Prediction of Egg Weight from External Egg Traits of Guinea Fowl Using Multiple Linear Regression and Regression Tree Methods

ABSTRACT

The study was done to predict egg weight from the external traits of the Guinea fowl egg using the statistical methods of multiple linear regression (MLR) and regression tree analysis (RTA). A total of 110 eggs from a flock of 23-week-old Guinea fowl were evaluated. Egg weight (EW) and external traits: eggshell weight (ESW), egg polar diameter (EPD), egg equatorial diameter (EED), egg shape index (ESI), and egg surface area (ESA) were measured. Descriptive statistics, Pearson correlation coefficients, and regression equations using the MLR were obtained; additionally, a RTA was done using the CHAID algorithm with the SPSS software (IBM ver. 22). EW presented positive correlations (p<0.0001) with ESA (r = 0.72), EPD (r = 0.65), and EED (r = 0.49). EW can be predicted through MLR using ESA as a predictor variable (R² = 72%). Predictive accuracy improves when adding EPD and EED traits to the model (R² = 75%). The RTA built a diagram using ESA, EED, and EPD as significant independent variables; of these, the most important variable was ESA (F = 50.295, df1 = 4, and df2 = 105; Adj. p<0.000) and the variation explained for EW was 74%. Likewise, the RTA showed that the highest egg weight (41.818 g) is obtained from eggs with a surface area > 59.03 cm² and a polar diameter > 5.10 cm. The proposed statistical methods can be used to reliably predict the egg weight of Guinea fowl.

INTRODUCTION

Currently, Guinea fowl has become a promising poultry species for rural farmers since it is an important source of food with good protein content such as meat and eggs. Furthermore, economic income can be obtained through the marketing of live Guinea fowl and their eggs (Kgwatalala et al., 2013). Compared with other poultry species such as the domestic hen, Guinea fowl are economically more viable for tropical regions, as they have good adaptability; because of this, greater attention to the potential of these birds and the quality of their products is required (Dzungwe et al., 2018).

Egg quality is a general term that refers to several parameters defined by its external and internal traits; therefore, its measurement is important in egg production for commercial purposes (Udoh et al., 2012). Egg weight is the most important trait in evaluating the external quality of eggs. Furthermore, it is in direct proportion to the weight of the albumen and the yolk, which are important parameters that determine the internal quality of eggs (Alkan et al., 2015), and thus, consumer acceptance (El-Tarabany, 2016). Also, EW has been shown to be related to the hatchability yield, embryo development, and chick weight at birth (Iqbal et al., 2016; Duman & Şekeroğlu, 2017). Because of the above and because egg weight is the primary...
criterion used in egg grading (small, medium, large and extra-large), adequate knowledge and prediction of EW can generate economic benefits for poultry producers and improve the methods of selection of eggs used in reproduction (Faridi et al., 2013; Okoro et al., 2017).

EW prediction can be successfully realized from external egg traits using different statistical methods (Khurshid et al., 2003; Orhan et al., 2016; Çelik et al., 2017). Predictive estimates and evaluation of the relationship between traits of interest are commonly performed using multiple linear regression analysis (MLR); however, these analyzes can be affected by problems of multicollinearity (high correlation between variables), causing errors in the interpretation of the results (Shafey et al., 2014). In face of this situation, it is important to complement the estimates done using MLR analysis with more efficient statistical procedures that avoid multicolinearity.

Regression tree analysis (RTA) is a non-parametric statistical method based on a tree diagram that has considerable advantages such as easy interpretation, assumptions of the distribution of the predictor variables are not required, being able to be applied using continuous dependent, nominal and ordinal variables, and not being affected by outliers (Mendeş & Akkartal, 2009; Topal et al., 2010). For the development of the regression tree, different data mining algorithms are used (CART, QUEST, CHAID, and exhaustive CHAID); however, previous studies have shown that predictive estimates using the CHAID algorithm showed models with better precision (Orhan et al., 2016; Çelik et al., 2017; Okoro et al., 2017).

Recently, studies using the RTA as a tool to predict traits of economic importance in animal science, such as body weight (Mendeş & Akkartal, 2009; Mohammad et al., 2012; Ali et al., 2015; Celik & Yilmaz, 2018), fleece weight (Eyduaran et al., 2016), weaning weight (Koc et al., 2017), and milk production (Eyduaran et al., 2013; Mikail & Bakir, 2019) have been increased; however, few studies have been done to evaluate and predict egg traits. Therefore, this study was carried out to predict egg weight from external traits of the Guinea fowl egg using the multiple linear regression and regression tree statistical methods.

**MATERIALS AND METHODS**

**Biological samples**

The care and handling of animals from which eggs was obtained was performed in accordance with the guidelines of official techniques of animal care and health in México (NOM-051-ZOO-1995).

For the study, a total 110 eggs obtained from a flock of Guinea fowl of first laying age (23 weeks), reared under natural environmental conditions using traditional management (Ruiz et al., 2014), were used in the experimental unit of the “Sustainable Tropical Animal Production” Academic Body of the Universidad Autónoma de Chiapas, located at the geographic coordinates of 19°8.64’ N and 98°16.55’ W, at an altitude of 522 masl. The region presents a subhumid warm climate with summer rains; A w2 (García, 2004). The mean annual temperature and total annual precipitation vary between 20-28 °C and 800-1200 mm, respectively (INEGI, 2017).

The eggs were collected in the early hours of the day (7:00-7:30 h) for one week. Each egg was labeled and stored at room temperature (18-27 °C) and average humidity (73.5 %) until their traits were measured.

**Estimation of egg traits**

Egg weight (EW) and egg shell weight (ESW) were recorded using an electronic scale (Medidata®) with a precision of 0.01 g. Egg polar diameter (EPD) and egg equatorial diameter (EED) were measured with a digital electronic caliper (Mitutoyo®) with an accuracy of 0.01 mm. Egg shape index (ESI) was calculated considering the EED/EPD x 100 ratio (Alkan et al., 2015). Egg surface area (ESA) was estimated using the following mathematical expression (Narushin, 2005):

\[
ESA (\text{cm}^2) = \left[3.155 - 0.0136 (\text{EPD}) + 0.0115 (\text{EED})\right] \times (\text{EPD})
\]

**Statistical analysis**

The external Guinea fowl egg traits data evaluated in the present study were analyzed using descriptive statistics, and Pearson correlation coefficients (r) between EW and the external egg traits were estimated. Linear regression models for EW were determined from the external egg traits with a multiple linear regression analysis using the stepwise option so that only the significant (p<0.05) predictor variables were included in the model.

The RTA was developed with the CHAID algorithm and the Bonferroni adjustment to obtain the adjusted p-values of F-values. This method performs an automatic pruning process and ignores the non-significant nodes, and uses the F significance test when analyzing a continuous dependent variable. The complete RTA methodology based on the CHAID algorithm has been previously described by Okoro et
al. (2017). A 10-fold cross-validation was used as the method for estimating prediction errors. In the RTA, the risk estimate is expressed as the variance within the subsets in the construction of the regression tree. The observed explained variation ($S_x^2$) in the dependent variable (EW) was estimated with the following equation (Mendes & Akkartal, 2009):

$$S_x^2 = (1 - S_e^2) \times 100$$

Where: $S_e^2$ is the unexplained variation in the dependent variable and is calculated as the risk/variance value of the root node ($S_e^2$).

All the statistical analyses of the data were done using the SPSS statistical software (IBM, ver. 22).

RESULTS AND DISCUSSION

The descriptive statistics of the evaluated external egg traits are shown in Table 1. The EW value determined in the present study (38.09 ± 3.21 g) was in contrast to the values determined in Guinea fowl raised in Nigeria (53.63 ± 0.15 g; Gwaza & Elkanah, 2017), Bosnia and Herzegovina (40.63 ± 0.27 g; Vekić et al., 2019), Turkey (40.14 ± 0.23 g; Alkan et al., 2013), and Poland (40.7 ± 0.54 g; Nowaczewski et al., 2008 and 40.8 ± 0.30 g; Bernacki et al., 2013). Previous studies reported that some factors such as age, body weight, and variety of birds (Oke et al., 2004; Kgwatalala et al., 2013), and laying management and intensity (Vekić et al., 2019), have important influence on the variation of egg weight in this poultry species; therefore, they should be considered in national programs to improve this trait.

Table 1 – Descriptive statistics of external egg traits of Guinea fowl.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg weight (g)</td>
<td>38.09±3.21</td>
<td>30.00</td>
<td>45.00</td>
<td>8.44</td>
<td></td>
</tr>
<tr>
<td>Egg polar diameter (cm)</td>
<td>4.90±0.18</td>
<td>4.40</td>
<td>5.60</td>
<td>3.85</td>
<td></td>
</tr>
<tr>
<td>Egg equatorial diameter (cm)</td>
<td>3.77±0.13</td>
<td>3.50</td>
<td>4.60</td>
<td>3.46</td>
<td></td>
</tr>
<tr>
<td>Egg shape index (%)</td>
<td>77.10±3.32</td>
<td>66.27</td>
<td>93.88</td>
<td>4.31</td>
<td></td>
</tr>
<tr>
<td>Egg shell weight (g)</td>
<td>7.16±1.06</td>
<td>5.00</td>
<td>10.00</td>
<td>14.88</td>
<td></td>
</tr>
<tr>
<td>Egg surface area (cm²)</td>
<td>57.97±3.41</td>
<td>51.90</td>
<td>70.80</td>
<td>5.89</td>
<td></td>
</tr>
</tbody>
</table>

SD: Standard deviation error; CV: Coefficient of variation.

The recorded EPD value = 4.90 ± 0.18 cm was lower than the values reported by Kgwatalala et al. (2013) in varieties: Pearl gray (5.09 ± 0.02 cm), Lavender (5.14 ± 0.02 cm), Royal purple (5.21 ± 0.03 cm), and White (5.01 ± 0.04 cm), while the reported EED values were similar to those found in this study (3.77 ± 0.13 cm), demonstrating that the length and width of the egg vary from the effect of the Guinea fowl variety used. The eggs evaluated in the present study presented an ESI = 77.10 ± 3.32%, which was consistent with the shape index observed in the eggs of Guinea fowl of the White (77.4 ± 0.36%) and Gray (75.6 ± 0.26) varieties (Bernacki et al., 2013).

Previously, Oke et al. (2004) found that, in Guinea fowl of the Pearl variety, the highest ESI value was reached at 40 weeks of age (81 ± 0.18%), thus the egg shape index has a significant effect on some quality characteristics of the egg (Duman et al., 2017), and so it should be considered in breeding programs in poultry.

The ESW value obtained, 7.16 ± 1.06 g, was higher than that found in eggs from Turkish (6.48 ± 0.080 g; Alkan et al., 2013), Polish (5.74 ± 0.06 - 5.70 ± 0.06 g; Bernacki et al., 2013), and Botswana (5.58 ± 0.08 - 6.33 ± 0.13 g; Kgwatalala et al., 2013) Guinea fowl. Likewise, the value recorded for ESA = 57.97 ± 3.41 cm² was higher than that reported in eggs of Turkish (55.69 ± 0.21 cm²; Alkan et al., 2013) and Polish (56.2 ± 0.49 cm²; Nowaczewski et al., 2008) Guinea fowl. These variations could be justified by the variety and genotype of Guinea fowl used in each study, since they have been reported to be factors that have an effect on ESW and ESA traits (Nowaczewski et al., 2008; Kgwatalala et al., 2013).

The Pearson correlation coefficients between EW and the external egg traits are shown in Table 2. EW showed positive correlations (p<0.0001) with the ESA (r = 0.72), EDP (r = 0.65), and EED (r = 0.49). Similar results were reported by Gwaza & Elkanah (2017) in French Guinea fowl, where egg weight showed high correlations with the egg length and egg width. This implies that as egg weight increases, so do egg length and egg width. In this sense, both traits can be used as selection criteria to improve egg weight (Dzungwe et al., 2018).

Table 2 – Correlation coefficients between external egg traits of Guinea fowl.

<table>
<thead>
<tr>
<th>Variable</th>
<th>EW</th>
<th>EPD</th>
<th>EED</th>
<th>ESI</th>
<th>ESW</th>
<th>ESA</th>
</tr>
</thead>
<tbody>
<tr>
<td>EW</td>
<td>1.00</td>
<td>0.65**</td>
<td>0.49**</td>
<td>-</td>
<td>-</td>
<td>0.72**</td>
</tr>
<tr>
<td>EPD</td>
<td>1.00</td>
<td>0.26*</td>
<td>-0.70**</td>
<td>-</td>
<td>-</td>
<td>0.82**</td>
</tr>
<tr>
<td>EED</td>
<td>1.00</td>
<td>0.49**</td>
<td>-0.75**</td>
<td>-</td>
<td>-</td>
<td>0.75**</td>
</tr>
<tr>
<td>ESI</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ESW</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ESA</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**p<0.0001; *p<0.05. EW: Egg weight; EPD: Egg polar diameter; EED: Egg equatorial diameter; ESI: Egg shape index; ESW: Egg shell weight; ESA: Egg surface area.

Prediction of egg weight using multiple linear regression

All the linear regression models developed to predict EW using the stepwise method were significant
(p<0.0001) and are shown in Table 3. EW can be adequately predicted using ESA as the predictor variable (R² = 72%; MSE = 3.21); however, predictive accuracy was improved by adding the EPD and EED traits to the model, since the value of the determination coefficient increased to 75%. These results are comparable with those obtained by Victoria & Dauda (2017) when

determining that the egg weight of Guinea fowl of the Pearl variety can be predicted with high precision (R² = 86.4%) using the width and length of the egg. For their part, Oke et al. (2004) found that the egg weight of Guinea fowl of the same variety can also be predicted using the body weight of the hens (R² = 69%).

Table 3 – Regression equations to predict egg weight using external egg traits of Guinea fowl.

<table>
<thead>
<tr>
<th>Eq. No</th>
<th>Equations</th>
<th>R²</th>
<th>MSE</th>
<th>Pr&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EW (g) = 0.93 (±0.06 **)×ESA</td>
<td>0.72</td>
<td>3.21</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>2</td>
<td>EW (g) = 5.69 (±3.50 ns)×EED + 0.79 (±0.10 **)×ESA</td>
<td>0.73</td>
<td>3.14</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>3</td>
<td>EW (g) = 132.81 (±48.26 ns)×EPD + 185.21 (±65.31 ns)×EED + -10.71 (±4.10 ns)×ESA</td>
<td>0.75</td>
<td>2.87</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Eq: equation; R²: Determination coefficient; MSE: Mean square error. EW: Egg weight; EPD: Egg polar diameter; EED: Egg equatorial diameter; ESA: Egg surface area. p-value; **p<0.0001; *p<0.05; ns: not significant.

**Prediction of egg weight using regression tree analysis**

The regression tree diagram developed using the CHAID algorithm to determine information on the predictor variables that significantly affected EW is shown in Figure 1. The significant independent variables included in the regression tree diagram were ESA, EED, and EPD, of which, the most important variable was ESA (F = 50.295, df1 = 4, and df2 = 105; Adj. p<0.0001). Initially, node 0, also called root node, grouped all the eggs evaluated in the present study (n = 110). The mean egg weight at node 0 was 38.054 g (S = 3.255). This node was divided into five secondary nodes based on the ESA variable; they showed an average EW range of 32.000 (node 1) to 40.349 g (node 5). Of the five nodes obtained in the regression diagram, nodes 1, 2, and 4 were terminal nodes. Node 1 grouped a total of 10 eggs with ESA ≤ 53.00 cm² and showed a mean EW of 32.000 g (S = 2.582). At node 2, 15 eggs were grouped with 53.00 < ESA ≤ 54.49 cm² and a mean EW of 35.000 g (S = 0.000). Node 3 included a group of 19 eggs with 54.49 < ESA ≤ 57.15 cm² and estimated mean EW of 37.368 g (S = 2.565), in parallel, this node branched out into two nodes (nodes 6 and 7) based on the ESA variable since it showed a significant effect on the EW that was grouped at node 3 (F = 17.895, df1 = 1 and df2 = 17; Adj. p<0.001). At node 6, a total of 13 eggs with EED ≤ 3.70 cm were grouped with an estimated mean EW of 36.154 g (S = 1.689) where

the largest group of eggs in the entire diagram (n = 43) was formed with ESA > 59.03 cm². Node 5 branched out into EPD-based nodes 8 and 9 because it was a significant variable (F = 14.893, df1 = 1, and df2 = 41; Adj. p<0.001). Node 8 grouped a total of 32 eggs with EPD ≤ 5.10 cm and generated a mean EW of 39.844 g (S = 0.884). Finally, at node 9, a group of 11 eggs with EPD > 5.10 cm was formed, which estimated a mean EW of 41.818 g (S = 2.523), which was the heaviest in the entire regression tree diagram. In this sense, the highest egg weight (41.818 g) is obtained from eggs with a surface area > 59.03 cm² and a polar diameter > 5.10 cm.

The explained observed variation for EW was calculated using the root node variance as: S_y² = 10.595. The unexplained variation was calculated using the risk value of the model as follows: S_y² = 2.758 / 10.595 = 0.26. With this, the explained variation for EW was: S_y² = (1- 0.26) x 100 = 0.74 = 74%. Previously, Orhan et al. (2016) also used the RTA based on the CHAID algorithm and determined that the most important independent variables to predict the egg weight of laying hens were albumin weight, yolk weight, and egg shell weight, with which the regression tree diagram was developed, and 99.988% of the variation of the evaluated dependent variable was explained.

**CONCLUSION**

The significant relationship found between EW and ESA, EPD and EED, suggests that it is possible to predict EW with sufficient accuracy from these external features of the egg through multiple regression analysis. Likewise, the regression tree analysis based on the CHAID algorithm proved to be convenient to
predict egg weight using the same external features of the egg as predictor variables. Both statistical methods can be used reliably for predictive estimates of egg weight as they showed similar accuracy. This study will serve as reference information for poultry producers and researchers in the creation of future national programs to improve traits of economic and biological importance such as egg weight in Guinea fowl.

**CONFLICT OF INTEREST**

The authors declare no conflict of interest.

**REFERENCES**


Prediction of Egg Weight from External Egg Traits of Guinea Fowl Using Multiple Linear Regression and Regression Tree Methods


Vekić M, Savić D, Cvijanović D. Changes of selected egg quality traits depending on the laying period of semi-intensively raised Guinea fowl hens. Agro-knowledge Journal 2019;20:141-149.