



*Arq. Bras. Med. Vet. Zootec.*, v.72, n.3, p.970-976, 2020

## Resizing the productive characteristics of meat-type quail through principal component analysis

[Redimensionamento das características produtivas de codornas de corte por meio da análise de componentes principais]

D.U. Tyska<sup>1</sup>, A.G. Gotuzzo<sup>1</sup>, N.J.L. Dionello<sup>1</sup>, R. Negri<sup>2</sup>,  
R.P. Della-Flora<sup>1</sup>, J.M. Germano<sup>1</sup>, F. Rutz<sup>1</sup>

<sup>1</sup>Aluno de pós-graduação - Faculdade de Agronomia Eliseu Maciel - Universidade Federal de Pelotas, RS

<sup>2</sup>Programa de pós-graduação - Universidade Federal do Rio Grande do Sul - Porto Alegre, RS

### ABSTRACT

This study aimed to establish criteria for eliminating redundant variables, to know the magnitude of the data relationship, and to provide information that helps researchers in the use of the technique to analyze and interpret production data and egg quality. The data used in this work was obtained from four successive generations of the quail lineage developed by the Department of Animal Science of the Federal University of Pelotas. The characteristics were measured from the 42nd day of age, when the egg production period began, until 126 days of production, obtaining three 28 day periods (cycles) in the four successive generations, totaling 545 females. Of the twelve original variables, only seven demonstrated potential to be maintained in future experiments, representing a 42% exclusion. The main philosophy of this study was the analysis of the studied variables and made possible the understanding of the relationship and the correlations.

Keywords: *coturnix coturnix coturnix*, discard of variables, egg quality, genetic, multivariate analysis

### RESUMO

*O objetivo do presente trabalho foi estabelecer critérios para eliminação de variáveis redundantes, conhecer a magnitude das relações dos dados, além de fornecer informações que auxiliem pesquisadores na utilização da técnica para analisar e interpretar dados de produção e qualidade de ovos de codornas. Os dados utilizados neste trabalho são provenientes de quatro gerações sucessivas da linhagem de codornas de corte desenvolvida pelo Departamento de Zootecnia da Universidade Federal de Pelotas. As características foram mensuradas do 42º dia de idade até o 126º dia de produção, totalizando 545 fêmeas. Das 12 variáveis originais analisadas, apenas sete demonstraram potencial para serem mantidas em experimentos futuros, representando uma exclusão de 42%. A análise de componentes principais foi efetiva para a redução das variáveis estudadas e possibilitou a compreensão da relação e correlação dessas.*

*Palavras-chave: análise multivariada, correlação genética, Coturnix coturnix coturnix, descarte de variáveis, qualidade de ovos*

### INTRODUCTION

Coturniculture has gained space within the poultry industry, because it allows a quick return of the invested capital to the producers, with the main products of this activity being the production of eggs and meat (Teixeira *et al.*, 2013). One of the bottlenecks of this production system is

obtaining genetic material, which for both egg production and meat production depend on importation (Dionello *et al.*, 2008), since Brazil does not have its own genetic material.

In search of a solution to this problem, some institutions have leveraged breeding programs for meat-type quail. The development of these strains

---

Recebido em 15 de abril de 2019

Aceito em 1 de outubro de 2019

E-mail: darilenetyska@gmail.com

occurs through the estimation of genetic parameters in order to guide selection decisions (Hidalgo *et al.*, 2007). Genetic evaluations are performed based on data collected through individual performance tests, and the estimation of genetic values via best linear unbiased prediction (BLUP). Based on these studies, the selection criteria are established for the next generations (Viana *et al.*, 2000), guaranteeing the renewal of herds with animals of superior genetic potential, adapted to the production systems in Brazil.

Among the characteristics measured in the programs, another problem arises, which should be measured within these breeding programs. One way to solve this problem is to define the characteristics that concentrate most of the total observed variation (Barbosa *et al.*, 2005), removing those that provide little or no additional information from the data analysis (Jolliffe, 1972). Multivariate statistics allows combining information of different characteristics simultaneously, preserving the biological information that these data provide, facilitating the selection and allowing the discrimination of superior populations, especially in the genetic context (Ledo *et al.*, 2003).

Among the multivariate analyzes, the principal component analysis (PCA) stands out because it allows the exclusion of redundant and difficult to measure characteristics, besides reducing the number of variables, allowing to explain the covariance structure and correlation between several characteristics simultaneously measured in one population (Baker *et al.*, 1988), where it should contribute to the interpretation of the relationships between variables and, consequently, decision making (Baker *et al.*, 1988; Roso and Fries, 1995).

Principal component analysis (PCA) is a classic technique in data analysis, being the most widely used multivariate technique to reduce the dimensionality of a data set (Shaker; Aziz, 2017). The main objective is to explain the structure of variance and covariance of the data, where each principal component is a linear combination of all original variables (Johnson and Wichern, 2002), these linear combinations are called principal components and are not correlated with each other (Hongyu *et al.*, 2015), retaining the maximum information of the total data variation. This study

aimed to establish criteria for elimination of redundant variables, to know the magnitude of data relations, providing theoretical basis for industry and researchers for interpretation of egg production and quality data.

## MATERIAL AND METHODS

The data come from four successive generations of the quail lineage (*coturnix coturnix coturnix*) developed by the Department of Animal Science/Eliseu Maciel College of Agronomy (FAEM)/ Federal University of Pelotas (UFPEl) in Pelotas, Rio Grande do Sul. The project was approved by the institution ethics committee under number 7252.

The animals were raised in the Teaching and Experimentation Laboratory Professor Renato Peixoto (LEEZO), where they were submitted to the selection of higher body weight at 28 days of age. The productive traits were measured in females from the 42nd day of life, when the egg production period begins, until 126 days, the end of production period. The data used come from the average of the values obtained in the three 28 day cycles, in four successive generations, with a total of 545 females evaluated in this period.

The eighteen variables studied were divided into two groups. Production variables: bird weight (BIRW), egg weight (EGGW), total feed intake (TFIN), feed conversion per dozen eggs (FDOZ), feed conversion per egg mass (CMAS), egg mass (MASS), total egg production (EGGP). Quality variables: albumen height (AHEI), length (LENG), bark thickness (BTHI), specific gravity (SGRA), width (WIDT), albumen weight (AWEI), shell weight (SWEI), yolk weight (YWEI), Haugh unit (HU) and morphological index (MORI). The color of the yolk cannot be analyzed in the PCA, because in obtaining it we used the Roche® colorimetric fan, which is a subjective method.

The multicollinearity test evaluates the linear dependence between the variables, avoiding the formation of singular or poorly conditioned matrices (Cruz and Carneiro, 2006). The Bartlett sphericity test tests the hypothesis of uncorrelated original variables, with the extreme case of  $R = I$ , where the main components would be the original variables (Hongyu *et al.*, 2015). Pearson's correlation coefficients ( $r$ ) for production characteristics and quality were estimated. From

the correlation matrix (R), data were generated for the analysis of the principal components, using the methodology explained in Hongyu *et al.* (2015).

To dispose of variables, the recommendation of Jolliffe (1972) was used, where the number of variables discarded should be equal to the number of components whose variance (self-value) is less than 0.7. The statistical analyses described were performed using the vegan statistical package, and the graphs (RcmdrPlugin. FactoMineR) of the Statistical Program R® version 3.3.0.

To estimate the heritability and genetic correlations of the traits, Bayesian inference and uni and multivariate analyses were used, following the animal model  $\gamma = X\beta + Z\alpha + e$  where:

$\gamma$  is the vector of observations;

$X$  is the fixed effects incidence matrix (bird generation) contained in the vector  $\beta$ ;

$\beta$  is the vector of the fixed effects;

$Z$  is the incidence matrix of additive genetic effects;

$\alpha$  is the additive genetic effects vector;

$e$  is the vector of random errors associated with each observation

It was further assumed that the fixed effect has uniform distribution and the inverted Wishart distribution variance components. For random effects normal distribution was assumed. The software GIBBSF90 and POSTGIBBSF90 (Misztal *et al.*, 2002) were used to obtain 5,000 samples of variance components to calculate heritability and genetic correlations for all traits.

The convergence monitoring of the chains generated by the Gibbs sampler was performed using Heidelberger and Welch (1983) diagnostic tests, available from the Convergence Diagnosis and Output Analysis (CODA), using the R software.

## RESULTS AND DISCUSSION

Following the proposed methodology, in the multicollinearity test, the variables FDOZ, CMAS, MASS, HU and MORI were removed because they are collinear, with variance inflation factor (Following the proposed methodology, in the multicollinearity test, the variables FDOZ, CMAS, MASS, HU and MORI were removed because they are collinear, with variance inflation factor (VIF) >10 (Cruz and Carneiro, 2006). The

Bartlett sphericity test, that presented P-value= 9,491215e<sup>-53</sup> for production and P-value= 0 for the quality data. In this case, confirming that there is correlation between the variables studied (Hongyu *et al.*, 2015).

The variables EGGW and BIRW have a moderate positive correlation of 56% (Table 1), where the 29% chance of the same genes influence the expression of these characteristics. One of the causes of this correlation is the pleiotropy that defines the process in which the same gene can affect two or more characteristics (Pereira, 2008). The heritabilities found suggest the possibility of genetic gains in these traits (Table 1, diagonal of the genetic correlation matrix) corroborating results by Della-Flora *et al.* (2012) in this same lineage.

As can be seen in the principal component analysis (Table 2), the original variable representing CPI is the EGGW, accounting for approximately 42% of the total variation alone, with a correlation between negative and moderate original characteristics except TFIN. CP2 retained approximately 28% variation with TFIN having the highest absolute value. CP3 concentrated 19% variation, maintaining the variable that represents this component (EGGP).

These three components together explained a high percentage of the total variation (89%), so they are interesting for animal evaluation and comparison purposes. The BIRW variable contributes with only 11% to the analysis and its eigenvector is lower 0.7 (Table 2) and its exclusion is a suggestion for future experiments.

The almost 90 ° angle (Figure 1) formed between the EGGW and TFIN variables shows the contrast between these variables that dominate the first PCs, and their low correlation. The approximation of the BIRW and EGGW axes show their medium-high correlation (0.56).

It is observed (Table 4) that the main components concentrate 83% of the preserved cumulative variance of the original data, that is, the egg quality data can be analyzed from these four components. Observing within PC1 the variables SWEI and LENG presented similar contributions within the component (Table 4), which can also be observed by the closed angle (Figure 2), showing the correlation between them and their redundancy.

*Resizing the productive...*

Table 1. Pearson Correlation Matrix (R) and Genetic Correlation Matrix of production variables, bird weight (BIRW), total egg production (EGGP), total feed intake (TFIN), egg weight (EGGW), of meat-type quail

Correlation matrix				
	BIRW	EGGP	TFIN	EGGW
BIRW	1			
EGGP	0.2	1		
TFIN	-0.06	0.17	1	
EGGW	0.56	0.22	0.02	1
Standard deviation	44.713	5.392	126.649	1.348
Genetic correlation matrix				
BIRW	0,44 (0.23-0.70)*			
EGGP	-0,14 (-0.28 – 0.04)*	0.32 (0.19 – 0.46)*		
TFIN	0.94 (0.86 – 0.98)*	0.21 (0.09 – 0.40)*	0.2 (0.08 – 0.31)*	
EGGW	0.29 (0.05 – 0.38)*	-0.19 (-0.26 – 0.02)*	0.1 (0.03 – 0.25)*	0.35 (0.20 – 0.43)*

Table 2. Principal component analysis of egg production variables, bird weight (BIRW), total egg production (EGGP), total feed intake (TFIN), egg weight (EGGW) of meat-type quails

Principal Component	CP1	CP2	CP3	CP4
Coefficients				
BIRW	-0.63939	0.2528	0.16481	-0.70718
EGGP	-0.40655	-0.4939	-0.76853	0.01191
TFIN	-0.06391	-0.81993	0.55912	-0.105
EGGW	-0.64947	0.14099	0.2638	0.69909
Variance ( $\lambda$ )	13.001	10.571	0.8709	0.6587
Total percentage change	0.423	0.279	0.19	0.109
Cumulative Variance	0.425	0.702	0.892	1

The SGRA and AHEI representing respectively poorly correlated CP2 and CP3, this contrast is evident from the near 90° angle (Figure 2). Miranda *et al.* (2013) using seven genetic groups of quail, evaluated egg quality variables, through (PCA) and concluded that the specific gravity and albumen weight are the characteristics selected because they present less redundancy among the others. The YWEI forms an acute angle (Figure 2)

with the remaining variables in CP4, showing its correlation with the variables WIDT, LENG, BTHI and AWEI, showing their redundancy.

After the analysis of principal components of the 12 original variables, variables that had low contribution were excluded, and only the remaining 7 are suggested to be kept in future experiments.

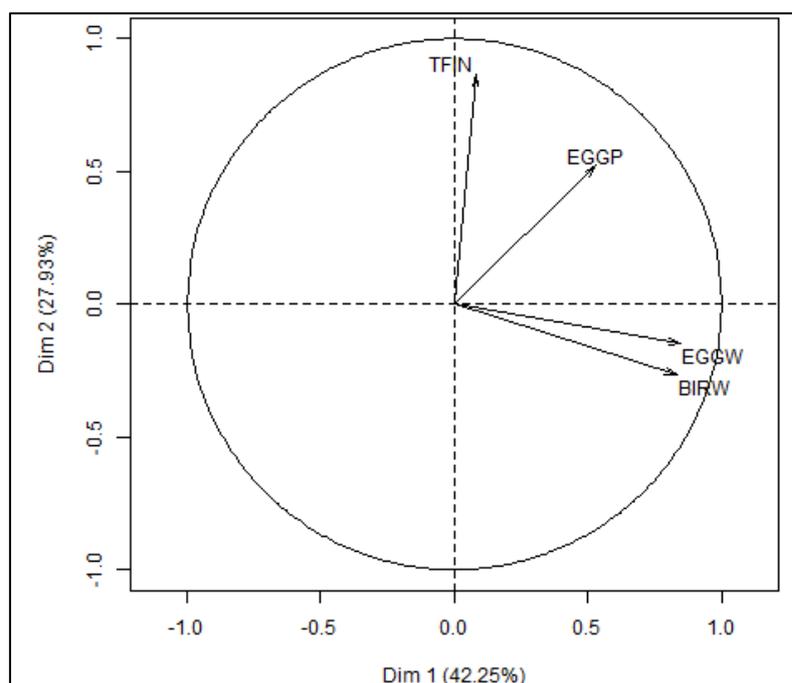


Figure 1. Dispersion of the first two main components of the production variables, bird weight (BIRW), total egg production (EGGP), total feed intake (TFIN), egg weight (EGGW) of broiler quail matrices.

Table 3. Correlation matrix of quality, albumen height (AHEI), egg length (LENG), shell thickness (BTHI), specific gravity (SGRA), egg width (WIDT), albumen weight (AWEI), shell weight (SWEI), egg yolk weight (YWEI), of the eggs of meat-type quails

Correlation matrix								
	AHEI	LENG	BTHI	SGRA	WIDT	AWEI	SWEI	YWEI
AHEI	1							
LENG	0.277	1						
BTHI	0.013	0.231	1					
SGRA	-0.08	0.11	0.624	1				
WIDT	0.354	0.611	0.207	0.114	1			
AWEI	0.126	0.544	0.31	0.169	0.598	1		
SWEI	0.367	0.503	0.575	0.457	0.415	0.491	1	
YWEI	0.029	0.525	0.137	0.021	0.466	0.358	0.434	1
Standard deviation	1.039	1.688	2.094	5.749	1.037	0.878	0.163	0.499
Genetic correlation matrix								
AHEI	0.39							
LENG	0.25	0.36						
BTHI	0.54	0.15	0.22					
SGRA	0.12	0.14	0.75	0.41				
WIDT	0.58	0.65	0.66	0.03	0.29			
AWEI	0.54	0.59	0.39	-0.21	0.93	0.34		
SWEI	0.48	0.56	0.85	0.51	0.74	0.46	0.72	
YWEI	0.88	0.75	0.38	0.49	0.48	0.62	0.39	0.28

Table 4. Principal component analysis of quality, albumen height (AHEI), egg length (LENG), shell thickness (BTHI), specific gravity (SGRA), width (WIDT), albumen weight (AWEI) variables, shell weight (SWEI), egg yolk weight (YWEI) of meat-type quail eggs

	Coefficients							
AHEI	-0.195	0.282	0.834	-0.142	0.004	-0.089	-0.017	-0.4
LENG	-0.421	0.244	-0.103	-0.022	0.199	0.841	0.067	-0.018
BTHI	-0.31	-0.542	0.046	-0.005	-0.166	-0.016	0.744	-0.164
SGRA	-0.227	-0.626	0.033	0.021	0.502	-0.012	-0.494	-0.242
WIDT	-0.411	0.278	0	0.309	0.541	-0.417	0.227	0.374
AWEI	-0.401	0.1	-0.185	0.627	-0.476	-0.11	-0.256	0.628
SWEI	-0.445	-0.18	0.199	-0.308	-0.401	-0.023	-0.283	-0.312
YWEI	-0.331	0.236	-0.466	-0.629	-0.001	-0.313	-0.027	-0.351
Variance ( $\lambda$ )	1.849	1.256	1.003	0.807	0.66	0.619	0.583	0.432
Total percentage change	0.428	0.197	0.126	0.081	0.054	0.048	0.043	0.023
Cumulative Variance	0.428	0.625	0.751	0.832	0.886	0.934	0.977	1

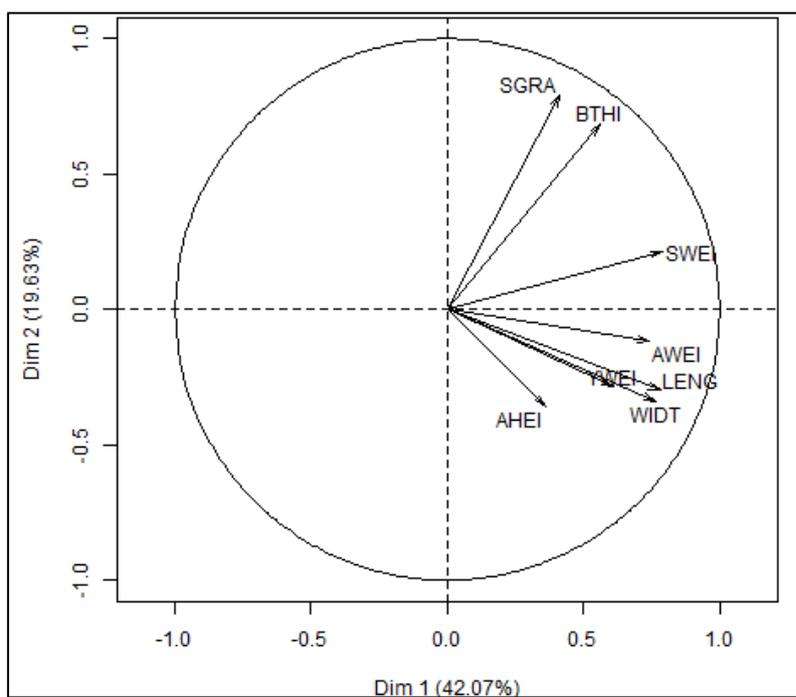


Figure 2. Scattering of the first two major components of the quality variables, albumen height (AHEI), egg length (LENG), shell thickness (BTHI), specific gravity (SGRA), width (WIDT), albumen weight (AWEI), shell weight (SWEI), egg yolk weight (YWEI) of meat-type quail eggs.

### CONCLUSION

The principal component analysis was effective to reduce the studied variables, suggesting a smaller number of variables to be evaluated in future studies, without significant loss of information, which would save time and labor.

### ACKNOWLEDGMENTS

To the Higher Education Personnel Improvement Coordination (CAPES), for the scholarship funding.

## REFERENCES

- BAKER, J.F.; STEWART, C.R.; CARTWRIGHT, T.C. Multiple regression and principal components analysis of puberty and growth in cattle. *J. Anim. Sci.*, v.66, p.2147-2158, 1988.
- BARBOSA, L.; LOPES, P.S.; REGAZZI, A.J. *et al.* Seleção de variáveis de desempenho de suínos por meio da análise de componentes principais. *Arq. Bras. Med. Vet. Zootec.*, v.57, p.805-810, 2005.
- CRUZ, C.D.; CARNEIRO, P.C.S. (Eds.). *Modelos biométricos aplicados ao melhoramento genético*. Viçosa: UFV, 2006. 480p.
- DELLA-FLORA, R.P.; GOTUZZO, A.; GERMANO, J.M. *et al.* Variabilidade genética de características produtivas e de qualidade de ovos de matrizes de codornas de corte. *Pubvet*, v.6, 2012.
- DIONELLO, N.J.L.; CORREA, G.S.S.; SILVA, M.A. *et al.* Estimativas da trajetória genética do crescimento de codornas de corte utilizando modelos de regressão aleatória. *Arq. Bras. Med. Vet. Zootec.*, v.60, p.454-460, 2008.
- HIDALGO, A.M.; MARTINS, E.N.; MICHELI, A. *et al.* Parâmetros genéticos para peso do ovo e peso corporal em codornas de postura. In: SIMPÓSIO INTERNACIONAL DE COTURNICULTURA, 3.; CONGRESSO BRASILEIRO DE COTURNICULTURA, 2., 2007, Lavras. *Anais...* Lavras: UFLA, 2007. p.219. (Resumo).
- HONGYU, K.; SANDANIELO, V.L.M.; OLIVEIRA-JUNIOR, G.J. Análise de componentes principais: resumo teórico, aplicação e interpretação. *Eng. Sci.*, v.5, p.83-90, 2015.
- JOHNSON, R.A.; WICHERN, D.W. (Eds.). *Applied multivariate statistical analysis*. New Jersey: Pearson, 2002. 773p.
- JOLLIFFE, I.T. Discarding variables in a principal component analysis. I: Artificial data. *J. Appl. Stat.*, v.21, p.160-163, 1972.
- LEDO, C.A.S.; FERREIRA, D.F.; RAMALHO, M.A.P. Análise de variância multivariada para os cruzamentos dialélicos. *Ciênc. Agrotec.*, v.27, p.1214-1221, 2003.
- MIRANDA, J.; FERREIRA, T.; ABREU, L. *et al.* Estudo da qualidade do ovo e descarte de variáveis em grupos genéticos de codornas utilizando análise multivariada. in X Simpósio Brasileiro de Melhoramento Animal, 2013, Uberaba. *Anais...* Uberaba, MG, 2013. (Resumo).
- MISZTAL, I.; TSURUTA, S.; STRABEL, T. *et al.* BLUPF90 and related programs (BGF90). WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier. *Proceedings...* Montpellier, France: [s.n.], 2002. CD-ROM. (Communication n° 28-07).
- NARDINO, M.; SOUZA, V.Q.; BARETTA, D. *et al.* Association of secondary traits with yield in maize F<sub>1</sub>'s. *Ciênc. Rural*, v.46, p.776-782, 2016.
- PAIVA, A.L.C.; TEIXEIRA, R.B.; YAHAKI, M. *et al.* Análise de componentes principais em características de produção de aves de postura. *Rev. Bras. Zootec.*, v.39, p.285-288, 2010.
- PEREIRA, J.C.C. (Ed.). *Melhoramento genético aplicado à produção animal*. Belo Horizonte: FEPMVZ, 2008. 618p.
- ROSO, V. M.; FRIES, L. A. Componentes principais em bovinos da raça Polled Hereford à desmama e sobreano. *Rev. da Soc. Bras. Zootec.*, v. 24, p.728-735, 1995.
- SHAKER, A.S.; AZIZ, S. R. Internal traits of eggs and their relationship to shank feathering in chicken using principal component analysis. *Poult. Sci. J.*, v.5, p.1-5, 2017.
- TEIXEIRA, B.B.; EUCLYDES, R.F.; TEIXEIRA, R.B. *et al.* Herdabilidade de características de produção e postura em matrizes de codornas de corte. *Ciênc. Rural*, v.43, p.361-365, 2013.
- VIANA, C.F.A.; SILVA, M.A. E.; PIRES, A.V. *et al.* Estudo da divergência genética entre quatro linhagens de matrizes de frangos de corte utilizando técnicas de análise multivariada. *Rev. Bras. Zootec.*, v. 29, p.1074-1081, 2000.
- WELCH, P.D. Simulation run length control in the presence of an initial transiente. *Oper. Res.*, v.31, p.1109-1144, 1983.