

## Principal components analysis applied to performance and carcass traits in the chicken

Luís Fernando Batista PINTO<sup>a\*</sup>, Irineu Umberto PACKER<sup>a</sup>,  
Cláudio Manoel Rodrigues DE MELO<sup>b</sup>, Mônica Corrêa LEDUR<sup>c</sup>,  
Luiz Lehmann COUTINHO<sup>a</sup>

<sup>a</sup> Animal Science Department, ESALQ/USP, Piracicaba, São Paulo State, Brazil

<sup>b</sup> Aquaculture Department, UFSC, Florianópolis, Santa Catarina State, Brazil

<sup>c</sup> EMBRAPA, Concórdia, Santa Catarina State, Brazil

(Received 30 May 2005 – Accepted 22 March 2006)

**Abstract** – In this study, the principal components methodology was used to analyze performance and carcass traits measured in a 3747 F2 experimental population of *Gallus gallus*. This technique allows us to reduce the number of variables considered in the evaluation of the animals, which may facilitate genetic programs. Performance traits were body weight at 35 and 42 days of age, and weight gain from 35 to 42 days of age; carcass traits were the following: weights of liver, heart, gizzard, wings, thighs, breast and lung. The five first principal components explained about 93.3% of the total variation, and the first component explained 66%. The first component was called ‘General Weight’, because the largest eigenvectors were associated with body weight at 35 and 42 days of age and liver, breast, wing and thigh weights. Heritability of the first principal component was  $0.23 \pm 0.05$ . The genetic gain of a selection index with the first two principal components was similar to the selection index gain with 10 original traits. The principal components methodology was efficient to evaluate the total variance in this group of correlated traits, allowing a drastic reduction in the number of traits to be included in the selection index for breeding purposes.

**body weight / broiler / multivariate analysis / poultry**

**Résumé** – Analyse en composantes principales des performances de croissance et des caractéristiques de carcasse chez le poulet. L’objectif de cette étude était de déterminer les composantes principales expliquant la variance des caractéristiques de carcasse et des mesures de développement chez *Gallus gallus*. Cette étude a permis de déterminer quelles étaient les caractéristiques essentielles de ces mesures au sein d’un programme d’amélioration génétique. Les critères retenus ont été le poids vif à 35 et 42 jours, le gain de poids entre 35 et 42 jours et plusieurs caractéristiques pondérales de carcasse (foie, coeur, gésier, ailes, cuisses, filet et poumons). Les cinq premières composantes principales ont permis d’expliquer 93,3 % de la variance totale, la première d’entre elles expliquant à elle seule 66 % de la variance totale. Nous l’avons nommée Poids Général, car les

---

\* Corresponding author: lfbpinto@esalq.usp.br

principaux vecteurs étaient associés au poids vif. L'héritabilité de la première composante principale était de 0,23. Le gain d'index de sélection avec deux composantes principales est apparu similaire au gain d'index de sélection obtenu à partir des 10 critères initialement retenus. En conclusion, cette analyse en composantes principales s'est montrée efficace pour estimer la variance totale d'un groupe de critères pris en compte pour la sélection. Cette approche diminue de manière importante le nombre de mesures à réaliser dans le cadre d'un programme d'amélioration génétique du poulet de chair.

## **analyse multivariée / oiseaux / poids vif / poulet**

### **1. INTRODUCTION**

Body composition and growth performance are important to assess the potential of development in broiler type chickens. Thus, body weight at various ages, weights of parts of the carcass, weight gain, carcass yield and feed efficiency are examples of variables that can indicate the usefulness of the chicken for commercial purposes.

A comparison of carcass composition and yield of 1957 versus 2001 broilers was conducted by Havenstein et al. [7], and the increase in breast meat was attributed mainly to genetic improvement. Reddish and Lilburn [12] observed larger breast weight in an improved strain for increased weight gain as compared with a control. Therefore, evaluation of carcass traits represents a primordial interest to supply information to the genetic breeding programs.

According to Morrison [11] the principal components analysis is a multivariate methodology that can be used with success when characteristics are correlated. This analysis transforms an original group of variables into another group, principal components, which are linear combinations of the original variables. The main advantage is the independence of these components. From the viewpoint of animal genetics and improvement, principal components simultaneously consider a group of attributes which may be interesting for selection purposes. Another important aspect is that each of the principal components explains a percentage of the total

variance. The first principal component explains the highest percentage of this variance.

There are only a few reports about the use of principal components analysis in chicken data. Ibe [8] analyzed the body weight of *Gallus gallus* at different ages, together with four body linear measurements: breast and thigh widths, and shank and keel lengths. In all ages the first two principal components explained at least 85% of the total variation. According to this author, the principal components could be used in the selection index to simplify them, because such an index would have few principal components in the place of all the original traits.

Abreu et al. [1] used the principal components technique to evaluate the combining ability of chicken strains for production traits. For this type of data the methodology was not efficient because it was not possible to explain a high percentage of the total variance with a reduced number of principal components. However, Abreu et al. [2], who worked with egg production and reproductive data, concluded that principal components analysis was efficient, since it was possible to explain 98% of the total variation with the first two components, and also reporting that selection based on components of high discriminatory value may bring satisfactory results.

Debut et al. [4] used principal component analysis for evaluating and understanding the relationships among meat quality indicators, such as muscle and meat quality traits, considering chickens from

two genotypes and different pre-slaughter stress conditions.

In this study, principal components analysis was applied to a set of growth and carcass traits of chickens in order to reduce the number of traits and understand their relationship for genetic and breeding purposes.

## 2. MATERIALS AND METHODS

### 2.1. Experimental animals

Data from 3747 birds from an experimental F2 population originating from reciprocal crossing between broiler and layer lines were used. Both lines were developed by the EMBRAPA Swine and Poultry Research Center, and have been under multi-trait selection for about 10 generations as reported by Figueiredo et al. [5, 6]. This population was raised for the purpose of quantitative trait loci mapping. The animals were raised as broilers, receiving water and feed ad libitum. The feed for the initial phase, from 1 to 21 days of age, contained 21% crude protein and 3150 kcal ME per kg; from 22 to 35 days, the chickens received feed for the growing phase, with 20% crude protein and 3200 kcal ME per kg, and in the final phase, from 36 to 41 days, the feed was composed of 18.5% crude protein and 3200 kcal ME per kg. Corn and soy crumb formed the feed base. After the weighting at 41 days, the chickens did not receive feed and they had their weights again measured at 42 days of age, a little before slaughter.

### 2.2. Traits evaluated

A total of 10 traits were considered in this analysis: body weight at 35 and 42 days of age, weight gain from 35 to 42 days of age, and the weights of the liver, heart, gizzard, wings, thighs, breast and lung.

### 2.3. Statistical analyses

The data were adjusted for the fixed effects of line, sex, hatch, and interactions line  $\times$  hatch, line  $\times$  sex, and sex  $\times$  hatch, with the GLM procedure of SAS [13].

The normality assumption was verified for each of the traits, and when data did not approximate to the normal distribution, it was transformed by square root or logarithmic scale.

According to Johnson and Wichern [9], principal components are linear combinations of the original variables and are estimated in such a way that the first principal component explains the largest percentage of the total phenotypic variance. The second principal component explains the second largest percentage and so on, until the whole variance is explained. Each of the original variables is associated to the principal component through an eigenvector, whose absolute value determines the importance of the traits for the principal component. Considering a  $p$ -variate system, the  $i$ th principal component is given in expression  $PC_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ij}X_j$ , with  $i = 1, 2, \dots, n$  principal components and  $j = 1, 2, \dots, p$  original variables, where  $a_{ij}$  is the  $j$ th component of the coefficient vector of the linear transformation, and  $X_j$  is the  $j$ th original variable. The eigenvectors and eigenvalues were estimated starting from the correlation matrix, because the variables presented very divergent variances. For the selection of the principal components, the criterion adopted was the percentage of approximately 90% of the total variance explained. The PRINCOMP procedure of SAS [13] was used for the principal component analysis.

Heritability of each of the principal components was estimated by the DXMUX program (Meyer, [10]), which fits an animal model, using a derivative-free algorithm and restricted maximum likelihood method.

**Table I.** Eigenvalues and proportion of total variance per principal components (PC).

Principal Components	Eigenvalues	Proportion of Total Variance	Accumulated Proportion
PC1	6.59707723	0.6597	0.6597
PC2	0.75749054	0.0757	0.7355
PC3	0.69741753	0.0697	0.8052
PC4	0.65991101	0.0660	0.8712
PC5	0.61846186	0.0618	0.9330
PC6	0.39269566	0.0393	0.9723
PC7	0.12582080	0.0126	0.9849
PC8	0.09477978	0.0095	0.9944
PC9	0.05628355	0.0056	0.9997
PC10	0.00006203	0.000006	1.0000

**Table II.** Pearson correlation among original traits and the first five principal components (PC).

	PC1	PC2	PC3	PC4	PC5
Liver weight	0.79	-0.03	-0.26	-0.03	0.05
Heart weight	0.62	0.24	0.24	-0.42	0.56
Gizzard weight	0.56	0.70	-0.35	0.24	-0.01
Wing weight	0.94	-0.09	-0.04	-0.03	-0.09
Thigh weight	0.95	-0.11	0.04	-0.05	-0.10
Breast weight	0.92	-0.17	-0.006	-0.08	-0.11
Lung weight	0.62	-0.14	0.28	0.64	0.34
Body weight at 35 days	0.91	-0.22	-0.28	-0.04	0.05
Body weight at 42 days	0.98	-0.08	-0.03	-0.04	-0.10
Weight gain from 35 to 42 days	0.68	0.29	0.53	-0.02	-0.39

The selection indexes for the principal components and original variables were obtained according to Van Vleck [15], by a SAS macro with comments presented by Saxton et al. [14]. In order to compare gains from these different indexes, economic values equal to one were used for each of the principal component and original trait. Due to scale differences between principal components and the original traits, the selection index gain was divided by its respective standard deviation.

### 3. RESULTS

While the first five principal components explained approximately 93.3% of

the total variance, the first principal component alone explained 66% (Tab. I). The correlation between the original traits and the first principal component were all positive and high (Tab. II), except for gizzard, heart and lung weights, and weight gain from 35 to 42 days of age. The first principal component has only positive eigenvectors, and represents a weighted average of the traits.

For the second, third, fourth and fifth principal components, the largest correlations were observed with gizzard (0.70), weight gain from 35 to 42 days of age (0.53), lung (0.64) and heart (0.56) (Tab. II). Those principal components explain about 7.56%, 6.97%, 6.6%, and 6.18% of the total variance (Tab. I),

**Table III.** Eigenvectors for each of the principal component (PC).

	PC1	PC2	PC3	PC4	PC5
Liver weight	0.31	-0.03	-0.31	-0.04	0.06
Heart weight	0.24	0.28	0.29	-0.52	0.71
Gizzard weight	0.22	0.80	-0.42	0.30	-0.01
Wing weight	0.37	-0.11	-0.05	-0.04	-0.11
Thigh weight	0.37	-0.13	0.05	-0.06	-0.12
Breast weight	0.36	-0.19	-0.01	-0.09	-0.13
Lung weight	0.24	-0.16	0.33	0.78	0.43
Body weight at 35 days	0.35	-0.26	-0.33	-0.05	0.06
Body weight at 42 days	0.38	-0.09	-0.03	-0.05	-0.13
Weight gain from 35 to 42 days	0.26	0.33	0.64	-0.03	-0.49

**Table IV.** Selection index gain (SIG) and standard deviation (SD) for original traits and indexes based on principal components.

Selection Index	SIG (SD)	SIG/SD
$\Delta G$	184.67 $\pm$ 104.09	1.77
$\Delta G1$	1.33 $\pm$ 0.76	1.75
$\Delta G2$	1.19 $\pm$ 0.68	1.75
$\Delta G3$	1.63 $\pm$ 0.93	1.75
$\Delta G4$	1.95 $\pm$ 1.11	1.76

$\Delta G$  – with the ten original traits;  $\Delta G1$  – with the first two principal components;  $\Delta G2$  – with the first three principal components;  $\Delta G3$  – with the first four principal components; and  $\Delta G4$  – with the first five principal components.

respectively, and their eigenvectors are positive and negative (Tab. III). According to Ibe [8], the principal components with positive and negative eigenvectors are related to the shape of the animals.

The heritabilities for the first five principal components were  $0.23 \pm 0.05$ ,  $0.42 \pm 0.08$ ,  $0.14 \pm 0.03$ ,  $0.19 \pm 0.04$  and  $0.32 \pm 0.06$ , respectively. The genetic gains obtained with the several selection indexes based on the principal components can be observed in the table (Tab. IV).

#### 4. DISCUSSION

The principal component analysis allowed better understanding of the complex

correlations among the traits and reduced the number of traits, using only the five first principal components, without loss of information.

The first five principal components explained together a high percentage of the total variance (Tab. I). Therefore, they are interesting for the purpose of evaluation and comparison of animals. Since the correlation between the principal components was zero, the selection of animals for any principal component will not cause correlated response in terms of other principal components.

The first principal component (PC1), which represents the size component, was named 'General Weight' because animals with larger values for the traits also presented larger values for the first principal component, whereas the second, third, fourth and fifth principal components are shape components, because they contrast animals in function of the traits.

The correlation coefficients (Tab. II) and eigenvectors (Tab. III) indicate the importance of each of the traits for the principal component. Almost all the original variables had high correlation with the first principal component, except for lung, heart, gizzard weights and weight gain from 35 to 42 days of age, indicating that these traits had a similar contribution to the first principal component. This was

confirmed by the eigenvectors, which were very similar for most traits.

On the contrary, the second principal component (PC2) showed very low and sometimes moderate correlation with original traits, except gizzard weight (0.70). Therefore, for this component, the important variable was just gizzard weight, which means that animals with the largest values for this component presented larger gizzards. This component was called 'Gizzard Weight'.

For similar reasons, the third (PC3), fourth (PC4) and fifth (PC5) components were called 'Lung Weight', 'Weight Gain' and 'Heart Weight', respectively, since they were strongly associated with just one of the original traits.

With these five principal components, it is possible to rank animals for general weight (PC1), gizzard weight (PC2), lung weight (PC3), weight gain (PC4) and heart weight (PC5), and thus select the animals based on a group of variables, rather than on isolated traits. According to Ibe [8] and Arnassom and Thorsteinsson [3], another interesting possibility is to obtain a selection index using principal components. In this case, the selection index would have only five weighted coefficients that would facilitate its estimation when compared to the index with 10 traits.

The heritabilities clearly indicate that selecting chickens according to the scores of these principal components may result in genetic gain. To confirm this hypothesis, we prepared a selection index according to Van Vleck [15]. With PC1 and PC2, the selection index standardized gain was 1.75, while it reached 1.77 when including all original traits. These results indicate that it is possible to obtain comparable genetic gain with the selection index composed of principal components and index of original traits. However, in spite of the potential benefit given by principal components for multi-trait evaluation, its use in animal breeding programs has been relatively

scarce, perhaps, partly because it is not always possible to explain high variance percentages with few principal components as it happened in this study. However, in circumstances when a good adjustment occurs, its practical use can bring great benefits to animal breeding programs because it can greatly facilitate the evaluations.

This study did not have the specific objective of obtaining principal components to be used in chicken breeding programs, especially because the data set here used comes from a reciprocal cross between broiler and layer lines. Our objective was to show that it is possible to obtain a linear combination of original traits in order to reduce the complexity of the selection index. Employing an index with only two principal components would be simpler than an index with 10 original traits. However, it is very important to remember that the principal components, in any study of this type, should explain a very high proportion of the total variance, to be sure that this total sample variance is fully exploited.

It can be concluded that the principal component analysis is an interesting tool for evaluating and understanding the total variance originated in a group of correlated traits, allowing for a drastic reduction in the number of traits to be considered in the selection index of poultry breeding programs

## ACKNOWLEDGEMENTS

The authors thank FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), and EMBRAPA Swine and Poultry Research Center, for resources and financial support.

## REFERENCES

- [1] Abreu V.M.N., Silva M.A., Cruz C.D., Figueiredo E.A.P., Abreu P.G., Capacidade de combinação de características produtivas de linhagens de matrizes de frangos de corte,

- usando a técnica de componentes principais, *Rev. Bras. Zootec.* 28 (1999) 250–257.
- [2] Abreu V.M.N., Silva M.A., Cruz C.D., Figueiredo E.A.P., Abreu P.G., Capacidade de combinação de características de produção de ovos de linhagens de matrizes de frangos de corte, usando análise de componentes principais, *Rev. Bras. Zootec.* 28 (1999) 955–959.
- [3] Árnason T.H., Thorsteinsson S.S., Genetics studies on carcass traits in iceland twin ram lambs II. Analysis of principal components and constructions of selection indexes, *Livest. Prod. Sci.* 8 (1982) 507–517.
- [4] Debut M., Berri C., Baéza E., Sellier N., Arnould C., Guémené D., Jehl N., Boutten B., Jégo Y., Beaumont C., Le Bihan-Duval B., Variation of chicken technological meat quality in relation to genotype and pre-slaughter stress conditions, *Poultry Sci.* 82 (2003) 1829–1838.
- [5] Figueiredo E.A.P., Rosa P.S., Scheuermann G.N., Jaenisch F.R.F., Schmidt G.S., Ledur M.C., Brentano L., Costa C.A.F., Genetic gain in body weight, feed conversion and carcass traits in White Plymouth Rock broiler strain Embrapa 021, in: *Proceedings of the 9th World Conference on Animal Production, Porto Alegre, Brazil, 2003* [Compact Disc].
- [6] Figueiredo E.A.P., Schmidt G.S., Ledur M.C., Avila V.S., Brum P.A.R., Fiorentin L., Jaenisch F.R.F., Genetic gain in egg production and egg weight in White Leghorn Embrapa 011, in: *Proceeding of the 9th World Conference on Animal Production, Porto Alegre, Brazil, 2003* [Compact Disc].
- [7] Havenstein G.B., Ferket P.R., Qureshi M.A., Carcass composition and yield of 1957 versus 2001 broilers when fed representative 1957 and 2001 broilers diet, *Poultry Sci.* 82 (2003) 1509–1518.
- [8] Ibe S.N., Measures of size and conformation in commercial broilers, *J. Anim. Breed. Genet.* 106 (1989) 461–469.
- [9] Johnson R.A., Wichern D.W., *Applied Multivariate Statistical Analysis*, 5th ed., Prentice Hall, Texas, 1998.
- [10] Meyer K., DFUNI, DXMUX, DFREML, Version 3.0  $\beta$ . User notes, Armidale, University of New England, 1998.
- [11] Morrison D.F., *Multivariate Statistical Methods*, McGraw-Hill Company, New York, 1976.
- [12] Reddish J.M., Lilburn M.S., A comparison of growth and development patterns in diverse genotypes of broilers. I. Male Broiler growth, *Poultry Sci.* 83 (2004) 1067–1071.
- [13] SAS Institute, *The SAS system for windows*, SAS Institute Inc., North Caroline, 2001 [Compact Disc].
- [14] Saxton A.M., Gibson G., Kang M.S., Tempelman R.J., Balzarini M.G., *Genetic Analysis of Complex Traits Using SAS*, SAS® Institute Inc., Cary, NC, 2004.
- [15] Van Vleck L.D., *Selection Index and Introduction to mixed model methods*, CRC Press, Boca Raton, 1993.

