğer kısmi regresyon katsayıları tanımlanabilir (Karagöz, 2016).

Çoklu regresyon analizinde çoklu doğrusal bağlantı sorunu görülmediğinde EKK yöntemine göre EKK tahmin edicileri veya modelde yer alan parametreler, matris işlemleri veya çok bilinmeyenli denklem sisteminde yok etme metodu kullanılarak tahmin edilebildiği gibi bu konuda geliştirilen istatistik yazılım programları ile çok kısa sürede bulunabilmektedir. Bu çalışmada amaçlanan türetilmiş bir veri seti kullanarak çoklu dorusal bağlantı sorunu varlığında kullanılması önerilen RR yönteminin EKK yöntemi ile karşılaştırılması ve RR yönteminin daha pratik bir şekilde kullanılması göstermek için RR analizi bir istatistik program ile çözümü yapılırken elle çözüm yöntemleri günümüzde pratik olmadığı için verilmemiştir.

Yapılan çalışmalarda çoklu doğrusal bağlantı sorunu gözlemlenmişse bu sorunu aşmak için uygulanan çeşitli yöntemler vardır. Çoklu dorusal bağlantıya neden olduğu düşünülen bağımsız değişkenlerin modelden çıkarılması veya bazı bağımsız değişkenlerin birleştirme yoluna gidilmesi şeklinde yollar izlenirken bunların yanı sıra ilk olarak Hoerl ve Kennard (1970a, 1970b) tarafından geliştirilmiş olan RR gibi yanlı tahmin yöntemlerinden birini kullanarak analizi sürdürmek, önerilen yöntemlerden biridir. Bu nedenle çoklu regresyon analizinde analize başlamadan önce EKK yöntemi için istenen varsayımların gerçekleşip gerçekleşmediği test edilmelidir.

Çoklu doğrusal regresyon analizinde EKK için gerekli olan varsayımlar sağlanmışsa yapılan parametre tahminleri yansız aksi durumda ise kısmi regresyon katsayılarının varyans ve kovaryansları artış gösterir bu durumda standart hatlarda büyük olacağından t testi sonuçları istatistik olarak önemsiz çıkar ($P > 0,05$). Bu durumda RR yöntemi ile tahminlere

küçük bir yanlılık sabiti ekleyerek kısmi regresyon katsayılarının varyans ve kovaryansları düşürülebilir. Varyans –kovaryans matrisinin köşegen değerlerine küçük bir yanlılık sabiti (k) ilave etmek dışında RR ve EKK tekniklerinin çözümü aynıdır. Sonuç olarak yansız tahminle yüksek varyans veya k oranında yanlı tahminle düşük varyans elde edilmesi tercih edilmelidir (Albayrak, 2005; Karagöz 2016; Orhunbilge, 1996, 2017).

En Küçük Kareler regresyon modeli matris notasyonu halinde verilmek istenirse ilgili modelin matris notasyonuna göre yazılımı eşitlik 2.7’de verilmiştir. Bu eşitliğe göre Y =bağımlı değişken, X =bağımsız değişken, B =tahmin edilecek kısmi regresyon katsayıları, ε =hata terimidir. Burada EKK tekniğinde tahminlenen regresyon katsayıları ise eşitlik 2.8’de gösterilmiştir (Albayrak, 2005).

$$Y=XB+\varepsilon \quad (2.7)$$

$$\hat{B}=(X'X)^{-1}X'Y \quad (2.8)$$

İlk olarak RR da tüm değişkenler ortalamalarından farkı alınıp standart sapmasına bölünerek standardize edilip sonra bulunan regresyon katsayıları tekrar eski haline dönüştürülmelidir. Çoklu doğrusal bağlantı sorunu tespit edilmiş ise RR yöntemi uygulanacaksa korelasyon matrisinin köşegen değerlerine küçük bir yanlılık sabiti (k) eklenerek yanlı standartlaştırılmış regresyon katsayıları eşitlik 2.9’da verildiği gibi hesaplanır.

$$\tilde{B}=(X'X+kI)^{-1}X'Y \quad (2.9)$$

Bu şekilde bağımsız değişken korelasyon matrisinin köşegen elemanlarına pozitif sabit eklenmesindeki amaç matris şartlı sayısının önemli ölçüde küçültülmesidir. Burada $k=0$ için RR çözümü EKK çözümüne eşdeğer olduğundan RR tahmini EKK tahmininin bir doğrusal dönüşümü olarak ifade edilir. Bu tahminin yanlılık büyüklüğü beklenen değeri ve kovaryans matrisi sırasıyla eşitlik 2.10 ve eşitlik 2.11’de verilmiştir (Albayrak, 2005).

$$E(\tilde{B}-B)=[(X'X+kI)^{-1}X'X-I]B \quad (2.10)$$

$$V(\tilde{B})=(X'X+kI)^{-1}X'X(X'X+kI)^{-1} \quad (2.11)$$

Çalışmada çoklu doğrusal bağlantı sorunu olduğunun işaretleri görülürse bağımsız değişkenler arasındaki korelasyonların yüksek olması veya VIF değerlerinin 10 değerini geçmesi RR yöntemi uygulanabilir. Böyle bir tercih yapılmış ise bu sefer k olarak bilinen yanlılık sabitinin belirlenmesi gereklidir. Bu sabitin belirlenmesi için yöntemlerden biri yanlı standart regresyon katsayıları ve k arasında hesaplanan yanlı regresyon grafiğini kullanmaktır (Ridge Trace). Bu grafikte yanlı standart regresyon katsayılarının durağanlaştığı yerde x eksenine bir dikme indirilerek k değeri seçilir. Diğer bir yöntem ise VIF değerinin 10 yaklaştığı yerden x eksenine bir doğru indirilerek oradaki k değeri alınabilir. K değere genelde 0-1 arasında olmakta ve bu değer 1 e yaklaştıkça yanlılık artarken varyans azalmaktadır (Albayrak, 2005; Hoerl ve Kennard 1970a, 1970b; Orhunbilge, 1996).

Bu çalışmada kısaca çoklu doğrusal bağlantı durumunda EKK yöntemine göre kullanılması önerilen RR yönteminin genel açıklaması baş kısımda yapılarak bu yöntem ile zootekni alanında yapılan çalışmalar ve vardıkları sonuçlar ise sonraki kısımda sunulmaya çalışılmıştır.

Schoeman, Aziz ve Jordaan (2002) yaptıkları bir melezleme çalışmasından elde edilen 17 bin 258 tane buzağaların süttten kesilme ağırlık kayıtlarına ilişkin verilerin çoklu doğrusal bağlantı problemini ve bunun doğrudan, maternal ve katkısız genetik etkilerini tahmini üzerindeki etkisini araştırmak için kullanılmıştır. Sonuçlar, kararsız tahminlere neden olan hem doğrudan hem de maternal etkiler arasında neredeyse bağımlı olduğunu göstermiştir. Araştırmacılar çalışmada ridge regresyon modelini kullanarak çoklu doğrusal bağlantı sorununu çözmeye çalışmışlardır ve çalışmada 5 farklı ırk Afrikaner (AF), Hereford, Angus, Simentaler (ST) ve Charolais (CH) kullanmışlar ve buzağı grupları Hereford ve Angus arasında hiçbir değişiklik yapmamışlardır. Araştırmacılar kullanmış oldukları verileri cinsiyet (erkek ve dişi), yaş (2 ila 9 yıl arasında), sürü yıl mevsimi (SYM) ve ırk grupları olarak sınıflandırmışlardır. Araştırmacılar SYM alt sınıflarının bazılarını az sayıda gözlem içerdiğinden dolayı aynı çiftlikten bir sonraki yıl aynı sezonda yer alan SYM ile bir araya getirmişlerdir. Araştırmacılar en küçük kareler varyans analiz yöntemi ilk model olarak SYM'nin sabit etkileri, cinsiyet, yaş ve ortak değişken olarak süttten kesme yaşı ve tüm tek yönlü etkileşimlerini kullanmışlar ve etkisi olmayan tüm etkileşimleri ($P>0,05$) daha sonra modelden çıkarmışlardır. Araştırmacılar farklı ırklar arasındaki doğrudan etkilerin korelasyon sayılarının negatif ve -0,03 ile -0,58 arasında olduğunu ve ayrıca Afrikan ırkının maternal genetik etkisi (MAF) ile Hereford ve Angus ırklarının maternal genetik etkisi (MHA) arasındaki korelasyon haricinde, maternal etkiler arasındaki hataların da 0,50 olarak bulmuşlardır. Ayrıca araştırmacılar doğrudan ve

maternal etkiler arasındaki ırk içi korelasyonların pozitif ve 0,48 ile 0,92 arasında değiştiğini, ırklar arasındaki ilişkilerin hepsinin negatif ve -0,01 ile -0,45 arasında değiştiğini bulmuşlardır. Bu nedenle araştırmacılar olası bir çoklu doğrusal bağlantı sorununun varlığını açıkça gösteren birkaç yüksek derecede korelasyonlu değişken gözlemlemiştir. Araştırmacılar bunlar arasında Charolais ırkının doğrudan etkisi (DCH) ve Simental ırkının doğrudan etkisi (DST) ile DCH ve Charolais ırkının maternal etkisi (MCH), MHA ve Hereford ırkının maternal etkisi (HM), Simental ırkının maternal etkisi (MST) ve MHA ve özellikle DAF ve MAF arasındaki varyans arttırıcı değerlerin (VIF) 1,43 ile 6,67 arasında değiştiğini görmüşlerdir. Araştırmacılar çoklu doğrusal bağlantı sorununu çözmek amacıyla RR modelini kullanarak modele yanlılık sabiti ($k=0,8$) eklemiştir. Bu sayede araştırmacılar ekledikleri k yanlılık sabiti sayesinde yaptıkları melezleme çalışmasındaki çoklu doğrusal bağlantı sorununu ortadan kaldırmış ve daha anlamlı bir biyolojik yoruma sahip daha kararlı ve doğru tahminler gözlemlemiştir. Araştırmacılar çoklu doğrusal bağlantı sorununu ortadan kaldırmak amacıyla RR modelini önermişlerdir.

Aktaş ve Yılmaz (2003) çoklu regresyon analizinde, bağımsız değişkenler arasında çoklu doğrusal bağlantı durumunda EKK yöntemi yerine modele konulan bağımsız değişkenleri elemine etmek yerine yanlı tahminler yapmasına karşılık bağımsız değişkenlerin regresyon katsayılarının standart hatalarını daha küçülten RR yöntemini ve son zamanlarda RR yöntemine alternatifi görülen Liu yönteminin de kullanılabileceğini ifade etmiştir.

Albayrak (2005) çoklu regresyon tekniklerinden EKK ve yanlı tahmin tekniklerinden olan RR ve PCR teknikleri kullanılarak insanların beden ağırlıkları (bağımlı değişken) çeşitli vücut ölçüleri (bağımsız değişkenler) olmak üzere tahmin etmeye çalışmıştır. Yapılan çalışmada asıl amacın regresyon tekniklerinin etkinliklerini karşılaştırmak olduğu ifade edilmiştir. Çalışmada bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu varlığı öngörülerek RR ve PCR tekniklerinin EKK tekniğine göre daha etkin olacağı bildirilmiştir. Sonuç olarak yapılan çalışmada çoklu doğrusal bağlantı sorunu tespit edilirse, EKK tekniği ile bulunan sonuçların yanlış olacağı ve yanlı tahmin metotlarından olan RR ve PCR regresyon yöntemlerinin kullanılması gerekliliği vurgulanmıştır.

RR ile yapılan diğer bir çalışmada Pimentel, Queiroz, Carvalheiro, ve Fries (2007) regresyon analizinde çoklu doğrusal bağlantı sorununu araştırmışlardır. Bunun için araştırmacılar regresyon analiz yöntemlerinden RR yöntemini kullanmışlardır. Araştırmacılar tropikal ve subtropikal bölgelerde yetiştirilen melez buzağuların performansını etkileyen

parametreleri bir katkı maddesi, egemenlik, melez gücü ve epistatik etkiler ve bunların etkileşimlerini çevre etkileşimlerine göre genotip model olarak kullanmışlardır. Araştırmacılar yaptıkları çalışmada Hereford ve Nelore ırklarının melezlenmesi sonucu 1974 ile 1998 yılları arasında doğan 109 bin 614 tane buzağı üzerinde çalışmış ve bunları Brezilya'da 29 farklı çiftlik içinde 4 bin 665 yaş grubuna dağıtmışlardır. Araştırmacılar buzağuların büyük bir bölümünü (%43,7) suni tohumlama ile saf ırk olan Nelore ve Hereford ırklarından elde etmişlerdir. Araştırmacılar çoklu doğrusal bağlantı problemini çözmek için varyans arttırıcı değerlere, korelasyon matris değerlerine ve öz değerlere bakmışlardır. Her bir ortak değişken için elde edilen VIF değerlerinin incelenmesi için 5 tip uygulama ile üç kriterin karşılaştırılmasını yapan araştırmacılar tahmin edilen değerleri saf Hereford'dan saf Nelore'ye kadar sekizde birlik bir kısımda altı serbestlikte dokuz genotip için hesaplamışlardır. Öngörülen tüm değerler, ortalama bir çevresel etkiye gelen sabit (0,61 kg/gün) ile eklenmiştir. Araştırmacılar yaptıkları analizde çoklu doğrusal bağlantı sorununu açıkça görmüşlerdir. 33 öz değerden dokuzu (9 çevresel + 8 genotipik + 16 etkileşim değeri) 0,001'den düşük ve korelasyon matrisinin tekillik ne kadar yaklaştığını gören araştırmacılar her 5 farklı metod için farklı VIF değerleri bulmuşlardır. Araştırmacılar tahminlenen parametrelerin her 5 yöntemde de maksimum VIF değerlerinin sırasıyla ilk 4 parametre için [1. yöntemde (371.906,4 ile 405,6), 2. Yöntemde (371.906,4 ile 404,5), 3. yöntemde (371.906,4 ile 400,8), 4. yöntemde (371.906,4 ile 354,9), 5. yöntemde ise (371.906,4 ile 387,9)] 300'den büyük olduğunu ancak sadece 5. parametreler için [1. yöntemde (299,6), 2. yöntemde (269,6), 3. yöntemde (282,9), 4. yöntemde (293,3) ve 5. yöntemde ise (296,9)] 300'denküçük olduğunu görmüşlerdir. Sonuç olarak araştırmacılar RR modelinin sadece değerlerin yorumlanmasında değil aynı zamanda tahmin amacıyla regresyon analizi yapıldığında, çoklu doğrusal bağlantı problemlerinin üstesinden gelmek için daha doğru bir regresyon modeli olduğunu önermişlerdir.

Topal, Eyduran, Yağanoğlu, Sönmez, Keskin (2010). Çalışmalarında 91 adet sazan balığından elde ettikleri karkas ağırlığı (KA) bağımlı değişken, çeşitli vücut ölçüleri [total boy (TB), standart boy (SB), baş boyu (BB), vücut yüksekliğini (VY)] ise bağımsız değişken olarak balıklarda çoklu doğrusal regresyon analizi yapmışlardır. Bu analizi yaparken de vücut ölçüleri arasında çoklu doğrusal bağlantı durumu nedeniyle EKK regresyonu yerine RR ve PCR analiz yöntemlerini kullanarak üç yöntemi karşılaştırmışlardır. Araştırmacılar çalışmada metotları karşılaştırırken belirleme katsayısı (R^2), hata kareler ortalamasının standart sapması (S), hata kareler ortalaması ve varyasyon katsayısı kriterlerinden yararlanmışlardır. Buna göre EKK yöntemine ait ($R^2=0,90$, $S=19,58$), ridge metoduna göre ($R^2=0,89$, $S=20,25$) ve temel

bileşenler regresyon yöntemine göre ise ($R^2=0,87$, $S=22,12$) olarak bulmuşlardır. Çalışmada araştırmacılar karkas ağırlığını 92,51 gr, total boyu 21,97 cm standart boyu 18,33 cm, baş boyu 4,16 cm vücut yüksekliğini 4,81 cm olarak bulmuşlardır. Bağımsız değişkenler arasındaki korelasyon katsayılarına bakıldığında total boy ile standart boy arasında ($r=0,97$), total boy ile baş boyu arasında ($r=0,93$), baş boyu ile vücut yüksekliği arasında ($r=0,97$) olarak yüksek korelasyon katsayıları bulmuşlardır. Ayrıca bağımsız değişkenlerin VIF değerleri ise total boy için 29,75 standart boy için 28,67 baş boyu için 22,16 vücut yüksekliği için 22,34 olarak sınır olarak kabul edilen 10 değerinden yüksek bulmuşlardır. Bunun yanında çoklu doğrusal bağlantının varlığını gösteren diğer kriterlere de baktıklarında bu kriterlerinde hepsinin çoklu doğrusal bağlantı sorununun varlığını gösterdiğini ifade etmişlerdir. Çalışmada her üç yöntem ile bulunan modeller ise EKK yöntemi için karkas ağırlığı= $-136,39 + 10,86 TB - 9,06 SB + 8,66 BB + 12,66 VY + 21,19$ yaş ($R^2=0,90$), temel bileşenler regresyon yöntemine göre ise karkas ağırlığı= $-147,09 + 1,45 TB - 1,38 SB + 19,84 BB + 14,15 VY + 11,33$ yaş ($R^2= 0,87$) ve de RR yöntemine göre ise ($k=0,005$) karkas ağırlığı= $-138,42 + 8,91 TB - 6,80 SB + 10,49 BB + 13,08 VY + 18,92$ yaş ($R^2=0,89$) olarak bulmuşlardır. Araştırmacılar sonuç olarak çoklu doğrusal bağlantı varlığında EKK analiz yöntemi yerine Ridge ve temel bileşenler yöntemlerinin tercih edilmesini önermişlerdir.

Rathert, Üçkardeş, Narin, Aksoy (2011) Japon bıldırcınlarında yumurta iç kalite özelliklerinin tahmin edilmesinde kullanılan çoklu doğrusal regresyon metotlarından olan En Küçük Kareler (EKK) ile temel bileşenler regresyon yöntemini karşılaştırmışlardır. Çalışmada bıldırcın yumurtalarının iç kalite özellikleri olarak ak yüksekliği, ak genişliği, ak uzunluğu, sarı uzunluğu ve sarı yüksekliğini ölçmüşlerdir. Araştırma 20-24 haftalık yaşlarda olan bıldırcınlardan elde edilen 104 tane yumurtada yapılmıştır. Yumurtaların ağırlığı (gr), yumurta uzunluğu (mm), yumurta genişliği (mm), kabuk kalınlığı (mm) ve kabuk ağırlığı (gr) olarak ölçülmüştür. Yumurta derinliği, yumurta uzunluğu, yumurta sarısı derinliği, yumurta ak uzunluğu ve yumurta ak derinliği dijital kumpas yardımıyla ölçülmüştür. Yumurta sarısı ve akı yüksekliği mikrometre yardımı ile ölçülerek yumurta kalite karakteristiği Haugh birimine göre hesaplanmıştır. Çalışmada yumurta ağırlığı (X_1), yumurta genişliği (X_2), yumurta uzunluğu (X_3) ve şekil indeksi (X_4) bağımsız değişkenler olarak kullanılmıştır. Bu değişkenlerin bağımlı değişken olan ak yüksekliği, ak genişliği, ak uzunluğu, sarı uzunluğu ve sarı yüksekliğine olan etkileri hem EKK yöntemi hem de temel bileşenler regresyon yöntemi kullanılarak tahmin edilmiştir. Araştırmacılar bağımlı ve bağımsız değişkenler arasındaki korelasyon katsayılarını hesapladığında bağımsız değişkenler arasındaki korelasyon

katsayısını yumurta ağırlığı (X_1) - yumurta genişliği (X_2) arasında 0,76 ($P<0,01$); yumurta ağırlığı (X_1) – yumurta uzunluğu (X_3) arasında ise 0,66 ($P<0,01$) olarak yüksek korelasyonlar bulmuşlardır. Özellikler arasında EKK tekniği ve temel bileşenler analizi ile hesaplanan varyans arttırıcı değerleri (VIF) en küçük kareler yönteminde 10’u geçtiği temel bileşenler yönteminde ise VIF değerlerinin 10’dan düşük olduğu görülmüştür. Benzer şekilde öz değerler ve koşul indeksine bakıldığında da korelasyon öz değerlerinin 0’a yakın koşul indeksinin 30’dan büyük olduğu değerler gözlenmiştir. Bu da bağımsız değişkenler arasında çoklu doğrusal bağlantı sorununun varlığını göstermiştir. Araştırmacılar yumurta iç kalite özelliklerini modellemek için 4 farklı bağımsız değişkeni 2 farklı yöntem kullanarak karşılaştırmışlardır. Ak ağırlığının tahmini için kullanılan modelde her iki yönteme ait belirleme katsayıları $R^2=0,39$ ($P<0,01$); Ak genişliği tahmini için kullanılan modelde EKK ve temel bileşenler yöntemine göre sırasıyla belirleme katsayıları $R^2=0,68 - 0,66$ olup ($P<0,01$); Ak uzunluğu tahmini için kullanılan modelde her iki yönteme ait belirleme katsayıları $R^2=0,62$ ($P<0,01$); Sarı yüksekliği tahmini için kullanılan modelde her iki yönteme ait belirleme katsayıları $R^2=0,29$ ($P<0,01$); Sarı genişliği tahmini için kullanılan modelde her iki yönteme ait belirleme katsayıları $R^2=0,35$ ($P<0,01$) olarak belirlemişlerdir. Araştırmacılar sonuç olarak kullandıkları bağımsız değişkenler ile en yüksek yumurta iç kalite özelliklerinden ak genişliğini modellemişlerdir ($R^2=0,66$). Bu çalışmanın sonucunda yumurta iç kalite özelliklerinin yumurta kabuğunu kırmadan dış kalite özellikleri kullanılarak modellenebileceği ve bağımsız değişkenler arasında çoklu doğrusal bağlantı sorununun ortaya çıkmasının muhtemel olduğu bu nedenle EKK yöntemi yerine çoklu doğrusal bağlantı sorununu ortadan kaldıran temel bileşenler regresyon yönteminin kullanılmasının daha doğru ve güvenilir tahmin eşitlikleri oluşturduğu bildirilmiştir.

Üçkardeş, Efe, Narin, Aksoy (2012) yaptıkları bir araştırmada araştırmacılar Japon bildircinlerinde yumurta iç kalite kriterlerinden biri olan yumurta ak indeksini En Küçük Kareler yöntemi yerine RR yöntemini kullanarak tahmin etmeye çalışmışlardır. Çalışmada RR’u seçme nedenleri olarak bağımsız değişkenler arasında yüksek korelasyon ilişkisine bağlı olarak çoklu bağlantı sorunu olduğunu bildirmişlerdir. Bu çalışmada Japon bildircinlerinden elde edilen toplamda 104 yumurta kullanılmış olup yumurtaların iç kalite özelliklerinden olan ak indeks değeri bağımlı değişken olarak (Y) bu değere tahmin etmek için ise bağımsız değişkenler olarak yumurta ağırlığı (X_1), yumurta genişliği (X_2), yumurta uzunluğu (X_3), Haugh birimi (X_4) ve şekil indeks değeri (X_5) olarak seçilmiştir. Çalışmanın yumurta materyali ise rastgele çiftleşmenin uygulandığı 20-24 haftalık yaşlardaki dişi hayvanlardan

elde edilmiştir. Çalışma sonucunda ak indeksini tahmin etmek için kullanılması önerilen model $Y = -11,743 + 0,201X_1 - 0,067X_2 - 0,081X_3 + 0,245X_4 - 0,008X_5$ olmuştur. Bu modele ait belirleme katsayısını ise $R^2=0,78$ olarak bildirmişlerdir. Bu tip çalışmalarda yumurta iç kalite özelliklerinin yumurtayı kırmadan belirlenebilmesi için bu tip dış kalite özelliklerini kullanarak iç kalite özelliklerini tahmini yaygın bir çalışma alanı bulmuştur. Her ne kadar ülkemizde ticari üretimde yumurtaların kalite özellikleri çok yaygın olarak üzerinde durulan parametreler olmamasına karşın bu tip çalışmalar gelecek için faydalı olacak niteliktedir. Araştırmacılar elde ettikleri veriler için öncelikle EKK analiz yöntemine göre çoklu regresyon analizi yapmışlar bu yönteme ait korelasyon katsayısı 0,88, belirleme katsayısı 0,79 ve düzeltilmiş belirleme katsayısı ise 0,78 olarak bulunmuştur. Bu yöntemle bulunan model için yapılan çoklu regresyon varyans analizi ise önemli bulunmuştur ($P<0,01$). Bununla beraber EKK yöntemi ile elde edilen regresyon parametreleri ve bu parametrelerin önem seviyelerine bakıldığında bazı değişkenlerin (X_2, X_3, X_4) önemsiz bulunmuştur. Çalışmada incelenen özelliklere ait bağımsız değişkenlerin korelasyon katsayıları yumurta ağırlığı ile yumurta genişliği arasında 0,76, yumurta genişliği ile yumurta uzunluğu arasında 0,76 ve VIF değerlerine bakıldığında yumurta genişliği için bu değer 319,63, yumurta uzunluğu için 241,45 ve şekil indeksi için 131,86 olarak bulunmuştur. Bu şekilde VIF değerlerinin 10'dan yüksek, bağımsız değişkenler arasında korelasyonların 0,70'den fazla ve tolerans değerlerinin 0'a yakın çıkması bağımsız değişkenler arasında çoklu doğrusal bağlantı varlığını göstermiştir. Bu değerlerin yanında öz değerlerin 0'a yaklaşması ve koşul indeksinin 30'dan büyük (44,26) olması da aynı sorunun bir işareti olarak gösterilmiştir. Çalışmada RR analizi için k yanlılık sabiti 0,03 olarak eşitlik yardımıyla bulunmuştur. Bu değer $k=0,03$ olarak alındığında RR yöntemine göre bulunan modelin korelasyon katsayısı 0,88 belirleme katsayısı 0,78 ve düzeltilmiş belirleme katsayısı 0,77 olup model istatistiki olarak anlamlı bulunmuştur ($P<0,01$). Çalışmada RR yöntemi kullanılarak elde edilen regresyon parametre değerleri ve bu değerlere ilişkin VIF değerlerine bakıldığında değerlerin 1'e yakın olarak düzeldiği görülmüştür. Bu çalışmada da görüldüğü gibi çoklu doğrusal bağlantının varlığı durumunda EKK regresyon tekniği yerine RR tekniğinin kullanılmasını önermişlerdir.

Arı ve Önder (2013) yaptıkları derleme çalışmasında elde edilmiş verilerin özelliğine göre hangi regresyon metodunun kullanılması gerektiğini açıklamaya çalışmışlardır. Buna göre cevap değişkeni kesikli olduğunda lojistik regresyon, verilerin çoğunda sıfır değeri var ise negatif binom regresyon, veriler kategorik ise poisson regresyon veya probit regresyon, bağımsız değişkenler arasında çoklu doğrusal bağlantı problemi var ise temel bileşenler

regresyonu ve Ridge regresyon, sağ kalım analizinde ise cox regresyon metotlarının kullanılması gerektiğini ifade etmişlerdir.

Öztürk (2014) çoklu regresyon analizinde, çoklu doğrusal bağlantı sorunu olan verilerde ridge regresyon ve en küçük kareler yöntemine örnek bir veri setinde etkinliklerini karşılaştırmıştır. Bunun için hayvan materyali olarak etçi piliç karkas ağırlığını bağımlı değişken ve bunu etkilediğini düşündüğü protein miktarı (gr/gün), yem tüketimi (kg), cinsiyet, sıcaklık (c^0), ışıklandırma (saat/gün) ve su tüketimini (ml/gün) ise bağımsız değişkenler olarak almıştır. Çalışmada etçi piliçlerde karkas ağırlığını tahmin için bağımsız değişkenler ile uygun çoklu regresyon modeli bulunmaya çalışılmıştır. Yapılan çalışmada ele alınan bağımlı ve bağımsız değişkenler arasındaki korelasyon katsayılarına bakıldığında özellikle bazı değerler için (sıcaklık-ışıklandırma $r=-0,98$; yem tüketimi-ışıklandırma $r=0,89$; protein ışıklandırma $r=0,97$; protein-sıcaklık $r=-0,99$; yem tüketimi-sıcaklık $r=-0,92$) olarak yüksek ilişkili bulunmuştur. Korelasyon katsayılarının ve bağımsız değişkenlere ait bulunan VIF değerlerinin (protein 145,93, sıcaklık 178,89, ışıklandırma süresi 36,00, yem tüketimi 14,72) olarak 10'dan büyük çıkması araştırmacı açısından çoklu doğrusal bağlantı varlığının kanıtı olmuştur. Ayrıca çalışmada bulunan korelasyon öz değerlerinin 0'a yakın ve koşul sayısının 1000'den fazla çıkması çoklu doğrusal bağlantının varlığını gösteren ekstra bulgular olarak belirtilmiştir. Çalışmada EKK yöntemine göre yapılan modellemede belirleme katsayısı 0,99 olarak oldukça yüksek bir değer olarak bulunmuştur ama parametre katsayılarına bakıldığında her ne kadar model önemli de bulursa (protein, yem tüketimi) gibi bağımsız değişkenler için t testi anlamlı bulunmamıştır. Bu da çalışmanın EKK yöntemine göre değil RR yöntemine göre yapılmasının daha isabetli olacağı sonucunu doğurmuştur. RR analizinde bulunan k yanlılık sabiti ise çalışmada ridge iz grafiğinden yararlanarak bu değer 0,03 olarak alınması ifade edilmiştir. Çalışmada k yanlılık sabiti 0,03 alındığında modelin belirleme katsayısı 0,97 olarak bulunmuştur. Aynı zamanda bağımsız değişkene ait VIF değerlerinin tamamı 10'un altına düşmüş ve çalışmada önerilen model istatistiki olarak önemlidir.

Karkas Ağırlığı=0,632 + 0,047 protein + 0,110 yem tüketimi + 0,112 cinsiyet - 0,015sıcaklık + 0,0002 ışıklandırma süresi - 0,0009 su tüketimi.

Sonuç olarak etçi piliçlerde karkas ağırlığına ait kestirim yapmak için ele alınan bağımsız değişkenler dikkate alındığında EKK yöntemi yerine RR yönteminin kullanılması gerektiğini vurgulanmıştır.

Shafey, Hussein, Mahmoud, Abouheif, Al-Batshan (2015) yaptıkları çalışmada bağımsız değişken olarak yumurta ölçümlerini [yumurta ağırlığını (YA) (X_1), yumurta genişliğini (YG) (X_2), yumurta şekil indeksini (YŞİ) (X_3), yumurta hacmini (YH) (X_4) ve yumurta yoğunluğunu (YY) (X_5)] ve bağımlı değişken olarak yumurta bileşenlerini [kabuk ağırlığını (KA), sarı ağırlığını (SA) ve ak ağırlığını (AA)] sırasıyla 32, 45 ve 59 haftalık tavuklardan alınan yumurtaları adimsal regresyon ve ridge regresyon modeli kullanarak çoklu doğrusal bağlantı sorununu araştırmışlardır. Araştırmacılar 32, 45 ve 59 haftalıkken her yaş grubundan 180 yumurta toplamışlardır. Alınan tüm yumurtalarda çoklu doğrusal bağlantı sorunlarına yol açan özellikler arasında anlamlı korelasyon olduğunu tespit etmişlerdir. Tavuk yaşı, yumurta özelliklerini ve yumurta özellikleri arasındaki ilişkilerin büyüklüğünü etkilediğini ifade etmişlerdir. Buna göre daha büyük yaşta üretilen yumurtalar daha küçük yaşta üretilenlerden önemli ölçüde ($P<0,01$) YA, YG, YH, SA ve AA'ya sahip olduğunu bulmuşlardır. Kullanılan adimsal regresyon modeli 32. Haftada üretilen yumurtalardaki çoklu doğrusal bağlantı sorununu hafifletmiş, koşul indeksi 30'dan büyük ve bir tahmini değişken olan YA, belirleme katsayısı R^2 %60 ila %99 arasında değişen ve tahminlenen yumurta bileşenlerine uygun bir modele sahip olduğunu bildirmişlerdir. 45. ve 59. haftalarda üretilen yumurtaların adimsal regresyon modelinde çoklu doğrusal bağlantı sorununun varyans artırıcı faktörü (VIF) 10'dan büyük olduğu ve dört farklı tahmini değişken YA, YG, YH ve YY için belirleme katsayısı R^2 %76 ila %99 arasında değişen yumurta bileşenlerini önemli ölçüde öngören bir modele sahip bulunmuştur. Yaşları 32, 45 ve 59. haftalarda üretilen yumurta karakteristikleri arasındaki korelasyon katsayılarına baktıklarında KA, SA ve AA ile YA arasındaki korelasyon katsayısı ($r=0,77-0,99$ $P<0,001$), KA, SA ve AA ile YH arasındaki korelasyon katsayısı ($r=0,60-0,96$ $P<0,001$) ve KA, SA ve AA ile YG arasındaki korelasyon katsayısı ($r=0,22-0,25$ $P<0,001$), bulmuşlardır. YY ve yumurta bileşenlerinden KA ve AA arasındaki korelasyon katsayıları ise sırasıyla ($r=0,20$ $P<0,01$ ve $r=0,28$ $P<0,001$) olarak görülmüştür. Bununla birlikte, araştırmacılar sonuçların üç farklı yaşta üretilen yumurtalar arasındaki korelasyonlarda farklılık gösterdiğini bulmuşlardır. Yumurta bileşenleri ile YG veya YY arasındaki korelasyon katsayıları 0,20 ($P<0,01$) ile 0,74 ($P<0,001$) arasında değişmiştir ve korelasyon katsayısının yumurtanın boyutu ile değiştiği görmüşlerdir. Araştırmacılar çalışmada 32, 45 ve 59 haftalık yaşlarda üretilen yumurtaların RR analiz sonucunda oluşan VIF değerlerine baktıklarında sırasıyla 4,9'dan 8,2'ye, 1,4'ten 8,7'ye ve 1,3'ten 4,1'e kadar değiştiğini görmüşlerdir. Her üç yaş için VIF değerlerinin 10'dan küçük olması Ridge regresyon metodunun çalışmada çoklu doğrusal bağlantı sorununu ortadan kaldırdığını göstermiştir. Sonuç olarak araştırmacılar RR yönteminin herhangi bir yaş grubunda

üretilen yumurtalar için çoklu doğrusal bağlantı sorununu çözmek ve tavukların yumurta bileşenlerini yumurta ölçümlerinden başarılı bir şekilde tahmin etmek için ideal bir metot olduğu sonucuna varmışlardır.

Akçay ve Sarıözkan (2015) RR metodunu kullanarak yumurta tavukçuluğundan elde edilen satış gelirini modellemek için yaş, yaşama gücü, yumurta ağırlığı ve yumurta verimine ait verileri kullanarak bir model oluşturmaya çalışmışlardır. Çalışmada bağımsız değişkenler arasındaki korelasyon katsayısının 0,70'den yüksek, koşul indeksi değerinin 1000'i aşması ve 3 bağımsız değişkene ait VIF değerlerinin 10'dan yüksek olarak bulunmuşlardır. Bu durumda çoklu doğrusal bağlantı sorunun var olduğu çalışmada gösterilmiştir. Bu nedenle çalışmada hem EKK regresyonuna ve buna alternatif olan ridge regresyonu ile modelleme yapılarak her iki metot karşılaştırılmıştır. Çalışmada EKK yöntemi uygulandığında modelin belirleme katsayısı $R^2=0,99$ ve RR metodu için yanlık sabit değeri olarak $k=0,01$ alınmış ve belirleme katsayısı $R^2=0,98$ olarak bulunmuştur. Araştırmacılar modelde yer alan parametrelerin regresyon katsayılarının standart hatalarının RR yönteminde EKK yöntemine göre daha küçük bulunmuşlardır. Sonuç olarak çoklu doğrusal bağlantı sorunu varlığında RR yönteminin EKK'ya göre daha küçük standart hatalı, daha tutarlı ve daha doğru tahminler sağladığı bildirilmiştir. Çalışmanın uygulanmasında toplam gelir bağımlı değişken ve yaşama gücü, yumurta verimi, yumurta ağırlığı ve yaş dört bağımsız değişken olarak belirlenmiştir. EKK yöntemine göre modelleme yapıldığında modelin istatistiki olarak anlamlı ($P<0,001$) belirleme katsayısı (R^2) ise 0,98 olarak bulunmasına karşın modelde yer alan parametrelerin standart hataları oldukça yüksek ve istatistiki olarak yumurta verimi dışında kalan tüm değişkenler önemsiz ($P>0,05$) olarak bulunmuştur. Bu nedenle bağımsız değişkenler arasındaki korelasyon katsayılarına bakıldığında yaşama gücü ile yumurta ağırlığı arasında -0,79, yaşama gücü ile yaş arasında -0,99 ve yumurta ağırlığı ile yaş arasında ise 0,79'luk yüksek korelasyon katsayıları görülmüştür. Ayrıca bağımsız değişkenlerden yaşama gücü, yumurta ağırlığı ve yaş değişkenine ait VIF değerleri sırasıyla 259,74; 10,98 ve 264,71 olup bu değerler 10'dan büyüktür. Bu nedenle çoklu doğrusal bağlantı sorunu olduğu gözlenmiştir. Bunun yanında korelasyonların öz değerlerine ve koşul indeksi değerine bakıldığında bazı öz değerlerin 0'a yakın çıkması ve bazı koşul indeksi değerinin 1463,53 olarak bulunması çoklu doğrusal bağlantının varlığını kesin olarak ortaya koymuştur. Çoklu doğrusal bağlantı varlığında EKK yöntemi yerine RR yöntemi uygulanması önerilmektedir bunun için RR yönteminde bilinmesi gereken yanlılık sabiti (k) belirlenmesi için ise yanlılık sabiti k ile yanlı standartlaştırılmış regresyon katsayısı arasındaki grafik kullanılarak bu katsayılarının durağanlaştığı noktadaki k

sabiti 0,01 olarak görülmüştür. Bunun yanında VIF grafiğine bakılarak da VIF değerinin 10'un altına düştüğü yerdeki k değeri k sabiti (k=0,01) olarak alınabileceğini göstermişlerdir. Çalışma sonucunda 10 bin başlık tavuk işletmesi için yumurta satış gelirini belirlemek için kullanılacak regresyon modelleri EKK ve RR metotlarına göre sırasıyla aşağıda verilmiştir.

$$Y=3961,3 + (-421,55X_1) + (103,6X_2) + (49,99X_3) + (-110,24X_4) \text{ ve}$$

$$Y=8669,47 + (-116,91X_1) + (100,61X_2) + (63,2X_3) + (-0,17X_4).$$

Sonuç olarak yapılan çalışmada incelenen özellikler arasında çoklu doğrusal bağlantı sorunu var ise EKK yerine buna alternatif olarak geliştirilen RR yöntemi kullanılmasının daha doğru ve tutarlı tahminler verebileceği ifade edilmiştir.

Çelik ve Söğüt (2015) ridge regresyon metodu kullanarak Japon bıldırcınlarında yumurta ak ağırlığını tahmin etmeye çalışmışlardır. Yumurta ak ağırlığını belirlemek için yumurta ağırlığı, şekil indeksi, ak indeksi, sarı indeksi, sarı ağırlığı, kabuk kalınlığı, kabuk ağırlığı, ak yüksekliği Haugh ünitesi özelliklerinden yararlanmışlardır. Çalışmada RR teknikleri ve EKK analiz tekniğine göre modelleme yapılmış RR tekniğinde kullanılan parametrelere ilişkin standart hatalar EKK yöntemine göre daha küçük bulunmuştur. RR analizine göre yapılan modelleme de modelin belirleme katsayısı 0,68 ve düzeltilmiş belirleme katsayısı ise 0,67 olarak bulunmuştur. Çalışmada bağımsız değişkenler arasında ki korelasyon katsayılarına bakıldığında bazı özellikleri arasında yüksek korelasyon bulunmuştur (yumurta ağırlığı ile sarı ağırlığı $r=0,77$, ak yüksekliği ile Haugh ünitesi $r=0,98$) ve aynı zamanda varyans arttırıcı faktör (VIF) değerine bakıldığında da bu değer bazı modellerde 10'dan büyük olduğu (199,26 ve 199,39) çalışmada görülmüştür. Çalışmada ak ağırlığı için EKK ve RR metodu kullanılarak iki yöntemle modelleme ve parametre tahminleri yapılmıştır. Buna göre EKK yönteminde modelde kullanılan parametrelerin t testine göre önemlilik analizi yapıldığında sarı indeksi, kabuk ağırlığı, ak yüksekliği ve Haugh birimine ait değişkenler için t testi önemsiz bulunmuştur. Çalışmada RR yöntemine göre k yanlılık sabiti 0,05 olarak alınmış, belirleme katsayısı 0,685 olarak bulunmuş ve VIF değerlerinin tamamının 10'dan küçük ve 1'e yakın olarak belirlenmiştir. Sonuç olarak yapılan çalışmada RR kullanılarak elde edilen model;

Ak Ağırlığı = -3,102 + 0,659 Yumurta Ağırlığı + 3,895 Şekil İndeksi + 3,961 Ak İndeksi + 0,786 Sarı İndeksi - 0,241 Sarı Ağırlığı + 3,512 Kabuk Kalınlığı + 0,615 Kabuk Ağırlığı + 0,05 Ak Yüksekliği - 0,003 Haugh Birimi olarak bulunmuştur.

Araştırmacılar çoklu doğrusal bağlantı durumunda RR yönteminin kullanılmasını önermişlerdir.

Kayaalp, Güney, Cebeci (2015) çoklu doğrusal regresyon analizi kullanarak model seçiminin nasıl yapılacağını göstermişlerdir. Çalışmalarında modelde yer alan bağımsız değişkenlerden hangilerinin modelde kullanılması gerektiğini klasik yöntemler ve adımsal yöntemler kullanarak karşılaştırmışlardır. Çalışmada adımsal yöntemler olarak ileriye doğru seçim, geriye doğru seçim ve adımsal yöntemi, klasik yöntemlerde ise değişken seçimini belirlemek için kullanılan bazı uyum kriterlerinden yararlanmışlardır. Bu uyum kriterleri belirleme katsayısı (R^2), düzeltilmiş belirleme katsayısı, hata kareler ortalaması ve cup değeri olmuştur. Çalışmada uygulama olarak ise yumurta ağırlığını bağımlı değişken, yumurta kabuk ağırlığı, yumurta ak indeksi ve yumurta sarı indeksi verileri ise bağımsız değişken olarak alınmış ve bunlara göre modelleme yapılmıştır. Çalışmanın sonucunda ise üç adımsal metod ile aynı model önerilmiş ve bağımsız değişkenlerin yumurta ağırlığını tahmin konusunda çok açıklayıcı olmadığı ifade edilmiştir. Araştırmacılar bağımlı değişken ile bağımsız değişkenler arasındaki basit doğrusal korelasyon katsayılarının yumurta ağırlığı ile kabuk ağırlığı arasında 0,65, yumurta ağırlığı ile ak indeksi arasında 0,27 ve yumurta ağırlığı ile sarı indeksi 0,35 olarak bulmuşlardır. Bu değerlerden de anlaşıldığı gibi modele girmeyi hak eden tekli bağımsız değişken kabuk ağırlığı olmuştur. Kabuk ağırlığı modele konarak yapılan tekli regresyon analizinde ($Y=29,52+5.07X_1$) şeklinde model elde edilmiş olup bu modeldeki parametreler istatistiki olarak anlamlı bulunmuştur ($P<0,01$). Buradan anlaşılmıştır ki X_1 değişken modelde kullanılmalı ve bunun yanında olması gereken diğer değişken ise kısmi korelasyon katsayılarına bakılarak X_1 sabitken y ile en yüksek korelasyonu veren X_2 değişkeni olduğuna karar verilmiştir. Bu aşamada çoklu doğrusal regresyon analizi ile X_1 ve X_2 değerleri modele konur ($Y=27,35+4,86X_1+0,58X_2$) modeli ve X_2 ve X_1 parametreleri istatistiki olarak anlamlı bulunduğundan modelde kalmasına karar verilir. Bu model X_3 bağımsız değişkeni eklendiği zaman bu değişkene ait önem testi ($P>0,05$) olduğundan modelde olmasına gerek olmadığı sonucuna varılmıştır. Geriye doğru seçim yapıldığında da ilk olarak tüm bağımsız değişkenler modele konur önemlilik seviyesi %10'dan büyük olan veya bu seviyesi en yüksek olan değişken modelden çıkarılarak analiz tekrar yapılır. Önemlilik seviyesi 0,10'un altında olan tüm değişkenler modelde tutulur bu çalışmada ileriye doğru seçim ve geriye doğru seçim sonuçları tamamen aynı bulunmuştur. Klasik yöntemle

göre modellerin belirleme katsayılarına bakıldığı zaman en yüksek belirleme katsayısının 0,46 olarak X_1 ve X_2 değişkeni kullanıldığı modelde görülmüştür. Bu model diğer modele göre (tekliye) daha yüksek belirleme katsayısına sahip olduğundan bu model araştırmacılar tarafından önerilmiştir. Bu çalışmada çoklu doğrusal regresyon yöntemleri karşılaştırılmış bir bağımsız değişkenin modele dahil olabilmesi için önem seviyesinin 0,05'ten küçük, modelden çıkarılabilmesi için ise önem seviyesinin 10'dan büyük olacak şekilde belirlenmesi uygun görülmüştür. Çalışmada son olarak yumurta ağırlığını belirlemek için bu modelde yer alan bağımsız değişkenler yerine daha başka bağımsız değişkenlerin (belirleme katsayısının düşük olduğundan dolayı) kullanılması gerektiği ifade edilmiştir.

Çelik vd. (2018) beyaz hindilerde yaptıkları çalışmada bağımsız değişken olarak karkas parça ağırlıklarını [uyluk ağırlığı (UA), göğüs ağırlığı (GA), kanat ağırlığı (KNA), sırt ağırlığı (SA), taşlık ağırlığı (TA), kalp ağırlığı (KPA) ve ayak ağırlığı (AA)], bağımlı değişken olarak ise bütün karkas ağırlığını (KA) almışlardır. Araştırmacılar çalışmada Bağımsız değişkenlerin bağımlı değişkene etkilerini ridge regresyonu ve faktöre dayalı regresyon ile analiz etmişlerdir. Bu amaçla çalışmada, 17 haftalık, 15 erkek, 15 dişi toplam 30 hindi karkası kullanmışlardır. Araştırmacılar en küçük kareler yöntemiyle (EKK) elde edilen karkas ağırlığı ve karkas değerleri arasındaki doğrusal ilişkinin %99,8 ve regresyon modelinin istatistiki olarak anlamlı olduğunu bildirmişlerdir ($P < 0,001$). Ayrıca araştırmacılar modelde bağımlı değişkenin standart hatasının yüksek olduğunu ve EKK yöntemiyle elde edilen regresyon ilişkisinin istatistiki olarak önemli olduğunu görmüşlerdir ($P < 0,01$). Araştırmacılar EKK yönteminde uyluk ağırlığı ve kanat ağırlığının VIF değerlerinin 10'dan yüksek olduğunu bulmuşlardır. Ayrıca korelasyon matrisi 0,90'dan yüksek korelasyon katsayılarına sahip olduğunu gören araştırmacılar hindilerde karkas ağırlığını tahmin etmek için ridge regresyon modelini kullanmışlardır. Araştırmacılar çalışmada RR modeline göre yaptıkları analizde karkas ağırlığı ile karkas parçaları arasındaki VIF değerlerinin 10'dan küçük, korelasyon katsayısının $r=0,994$, belirleme katsayısının $R^2=0,988$ ve düzeltilmiş belirleme katsayısının $R^2_d=0,984$ olduğunu bulmuşlardır. Araştırmacılar RR modeli ile yapılan analiz sonucunda tahmin edilen parametre sonuçlarının standart hatalarının ve VIF değerlerinin EKK yöntemine göre önemli ölçüde düştüğünü görmüşlerdir. Bu sonuçlara bağlı olarak araştırmacılar çalışmada RR yönteminin belirleme katsayısı (R^2), faktöre dayalı regresyon analizine göre daha yüksek olduğundan ve karkas ağırlığını daha iyi açıkladığından RR yönteminin kullanılmasının daha uygun olduğunu önermişlerdir.

Yılmaz (2018) bir kabuklu türü olan Bivalv (*Abra alba*) türünde Ridge regresyon yöntemi kullanarak balıklarda çeşitli vücut ölçüleri (kabuk, boy, en, kalınlık ve cup indeksi) kullanarak kabuk ağırlığını tahmin etmeye çalışmıştır. Bu çalışmada araştırmacı kabuk ağırlığını bağımlı değişken ve çeşitli vücut ölçülerini bağımsız değişken olarak almıştır. Çalışmada vücut ölçüleri arasında çoklu doğrusal bağlantı olarak ifade edilen bir ilişkiyi saptadığı için çoklu doğrusal regresyon analizini en küçük kareler yöntemine göre değil de bu yöntemin alternatifini olan Ridge regresyonunun kullanmıştır. Çalışmada 136 birey kullanılmış olup değişkenler arasında doğrusal ilişki sağlanması için verilere logaritmik dönüşüm uygulanmıştır. Çalışmada 136 bireye ait kabuk ağırlığı (gr), kabuk boy (mm), kabuk en (mm), kabuk kalınlığı (mm) ve cup indeksi (CI) ölçümleri yapılmıştır. Bu değişkenlere ait çeşitli tanımlayıcı istatistikler hesaplanmıştır. Çoklu doğrusal regresyon analizi için EKK yöntemi kullanıldığında korelasyon katsayısı 0,92, belirleme katsayısı 0,85 ve standart hata 0,24 olarak bulunmuştur. EKK yöntemine göre yapılan varyans analizi sonucunda modelin anlamlılık düzeyi istatistik olarak önemli bulunmuştur ($P < 0,05$). EKK yöntemine göre yapılan regresyon analizi sonuçlarına bakıldığında her ne kadar model anlamlı çıksa da kabuk boyu, kabuk kalınlığı ve CI değişkenlerine ait regresyon katsayılarının t testi sonuçları istatistiki olarak anlamlı bulunmamıştır. Ayrıca bağımsız değişkenler arasındaki korelasyon katsayılarına bakıldığında kabuk eni ile kabuk boyu arasında 0,91'lik, kabuk kalınlığı ile kabuk boyu arasında 0,75'lik, kabuk kalınlığı ile kabuk eni arasında 0,83'lük ve Cup indeksi ile kabuk kalınlığı arasında 0,59 olarak yüksek korelasyonlar hesaplanmıştır. Bu değerlerin yüksek olması bağımsız değişkenler arasında çoklu doğrusal bağlantının varlığını göstermiştir. Bunun yanında hesaplanan varyans arttırıcı (VIF) değerinin kabuk boyu, kabuk eni, kabuk kalınlığı için sırasıyla 10,27; 11,58 ve 25,12 olması bunun yanında hesaplanan öz değerlerin kabuk eni (0,07) ve kabuk kalınlığı (0,02) için 0'a yakın olması ve koşul indeksi değerinin kabuk kalınlığı için 113,20 yani 30'dan büyük çıkması saptandığı için çoklu doğrusal bağlantının varlığı kesin olarak gösterilmiştir. Çoklu doğrusal bağlantının varlığından dolayı araştırmacı RR yöntemi kullanmış bunun içinde gerekli olan yanlılık sabiti olarak bilinen k sabitini RR katsayılarının durağanlaştığı ve bu katsayılara bağlı olarak VIF değerlerinin 1'e yaklaştığı noktaya denk gelen 0,1 değerini k sabiti olarak almıştır. Bu şekilde RR ile yapılan analiz sonucunda ise EKK yöntemine göre bulunan değerler ile farklılık göstermiş RR modeli istatistik olarak ($P < 0,05$) anlamlı bulunmuş. VIF değerleri ve RR ile elde edilen standart hata değerlerinde önemli bir azalış görülmüştür. Çalışmada iki metot karşılaştırıldığında belirleme katsayıları EKK yöntemine göre 0,85 ve RR yöntemine göre 0,81 olarak bulunmuştur. RR yöntemi belirleme katsayısı için çok büyük bir farklılık göstermeden anlamlı bir model

oluşturmuştur. Bağımsız değişkenler için hesaplanan VIF değerleri kabuk boyu için 10,27'den 1,52'ye, kabuk eni için 11,58'den 1,56'ya kabuk kalınlığı için 25,12'den 1,11'e ve cup indeksi için 8,55'den 0,89'a düşmüştür. Çalışmada RR ile elde ettiği modeli $KA = -3,966 + 1,205KB + 1,942KE + 0,741KK - 1,003CI$ şeklinde ve EKK yöntemine göre ise $KA = -4,215 + 0,850KB + 2,748KE + 0,436KK - 10,780CI$ olarak modellemiştir. Araştırmacı bu iki yöntemi karşılaştırmış çalışmasında RR yönteminin kullanılması ile çoklu doğrusal bağlantı sorununun aşılabacağı ve bu şekilde yapılan modelleme ile kabuk ağırlığının tahmininin daha doğru olacağını ifade etmiştir.

Çiftsüren ve Akkol (2018) yaptıkları çalışmada yumurta kalitesini hem dış hem de iç kalite özelliklerini iki bölümde incelemişlerdir. Bu çalışmada araştırmacılar bağımlı değişken olarak iç kalite özelliklerinden yumurta ak ağırlığını (gr) ve yumurta sarı ağırlığını (gr), bağımsız değişken olarak ise yumurta dış kalite özelliklerinden yumurta ağırlığını (gr), yumurta kabuk ağırlığını (gr), yumurta genişliğini (mm), yumurta uzunluğunu (mm) ve şekil indeksini kullanmışlardır. Çoklu doğrusal regresyon analizi en küçük kareler yöntemine göre yapıldığında bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olduğu ve bu sorun çıkan sonuçları doğrudan etkilediğinden dolayı bunu önlemek için Ridge regresyonu, en az mutlak seçilme operatör (LASSO) ve Elastik Net (EN) analizlerini ayrı ayrı kullanmışlardır. Yumurta genişliği ile yumurta ağırlığı ve yumurta genişliği ile şekil indeksi arasındaki korelasyon katsayılarına bakıldığında sırasıyla 0,371 ve 0,806 ($P < 0,01$), yumurta uzunluğu ile yumurta ağırlığı ve yumurta uzunluğu ile şekil indeksi arasındaki korelasyon katsayılarına bakıldığında sırasıyla 0,654 ve -0,529 ($P < 0,01$) ve yumurta ağırlığı ile yumurta kabuk ağırlığı arasındaki korelasyon ise 0,183 ($P < 0,05$) bulmuşlardır. Araştırmacılar bu sonuçları incelediğinde bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olduğunu görmüşlerdir. Varyans arttırıcı değerlere bakıldığında yumurta genişliği, yumurta uzunluğu ve şekil indeksinin sırasıyla 872,7; 416,4 ve 1197,2 çok yüksek ayrıca araştırmacılar tolerans değerine baktıklarında ise sırasıyla 0,00115; 0,00240 ve 0,00084 0'a çok yakın olduğunu görmüşlerdir. Bu araştırmada yine araştırmacılar, öz değerlerin sıfıra çok yakın olduğunu ($0,018$ 'den $6,18 \times 10^{-7}$ 'ye kadar değiştiğini) ve koşul indeksinin ise çok yüksek olduğunu ($17,98$ 'den $3109,37$ 'ye kadar değiştiğini) görmüşlerdir. Yapılan tüm analizler için bulunan tahminlerin tamamı önemli bulunmuştur ($P < 0,01$). Araştırmacılar Ridge regresyonu ile yapılan analizde yumurta ağırlığı ve yumurta kabuk ağırlığı dışında kalan değişkenlerde regresyon parametrelerinde en küçük kareler yöntemine göre önemli bir düşüş olduğunu bulmuşlardır. Araştırmacılar LASSO ve EN analiz sonuçlarını değerlendirdiklerinde yumurta genişliği, şekil

indeksi ve yumurta kabuk ağırlığı regresyon katsayılarının yumurta sarısı ağırlığı için sıfıra düştüğünü, yumurta genişliği yumurta uzunluğu ve şekil indeksi regresyon katsayılarının da yumurta ak ağırlığı için sıfıra düştüğünü bulmuşlardır. Araştırmacılar yumurta ak ağırlığı ve yumurta sarı ağırlığı tahmini için EKK ve RR yönteminde 5 tane bağımsız değişken ve LASSO ile EN için ise 2 tane bağımsız değişken önermişlerdir. Araştırmacılar Ridge regresyon LASSO ve EN analizlerinde yumurta sarısı ağırlığı için düzeltilmiş belirleme katsayısı (R^2_d) sırasıyla yüzde (%) 58,34; 59,17 ve 59,11 olarak, yumurta ak ağırlığı için düzeltilmiş belirleme katsayısı (R^2_d) sırasıyla yüzde olarak 75,60; 75,94 ve 74,81 bulmuşlardır. Araştırmacılar yapılan bu çalışmada RR analizi, LASSO ve EN metotların en küçük kareler yöntemine karşı kullanılması araştırmanın doğruluğu için daha doğru olacağını ifade vermişlerdir.

Çankaya, Eker, Abacı (2019). Regresyon analizinde çoklu doğrusal bağlantı probleminin varlığında EKK, RR ve Temel bileşenler regresyonu yaklaşımlarını kullanarak bu üç metodu birbiriyle karşılaştırmışlardır. Bunun için 85 Karayaka kuzusunda canlı ağırlık ile bu kuzulardan alınan çeşitli vücut ölçüleri (cidago yüksekliği, sağrı yüksekliği, vücut uzunluğu, göğüs derinliği, göğüs çevresi, göğüs genişliği, ön, orta ve arka yumru yüksekliği) alınmıştır. Bu çalışmada CA bağımlı değişken ve alınan vücut ölçüleri ise canlı ağırlığı modellemek için kullanılan bağımsız değişkenler olarak tanımlanmıştır. Çalışmanın sonucunda bağımsız değişkenler arasında yüksek korelasyonların görülmesi çoklu doğrusal bağlantı sorununu meydana çıkarmış ve bu sorunu gidermek için tercih edilen RR ve PCR metotlarından RR metodu en küçük hata kareler ortalamasına ve en yüksek belirleme katsayısına sahip olduğu görülmüştür. Bu çalışmada ayrıca PCR metodunun Karayaka kuzularında süttten kesim ağırlığını tahmin etmek için kullanılabilen en uygun modeli oluşturduğu ifade edilmiştir. Yapılan çalışmada bağımsız değişkenler arasındaki korelasyon katsayıları özellikle cidago yüksekliği ve arka yumru yüksekliği arasında 0,79 ($P<0,01$), ön yumru yüksekliği ile orta yumru yüksekliği arasında ise 0,96 ($P<0,01$) olarak hayvanların özellikle yükseklik ölçüleri arasında yüksek korelasyonların bulunması neticesinde çoklu doğrusal bağlantının varlığı anlaşılmıştır. Ayrıca EKK yöntemine göre yapılan parametre tahminlerinde parametrelerin önem testleri önemsiz ve VIF değerleri özellikle yükseklik ölçüleri için (14,06 ve 16,47) 10'dan büyük çıkmıştır. Bunun yanında koşul indeks değerinin 30'dan büyük ve öz değerlerin 0'a yaklaşması aynı şekilde çoklu doğrusal bağlantı sorununun varlığını ortaya çıkarmıştır. Yapılan RR analizinde yanlılık sabiti ($k=0,01$) olarak bulunmuş ve her üç metodun karşılaştırılmasında EKK, RR ve PCR metotları için hata kareler

ortalaması sırasıyla 0,039; 0,039 ve 0,042; belirleme katsayıları 0,63; 0,62 ve 0,62; düzeltilmiş belirleme katsayıları 0,59; 0,58 ve 0,54; varyasyon katsayıları 0,19; 0,06 ve 0,06 olup her üç metotta istatistik olarak anlamlı bulunmuştur. Araştırmacılar çoklu doğrusal bağlantı sorunu olduğunda EKK yöntemi ile çoklu doğrusal regresyon analizi yerine bu durumda Ridge regresyon (RR) veya temel bileşenler regresyon (PCR) yöntemlerinden birinin kullanılmasını önermişlerdir.

Bu çalışmada çoklu doğrusal regresyon analizinde modelleme yaparken öncelikle bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olup olmadığı, uygun yöntemler ile kontrol edilerek bu sorunun varlığı halinde, sorunu gideren yöntemlerden biri olan RR yönteminin nasıl uygulanacağı somut bir örnek ile gösterilmiştir. Çalışmada RR yönteminin uygulama yöntemlerinin yanında zootekni alanında bu konuda yapılan çalışmalar ve sonuçları da araştırmacılarının dikkatini çekmek için ayrıntılı olarak gösterilmiştir.

3. MATERYAL ve YÖNTEM

3.1. Araştırma Materyali

Çalışmada ele alınan konuyu daha iyi açıklayabilmek ve örneklendirebilmek amacıyla basit bir örnek veri seti oluşturulmuştur. Bu çalışma için, çoklu doğrusal regresyon analiz yöntemlerinden olan Ridge Regresyon (RR) ve En Küçük Kareler (EKK) analiz yöntemine göre bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olacak şekilde oluşturulan ve 48 baş hayvandan elde edilmiş özel bir işletmeden alınan ölçümler kullanılmıştır. Bu örnek veri setinin belirlenmesinde ise özel bir çiftlikte benzer şartlarda yetiştirilen ve özel bir mezbahada kesilen 2-3 yaşlı Esmer erkek hayvanların kesim öncesi alınan çeşitli vücut ölçülerine ve kesim sonrası alınan sıcak karkas ağırlıklarına ait ortalama değerlerinden yararlanılmıştır. Bu örnek veri setinde hayvanların sıcak karkas ağırlıkları bağımlı değişken ve bu hayvanlardan alınan çeşitli vücut ölçüleri; Kuyruk Sokumu Yüksekliği (KSY), Sağrı Yüksekliği (SY), Cidago Yüksekliği (CY), Vücut Uzunluğu (VU) ve Göğüs Çevresi (GÇ) bağımsız değişken olarak alınmıştır. Hayvanların;

Kuyruk Sokumu Yüksekliği; kuyruk sokumu ile yer düzlemi arasındaki dikey yükseklik,

Sağrı Yüksekliği; sağrı ile yer düzlemi arasındaki dikey yükseklik,

Cidago Yüksekliği; cidagonun en yüksek yerinden yer düzlemine kadar olan dikey yükseklik olup ölçümler ölçü bastonu ile alınır.

Vücut Uzunluğu; omuz ucu (*Articulus humeri*) ile oturak yumrusu (*Tuber ichii*) arasındaki mesafe ve

Göğüs Çevresi; cidago üzerinden göğüs kemiğini dolaşan çevre ölçüsü olup ölçü şeridi ile alınır. Kullanılan veri setinde sıcak karkas ağırlığı (kg) ve vücut ölçüleri ise (cm) olarak çalışmada kullanılmıştır.

3.2. Araştırma Yöntemi

Çalışmada kullanılan örnek sığırlarda karkas ağırlığının tahmini için vücut ölçülerinden tahmin yapmayı amaçlayacak şekilde elde edilen tahmini verilerde yapılmıştır. Bu amaçla üretilen verilerin önce tanımlayıcı istatistikleri, değişkenler arasında korelasyon

katsayıları ve RR yöntemi ile EKK analiz yöntemine göre elde edilen parametre tahminleri ve bu parametrelerden oluşan modeller verilmiştir.

Çok değişkenli doğrusal regresyon modelinde bir bağımlı değişken (Y) ve bu bağımlı değişkeni açıkladığı düşünülen iki yada daha fazla bağımsız değişken (X_1, X_2, \dots, X_n) bulunur. Bu modelin regresyon denklemi eşitlik 3.1'de verilmiştir.

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_n X_{ni} + \varepsilon \quad (3.1)$$

Kullanılan veri setinde çoklu doğrusal bağlantı sorununun belirlenmesi için VIF değeri ve koşul sayısı (CN) değerlerinden yararlanılmıştır. Bu değerler eşitlik 3.2 ve 3.3'de gösterilmiştir (Orhunbilge, 1996; Albayrak, 2005; Orhunbilge, 2017).

$$(VIF)_k = (1 - R_k^2)^{-1} \quad (3.2)$$

$$CN = \left(\frac{\lambda_{max}}{\lambda_{min}} \right) \quad (3.3)$$

Bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu varlığı halinde RR ve EKK yöntemiyle iki modelin karşılaştırılması örnek veri setinde üzerinden gösterilmiştir. Ayrıca çoklu doğrusal bağlantı sorunu bulunan veri seti NCSS programı kullanılarak sonuçlar elde edilmiş ve EKK ve RR yöntemleri ile çoklu doğrusal regresyon analizi yapılarak bulunan çok değişkenli doğrusal regresyon denklemleri birbiri ile karşılaştırılmıştır (NCSS, 2001).

4. ARAŞTIRMA BULGULARI ve TARTIŞMA

Et sığırcılığında karkas ağırlığının hayvanın kesilmeden tahmin edilmesi önemli bir yetiştirici pratiğidir. Eğer yetiştirici karkas ağırlığını hayvan kesilmeden çeşitli vücut ölçüleri kullanarak doğru bir yöntem ve doğru bağımlı değişkenleri kullanarak modellenenirse bu sayede çok önemli bir bilgiyi hayvanı kesmeden önce elde etmiş olacaktır. Bu amaçla örnek olarak yapılan çalışmada elde edilen sonuçlar sunulmuştur.

Bu çalışmada sunulan konuya örnek teşkil etmesi için bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olacak şekilde oluşturulan bir örnek veri seti kullanılmıştır. Bu örnek veri seti Esmir sığırlara ait karkas ağırlıkları ve çeşitli vücut ölçüleridir. Çalışmada karkas ağırlıkları (KA) bağımlı değişken ve kullanılan çeşitli vücut ölçüleri; kuyruk sokumu yüksekliği (KSY), sağrı yüksekliği (SY), cidago yüksekliği (CY), vücut uzunluğu (VU) ve göğüs çevresi (GÇ) ölçüleri bağımsız değişken olarak alınmıştır.

Çalışmada aralarında çoklu doğrusal bağlantı sorunu bulunan veri setine ilişkin faktörlerin ortalama ve standart hataları Çizelge 4.1’de sunulmuştur.

Çizelge 4.1. Kullanılan örnek veri setini oluşturan hayvanların karkas ağırlığı ve çeşitli vücut ölçülerine ait tanımlayıcı istatistikleri

Değişken ¹	N	Ortalama ± Standart Hata	Minimum	Maksimum
KSY (cm)	48	127,60±4,67	117	135,50
SY (cm)	48	129,35±4,68	119	137
CY (cm)	48	126,89±4,64	116,50	134,50
VU (cm)	48	125,33±13,02	107	203
GÇ (cm)	48	192,25±7,09	177	215
KA (kg)	48	294,46±27,65	227,80	353,60

¹KSY: kuyruk sokumu yüksekliği, SY: sağrı yüksekliği, CY: cidago yüksekliği, VU: vücut uzunluğu, GÇ: göğüs çevresi, KA: karkas ağırlığı

Çizelge 4.1.’e göre karkas ağırlığına ait ortalama ve standart hata değerleri sırasıyla 294,46±27,65 kg olarak bulunmuştur. Cidago yüksekliği, sırt yüksekliği ve kuyruk sokumu

yüksekliği ise sırasıyla 126,89±4,64; 129,35±4,68; 127,60±4,67 cm, vücut uzunluğu ise 125,33±13,02 cm bulunurken göğüs çevresi ise 192,25±7,09 cm bulunmuştur.

Anonim (1989) Alman Tarım Bakanlığı Almanya’da yetiştirilen ve sisteme kayıtlı olan Esmer, Simental ve Holstein sığırlarında sırasıyla cidago yüksekliğini 133, 133, 136 cm, göğüs çevresini 187, 193-195, 198 cm ve canlı ağırlığı ise 511, 547-552, 538 kg olarak bildirmiştir. Ulsan, Solmaz ve Ekici (1996) Kars ilinde kesilen 2-3 yaşlı Esmer ve Simental sığırlarında cidago yüksekliğini sırasıyla 121-138 cm, göğüs çevresini 177-201 cm ve sıcak karkas ağırlığını 177-355 kg olarak bildirmişlerdir.

Hayvanların bağımlı değişken karkas ağırlığı (KA) ile bağımsız değişken olarak kabul edilen çeşitli vücut ölçüleri arasındaki korelasyon katsayıları ve önem testi sonuçları Çizelge 4.2’de verilmiştir. Bu tabloya göre en yüksek korelasyon katsayıları kuyruk sokumu yüksekliği (KSY) ile sırt yüksekliği (SY) arasında 0,99 (P<0,01), kuyruk sokumu yüksekliği (KSY) ile cidago yüksekliği (CY) arasında 0,99 (P<0,01), cidago yüksekliği (CY) ile sırt yüksekliği (SY) arasında 0,99 (P<0,01), vücut uzunluğu (VU) ile kuyruk sokumu yüksekliği (KSY) arasında 0,50 (P<0,01), vücut uzunluğu (VU) ile sırt yüksekliği arasında (SY) 0,49 (P<0,01), vücut uzunluğu (VU) ile cidago yüksekliği (CY) arasında 0,49 (P<0,01) bulunmuştur. Çizelge 4.2’de bunlar dışında bulunan değişkenlere ait korelasyon katsayıları ise istatistik olarak önemsiz bulunmuştur.

Çizelge 4.2. Vücut ölçüleri arasındaki korelasyon katsayıları

	KSY	SY	CY	VU	GÇ
SY	0,998**				
CY	0,996**	0,998**			
VU	0,502**	0,494**	0,491**		
GÇ	0,363 ^{öd}	0,364 ^{öd}	0,355 ^{öd}	0,214 ^{öd}	
KA	-0,088 ^{öd}	-0,091 ^{öd}	-0,100 ^{öd}	-0,122 ^{öd}	-0,098 ^{öd}

Not:**= P<0,01; öd= Önemli Değil (P>0,05) KSY: kuyruk sokumu yüksekliği, SY: sağrı yüksekliği, CY: cidago yüksekliği, VU: vücut uzunluğu, GÇ: göğüs çevresi, KA: karkas ağırlığı

Çalışmada bir bağımlı değişken (karkas ağırlığı) ve beş tane bağımsız değişken (vücut ölçüleri) kullanılarak en küçük kareler yöntemi ve ridge regresyon yöntemi ile elde edilen

analiz sonuçları tablolar ve şekiller verilerek gösterilmiştir. Buna göre Çizelge 4.3’de en küçük kareler analiz yöntemine göre elde edilen bağımsız değişkenlere ait varyans arttırıcı faktör (VIF) değerleri sunulmuştur.

Çizelge 4.3. EKK analizi sonucu bağımsız değişkenlere ait varyans arttırıcı (VIF) değerleri

Bağımsız Değişkenler (cm)	VIF
KSY	365,9900
SY	739,0477
CY	277,2151
VU	1,3769
GÇ	1,1818

KSY: kuyruk sokumu yüksekliği, SY: sağrı yüksekliği, CY: cidago yüksekliği, VU: vücut uzunluğu, GÇ: göğüs çevresi,

Çizelge 4.3’de görüldüğü gibi en küçük kareler analiz yöntemine göre yapılan çoklu doğrusal regresyon analizi sonucunda varyans arttırıcı faktör olarak bilinen (VIF) değerlerinin sırasıyla kuyruk sokumu yüksekliği (KSY) için 365,99, sağrı yüksekliği (SY) için 739,04, cidago yüksekliği (CY) için 277,21 vücut uzunluğu (VU) için 1,37 ve göğüs çevresi (GÇ) için 1,18 olarak hesaplanmıştır. Çoklu doğrusal regresyon analizinde VIF değerlerinin 10 ve daha büyük bulunması bağımsız değişkenler arasında çoklu doğrusal bağlantının varlığının bir göstergesidir. Bu durumda en küçük kareler analiz yöntemi yerine ridge regresyon yöntemi olarak bilinen metodun kullanılması daha doğru bir yöntem olduğu bilinmektedir.

Çalışmada elde edilen korelasyonların özdeğerleri ve koşul sayıları (CN) Çizelge 4.4’de sunulmuştur. Bulunan özdeğerlerin sıfıra yakın çıkması veya koşul sayılarının binden büyük çıkması değişkenler arasında çoklu doğrusal bağlantı sorununun varlığını gösteren diğer önemli göstergelerdir. Çizelge 4.4’de gösterildiği gibi son dört özdeğer sırasıyla 0,82; 0,67; 0,003; 0,0008 olarak bulunmuştur bu değerlerin sıfıra yakın olması çoklu doğrusal bağlantı varlığını gösteren diğer bir kriterdir. Bunun yanı sıra Çizelge 4.4’de verilen koşul sayısı değerlerine bakıldığında ise beşinci sırada bulunan koşul sayısı değerinin (CN) 3895,75 olarak binden büyük bir değer olarak bulunması çoklu doğrusal bağlantı sorununun varlığına işaret eden diğer bir göstergedir. Bağımsız değişkenler arasında çoklu doğrusal bağlantının varlığını tespit etmek için tek bir kritere göre değil de daha fazla kritere bakarak kararın verilmesi daha doğru bir yaklaşım olacaktır.

Çizelge 4.4. Korelasyon katsayılarının özdeğerleri

No.	Özdeğer	Yüzde Frekans %	Eklemeli Frekans	Koşul Sayısı (CN)
1	3,489964	69,80	69,80	1,00
2	0,826089	16,52	86,32	4,22
3	0,679286	13,59	99,91	5,14
4	0,003765	0,08	99,98	926,83
5	0,000896	0,02	100,00	3895,75

Bu durumda Ridge regresyon yönteminin kullanılması gerekliliği ortaya çıkmıştır. Bunun içinde Ridge regresyon yönteminde $X'X$ korelasyon matrisinin köşegen elemanlarına eklenecek olan k yanlılık sabitinin bulunması gereklidir. Bu sabit pozitif ve 0 ile 1 arasında olur. Yanlılık sabiti 0 alınırsa model tamamen EKK yöntemiyle aynı sonucu verir ($0 \leq k \leq 1$). Yanlılık sabiti k değeri artıca varyans küçülür. Varyans küçüldükçe de regresyon katsayılarının standart hataları küçülecektir. Bu sayede EKK yönteminde karşılaşılan modelin anlamlı olmasına karşılık modeldeki bağımsız değişkenlerin bazılarının önemsiz bulunması sorunu aşılmış olacaktır. Ayrıca EKK yönteminde karşılaşılan modelin anlamlı olmasına karşılık modeldeki bağımsız değişkenlerin bazılarının önemsiz bulunması çoklu doğrusal bağlantı sorunun göstergelerinden biri olarak kabul edilir.

Bu çalışmada k yanlılık sabitinin belirlenmesi işleminde çeşitli yöntemlerden yararlanılmış ve örnek çalışmada bu değer $k=0,014$ olarak bulunmuştur. Bunun anlamı %1,4 lük bir sapma ile bağımlı değişkenin kendini açıkladığı kabul edilen bağımsız değişkenler ile modellendiğidir.

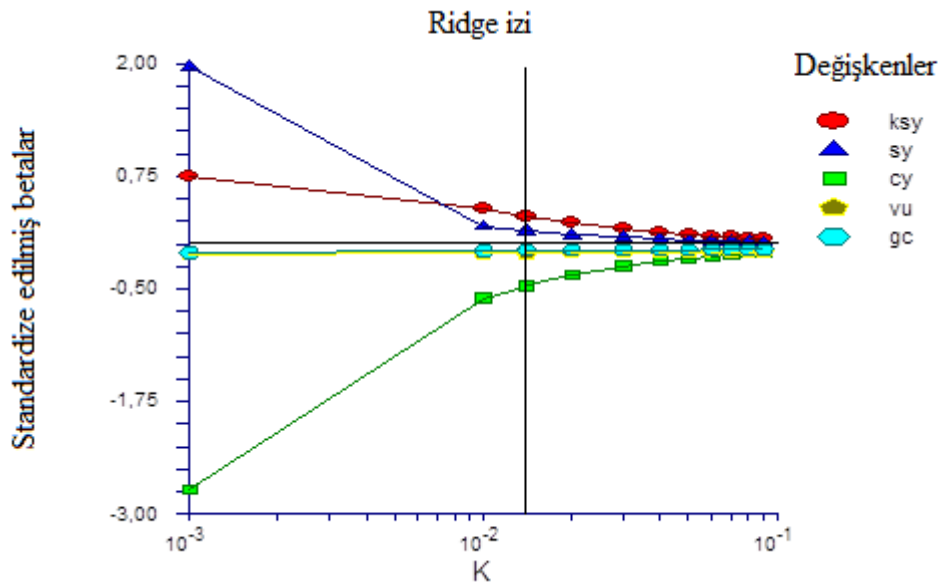
Yanlılık sabiti olan k değerinin belirlenmesinde kullanılan yöntemlerden biri, çeşitli k değerleri alınarak hesaplanan VIF değerlerine bakılmasıdır. Çalışmada farklı k değerleri için bulunan VIF değerleri Çizelge 4.5.'de verilmiştir. Çizelge 4.5.'e göre VIF değerinin 10 değerinin altına indiği k değeri 0,014 olarak bulunmuştur.

Çizelge 4.5. Çeşitli k-değerleri için elde edilen standardize edilmiş Ridge regresyon (RR) katsayıları

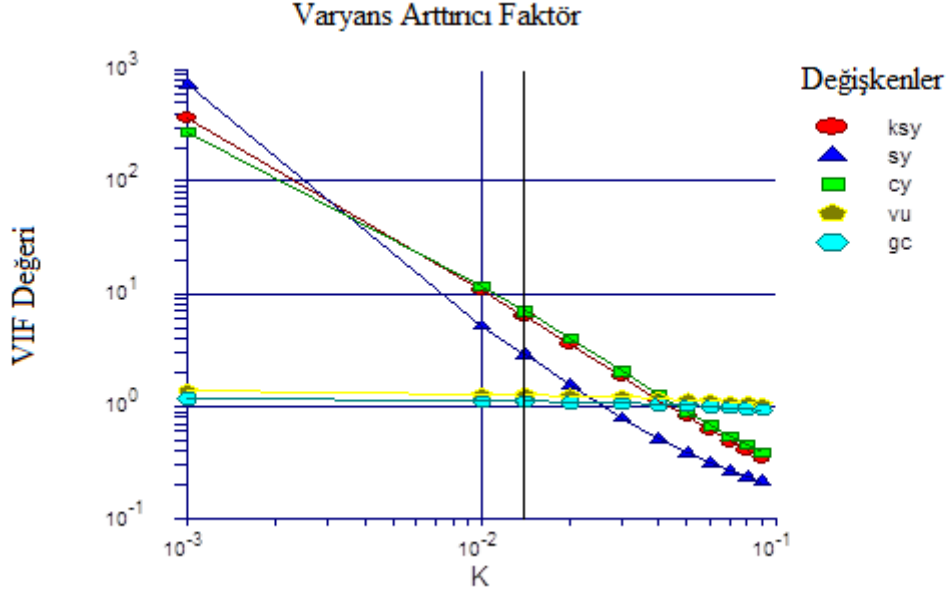
k	KSY	SY	CY	VU	GÇ	R ²	Max VIF
0,00000	0,7494	1,9692	-2,7235	-0,1124	-0,0962	0,0517	739,0477
0,01000	0,3958	0,1937	-0,6054	-0,1025	-0,0748	0,0282	11,8899
0,01400	0,3084	0,1421	-0,4679	-0,1007	-0,0733	0,0265	7,1448
0,01408	0,3071	0,1413	-0,4658	-0,1007	-0,0732	0,0264	7,0812
0,02000	0,2306	0,1006	-0,3499	-0,0989	-0,0718	0,0249	4,0389
0,03000	0,1609	0,0664	-0,2475	-0,0967	-0,0701	0,0234	2,0707
0,04000	0,1225	0,0485	-0,1923	-0,0951	-0,0690	0,0226	1,2917
0,05000	0,0981	0,0374	-0,1579	-0,0936	-0,0680	0,0220	1,1577
0,06000	0,0813	0,0298	-0,1344	-0,0923	-0,0671	0,0216	1,1274
0,07000	0,0689	0,0244	-0,1173	-0,0911	-0,0663	0,0212	1,0983
0,08000	0,0595	0,0202	-0,1043	-0,0900	-0,0656	0,0209	1,0703
0,09000	0,0520	0,0168	-0,0942	-0,0889	-0,0649	0,0207	1,0435

k: yanlılık sabiti, KSY: kuyruk sokumu yüksekliği, SY: sağrı yüksekliği, CY: cidago yüksekliği, VU: vücut uzunluğu, GÇ: göğüs çevresi, R²: belirtme sayısı, VIF: varyans arttırıcı faktör

Yanlılık sabiti olarak bilinen k değerinin bulunma yollarından bir diğer yöntemde Ridge izi olarak bilinen grafiği kullanmaktır. Örnek çalışmada bulunan Ridge izi grafiği Şekil 4.1.'de sunulmuştur. Şekil 4.1.'de görüldüğü gibi VIF değeri yanlı standart regresyon katsayılarının durağanlaştığı yerden seçilir. Tüm standart regresyon katsayılarının durağanlaştığı noktaya karşılık gelen k yanlılık sabiti 0,01 ile 0,1 arasında bir yerde olacağı Şekil 4.1.'de görülmektedir.



Şekil 4.1. Ridge (trace) iz grafiği yardımı ile k-sabitinin belirlenmesi



Şekil 4.2. Varyans Arttırıcı Faktör (VIF) grafiği yardımı ile k-sabitinin belirlenmesi

Yanlılık sabiti olan k değerinin bulunma yollarından bir diğeri de VIF faktörü ile k yanlılık sabitinin beraber gösterildiği grafikten yararlanmaktır. Bu çalışmaya ait bu grafik Şekil 4.2.'de gösterilmiştir. Şekil 4.2.'de görüldüğü gibi VIF değeri yüksek olan KSY ve SY bağımsız değişkenlerinin %1,4 sapma ile 10'un altına düşmüşlerdir. Yanlılık sabitinin (k) belirlenmesinde VIF değerinin 10 değerinin altına indiği noktanın alınması veya tüm bağımsız değişkenlerin durağanlaştığı noktaya karşılık gelen noktanın k sabiti olarak alınabilir. Buna göre k yanlılık sabiti 0,01 ile 0,1 arasında bir yerde olacağı Şekil 4.2.'de görülmektedir.

Çalışmada konunun daha açık bir şekilde anlaşılması için somut bir örnek olarak oluşturulan veri setinden elde edilen analiz sonuçları Çizelge 4.6. ve Çizelge 4.7.'de sunulmuştur. Çizelge 4.6.'da yanlılık sabiti olan k=0,014 alınarak uygulanan Ridge regresyon analizine göre (k=0,014) elde edilen regresyon katsayıları, standart hataları ve VIF değerleri verilmiştir. Bu modele göre karkas ağırlığı bağımlı değişken ve ele alınan vücut ölçüleri bağımsız değişken olmak üzere beş farklı bağımsız değişkenin beş tanesi birden modele dahil edildiğinde elde edilen model;

$KA = 388,0825 + 1,826474 * KSY + 0,8392559 * SY - 2,78606 * CY - 0,2139673 * VU - 0,2855236 * GÇ$ şeklinde olmuştur. Ridge regresyon yöntemine göre ele alınan bağımsız değişkenlerin VIF değerleri ise beklenildiği gibi 10 değerinin altında bulunmuştur.

Çizelge 4.6. Ridge Regresyon katsayıları, standart hataları ve VIF değerleri (k=0,014)

Bağımsız Değişkenler	Regresyon Katsayıları	Standart Hata	VIF
Sabit (a)	388,082		
KSY	1,826	2,281	6,404
SY	0,839	1,529	2,893
CY	-2,786	2,423	7,144
VU	-0,213	0,365	1,278
GÇ	-0,285	0,626	1,115

KSY: kuyruk sokumu yüksekliği, SY: sağrı yüksekliği, CY: cidago yüksekliği, VU: vücut uzunluğu, GÇ: göğüs çevresi, VIF: varyans arttırıcı Sabit (a): kesme noktası

Çalışmada ayrıca EKK yöntemi ve RR yöntemi ile elde edilen parametrelerin regresyon katsayılarının standart hataları Çizelge 4.7.'de sunulmuştur. Çizelge 4.7.'e göre EKK yöntemi ve RR yöntemi ile elde edilen parametrelerin regresyon katsayılarının standart hataları sırasıyla KSY için (17,02-2,028), SY için (24,12-1,52), CY için (14,89-2,42), VU için (0,37-0,36) ve GÇ için (0,63-0,62) olarak bulunmuştur. Bu değerler göre RR yöntemi ile bulunan regresyon katsayılarının standart hataları EKK yöntemine göre RR yönteminden beklenildiği gibi daha düşük bulunmuştur.

Çalışmada RR ve EKK yöntemiyle bulunan regresyon katsayıları ve standart hata değerleri karşılaştırmalı olarak Çizelge 4.7.'de gösterilmiştir. Çizelge 4.7.'de verildiğine göre her iki yöntemle bulunan sabit (a kesme noktası) sırasıyla 388,082 ve 383,657 olarak hesaplanmıştır. Benzer şekilde KSY olarak isimlendirilen bağımsız değişkene ait regresyon katsayısı ve standart hata değerleri her iki yöntem için sırasıyla $1,826 \pm 2,28$ ve $4,437 \pm 17,022$ olmuşken, SY bağımsız değişkeni için sırasıyla $0,839 \pm 1,529$ ve $11,629 \pm 24,124$ olarak hesaplanmıştır.

Çizelge 4.7. Ridge regresyon (RR) ve En Küçük Kareler (EKK) yönteminin karşılaştırılması

Bağımsız Değişkenler	Bağımsız Değişkenler	En Küçük Kareler Regresyon Katsayıları	En Küçük Kareler Standart Hata	Ridge Standart Hata
Sabit (a)	388,082	383,657		
KSY	1,826	4,437	17,022	2,281
SY	0,839	11,629	24,124	1,529
CY	-2,786	-16,216	14,896	2,423
VU	-0,213	-0,238	0,374	0,365
GÇ	-0,285	-0,374	0,636	0,626

KSY: kuyruk sokumu yüksekliği, SY: sağrı yüksekliği, CY: cidago yüksekliği, VU: vücut uzunluğu, GÇ: göğüs çevresi, Sabit (a): kesme noktası

Çalışmada tek bir bağımlı değişken (KA) ve ele alınan beş farklı vücut ölçüsünün (KSY, SY, CY, VU ve GÇ) beşi bağımsız değişken olarak tamamı alınarak modelleme yapılmıştır. Uygulamada bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu varlığı halinde çoklu doğrusal regresyon yöntemlerinden biri olan Ridge regresyon analizine göre elde edilen model ($KA=388,08 + 1,82*KSY + 0,83*SY - 2,78*CY - 0,21*VU - 0,28*GÇ$) ve bu soruna rağmen kullanılan EKK yöntemine göre elde edilen model ($KA=383,65 + 4,43*KSY + 11,62*SY - 16,21*CY - 0,23*VU - 0,37*GÇ$) şeklinde bulunmuştur. Çalışmada kullanılan modellerin anlamlılıkları karşılaştırıldığında ise her iki modelinde belirleme katsayısı düşük olup ($R^2 < 0,25$) bir bağımlı ve beş farklı bağımsız değişkenin beraber kullanıldığı model ele alınan özellik için istatistik olarak çok anlamlı bulunmamıştır. Buna rağmen RR yönteminin kullanılması beklenildiği gibi VIF değerlerini azalttığı, çoklu doğrusal bağlantı sorunundan kaynaklanan bağımsız değişkenlere ait standart hata değerlerini EKK yöntemine kıyasla düşürdüğü örnek üzerinde görülmüştür.

Yapılan bir çalışmada Vovesny (1982) Çekoslovakya'da yetiştirilen çeşitli melez sığırlarda yapılan bir çalışmada karkas özellikleri ile cidago yüksekliği, vücut uzunluğu gibi fenotipik özellikler arasındaki korelasyon katsayılarının yüksek olmadığını bildirmiştir. Ulusan, Solmaz ve Ekici (1996), Kars ilinde kesilen Esmer ve Simental sığırlarında beden ve karkas özellikleri arasındaki ilişkileri korelasyon katsayıları ile araştırmışlardır. Çalışmalarında fenotipik korelasyonları en yüksek canlı ağırlık ile but çevresi arasında ($R=0,90$) ve en düşük arka incik çevresi ile but çevresi arasında ($R=0,04$) olarak bulmuşlardır. Çalışmalarında canlı ağırlık-cidago yüksekliği arasındaki korelasyon katsayısını 0,37 ve karkas ağırlığı-gövde derinliği arasındaki korelasyon katsayısını ise 0,25 olarak bildirmiştir. Araştırmacılar tüm fenotipik korelasyon katsayılarının değişim aralığını ise -0,08 ile 0,90 arasında olduğunu ifade etmişlerdir.

Zootekni alanında EKK yöntemi RR yöntemini karşılaştırmak için yapılan çeşitli çalışmalarda da çoklu doğrusal bağlantı sorunu varlığında, RR yönteminin kullanılması gerekliliği çeşitli araştırmacılar tarafından önerilmiştir. Araştırmacıların araştırma sonuçları ise özet olarak aşağıda sunulmuştur.

Schoeman, Aziz ve Jordaan (2002) beş farklı sığır ırkında (Afrikan, Hereford, Angus, Simentaler ve Charolais) en küçük kareler varyans analiz yönteminde ilk model olarak cinsiyet, yaş ve ortak değişken olarak süttten kesme yaşı ve tüm tek yönlü ilişkileri

araştırmışlardır. Araştırmacılar çoklu doğrusal bağlantı sorununun varlığını açıkça gösteren birkaç yüksek derecede korelasyonlu değişken gözlemişlerdir. Araştırmacılar çoklu doğrusal bağlantı sorununu çözmek amacıyla Ridge regresyon modelini kullanarak modele yanlılık sabiti ($k=0.80$) eklemişlerdir. Bu sayede araştırmacılar çoklu doğrusal bağlantı sorununu ortadan kaldırmış ve daha anlamlı bir biyolojik yoruma sahip daha kararlı ve doğru tahminler elde etmişlerdir. Araştırmacılar çoklu doğrusal bağlantı sorununu ortadan kaldırmak amacıyla ridge regresyon modelini önermişlerdir.

Ridge regresyonu ile yapılan diğer bir çalışmada Pimentel, Queiroz, Carvalheiro, ve Fries (2007) regresyon analizinde çoklu doğrusal bağlantı sorununu araştırmışlardır. Bunun için araştırmacılar regresyon analiz yöntemlerinden Ridge regresyon yöntemini kullanmışlardır. Araştırmacılar tropikal ve subtropikal bölgelerde yetiştirilen melez buzağuların performansını etkileyen parametreleri ve bunların çevreyle olan ilişkilerini araştırmışlardır. Araştırmacılar çoklu doğrusal bağlantı problemini çözmek için varyans arttırıcı değerlere, korelasyon matris değerlerine ve öz değerlere bakmışlardır. Araştırmacılar yaptıkları analizde çoklu doğrusal bağlantı sorununu belirlemişlerdir. Sonuç olarak araştırmacılar Ridge regresyon modelinin sadece değerlerin yorumlanmasında değil aynı zamanda tahmin amacıyla regresyon analizi yapıldığında, çoklu doğrusal bağlantı problemlerinin üstesinden gelmek için daha doğru bir regresyon modeli olduğunu önermişlerdir.

Topal, Eyduran, Yağanoğlu, Sönmez ve Keskin (2010) çalışmalarında sazan balığından elde ettikleri karkas ağırlığı bağımlı değişken, çeşitli vücut ölçüleri (total boy, standart boy, baş boyu, vücut yüksekliğini ise bağımsız değişken olarak balıklarda çoklu doğrusal regresyon analizi yapmışlardır. Bu analizi yaparken de vücut ölçüleri arasında çoklu doğrusal bağlantı durumu nedeniyle EKK regresyonu yerine ridge ve temel bileşenler regresyon analiz yöntemlerini kullanarak üç yöntemi karşılaştırmışlardır. Bunun yanında çoklu doğrusal bağlantının varlığını gösteren diğer değerlere de baktıklarında bu değerlerin hepsinin çoklu doğrusal bağlantı sorununun varlığını gösterdiğini ifade etmişlerdir. Araştırmacılar sonuç olarak çoklu doğrusal bağlantı varlığında en küçük kareler analiz yöntemi yerine Ridge ve temel bileşenler yöntemlerinin tercih edilmesini önermişlerdir.

Rathert, Üçkardeş, Narin ve Aksoy (2011) Japon bildircinlerinde yumurta iç kalite özelliklerinin tahmin edilmesinde kullanılan çoklu doğrusal regresyon metotlarından olan en küçük kareler ile temel bileşenler regresyon yöntemini karşılaştırmışlardır. Çalışmada yumurta ağırlığı, yumurta genişliği, yumurta uzunluğu ve şekil indeksi bağımsız değişkenler

olarak kullanılmıştır. Bu değişkenlerin bağımlı değişken olan ak yüksekliği, ak genişliği, ak uzunluğu, sarı uzunluğu ve sarı yüksekliğine olan etkileri hem EKK yöntemi hem de temel bileşenler regresyon yöntemi kullanılarak tahmin edilmiştir. Araştırmacılar özellikler arasında EKK tekniği ve temel bileşenler analizi ile hesaplanan varyans artırıcı değerleri (VIF) en küçük kareler yönteminde 10'u geçtiğini gözlenmiştir bu da bağımsız değişkenler arasında çoklu doğrusal bağlantı sorununun varlığını göstermiştir. Araştırmacılar sonuç olarak yumurta iç kalite özelliklerinin yumurta kabuğunu kırmadan dış kalite özellikleri kullanılarak modellenebileceği ve bağımsız değişkenler arasında çoklu doğrusal bağlantı sorununun ortaya çıkmasının muhtemel olduğunu bu nedenle EKK yöntemi yerine çoklu doğrusal bağlantı sorununu ortadan kaldıran temel bileşenler regresyon yönteminin kullanılmasının daha doğru ve güvenilir tahmin eşitlikleri oluşturduğunu bildirmiştir.

Üçkardeş, Efe, Narin ve Aksoy (2012) araştırmacılar Japon bildirimlerinde yumurta iç kalite kriterlerinden biri olan yumurta ak indeksini EKK yöntemi yerine RR yöntemini kullanarak tahmin etmeye çalışmışlardır. Çalışmada RR'ü seçme nedenleri olarak bağımsız değişkenler arasında yüksek korelasyon ilişkisine bağlı olarak çoklu doğrusal bağlantı sorunu olduğunu bildirmişlerdir. Bu çalışmada yumurtaların iç kalite özelliklerinden olan ak indeks değeri bağımlı değişken olarak (Y) bu değere tahmin etmek için ise bağımsız değişkenler olarak yumurta ağırlığı (X_1), yumurta genişliği (X_2), yumurta uzunluğu (X_3), Haugh birimi (X_4) ve şekil indeks değeri (X_5) olarak seçilmiştir. Çalışma sonucunda ak indeksini tahmin etmek için kullanılması önerilen model $Y = -11,743 + 0,201X_1 - 0,067X_2 - 0,081X_3 + 0,245X_4 - 0,008X_5$ olmuştur. Çalışmada RR analizi için k yanlılık sabiti 0,03 olarak alınmıştır. Bu değer $k=0,03$ olarak alındığında RR yöntemine göre bulunan modelin belirleme katsayısı 0,78 olup model istatistiki olarak anlamlı bulunmuştur ($P < 0,01$). Bu çalışmada da görüldüğü gibi çoklu doğrusal bağlantının varlığı durumunda EKK regresyon tekniği yerine RR tekniğinin kullanılmasını önermişlerdir.

Öztürk (2014) çoklu regresyon analizinde çoklu doğrusal bağlantı sorunu olan verilerde RR ve EKK yöntemine örnek bir veri setinde etkinliklerini karşılaştırmıştır. Bunun için hayvan materyali olarak etçi piliç karkas ağırlığını bağımlı değişken ve bunu etkilediğini düşündüğü protein miktarı (gr/gün), yem tüketimi (kg), cinsiyet, sıcaklık ($^{\circ}C$), ışıklandırma (saat/gün) ve su tüketimini (ml/gün) ise bağımsız değişkenler olarak almıştır. Yapılan çalışmada ele alınan bağımlı ve bağımsız değişkenler arasındaki korelasyon katsayılarına bakıldığında yüksek bulunmuştur. Korelasyon katsayılarının ve bağımsız değişkenlere ait bulunan VIF değerlerinin olarak 10'dan büyük çıkması araştırmacı açısından çoklu doğrusal

bağlantı varlığının kanıtı olmuştur. Çalışmada RR analizinde bulunan k yanlılık sabiti ise 0,03 olarak alınması ifade edilmiştir. Çalışmada k yanlılık sabiti 0,03 alınmış ve VIF değerlerinin tamamı 10'un altına düşmüştür. Çalışmada önerilen model istatistiki olarak önemli olup Karkas Ağırlığı=0,63 + 0,04 protein + 0,11 yem tüketimi + 0,11 cinsiyet - 0,01 sıcaklık + 0,0002 ışıklandırma süresi - 0,0009 su tüketimi. Sonuç olarak etçi piliçlerde karkas ağırlığına ait kestirim yapmak için ele alınan bağımsız değişkenler dikkate alındığında EKK yöntemi yerine RR yönteminin kullanılması gerektiğini vurgulanmıştır.

Shafey, Hussein, Mahmoud, Abouheif, Al-Batshan (2015) tavuklardan alınan yumurtaları adimsal regresyon ve Ridge regresyon modeli kullanarak çoklu doğrusal bağlantı sorununu araştırmışlardır. Yaptıkları çalışmada bağımsız değişken olarak yumurta ölçümlerini (yumurta ağırlığını, yumurta genişliğini, yumurta şekil indeksini, yumurta hacmini ve yumurta yoğunluğunu) ve bağımlı değişken olarak yumurta bileşenlerini (kabuk ağırlığını, sarı ağırlığını ve ak ağırlığını) almışlardır. Alınan tüm yumurtalarda çoklu doğrusal bağlantı sorunlarına yol açan özellikler arasında anlamlı korelasyon olduğunu tespit etmişlerdir. Araştırmacılar çalışmada 32, 45 ve 59 haftalık yaşlarda üretilen yumurtaların ridge regresyon analiz sonucunda oluşan VIF değerlerine baktıklarında sırasıyla 4,9'dan 8,2'ye, 1,4'ten 8,7'ye ve 1,3'ten 4,1'e kadar değiştiğini görmüşlerdir. Her üç yaş için VIF değerlerinin 10'dan küçük olması ridge regresyon metodunun çalışmada çoklu doğrusal bağlantı sorununu ortadan kaldırdığını göstermiştir. Sonuç olarak araştırmacılar RR yönteminin herhangi bir yaş grubunda üretilen yumurtalar için çoklu doğrusal bağlantı sorununu çözmek ve tavukların yumurta bileşenlerini yumurta ölçümlerinden başarılı bir şekilde tahmin etmek için ideal bir metot olduğu sonucuna varmışlardır.

Akçay ve Sarıözkan (2015) Ridge regresyon metodunu kullanarak yumurta tavukçuluğundan elde edilen satış gelirini modellemek için yaş, yaşama gücü, yumurta ağırlığı ve yumurta verimine ait verileri kullanarak bir model oluşturmaya çalışmışlardır. Çalışmada bağımsız değişkenler arasındaki korelasyon katsayısının 0,70'den yüksek, koşul indeksi değerinin 1000'i aşması ve 3 bağımsız değişkene ait VIF değerlerinin 10'dan yüksek olarak bulmuşlardır. Bu durumda çoklu doğrusal bağlantı sorunun var olduğu çalışmada gösterilmiştir. Bu nedenle çalışmada hem EKK regresyonuna ve buna alternatif olan Ridge regresyonu ile modelleme yapılarak her iki metot karşılaştırılmıştır. Araştırmacılar modelde yer alan parametrelerin regresyon katsayılarının standart hatalarının RR yönteminde EKK yöntemine göre daha küçük bulmuşlardır. Sonuç olarak çoklu doğrusal bağlantı sorunu

varlığında RR yönteminin EKK'ya göre daha küçük standart hatalı, daha tutarlı ve daha doğru tahminler sağladığı bildirilmiştir.

Çelik ve Söğüt (2015) Ridge regresyon metodu kullanarak Japon bıldırcınlarında yumurta ak ağırlığını tahmin etmeye çalışmışlardır. Yumurta ak ağırlığını belirlemek için yumurta ağırlığı, şekil indeksi, ak indeksi, sarı indeksi, sarı ağırlığı, kabuk kalınlığı, kabuk ağırlığı, ak yüksekliği Haugh ünitesi özelliklerinden yararlanmışlardır. Çalışmada RR teknikleri ve EKK analiz tekniğine göre modelleme yapılmış RR tekniğinde kullanılan parametrelere ilişkin standart hatalar EKK yöntemine göre daha küçük bulunmuştur. Çalışmada bağımsız değişkenler arasındaki korelasyon katsayılarına bakıldığında bazı özellikleri arasında yüksek korelasyon bulunmuştur. Aynı zamanda varyans arttırıcı faktör (VIF) değerine bakıldığında da bu değer bazı modellerde 10'dan büyük olduğu çalışmada görülmüştür. Sonuç olarak yapılan çalışmada RR kullanılarak elde edilen model Ak Ağırlığı = -3,102 + 0,659Yumurta Ağırlığı + 3,895Şekil İndeksi + 3,961Ak İndeksi + 0,786Sarı İndeksi - 0,241Sarı Ağırlığı + 3,512Kabuk Kalınlığı + 0,615Kabuk Ağırlığı + 0,05Ak Yüksekliği - 0,003 Haugh Birimi olarak bulunmuştur. Araştırmacılar çoklu doğrusal bağlantı durumunda RR yönteminin kullanılmasını önermişlerdir.

Çelik vd. (2018) beyaz hindilerde yaptıkları çalışmada bağımsız değişken olarak karkas parça ağırlıklarını (uyluk ağırlığı, göğüs ağırlığı, kanat ağırlığı, sırt ağırlığı, taşlık ağırlığı, kalp ağırlığı ve ayak ağırlığı), bağımlı değişken olarak ise bütün karkas ağırlığını almışlardır. Araştırmacılar çalışmada bağımsız değişkenlerin bağımlı değişkene etkilerini ridge regresyonu ve faktöre dayalı regresyon ile analiz etmişlerdir. Araştırmacılar en küçük kareler yöntemiyle elde edilen karkas ağırlığı ve karkas değerleri arasındaki doğrusal ilişkinin %99,8 ve regresyon modelinin istatistikî olarak anlamlı olduğunu bildirmişlerdir (P<0,001). Araştırmacılar EKK yönteminde uyluk ağırlığı ve kanat ağırlığının VIF değerlerinin 10'dan yüksek olduğunu bulmuşlardır. Araştırmacılar RR modeli ile yapılan analiz sonucunda tahmin edilen parametre sonuçlarının standart hatalarının ve VIF değerlerinin EKK yöntemine göre önemli ölçüde düştüğünü görmüşlerdir. Bu sonuçlara bağlı olarak araştırmacılar çalışmada RR yönteminin belirleme katsayısı (R²), faktöre dayalı regresyon analizine göre daha yüksek olduğundan ve karkas ağırlığını daha iyi açıkladığından RR yönteminin kullanılmasının daha uygun olduğunu önermişlerdir.

Yılmaz (2018) bir kabuklu türü olan Bivalv (*Abra alba*) türünde Ridge regresyon yöntemi kullanarak balıklarda çeşitli vücut ölçüleri (kabuk, boy, en, kalınlık ve cup indeksi)

kullanarak kabuk ağırlığını tahmin etmeye çalışmıştır. Bu çalışmada araştırmacı kabuk ağırlığını bağımlı değişken ve çeşitli vücut ölçülerini bağımsız değişken olarak almıştır. Çalışmada hesaplanan varyans arttırıcı (VIF) değerinin kabuk boyu, kabuk eni, kabuk kalınlığı için sırasıyla 10,27; 11,58 ve 25,12 olması bunun yanında hesaplanan öz değerlerin kabuk eni (0,07) ve kabuk kalınlığı (0,02) için 0'a yakın olması ve koşul indeksi değerinin kabuk kalınlığı için 113,20 yani 30'dan büyük çıkması saptandığı için çoklu doğrusal bağlantının varlığı kesin olarak gösterilmiştir. Çoklu doğrusal bağlantının varlığından dolayı araştırmacı RR yöntemi kullanmıştır. Bu şekilde RR ile yapılan analiz sonucunda ise EKK yöntemine göre bulunan değerler ile farklılık göstermiş RR modeli istatistik olarak ($P < 0,05$) anlamlı bulunmuş. VIF değerleri ve RR ile elde edilen standart hata değerlerinde önemli bir azalış görülmüştür. Çalışmada iki metot karşılaştırıldığında belirleme katsayıları EKK yöntemine göre 0,85 ve RR yöntemine göre 0,81 olarak bulunmuştur. Araştırmacı sonuç olarak RR yönteminin kullanılması ile çoklu doğrusal bağlantı sorununun aşılacağı ve bu şekilde yapılan modelleme ile kabuk ağırlığının tahmininin daha doğru olacağını ifade etmiştir.

Çiftsüren ve Akkol (2018) yaptıkları çalışmada yumurta kalitesini hem dış hem de iç kalite özelliklerini iki bölümde incelemişlerdir. Bu çalışmada araştırmacılar bağımlı değişken olarak iç kalite özelliklerinden yumurta ak ağırlığını (gr) ve yumurta sarı ağırlığını (gr), bağımsız değişken olarak ise yumurta dış kalite özelliklerinden yumurta ağırlığını (gr), yumurta kabuk ağırlığını (gr), yumurta genişliğini (mm), yumurta uzunluğunu (mm) ve şekil indeksini kullanmışlardır. Çoklu doğrusal regresyon analizi en küçük kareler yöntemine göre yapıldığında bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olduğu ve bu sorun çıkan sonuçları doğrudan etkilediğinden dolayı bunu önlemek için Ridge regresyonu, en az mutlak seçilme operatör (LASSO) ve Elastik Net (EN) analizlerini ayrı ayrı kullanmışlardır. Araştırmacılar bu sonuçları incelediğinde bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olduğunu görmüşlerdir. Araştırmacılar yumurta ak ağırlığı ve yumurta sarı ağırlığı tahmini için en küçük kareler yöntemi ve Ridge regresyon yönteminde 5 tane bağımsız değişken ve LASSO ile EN için ise 2 tane bağımsız değişken önermişlerdir. Araştırmacılar yapılan bu çalışmada Ridge regresyon analizi, LASSO ve EN metotların en küçük kareler yöntemine karşın kullanılmasının araştırmanın doğruluğu için daha doğru olacağını ifade vermişlerdir.

Çankaya, Eker, ve Abacı (2019) regresyon analizinde çoklu doğrusal bağlantı probleminin varlığında EKK, RR ve Temel bileşenler regresyonu yaklaşımlarını kullanarak bu üç metodu birbiriyle karşılaştırmışlardır. Bunun için Karayaka kuzusunda canlı ağırlık ile

bu kuzulardan alınan çeşitli vücut ölçüleri (cidago yüksekliği, sağrı yüksekliği, vücut uzunluğu, göğüs derinliği, göğüs çevresi, göğüs genişliği, ön, orta ve arka yumru yüksekliği) alınmıştır. Bu çalışmada CA bağımlı değişken ve alınan vücut ölçüleri ise canlı ağırlığı modellemek için kullanılan bağımsız değişkenler olarak tanımlanmıştır. Çalışmanın sonucunda bağımsız değişkenler arasında yüksek korelasyonların görülmesi çoklu doğrusal bağlantı sorununu meydana çıkarmış ve bu sorunu gidermek için tercih edilen RR ve PCR metotlarından RR metodu en küçük hata kareler ortalamasına ve en yüksek belirleme katsayısına sahip olduğu görülmüştür. Araştırmacılar çoklu doğrusal bağlantı sorunu olduğunda EKK yöntemi ile çoklu doğrusal regresyon analizi yerine bu durumda Ridge regresyon (RR) veya temel bileşenler regresyon (PCR) yöntemlerinden birinin kullanılmasını önermişlerdir.



5. SONUÇ ve ÖNERİLER

Bu çalışmada, çoklu doğrusal regresyon yöntemi kullanılarak tek bir bağımlı değişken ve bu bağımlı değişkeni açıkladığı düşünülen çok sayıda bağımsız değişken yardımı ile bağımlı değişkenin açıklanmasında kullanılan EKK ve RR yöntemlerinin karşılaştırılması yapılmıştır. Çoklu doğrusal regresyon analizi için bazı varsayımların gerçekleştiği varsayılarak araştırmacılar arasında çok sık olarak En Küçük Kareler (EKK) analiz yöntemi kullanılmıştır. Eğer bu varsayımlar gerçekleşmediği ve bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olduğunda yanlış tahmin metotlarından biri olan Ridge Regresyon (RR) yönteminin kullanılmasının gerekliliği bu çalışma ile vurgulanmıştır. Çalışmada çoklu doğrusal bağlantı sorunu, çoklu doğrusal bağlantı sorununun belirlenme yolları, bu durumda kullanılan yöntemlerden biri olan RR yönteminin ve bu yöntemde kullanılan yanlışlık sabiti olan (k) değerinin belirlenme yöntemleri belirtilerek bu yöntemin nasıl uygulandığı ve bulunan değerlerin nasıl yorumlandığı açıklanmıştır.

Çalışmada çoklu doğrusal regresyon analiz yöntemlerinden biri olan RR yönteminin, bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olduğu zamanlarda EKK yöntemine göre tercih edilerek kullanılmasını gerekliliği yapılan araştırma çalışmalarında gösterilmiştir. Çalışmanın daha somut ve anlaşılır olabilmesi için basit bir örnek model ile örneklendirilmiştir. Bu örnek veri seti için Esmer sığırların karkas ağırlıkları ve çeşitli vücut ölçüleri kullanılmıştır. Çalışmada bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu varlığı halinde çoklu doğrusal regresyon yöntemlerinden biri olan Ridge regresyon analizine göre elde edilen model ise $KA=388,08 + 1,82*KSY + 0,83*SY - 2,78*CY - 0,21*VU - 0,28*GÇ$ şeklinde bulunmuştur. Bu durumda EKK yöntemi ile elde edilen model ise $KA=383,65 + 4,43*KSY + 11,62*SY - 16,21*CY - 0,23*VU - 0,37*GÇ$ şeklinde olmuştur. Yapılan çalışmada her iki modelinde belirleme katsayısı düşük olup ($R^2 < 0,25$), bir bağımlı ve beş farklı bağımsız değişkenin beraber kullanıldığı model ele alınan özellik için istatistik olarak anlamlı bulunmamıştır. Buna rağmen RR yönteminin kullanılması beklenildiği gibi VIF değerlerini azalttığı, çoklu doğrusal bağlantı sorunundan kaynaklanan bağımsız değişkenlere ait regresyon katsayılarının standart hata değerlerini, EKK yöntemine göre daha da düşürdüğü görülmüştür.

Zootekni alanı içinde yapılan çeşitli çalışmalarda, EKK ile RR yöntemi karşılaştırılmaları yapılmıştır. Bu çalışmalar olabildiğince ayrıntılı olarak tez içinde verilmeye

alıřılmıřtır. Bu řekilde bađımlı deđiřkenlerin ok sayıda bađımsız deđiřken tarafından modellenmesi yapılacak alıřmalar iin kullanılacak yntem seimi, isabetin artırılması iin olduka nemlidir. Dođru yntemin seiminde, ncelikle oklu dođrusal bađlantı sorununun varlıđının belirlenmesi ve dođru yntemin kullanılmasının ne derece nemli olduđu bu konuda arařtırcıların daha dikkatli ve seici davranmalarının gerekliliđi alıřmamızda ve yapılan birok alıřmada vurgulanmıřtır.



6. KAYNAKLAR

- Akçay, A., Sariözkan, S. (2015). Yumurta tavukçuluğunda gelirin ridge regresyon analizi ile tahmini. *Ankara Üniversitesi Veteriner Fakültesi Dergisi*, 62, 69-74.
- Aktaş, C., Yılmaz, V. (2003). Çoklu bağıntılı modellerde liu ve ridge regresyon kestiricilerinin karşılaştırılması. *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi*, 4(2), 189-194.
- Albayrak, A. S. (2005). Çoklu doğrusal bağlantı haşinde en küçük kareler tekniğinin alternatif yanlı tahmin teknikleri ve bir uygulama. *ZKÜ Sosyal Bilimler Dergisi*, 1(1), 105-126.
- Anonim. (1989). German federal Republic, Bayerische Landesanstalt Fur Tierzueht Grub. 28, 36-38, Deutschland (*Anim. Breed. Abst.* 59,3419,1990).
- Arı, A., Önder, H. (2013). Farklı veri yapılarında kullanılabilecek regresyon yöntemleri. *Anadolu Tarım Bilimleri Dergisi*, 28(3), 168-174.
- Büyükuysal, M. Ç., Öz, İ. İ. (2016). Çoklu doğrusal bağıntı varlığında en küçük kareler alternatif yaklaşım: Ridge regresyonu. *Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi*, 6(2), 110-114.
- Çankaya, S., Eker, S., Abacı, S. H. (2019). Comparison of least squares, ridge regression and principal component approaches in the presence of multicollinearity in regression analysis. *Turkish Journal of Agriculture Food Science and Technology*, 7(8), 1166-1172.
- Çelik, Ş., Söğüt, B. (2015). Estimation of the egg albumen weight in the japanese quails using ridge regression method. *International Journal of Science and Research*, 2319-7064.
- Çelik, Ş., Şengül, T., Söğüt, B., İnci, H., Şengül, A. Y., Kayaokay, A., Ayaşan, T. (2018). Analysis of variables affecting carcass weight of white turkeys by regression analysis based on factor analysis scores and ridge regression. *Brazilian Journal Poultry Science*, 273-280.
- Çiftsüren, M. N., Akkol, S. (2018). Prediction of internal egg quality characteristics and variable selection using regularization methods: Ridge, LASSO and Elastic Net. *Arch. Anim. Breed*, 61, 279-284.

- Eduardo, da Cruz, Gouveia, Pimentel., Sandra, Aidar, de Queiroz., Roberto, Carneiro., Luiz, Alberto, Fries. (2007). Use of ridge regression for prediction of early growth performance in crossbred calves. *Genetics and Molecular Biology*, 30,3, 536-544.
- Hoerl, A. E., Kennard, R. W. (1970a). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-56.
- Hoerl A. E., Kennard, R. W. (1970b). Ridge regression: applications to nonorthogonal problems, *Technometrics*, 12(1), 69-82.
- Karagöz, Y. (2016). SPSS ve AMOS 23 uygulamalı istatistiksel analizler, Nobel Kitabevi, yayın no:1640, ISBN:978-605-320-547-0, İstanbul.
- Kayaalp, G. T., Güney, Çelik, M., Cebeci, Z. (2015). Çoklu doğrusal regresyon modelinde değişken seçiminin zootekniye uygulanışı. *Çukurova Üniversitesi Ziraat Fakültesi Dergisi*, 30(1), 1-8.
- NCSS Inc.: NCSS User Guide (2001), Kaysville, NCSS Inc.
- Orhunbilge, N. (1996). Uygulamalı regresyon ve korelasyon analizi. Avcıol Basın Yayınevi, 1. Basım, 261 S, ISBN:9754044562, İstanbul.
- Orhunbilge, N. (2017). Uygulamalı regresyon ve korelasyon analizi. *Nobel Yayınevi, gözden geçirilmiş 3. Basım, VIII, 386 S, ISBN:978-605-320-644-6, İstanbul.*
- Öztürk, İ. (2014). Hayvansal üretim verilerinde çoklu bağlantı probleminin yanlı regresyon yöntemi ile çözümü. *KSÜ Doğa Bilimleri Dergisi*, 17(3).
- Rathert Çiçek, T., Üçkardeş, F., Narin. D., Aksoy, T. (2011). comparison of principal component regression with the least square method in prediction of internal egg quality characteristics in japanese quail. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 17(5), 687-692.
- Schoeman, S. J., Aziz, M. A., Jordaan, G. F. (2002). The influence of multicollinearity on crossbreeding parameter estimates for weaning weight in beef cattle. *South African Journal of Animal Science*,32(4), 239-246.
- Shafey, T. M., Hussein, E. S., Mahmoud, A. H., Abouheif, M. A., Al-Batshan, H. A. (2015). Managing collinearity in modeling the effect of age in the prediction of egg

components of laying hens using stepwise and ridge regression analysis. *Brazilian Journal Poultry Science*, 473-482.

Soysal, M. İ. (2012). Biyometrinin prensipleri, genel yayın no:10 Ders Notu Yayın No:3 Sayfa No:351

Topal, M., Eyduran, E., Yağanoğlu, A., Sönmez, A. Y., Keskin, S. (2010). Çoklu doğrusal bağlantı durumunda ridge ve temel bileşenler regresyon analiz yöntemlerinin kullanımı. *Atatürk Üniversitesi Ziraat Fakültesi Dergisi*, 41(1), 53-57.

Uluslan, H. O. K., Solmaz, R., Ekici, Z. (1996). Besi sığırlarında beden ve karkas özellikleri ile aralarındaki ilişkiler. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 2 (1), 7-12.

Üçkardeş, F., Efe, E., Narin, D., Aksoy, T. (2012). Japon bıldırcınlarında yumurta ak indeksinin ridge regresyon yöntemiyle tahmin edilmesi. *Akademik Ziraat Dergisi*, 1 (1), 11-20.

Vovesny, V. (1982). Linear dependence between body conformation and meat performance indices in slaughter cattle. *Zivacisna Vyroba*, 27 (10),

Yılmaz, E. C. (2018). Bir Bivalv türünün (*Abra alba* (W.Wood, 1802) kabuk ağırlığının Ridge regresyon yöntemi ile tahmini. *KSÜ Tarım ve Doğa Dergisi*, 21(4), 593-598.

ÖZGEÇMİŞ

Ben, Metehan Sarp ESKİTOROS. 1992 yılında Ankara ili, Altındağ ilçesinde doğdum. İlk, orta ve lise öğrenimimi Ankara'nın Mamak ilçesinde, lisans öğrenimimi ise Tekirdağ Namık Kemal Üniversitesi Ziraat Fakültesi'nde tamamladım. Lisans eğitimimden sonra lisansüstü eğitim için Tekirdağ Namık Kemal Üniversitesi Fen Bilimleri Enstitüsü Zootekni Anabilim Dalında yüksek lisans eğitimine başladım. Yüksek lisans eğitimimin ikinci yılında Erasmus+ programı aracılığıyla bir sene Macaristan'da Szent Istvan Üniversitesi'nde eğitim gördüm.

