Turkish Journal of Agricultural and Natural Sciences 12 (4): 1166–1174, 2025
https://doi.org/10.30910/turkjans.1757995

TÜRK
TARIM ve DOĞA BİLİMLERİ
DERGİSİ

TURKISH
JOURNAL of AGRICULTURAL
and NATURAL SCIENCES

www.dergipark.gov.tr/turkjans
Original Article

Comparative Evaluation of Some Penalty Regression Techniques in the Multicollinearity Problem

Tolga Tolun[1]✉

[1]Kahramanmaraş Sütçü İmam University, Faculty of Agriculture, Department of Bioengineering and Sciences, 46100, Kahramanmaraş, TÜRKİYE

[1]https://orcid.org/0000-0003-4081-1222

✉: tolga_tolun@hotmail.com

ABSTRACT

This study aimed to predict an important biological trait, such as egg albumen height, in the presence of multicollinearity problem using some external quality parameters (egg weight, width, length, shape index, Haugh unit). Although a high coefficient of determination (R^2=0.995) was obtained in the multiple regression model generated using the Classical Least Squares Method (LSM), serious multicollinearity problem was detected among the independent variables, negatively impacting the model's reliability. To address this issue, LASSO and Liu regression techniques were applied; in models generated with both methods, the explanatory factor R^2 decreased to approximately 89 %, but the multicollinearity problem was effectively mitigated. Comparisons also showed that the Liu regression model outperformed the LASSO model in terms of information criteria (AIC, cAIC, BIC). The results show that regression methods with penalty terms provide reliable and consistent estimates in data sets with multicollinearity problems, and these techniques are recommended for the analysis and modeling of biological data.

Key words: Multicollinearity problem, LASSO and Liu regression methods, MATLAB, Egg albumen height.

Çoklu Doğrusal Bağlantı Sorununda Bazı Ceza Terimli Regresyon Tekniklerinin Karşılaştırmalı Değerlendirmesi

ÖZ

Bu çalışmada, çoklu doğrusal bağlantı probleminin varlığında yumurta ak yüksekliği gibi önemli bir biyolojik özelliğin, bazı dış kalite parametreleri (yumurta ağırlığı, eni, boyu, şekil indeksi, Haugh birimi) kullanılarak tahmin edilmesi amaçlanmıştır. Klasik En Küçük Kareler Yöntemi (LSM) ile oluşturulan çoklu regresyon modelinde yüksek belirleyicilik katsayısı (R^2=0.995) elde edilmesine rağmen, bağımsız değişkenler arasında ciddi çoklu doğrusal bağlantı sorunu tespit edilmiştir ve bu durum modelin güvenilirliğini olumsuz etkilemiştir. Bu sorunu çözmek için LASSO ve Liu regresyon teknikleri uygulanmış; her iki yöntemle oluşturulan modellerde açıklama oranı R^2 yaklaşık % 89'a gerilemiş, ancak çoklu doğrusal bağlantı problemi etkin biçimde azaltılmıştır. Yapılan karşılaştırmalar da ise, bilgi kriterleri (AIC, cAIC, BIC) açısından Liu regresyon modelinin LASSO modeline göre daha üstün performans sergilediğini göstermiştir. Sonuçlar, ceza terimli regresyon yöntemlerinin çoklu doğrusal bağlantı probleminin bulunduğu veri setlerinde güvenilir ve tutarlı tahminler sağladığını ortaya koymakta olup, biyolojik verilerin analizi ve modellemesinde bu tekniklerin tercih edilmesi önerilmektedir.

Anahtar kelimeler: Çoklu doğrusallık sorunu**,** LASSO ve Liu regresyon yöntemi, MATLAB, Yumurta ak yüksekliği.

INTRODUCTION

In biological and agricultural research, the multicollinearity problem significantly reduces model reliability and prediction accuracy, especially in regression analyses (Montgomery et al., 2012). This problem arises when there is a high correlation between independent variables, increasing the variance of regression coefficients and reducing statistical significance (James et al., 2013). Therefore, the presence of multicollinearity problem, especially in biological data, complicates model interpretation and negatively affects prediction performance (Zhang & Lu, 2007).

In egg quality research, accurate and reliable modeling of the relationships between external quality parameters and internal quality indicators is critical for improving product quality and optimizing studies in the poultry sector (Amerah et al., 2014). Egg albumen height is an important biological parameter considered an indicator of egg freshness and quality, and its estimation using external quality measurements is valuable from both practical and economic perspectives. Furthermore, other external quality traits such as shell thickness, shell weight, egg weight, and shape index also hold substantial economic value. Taking these parameters into account not only enhances the accuracy of internal quality predictions but also contributes to evaluating factors related to transportability, marketability, and shelf life. (Samli et al., 2005).

The Classical Least Squares Method (LSM) is one of the fundamental techniques widely used in regression analysis; however, the presence of multicollinearity can lead to problems such as overfitting in the model and high variance in the estimates (Hoerl & Kennard, 1970). To overcome this problem, techniques with penalty terms, such as LASSO (Tibshirani, 1996) and Liu regression (Liu, 1993), which are regularization methods, have been developed in the literature. LASSO regression method provides the advantage of variable selection by reducing some of the regression coefficients to zero, while Liu regression method increases model stability through parameter minimization (Friedman et al., 2010; Liu, 1993).

This study was conducted to estimate egg albumen height using external quality parameters obtained from Hy-Line hybrid hen eggs and to compare the performance of LASSO and Liu regression techniques against multicollinearity problem. Thus, it is aimed to contribute to the effectiveness and applicability of regression methods with penalty terms in biological data.

MATERIALS AND METHODS

The animal material for the study consisted of Hy-Line native layer hybrid hens developed by the Ankara Poultry Research Institute. A total of 152 egg samples were obtained from 53-week-old Hy-Line hybrids raised in cage system number 3 at the Çukurova University Research and Application Farm. Data obtained from internal and external egg quality analyses were evaluated at the Poultry Laboratory of the Animal Science Department of the Çukurova University Faculty of Agriculture.

Egg width and length measurements were taken with a digital caliper with a precision of ±0.01 mm. Egg weight was determined using an electronic scale with a precision of 0.01 g. The eggs were broken individually on a glass breaking bench, and the dimensions of the yolk and albumen (height, width, and length) were recorded with a digital caliper with a precision of ±0.01 mm. Additionally, the egg albumen and yolk were collected in separate petri dishes and weighed using an electronic scale with a precision of 0.01 g. Based on the collected measurements, egg quality indicators such as shape index, albumen index, yolk index and Haugh unit were calculated using the following formulas (Yannakopoulos & Tserveni-Gousi, 1986; Kaya & Aktan, 2011; Olawumi & Christiana, 2017):

Albumen Index= Albumen height (mm) / Average of Albumen Width and Length (mm)*100 (1)

Haugh unit= 100 log (Albumen Height +7.57-1.7 x Egg Weight $^{0.37}$) (2)

Yolk Index = Yolk height (mm) / Yolk Width (mm)*100 (3)

Shape Index Egg Width (mm)/ Egg Length (mm)*100 (4)

In the study, a prediction model was developed using albumen height as the dependent variable and egg weight, width, length, shape index, and Haugh unit as independent variables. Multiple regression analysis was applied to evaluate the model's performance and fit, and the analyses were conducted using the SPSS 26 statistical package.

Multiple regression analysis is a statistical method that allows a dependent variable to be predicted using multiple independent variables. In this analysis, the basic model equation is expressed as follows (Üçkardeş et al., 2012):

$$Y = \delta_0 + \delta_1 X_1 + \cdots + \delta_n X_n + \varepsilon \quad (5)$$

In this equation, Y represents the dependent variable, δ_0 represents the constant coefficient, $\delta_1 - \delta_n$ represents the regression coefficients, and ε: represents the error term.

In multiple regression analysis, certain assumptions must be met to ensure model validity and reliability. One of these assumptions is the absence of a high-level linear relationship between the independent variables. In other words, each independent variable is expected to provide unique and independent information in the model. However, if there is a high correlation between the independent variables, this leads to the problem of multicollinearity. Multicollinearity problem can reduce the statistical significance of the regression coefficients and weaken the predictive power of the model (Maxwell, 2000; Montgomery et al., 2001).

The most commonly used methods for identifying multicollinearity problems can be listed as follows:

Simple correlation coefficient: A correlation of 0.90 or higher among independent variables indicates multicollinearity problem (Çokluk, et al., 2014).

Variance increment factor (VIF): This method allows assessing the degree of linear relationship between each independent variable and other independent variables and is one of the key criteria for identifying multicollinearity problem. VIF indicates the extent to which the variance of an independent variable is affected by its linear relationship with other independent variables. As the VIF value increases, the linear relationship of that variable with other independent variables becomes stronger. In the literature, a VIF value of 10 or higher indicates the presence of multicollinearity problem in the model (Albayrak, 2005). The VIF value is calculated using the following formula:

$$\mathrm{VIF}=\frac{1}{1-\mathrm{R}^2} \tag{6}$$

Tolerance value (TV): It is an indicator that measures the relationship between each independent variable and the other independent variables. TV is calculated by subtracting the coefficient of determination (R^2) from 1 (Albayrak, 2005).

$$\mathrm{TV}=1-\mathrm{R}^2 \tag{7}$$

Because each of these methods assesses the presence of multicollinearity problem from different perspectives, using them together yields more reliable results. Various solution methods exist in the literature to mitigate the problem of multicollinearity. These include variable removal, combining correlated variables, adding new variables to the model, and regularization techniques, particularly LASSO and Liu regression methods. In this study, LASSO and Liu regression methods were chosen to address the problem of multicollinearity.

LASSO and Liu regression methods are statistical techniques developed to overcome the problem of multicollinearity and are used instead of classical LSM. The LSM method can cause large variances in the estimates when there is a high correlation between independent variables, negatively affecting the reliability of the model.

These methods stabilize the model by adding penalty terms to the regression coefficients to address the problem of multicollinearity between highly correlated explanatory variables. LASSO regression method reduces some coefficients to zero by selecting variables, while Liu regression method increases the stability of the estimates by establishing a balance between LSM and the penalty term. These techniques effectively address the problem of multicollinearity by reducing the prediction variance and increasing model stability. However, due to the use of penalty terms, the interpretability of the coefficients can be somewhat more complex compared to classical regression models. Despite this, the established stable structure significantly improves the overall performance of the model. Therefore, LASSO and Liu regression methods are widely preferred in statistical modeling applications when multicollinearity problem is evident (Albayrak, 2005).

RESULTS AND DISCUSSION

Within the scope of internal and external quality analyses conducted on egg samples from Hy-Line chickens, egg albumen and yolk indices and height parameters were evaluated as dependent variables. Based on the modeling studies, egg albumen height with a coefficient of determination (R^2) of 0.75 or higher was selected as the most appropriate dependent variable. Therefore, a regression model was created that could significantly predict egg albumen height from external quality measurements, and the model was determined as follows.

$$\hat{\mathrm{Y}}=\widehat{\delta_0}+\widehat{\delta_1}\mathrm{X}_1+\widehat{\delta_2}\mathrm{X}_2+\widehat{\delta_3}\mathrm{X}_3+\widehat{\delta_4}\mathrm{X}_4+\widehat{\delta_5}\mathrm{X}_5\hat{\mathrm{e}} \tag{8}$$

$\hat{Y}$ = Egg albumen height
X_1 = Egg weight
X_2 = Egg width
X_3= Egg length
X_4= Shape index
X_5= Haugh unit

The LSM method was then applied to determine the external quality characteristics affecting egg albumen height. The findings demonstrating the overall significance of the model are presented in Table 1. The regression coefficients for the independent variables and their statistical significance levels are reported in Tables 2 and 3, respectively. Furthermore, the VIF value and TV results calculated to assess multicollinearity among the independent variables are presented in Table 4.

Table 1. Goodness of fit and standard error values of the multiple regression equation obtained by the LSM.

R	R Square	Adjusted R Square	Std. Error of the Estimate
0.998	0.995	0.995	0.072

Table 2. Variance analysis value of the multiple regression equation obtained by the LSM.

	Sum of Squares	df	Mean Square	F	Sig.
Regression	155.130	4	31.026	5890.707	0.000
Residual	0.769	146	0.005		
Total	155.899	151			

Table 3. Parameter estimate and significance level of the multiple regression equation obtained by the LSM.

	Regression Coefficient	Std. Error of the Estimate	t	p
$\widehat{\delta_0}$	-10.537	3.696	-2.851	0.005
$\widehat{\delta_1}$	0.053	0.004	12.914	0.000
$\widehat{\delta_2}$	0.044	0.090	0.486	0.627
$\widehat{\delta_3}$	-0.036	0.064	-0.554	0.580
$\widehat{\delta_4}$	-0.023	0.050	-0.448	0.655
$\widehat{\delta_5}$	0.193	0.001	168.268	0.000

Table 4. The relationship between the independent variables of the multiple regression equation obtained by the LSM, VIF values and TV.

Correlation Matrix							
Variables	Egg Weight	Egg Width	Egg Length	Shape Index	Haugh Unit	TV	VIF
Egg Weight	1					0.059	17.037
Egg Width	0.923^{**}	1				0.002	505.563
Egg Length	0.855^{**}	0.699^{**}	1			0.001	699.819
Shape Index	-0.141	0.163^{*}	-0.591^{**}	1		0.003	375.543
Haugh Unit	-0.143	-0.146	-0.160^{*}	0.057	1	0.968	1.034

As a result of examining Table 1, it is seen that the multiple regression model created with the LSM method explains the relationship between egg albumen height and determined external quality traits at a rate of 99 %. This finding reveals that external quality traits have a significant effect on egg albumen height. However, the high explanatory power of the model also creates a situation that may indicate a multicollinearity problem. Indeed, the statistical analysis results presented in Table 2 show that the overall significance of the regression was achieved ($p<0.05$) and the model was statistically significant. On the other hand, when the significance tests of the regression coefficients in Table 3 are examined, it is seen that only the coefficients $\hat{\delta}_1$and $\hat{\delta}_5$ are statistically significant while the other variables did not reach the significance level ($p>0.05$). This indicates that there are high level linear relationships between the independent variables. As a matter of fact, when Table 4 is examined in order to support this situation, it is seen that especially between egg weight and width and length at a rate of 92 % and 85 %, respectively; Additionally, a highly significant correlation of 69 % was found between egg width and length ($p < 0.05$). Furthermore, the fact that most of the reported VIF values were 10 and above confirms that the model has a serious multicollinearity problem. To definitively identify this problem;

VIF value: $\frac{1}{1-R^2}$ (9)

with the help of the formula,

$$\text{VIF value} = \frac{1}{1-0.99} = \frac{1}{0.01} = 100 \quad (10)$$

calculated.

When egg albumen height was considered as the dependent variable, a VIF value of 10 or higher (VIF ≥ 10) for the variables included in the model was a clear indicator of multicollinearity problem. Therefore, to eliminate multicollinearity problem, the Lasso and Liu regression methods, widely used in the literature, were implemented with the support of the MATLAB programming package.

In the multiple regression model re-estimated with the Lasso and Liu regression methods, the explanatory power of the linear relationship between egg albumen height and external quality characteristics decreased to 89 %. This demonstrates that the multicollinearity problem was effectively mitigated. Furthermore, graphs between the actual and estimated values obtained with the Lasso and Liu regression methods are shown in Figures 1 and 2, while parameter estimates, significance levels, VIF values, and TV are presented in Tables 5 and 6.

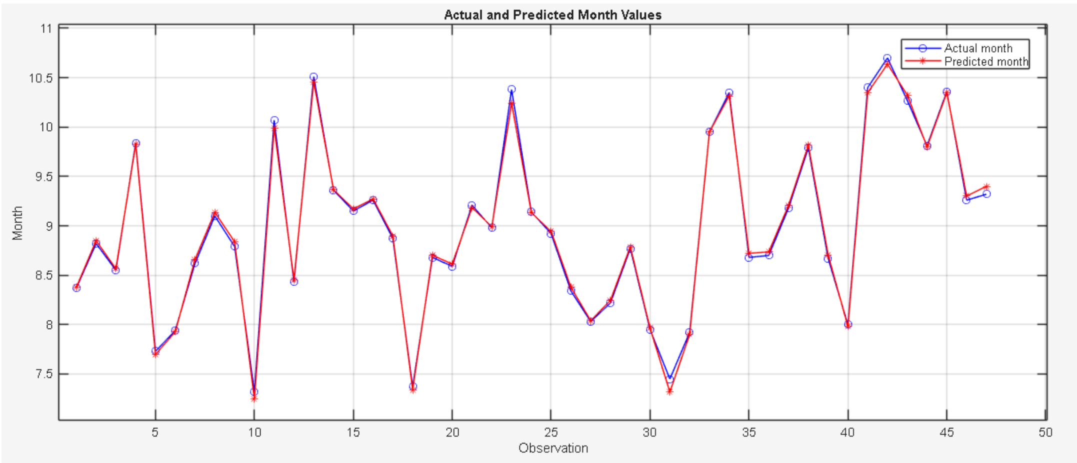

Figure 1. Actual and estimated values obtained by the Lasso regression method.

Table 5. Parameter estimates of the multiple regression equation obtained by the LASSO regresyon method, VIF values and TV.

	Regression Coefficient	Std. Error of the Estimate	VIF	TV
$\widehat{\delta_0}$	-13.508	-	-	-
$\widehat{\delta_1}$	0.047	0.008	9.466	0.001
$\widehat{\delta_2}$	0.0164	0.110	4.937	0.006
$\widehat{\delta_3}$	0	-	-	-
$\widehat{\delta_4}$	0	-	-	-
$\widehat{\delta_5}$	0.200	0.001	1.013	0.987

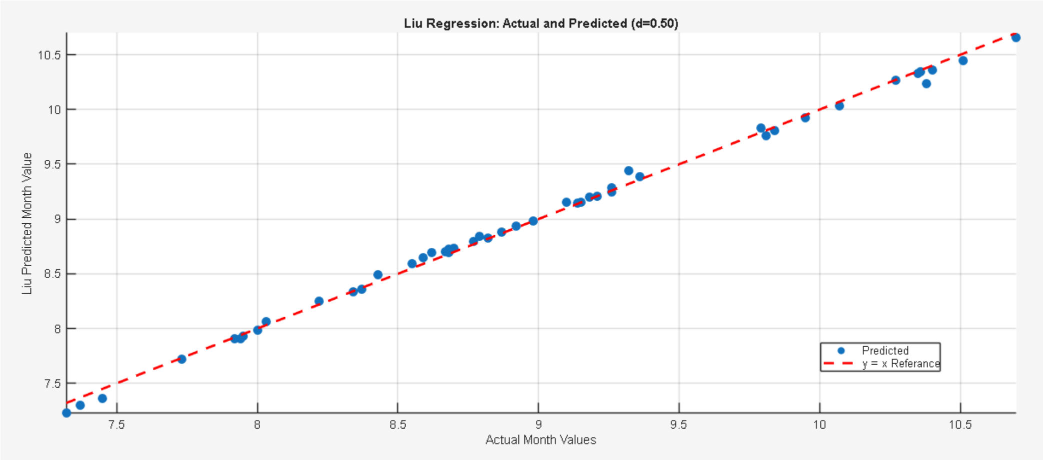

Figure 2. Actual and estimated values obtained by Liu regression method.

Table 6. Parameter estimates of the multiple regression equation obtained by the Liu regresyon method, VIF values and TV.

	Regression Coefficient	Std. Error of the Estimate	VIF	TV
$\widehat{\delta_0}$	-0.0001	-	-	-
$\widehat{\delta_1}$	0.071	0.009	9.514	0.002
$\widehat{\delta_2}$	0.214	0.124	6.125	0.004
$\widehat{\delta_3}$	-0.218	0.082	5.398	0.008
$\widehat{\delta_4}$	-0.145	0.052	3.129	0.110
$\widehat{\delta_5}$	0.200	0.001	1.013	0.988

An examination of Tables 5 and 6 reveals that the parameter estimates obtained with the Lasso and Liu regression methods differ significantly from those obtained with the LSM method. Lasso and Liu regression methods are parametric regularization techniques developed to reduce multicollinearity problem and yield more reliable estimates.

Lasso regression method allows variable selection by applying an L1 penalty term, reducing some regression coefficients to zero; this feature facilitates model simplification and interpretation. Liu regression method, on the other hand, allows parameter minimization and offers more flexible solutions by using an additional deviation parameter. In models estimated with both methods, the effects of multicollinearity problem were significantly reduced, and the standard errors of the parameter estimates were lowered.

The decrease in the linear relationship explanatory factor to 89 % in both models indicates that the model does not exhibit overfitting and provides more consistent results. As a result, Lasso and Liu regression methods produced models with more robust, reliable, and higher predictive power compared to the classical LSM method. The models estimated with these methods are given below.

LASSO regression method:

$$Y = -13.508 + 0.047X_1 + 0.016X_2 + 0X_3 + 0X_4 + 0.200X_5 \tag{11}$$

Liu regression method:

$$Y = -0.0001 + 0.071X_1 + 0.214X_2 - 0.218X_3 - 0.145X_4 + 0.200X_5 \tag{12}$$

Additionally, to determine which of the two regression methods applied in the case of multicollinearity problem is more effective than the other, several comparison criteria commonly used in the literature were evaluated. In this context, the coefficient of determination (R^2) was calculated, along with the Akaike Information Criterion (AIC), the corrected Akaike Information Criterion (cAIC), and the Bayesian Information Criterion (BIC), which are frequently used information criteria in model comparisons. The evaluation results based on these information criteria are presented in Table 7.

Table 7. Comparison criteria for the LASSO and Liu regression methods.

Regression	R^2	AIC	cAIC	BIC
LASSO	0.898	-149.689	-147.589	-138.588
Liu	0.890	-151.492	-150.540	-144.092

When the comparison criteria in Table 7 are examined, it is seen that the Liu regression method has the lowest values compared to the other method when all criteria are taken into account. This finding reveals that when evaluated in the context of multicollinearity problems, Liu regression method performs more effectively than the Lasso regression method.

According to the LSM analysis results presented in Tables 1, 2, and 3, it was determined that most of the parameters in the estimated regression model were not statistically significant ($p > 0.05$). Furthermore, the standard error value for the constant term δ_0 was found to be relatively high compared to other parameters. These findings are consistent with the studies conducted by Küçük (2019) and Özbey (2012) in the literature and are noteworthy in terms of revealing the limitations of the LSM method in similar data structures.

An examination of Table 4 reveals significant findings regarding the relationships between the independent variables. A positive and significant correlation was found between egg weight and egg width and length at 92 % and 85 %, respectively, and between egg width and length at 69 % ($p < 0.05$). These findings are consistent with the findings of Akbaş et al. It is parallel to the results of the studies conducted by (1996), Aktan (2004), Kul & Şeker (2004), Alkan et al. (2010) and Gök et al. (2023).

According to the analysis results in Tables 5 and 6, the Lasso and Liu regression methods applied to address the multicollinearity problem provided lower parameter estimates compared to the classical LSM method. Furthermore, considering the comparison criteria presented in Table 7, it was possible to determine which of these three regression models was more effective than the others.

The findings obtained are as follows; Albayrak (2005), Topal et al. (2010), Üçkardeş et al. (2012), Akçay & Sarıözkan (2015), Çiftsüren (2017), Yalçınöz & Şahin (2020), Demir et al. (2021), Tolun et al. (2023), Yavuz et al. (2023), Çetenak et al. (2024), Tolun et al. (2024) is also compatible and consistent with the results of the studies conducted by Gök & Kurşun (2025a), Gök & Kurşun (2025b) and Kurşun & Gök (2025).

CONCLUSION

This study aimed to predict egg albumen height using some external quality parameters in egg samples from Hy-Line chickens. A multiple regression model with very high explanatory power was created between egg albumen height, selected as the most suitable dependent variable, and external quality measurements. However, serious multicollinearity problem was detected in the model. This reduced the reliability of parameter estimates obtained with the classical LSM.

Lasso and Liu regression methods were applied to address the multicollinearity problem. Although the models generated with these methods exhibited a slight decrease in the explanatory power (R^2 approximately 89 %), it was observed that the model was free of overfitting, and the parameter estimates became more consistent and significant. In the comparisons, the information criteria (AIC, cAIC, BIC) values were found to be lower in the Liu regression method compared to the Lasso regression method, demonstrating that the Liu regression method provides more effective and reliable model performance.

In light of these results, the use of parametric adjustment methods robust to multicollinearity problem is recommended to model the effect of egg external quality traits on egg albumen height. The Liu regression method, in particular, should be preferred because it increases the reliability of model parameter estimates and improves the overall performance of the model.

In future studies, comparing different parametric and nonparametric regression techniques, as well as testing model performance with different chicken breeds and environmental conditions, will significantly contribute to egg quality assessment studies.

Declaration of interests

The author declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Author Contributions

Tolga TOLUN **:**Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; software; writing— original draft; writing—review and editing.

ORCID
Tolga TOLUN **http://orcid.org/0000-0003-4081-1222**

Article History	
Submission received:	04.08.2025
Revised:	11.10.2025
Accepted:	14.10.2025

REFERENCES

Akbaş, Y., Altan, O., & Koçak, C. (1996). Effects of hen's age on external and internal egg quality characteristics. *Turkish Journal of Veterinary Animal Science*, *20*, 455-460. http://dx.doi.org/10.55730/1300-0128.4074

Akçay, A., & Sarıözkan, S. (2015). Yumurta tavukçuluğunda gelirin Ridge Regresyon analizi ile tahmini. *Ankara Üniversitesi Veterinerlik Fakülte Dergisi*, *62*, 69-74.

Albayrak, S. A. (2005). Çoklu bağlantı halinde en küçük kareler tekniğinin alternatifi yanlı tahmin teknikleri ve bir uygulama. *Zonguldak Kara Elmas Üniversitesi Sosyal Bilimler Dergisi*, *1*, 105-126.

Alkan, S., Karabağ, K., Galiç, A., Karslı, T., & Balcıoğlu, M. S. (2010). Effects of selection for body weight and egg production on egg quality traits in Japanese quails (Coturnix coturnix japonica) of different lines and relationships between these traits. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 16(2), 239-244. http://dx.doi.10.9775/kvfd.2009.633

Aktan, S. (2004). Determination of some internal and external quality traits and their relationships in quail eggs by digital image analysis. *Animal Production 45*(1), 7-13.

Amerah, A. M., Ravindran, V., Lentle, R. G., & Thomas, D. G. (2007). Feed particle size: Implications on the digestion and performance of poultry. *World's Poultry Science Journal*, *63*(3), 439-455. https://doi.org/10.1017/S0043933907001560

Çetenak, T., Gök, İ., Yavuz, E., & Şahin, M. (2024). Statistical models and evaluation criteria used in poultry farming. *Black Sea Journal of Agriculture*, 7(6), 710-719. http://dx.doi.10.47115/bsagriculture.1532659

Çiftsüren, M. N. (2017). Yumurta iç ve dış kalite özellikleri arasındaki ilişkiyi belirlemede Ridge ve Lasso regülarizasyon yöntemlerinin karşılaştırılması. (Master Thesis) Yüzüncü Yıl Üniversity, The Institute of Natural Sciences, Van, Türkiye.

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2014). Sosyal bilimler için çok değişkenli istatistik SPSS ve lisrel uygulamaları [3.Basım]. Ankara: Pegem Akademi.

Demir, Y., Keskin, S., & Çavuşoğlu, Ş. (2021). Doğrusal olmayan temel bileşenler analizinin tanıtımı ve uygulanabilirliği. *Kahramanmaraş Sütçü İmam Üniversitesi Tarım Ve Doğa Dergisi*, *24*(2), 442-450.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1-22. https://doi.org/10.18637/jss.v033.i01 arXiv+7jstatsoft.org+7PubMed+7

Gök İ., Şahin M., & Tolun T. (2023). Determination of Impact Size by Canonic Correlation Analysis of the Factors Affecting the Buying or Selling Agricultural Lands. *Cumhuriyet Science Journal*, *44*(2),411-417. https://doi.org/10.17776/csj.1139858

Gök, İ., & Kurşun, K. (2025a). Prediction model of albumen index and height in Japanese quail eggs via external quality characteristics. *International Journal of Agriculture Environment and Food Sciences*, *9*(2), 493-501. https://doi.org/10.31015/2025.2.22

Gök, İ., & Kurşun, K. (2025b). Comparative Estimation Models of Egg Albumen Index in Atak-S Hens with Ridge and Principal Component Regression Methods . *Turkish Journal of Agriculture - Food Science and Technology*, *13*(9), 2615–2622. https://doi.org/10.24925/turjaf.v13i9.2615-2622.7699.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67. https://doi.org/10.1080/00401706.1970.10488634

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R. Springer.

Kaya, E., & Aktan S. (2011). Flock age and hatching egg storage period in Japanese quail: 1. Effects on dark albumen traits. *Süleyman Demirel University Journal of Agriculture Faculty*, *6*(2), 30-38, ISSN 1304-9984.

Kul, S., & Şeker, I. (2004). Phenotypic correlations between some external and internal egg quality traits in the Japanese quail (*Coturnix coturnix* japonica). *International Journal of Poultry Science*, *3*(6), 400-405. http://dx.doi.org/10.3923/ijps.2004.400.405

Kurşun, K., & Gök, İ. (2025). Prediction model of albumen ındex in duck eggs via external egg quality characteristics in case of multicollinearity. *Black Sea Journal of Agriculture*, *8*(4), 525-532.
Küçük, A. (2019). Doğrusal regresyonda Rıdge, Liu ve lASSO tahmin edicileri üzerine bir çalışma. (Master Thesis) Hacetepe Üniversity, The Institute of Natural Sciences, Ankara, Türkiye.
Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods,* 22(2), 393-402.
Maxwell, S. E. (2000). Sample size in multiple regression analysis. *Psychological Methods*, *5*(4), 434-458. https://psycnet.apa.org/doi/10.1037/1082-989X.5.4.434
Montgomery, D. C., Peck E. A., & Vining, G. G. (2001). Introduction to Linear Regression Analysis, 3rd Edition, John Wiley & Sons, New York.
Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to Linear Regression Analysis. John Wiley & Sons.
Olawumi, S., & Chiristiana, B. (2017). Phenotypic correlations between external and internal egg quality traits of Coturnix quails reared under intensive housing system. *Journal of Applied Life Sciences International*, *12*(3), 1-6. https://doi.org/10.9734/JALSI/2017/33802
Özbey, F. (2012). Çoklu iç ilişki olması durumunda lineer regresyon modellerinin kullanımı ile kestirim ve öngörü. (Ph. D. Thesis), Çukurova Üniversity, The Institute of Natural Sciences, Adana, Türkiye.
Samli, H. E., Okur, A. A., & Senkoylu, N. (2005). Effects of storage time and temperature on egg quality in old laying hens. *Journal of Applied Poultry Research, 14*(3), 548–553. https://doi.org/10.1093/japr/14.3.548
Topal, M., Eyduran, E., Yağanoğlu, M., Sönmez, A.Y., & Keskin, S. (2010). Çoklu doğrusal bağlantı durumunda Ridge ve Temel Bileşenler regresyon analiz yöntemlerinin kullanımı. *Atatürk Üniversitesi Ziraat Fakültesi Dergisi*, (41), 53-57. ISSN : 1300-9036.
Tolun, T., Yavuz, E., Şahin, M., & Gök, İ. (2023). Modeling egg curves ın partridges. *Black Sea Journal of Agriculture*, 6(1),21-25. https://doi.org/10.47115/bsagriculture.1139272
Tolun, T., Gök, İ., & Şahin, M. (2024). Modeling of some egg characteristics in henna partridges. *Black Sea Journal of Agriculture*, *7*(6), 729-742). https://doi.org/10.47115/bsagriculture.1555738
Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B (Methodological), 58(*1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x
Üçkardeş, F., Efe, E., Narinç, D., & Aksoy, T. (2012). Japon bıldırcınlarında yumurta ak indeksinin Ridge regresyon yöntemiyle tahmin edilmesi. *Akademik Ziraat Dergisi*, *1,*11-20. ISSN: 2147-6403.
Yannakopoulos, A. L., & Tserveni-Gousi, A. S. (1986). Qualitycharacteristics of quaileggs. *British PoultryScience, 27*, 171-176. . https://doi.org/10.4236/ojas.2021.112016
Yalçınöz, E., & Şahin, M. (2020).Yumurtacı tavuklarda yumurta verim eğrilerinin modellenmesi. *KSÜ Tarım ve Doğa Dergisi, 23*(5),1373-1378 https://doi.org/10.18016/ksutarimdoga.vi.631937
Yavuz, E., Abacı, S. H., Erensoy, K., & Şahin, M. (2023). Modeling of ındividual egg weights of Lohmann-Brown layer hens. *Turkish Journal of Veterinary & Animal Sciences*, *47*(3), 229-335. https://doi.org/10.55730/1300-0128.4290
Zhang, H., & Lu, W. (2007). Adaptive Lasso for Cox's proportional hazards model. *Biometrika, 94*(3), 691-703. https://doi.org/10.1093/biomet/asm056