

Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

Análise genética de características de produção e qualidade de ovos de linhagens de aves caipiras (*Gallus gallus*) utilizando modelos de regressão e multivariados

Luiz Firmino dos Santos Júnior

Dissertação apresentada para obtenção do título de Mestre em Ciências. Área de concentração: Ciência Animal e Pastagens

Piracicaba
2022

Luiz Firmino dos Santos Júnior
Zootecnista

**Análise genética de características de produção e qualidade de ovos de linhagens de aves
caipiras (*Gallus gallus*) utilizando modelos de regressão e multivariados**
versão revisada de acordo com a Resolução CoPGr 6018 de 2011

Orientador:
Prof. Dr. **GERSON BARRETO MOURÃO**

Dissertação apresentada para obtenção do título de
Mestre em Ciências. Área de concentração: Ciência
Animal de Pastagens

Piracicaba
2022

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP**

Santos Júnior, Firmino dos

Análise genética de características de produção e qualidade de ovos de linhagens de aves caipiras (*Gallus Gallus*) utilizando modelos de regressão e multivariados / Luiz Firmino dos Santos Júnior. - - versão revisada de acordo com a Resolução CoPGr 6018 de 2011. - - Piracicaba, 2022.

48 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Análise fatorial exploratória 2. Avicultura alternativa 3. Medidas longitudinais 4. Polinômios de Legendre 5. Regressão aleatória I. Título

DEDICATÓRIA

Aos meus pais, Luiz e Izabel, por serem as pessoas que mais acreditaram em mim e por sempre me apoiarem e serem meu alicerce!

Dedico

AGRADECIMENTOS

À Escola Superior de Agricultura “Luiz de Queiroz” – Universidade de São Paulo e também ao Departamento de Zootecnia por terem me dado a oportunidade de poder realizar este trabalho.

Ao Professor Doutor Gerson Barreto Mourão pela orientação, oportunidade de poder contribuir com seu grupo de estudos e pela confiança para realização deste estudo.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES pela concessão da bolsa de estudos.

Aos meus companheiros de grupo Paola Boscollo, Sofia, Izally, Giovanni, Fabrício, Leonardo Mazzero, Brayan, Mayara, Felipe, Laura Abreu, Letícia, Robson, Beatriz, Laura, Karoline, Elizangela e Camila pelo companheirismo, conversas, ajuda nos momentos oportunos, pelo trabalho em equipe, pelas trocas de conhecimento, conversas e momentos de descontração.

Aos meus amigos de fé Maysa, Jaqueline, Jennifer, Letícia, Ludmilla, Pollyanna, Fábio, Carol e todos mais que compartilharam momentos comigo em Piracicaba.

Aos companheiros de consulta e trocas de experiências científicas em geral Carol, Juliana, Eula e Beatriz.

Aos meus pais por sempre acreditarem em mim, serem meu suporte emocional e por toda cumplicidade e amor que temos, sobrinhos, irmãos e familiares.

E acima de tudo, quero agradecer a Deus, que me deu vida, direcionamento, inspiração e me permitiu estar vivendo esta oportunidade.

MUITO OBRIGADO!

BIOGRAFIA

Eu nasci no dia 24 de março de 1992, às 20 horas. Sou filho de Luiz Firmino dos Santos e Izabel Braga Antônio dos Santos, duas pessoas maravilhosas e bem simples, que sempre me ensinaram muito e me incentivaram a busca por conhecimento.

Desde criança já gostava muito de animais e dizia que queria ser cientista, algo que começou a se concretizar no ano de 2011, quando ingressei na faculdade para cursar Zootecnia, na Universidade Estadual Paulista “Júlio de Mesquita Filho” em Ilha Solteira.

Ainda na graduação, no ano de 2014, tive o privilégio de fazer um intercâmbio na Universidade de Salamanca na Espanha, e posteriormente em 2016, pude novamente experienciar o estudo no exterior, agora na Universidade de Santiago do Chile, no Chile.

Concluí minha graduação em 2018 e já em 2019 decidi que queria fazer mestrado na Escola Superior de Agricultura “Luiz de Queiroz”, onde tenho o privilégio de agora poder escrever este trabalho científico.

EPÍGRAFE

*“Se você não sabe para onde ir,
Qualquer caminho serve”.*

Lewis Carroll

*“Dificuldades preparam pessoas comuns
Para destinos extraordinários.”*

C.S. Lewis

*A objetividade da pesquisa científica só começa
depois da escolha das variáveis e das
metodologias de análise, antes disto à atividade
científica é completamente subjetiva.*

J. M. M. Neto

SUMÁRIO

RESUMO	8
ABSTRACT	9
1. INTRODUÇÃO.....	11
Referências	11
2. REVISÃO DE LITERATURA	13
2.1. Avicultura alternativa	13
2.2. Melhoramento e parâmetros genéticos de aves	14
2.3. Modelos de regressão aleatória.....	16
2.4. Análise fatorial exploratória	18
Referências	20
3 ANÁLISE GENÉTICA DE CARACTERÍSTICAS DE PRODUÇÃO E QUALIDADE DE OVOS DE LINHAGENS DE AVES CAPIRAS (<i>Gallus gallus</i>) UTILIZANDO MODELOS DE REGRESSÃO E MULTIVARIADOS.....	25
Resumo	25
Abstract.....	25
3.1. Introdução.....	26
3.2. Material e Métodos.....	28
3.2.1. Linhagens e dados de produção de ovos.....	30
3.2.2. Avaliação dos ovos.....	31
3.2.3. Parâmetros genéticos pelo modelo de regressão aleatória.....	31
3.2.4. Análise fatorial exploratória.....	33
3.3. Resultados.....	32
3.3.1. Estatística descritiva.....	33
3.3.2. Parâmetros genéticos.....	35
3.3.3. Análise factorial exploratória.....	38
3.4. Discussão	41
3.5. Conclusão	45
Referências	46

RESUMO

Análise genética de características de produção e qualidade de ovos de linhagens de aves caipiras (*Gallus gallus*) utilizando modelos de regressão e multivariados

A avicultura de postura alternativa é um dos mercados com maior potencial de crescimento e a agricultura familiar, onde está inserida a produção alternativa de ovos, é responsável por mais de 70% dos alimentos que chegam à mesa dos brasileiros. Na avicultura alternativa de postura, são utilizadas linhagens de galinhas caipiras criadas com acesso a pasto ou piquetes. Uma das principais vantagens deste sistema, é que ao final do ciclo de postura dessas aves, elas podem ser abatidas com peso corporal superior às aves de postura comerciais. Porém, ainda não existem muitos estudos no sentido de selecionar e estudar linhagens brasileiras caipiras para postura dentro dos programas de melhoramento genético de aves. O objetivo deste trabalho foi analisar os dados de produção e qualidade de ovos de 732 aves de linhagens caipiras, por meio de modelos de regressão e multivariados, para estimar componentes de variância e realizar uma avaliação genética das linhagens Caipirinha, Caipirão e Azul da ESALQ, 7P, Carijó Barbada, Baianinha, Perna Longa e Sooho. Medidas repetidas ao longo do tempo ou dados longitudinais são como regra dentro dos estudos com melhoramento genético, e dentre algumas das classes de métodos mais adequados para estes estudos, temos a classe não paramétrica, que usa funções de covariância baseadas em funções flexíveis como os polinômios por exemplo, o que vai gerar o método de regressão aleatória (RRM). Os RRM são utilizados quando se deseja modelar a forma ou trajetória de observações que foram realizadas ao longo do tempo (dados longitudinais). Como o avanço científico, cada vez mais características de interesse dentro da avicultura de postura têm sido consideradas relevantes, o que faz com que a estatística multivariada tenha certas limitações ao analisar todas essas características, criando modelos muito complexos, de difícil interpretação e também obtendo um ganho genético menor, já que esse ganho é reduzido à medida que se aumenta a quantidade de características em um índice de seleção. Uma alternativa para solucionar esse impasse é a utilização de metodologias multivariadas, como a análise fatorial exploratória (AFE), que permitam que se identifique um pequeno grupo de variáveis que possam analisar um grande número de características. O modelo de regressão aleatória utilizando polinômios de Legendre de 2ª ordem se ajustou bem aos dados de produção de ovos de galinhas caipiras, descrevendo bem a curva de produção ao longo das idades de 120 a 456 dias das aves. A herdabilidade apresentou comportamento decrescente inicialmente ao longo da trajetória das idades, enquanto que as variâncias genética aditiva e de ambiente permanente apresentaram o mesmo comportamento no início, mas ao longo do tempo voltaram a crescer. A linha Caipirinha da ESALQ foi a que apresentou os melhores valores genéticos estimados (EBV). A análise fatorial exploratória utilizando dados de valores genéticos estimados para as características de produção e qualidade de ovos com dois fatores apresentou uma melhor explicação das porcentagens das variâncias que ocorrem 7 nas variáveis originais deste estudo quando comparada com a AFE com dados fenotípicos, mostrando que o fator genético dá mais importância para uma maior quantidade de características. A variável altura de albúmen não apresentou importância neste estudo na composição de nenhum dos fatores. No final da análise, das 7 variáveis iniciais, apenas 6 foram importantes para compor os dois fatores da AFE, sendo elas produção de ovos, idade ao primeiro ovo, espessura de casca, peso da gema, peso da casca e peso do ovo, sendo que altura de albúmen não apresentou relevância para compor nenhum dos fatores.

Palavras-chave: Análise fatorial exploratória, Avicultura alternativa, Medidas longitudinais, Polinômios de Legendre, Regressão aleatória

ABSTRACT

Genetic analysis of production traits and egg quality of free-range bird (*Gallus gallus*) strains using regression and multivariate models

Alternative egg production is one of the markets with the greatest growth potential and family farming, where alternative egg production is inserted, is responsible for more than 70% of the food that reaches the Brazilian table. In alternative laying poultry, strains of free-range chickens raised with access to pasture or paddocks are used. One of the main advantages of this system is that at the end of the laying cycle of these birds, they can be slaughtered with higher body weight than commercial laying birds. However, there are still not many studies in the sense of selecting and studying Brazilian redneck lines for laying within the poultry genetic improvement programs. The objective of this work was to analyze data on production and egg quality of 732 free-range strains, using regression and multivariate models, to estimate variance components and carry out a genetic evaluation of the Caipirinha, Caipirão and Azul strains from ESALQ, 7P, Carijó Barbada, Baianinha, Long Leg and Sooho. Repeated measurements over time or longitudinal data are as a rule within studies with genetic improvement, and among some of the most suitable classes of methods for these studies, we have the nonparametric class, which uses covariance functions based on flexible functions such as polynomials for example, what will generate the random regression method (RRM). RRM's are used when you want to model the shape or trajectory of observations that have been carried out over time (longitudinal data). As scientific advances, more and more characteristics of interest within the poultry industry have been considered relevant, which makes multivariate statistics have certain limitations when analyzing all these characteristics, creating very complex models, difficult to interpret and also obtaining a lower genetic gain, as this gain is reduced as the number of traits in a selection index increases. An alternative to solve this impasse is the use of multivariate methodologies, such as exploratory factor analysis (EFA), which allow the identification of a small group of variables that can analyze a large number of characteristics. The random regression model using 2nd order Legendre polynomials fitted well to the egg production data of free-range hens, describing well the production curve over the ages of 120 to 456 days of the birds. Heritability showed a decreasing behavior initially along the trajectory of the ages, while the additive genetic and permanent environment variances showed the same behavior at the beginning, but over time they increased again. ESALQ's Caipirinha line presented the best estimated genetic values (EBV). Exploratory factor analysis using estimated breeding value data for egg production and egg quality traits with two factors presented a better explanation of the percentages of variances that occur in the original variables of this study when compared with the AFE with phenotypic data, showing that the Genetic factor gives more importance to a greater amount of traits. The albumen height variable was not important in this study in the composition of any of the factors. At the end of the analysis, of the 7 initial variables, only 6 were important to compose the two EFA factors, namely egg production, age at first egg, shell thickness, yolk weight, shell weight and egg weight, being which albumen height was not relevant to compose any of the factors.

Keywords: Exploratory factor analysis, Alternative poultry farming, Longitudinal measurements, Legendre polynomials, Random Regression

1. INTRODUÇÃO

O ovo é um dos alimentos mais consumidos e nutricionalmente completos do mundo, com produção mundial em torno de 116 milhões de toneladas em 2020, totalizando pouco mais de 2,2 trilhões de ovos vindos de quase 11 bilhões de poedeiras. Dentre os maiores produtores, o Brasil é o sexto no geral e segundo maior do continente Americano, com uma produção de 3,26 milhões de toneladas (57,2 bilhões de ovos) e um plantel de aves de 252 milhões (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS - FAOSTAT, 2020).

Dessas milhões de toneladas produzidas, grande parte provém da avicultura industrial, que recebeu esse nome a partir de meados da década de 50, onde ocorreram grandes evoluções científicas nas áreas de melhoramento genético animal e nutrição animal. Nesse período chegaram as primeiras gaiolas no Brasil, algo que caracterizaria a forma com que os ovos seriam produzidos em grande escala a partir dali (MAIA, 1997).

Nesta mesma década, nos Estados Unidos da América, a avicultura alternativa de postura era um modelo de produção, onde foi realizado um trabalho bastante intenso de melhoramento genético com linhagens alternativas, conhecidas como galinhas caipiras, coloniais ou de capoeira, aves produtoras de ovos da espécie *Gallus gallus domesticus*, que ao final de seu ciclo produtivo são destinadas a produção de carne. Porém, quando este modelo chegou ao Brasil, estávamos no período pós guerra, onde a transformação da cadeia produtiva de ovos estava sofrendo alterações para chegar aos moldes de como é produzida até hoje. Isso fez com que aqui, a avicultura alternativa, como o próprio nome menciona, servisse apenas como fonte extra de renda para pequenos produtores (MAIA, 1997; MUIR et al., 2008; PUTMAN et al., 2017).

Ainda assim, a avicultura alternativa têm crescido, pois as pessoas têm buscado por produtos que tenham sido produzidos sob práticas que promovam melhor conservação do meio ambiente, bem estar dos animais envolvidos e que sejam mais naturais e saudáveis. Alguns desses produtos, como os orgânicos, também poderiam ser produzidos sem o uso de antibióticos, promotores de crescimento e outros produtos que possam deixar resíduo no ovo ou carne. A demanda por esses produtos em geral, tem refletido nas legislações ou regras que alguns mercados importadores têm exigido de seus fornecedores (dentre os quais se enquadra o Brasil), como na Europa, onde o número de varejistas que vendem apenas ovos de aves criadas em sistema com acesso a pasto tem aumentado (DEMATTÊ FILHO; MARQUES, 2011).

Apesar dessa demanda, áreas como exportação e melhoramento genético de produtos nacionais ainda precisam ser bastante desenvolvidas e estudadas, considerando que a maior parte do material genético que é utilizado na avicultura de postura brasileira é importada de empresas localizadas principalmente na Europa e nos Estados Unidos. São poucas as linhagens desenvolvidas no Brasil, e o material genético puro das linhagens de aves importadas são exclusivos dessas empresas internacionais, as únicas capazes de gerar o animal híbrido industrial que comercializam. Considerando essa demanda por mais estudos, especialmente na área de genética, o objetivo deste trabalho foi analisar os dados de produção e qualidade de ovos de linhagens caipiras, por meio de modelos estatísticos multivariados e univariados, para estimar componentes de variância e realizar uma avaliação genética dessas linhagens.

Referências

- DEMATTE FILHO, L. C.; MARQUES, P. E. M. Dinâmica tecnológica da cadeia industrial da avicultura alternativa: multifuncionalidade, desenvolvimento territorial e sustentabilidade. *Segurança Alimentar e Nutricional*, v. 18, n. 2, p. 1-11, 2011.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAOSTAT). *Livestock Primary*. 2020. Disponível em: <<http://www.fao.org/faostat/en/#data/QL>>. Acesso em: 26 mar. 2022.
- MAIA, G. A. R. Avicultura Alternativa: Carne e Ovos pelo Sistema de Pastejo. *Artigos Técnicos da Sociedade Nacional de Agricultura*, Brasília, ano, v. 100, p. 1-6, 1997.
- MUIR, W. M. et al. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proceedings of the National Academy of Sciences*, v. 105, n. 45, p. 17312-17317, 2008.
- PUTMAN, B. et al. A retrospective analysis of the United States poultry industry: 1965 compared with 2010. *Agricultural Systems*, v. 157, p. 107-117, 2017.

2. REVISÃO DE LITERATURA

2.1 Avicultura alternativa

A avicultura de postura alternativa é um dos mercados com maior potencial de crescimento, principalmente na geração de linhagens brasileiras de postura e o estudo das mesmas, que ainda são bem poucas, sendo a grande maioria do plantel de poedeiras utilizadas no Brasil, importadas de empresas internacionais. A agricultura familiar, onde está inserida a produção alternativa de ovos, é a principal responsável pelos alimentos que chegam à mesa das pessoas, e mesmo no contexto atual da pandemia do covid-19, a produção familiar não parou (ALTIERI; NICHOLLS, 2020; BERKHOFF et al., 2020; SILVA et al., 2021). O ovo é um dos alimentos que teve seu consumo e procura aumentados nesse período, principalmente por ser um alimento completo e saudável, perdendo apenas para o leite materno, e por ser uma fonte de proteína mais acessível que a carne, considerando a queda do poder aquisitivo, assim, as pessoas por se preocuparem com sua saúde e com o impacto do covid-19, têm preferido este alimento no lugar de outras fontes de proteína (HIRVONEN; BRAUW; ABATE, 2021).

A seleção e melhoramento genético de linhagens brasileiras para produção de ovos em sistema alternativo aumentaria a oferta de produtos para este método de produção para produtores interessados, geraria um produto com maior valor agregado e tornaria o Brasil mais competitivo neste setor. Com a crescente demanda por produtos alternativos e orgânicos, reflexo do perfil de um consumidor mais informado sobre os processos industriais da cadeia de produção de ovos, a qualidade de um produto passou a ser atrelada também a como ele foi gerado, e não apenas em sua composição nutricional e inocuidade (BLOKHUIS et al., 2003).

Na avicultura alternativa de postura, são utilizadas linhagens de galinhas caipiras criadas com acesso a pasto ou piquetes desde que foram introduzidas no Brasil na época da colonização. Nesses espaços elas podem desenvolver seus hábitos e comportamentos naturais da espécie, não ficando toda sua vida produtiva dentro de gaiolas, o que promove melhoria na qualidade de vida desses animais. As linhagens utilizadas na avicultura comercial de postura foram selecionadas para terem alto desempenho produtivo em gaiola, onde geralmente toda a alimentação e ambiente são controlados, portanto, para o sistema alternativo, onde as aves possuem acesso a piquetes nas 24 horas do dia ou em parte do dia, é necessário que se utilize uma linhagem com genótipo mais resistente para este sistema, sendo capaz de se adaptar a essas condições, característica presente em aves caipiras (BOLLING et al., 2003; HELLMEISTER FILHO et al., 2003).

Dentre as linhagens brasileiras, algumas foram selecionadas para este tipo de sistema, demonstrando rusticidade ideal para criação ao ar livre e boa produção, como a Caipirinha da ESALQ, Caipirão da ESALQ, Carijó Barbada, etc. Porém, ainda não existem muitos estudos no sentido de selecionar e estudar linhagens brasileiras caipiras para postura e que apresentem boa produtividade e, unindo isso a crescente procura por produções alternativas, uma alternativa para atender a essas duas demandas seria a seleção e melhoramento genético de mais linhagens de poedeiras caipiras e mais estudos sobre as mesmas, para serem inseridas mais possibilidades no mercado de ovos.

A Escola Superior de Agricultura “Luiz de Queiroz” - ESALQ vem trabalhando desde a década de 80 na seleção de aves caipiras para corte e postura. Dentre as linhagens que mais chamaram atenção nos processos de seleção e melhoramento genético, a linhagem Azul da ESALQ se destacou por sua postura e fenótipo azul para casca dos ovos, enquanto que a linhagem Caipirinha da ESALQ se destacou por apresentar boa postura (240 ovos por ciclo) e também peso adulto de 2,42 kg, podendo ser utilizada para dupla aptidão.

Ambas as linhagens possuem crescimento lento e cor de plumagem variada, e foram desenvolvidas por meio de seleção massal, inicialmente dentro do projeto “Frango Feliz” desenvolvido pela ESALQ. Este projeto teve início na década de 1997, com o intuito de melhorar geneticamente linhagens caipiras nacionais e, com o fim dele, as aves foram doadas para o Setor de Melhoramento de Genético de Aves da ESALQ, onde continuaram sendo selecionadas de acordo com suas aptidões, estando atualmente na sexta geração de seleção (SAVINO et al., 2007).

Um das principais vantagens deste sistema, é que ao final do ciclo de postura dessas aves, elas podem ser abatidas com peso corporal superior às aves de postura comerciais, o que pode ser um atrativo e uma segunda possibilidade de renda com esses animais para os produtores. Assim, seria possível ter aves adaptadas a um sistema de produção orgânico, selecionadas para uma boa postura dentro desse tipo de sistema e que ainda poderiam ser aproveitadas para corte no final de suas vidas produtivas (FARUKE et al., 2013; DIKMEN et al., 2016).

2.2 Melhoramento e parâmetros genéticos de aves

No ano de 1918 aconteceu algo muito importante para os estudos da área de genética, Fisher demonstrou que os princípios mendelianos, que antes geravam muitas discussões dividindo pesquisadores entre seguidores de Mendel e aqueles que acreditavam na lei dos ancestrais, se aplicavam tanto para as características qualitativas quanto para as quantitativas. Alguns anos depois, J. Lush publicou seu livro *Programas de Melhoramento Animal*, baseado nos trabalhos de Sewall Wright, fazendo com que se difundisse a partir daquele momento o Melhoramento Animal como uma ciência dentro das escolas de agricultura (PROVINE, 1977; HILL, 2014). A partir dessa época, começaram a surgir empresas de melhoramento de milho, e na sequência também as de melhoramento de aves, que baseando seus programas nos realizados pelas primeiras citadas, cometeram vários erros e acertos, levando-se em consideração o que se conhecia até aquele momento. Nas décadas seguintes houve um avanço grande dessas empresas, até culminar na realidade que temos mais atualmente, onde a genética avícola está concentrada e conseqüentemente sendo controlada por um pequeno grupo de grandes empresas não brasileiras, o que causa uma dependência de quase 100% do Brasil em material genético produzido no exterior (HILL, 2014).

O início de uma evolução mais pronunciada na avicultura brasileira começou depois da década de 60, período este onde começaram a serem introduzidos no Brasil materiais genéticos importados, de aves com especialidade em produção de ovos e de carne, resultado de avanços científicos deste setor, especialmente na área de melhoramento genético. Em meados desse período, logo que se regulamentou a entrada do material genético apenas de avós, a estrutura da avicultura brasileira adaptou-se aos moldes da americana, onde se tinha granjas com avós, matrizes e por fim os produtores comerciais (COSTA; PETIT; BITTENCOURT, 2008).

Os melhoristas viram no baixo custo de produção, somado às mudanças nos hábitos de consumo e crescimento da população, um potencial a ser explorado nas aves de postura, e, com a evolução econômica que começou a ocorrer, aumento da competitividade de mercado e uma dependência do material genético que precisava ser importado, as pesquisas em melhoramento genético com o objetivo de se selecionar e desenvolver linhagens com alta produção e comercialmente competitivas, começaram a ocorrer no país (HORST & MATHUR, 1992).

O cerne de qualquer programa de melhoramento genético para aves de postura, é selecionar e reproduzir indivíduos que produzam bastante ovos, que estes ovos sejam de qualidade e que a conversão alimentar das aves seja eficiente, e considerando o contexto da avicultura alternativa, que estes animais também tenham uma boa adaptabilidade a ambientes variados e resistência a doenças (KETELAERE et al., 2002).

Inicialmente as poedeiras são selecionadas para reprodução, para melhorar características que venham aumentar o lucro na produção de ovos, onde essas aves são estudadas e gerações discretas são geradas em intervalos de tempo geralmente curtos, o que resultará em ganhos genéticos por período de tempo. Nas linhagens de ovos brancos, é bastante comum se realizar a seleção dentro de linha e a seleção recorrente recíproca, a partir do uso de informações das linhas puras e das cruzadas, onde o valor genético será determinado pelo uso de uma função bioeconômica contendo as características produção de ovos, peso do ovo, conversão alimentar e peso corporal (CAVERO et al., 2011).

As características reprodutivas e de desempenho possuem uma relevância importante para os programas em melhoramento genético, especialmente quando se pensa em selecionar várias características ao mesmo tempo. Também é importante conhecer e estimar os parâmetros genéticos dessas características economicamente importantes, pois eles vão direcionar em como a seleção será aplicada, e são estas estimativas que servirão como ferramental indispensável para o alcance de altos resultados em ganhos genéticos na população e consequentemente contribuirão para uma melhor eficiência na seleção das linhagens de aves de postura (LEDUR et al., 1993).

A condução de programas de melhoramento genético e o delineamento dos mesmos dependem da estimativa dos parâmetros genéticos para que o melhorista possa entender e quantificar o quanto da expressão do fenótipo da ave para uma determinada característica (selecionada e estudada) é influenciada pela genética desse animal, e também, essas estimativas permitem mensurar qual é o grau do relacionamento existente entre uma característica escolhida com alguma outra, o relacionamento entre duas ou mais medidas de uma mesma característica coletadas em períodos diferentes ao longo da vida do animal (medidas longitudinais), e assim, a soma de todas essas informações irá auxiliar na escolha de qual método de seleção será aplicado, permitindo uma estimativa da resposta da população estudada em relação à seleção (LEDUR et al., 1993).

A característica mais importante na avicultura de postura é a produção de ovos, e, apesar de a qualidade do ovo estar recebendo bastante interesse por parte do consumidor, que está mais preocupado com essa questão hoje do que no passado, a produção ainda é a variável mais estudada em programas de seleção, tendo seu progresso sido um sucesso, o que fez com que a quantidade de ovos produzidos por uma ave mais do que dobrasse da década de 50 até os dias atuais, chegando fácil a valores acima de 250 ou 300 ovos em todo seu ciclo de postura (WOLC et al., 2007).

Parte do sucesso na seleção para produção de ovos se deve ao fato de que a característica, por ter sido bastante estudada, em todo seu processo desde a formação do ovo até a sua postura, foi mais bem compreendida com os avanços científicos, e graças a esse bom resultado da seleção, o limite biológico de um ovo por dia já está próximo de acontecer (WOLC et al., 2007; KETELAERE et al., 2002).

Os modelos empregados na avaliação genética assim como as definições de características têm sofrido variações no decorrer do tempo. Inicialmente a seleção era baseada em fenótipos, até que os índices de seleção foram introduzidos. Métodos baseados na análise de variância (ANOVA) foram os primeiros a serem utilizados para se estimar os componentes de variância, até que foram substituídos pela metodologia de máxima verossimilhança restrita (REML) que é baseada no modelo animal (WOLC et al., 2007).

Uma das estimativas mais importantes é a herdabilidade de uma característica, que no sentido restrito (h^2), pode ser expressa como sendo a razão entre a variância genética aditiva e a variância fenotípica. Valores de herdabilidade baixos indicam que a maior parte da variação observada no fenótipo de uma população para uma determinada característica é influenciada por diferenças no manejo, condições ambientais no geral e aos efeitos não aditivos dos genes, enquanto que valor de herdabilidade altos, indicam que é decorrente do efeito aditivo dos genes a

maior parte da variação fenotípica observada nos indivíduos de uma população para certa característica (FALCONER et al., 1996).

Existe algo a se considerar quando se fala em herdabilidade da característica produção de ovos, que é o quanto esta estimativa pode mudar ao longo do período de postura, portanto, pegar um único valor de produção acumulada de ovos durante todo o ciclo de postura e assumir que isso descreve bem esse período não é uma verdade. Uma alternativa para contornar essa questão e poder se alcançar uma melhor eficiência da seleção genética para a característica de produção de ovos é o uso de modelos longitudinais, que são capazes de considerar de forma mais correta as variâncias e covariâncias genéticas ao longo do tempo de postura da ave (WOLC et al., 2011; BEGLI et al., 2021).

2.3 Modelos de regressão aleatória

Quando uma característica é avaliada repetidas vezes ao longo da vida de um organismo, seja ele planta ou animal, essa característica é denominada como infinitamente dimensional, no sentido em que cada unidade de tempo ou idade, essa característica poderá ser avaliada, gerando um conjunto de dados multidimensional. Esse tipo de análise é interessante quando se quer prever valores dos indivíduos em um determinado ponto no tempo ou então em todos os pontos, assim como também na identificação de uma estrutura de variância parcimoniosa ao longo do tempo (RESENDE; SILVA; AZEVEDO, 2014).

Medidas repetidas ao longo do tempo ou dados longitudinais são como regra em estudos com melhoramento genético de animais e plantas perenes, onde o BLUP (melhor predição linear não viesada) com os modelos que utilizam medidas repetidas pode ser ajustado a partir do uso de diferentes parametrizações da estrutura de correlação dos fenótipos ao longo do tempo (MEYER, 2005; MRODE, 2005).

Dentre algumas das classes de métodos mais adequados ao estudo de dados longitudinais ou medidas repetidas, temos a classe não paramétrica, que usa funções de covariância baseadas em funções flexíveis como os polinômios, por exemplo, o que vai gerar o método de regressão aleatória, e nesta classe, pode-se basear do uso de polinômios de Legendre (BAILEY; SRINATH, 1996; RESENDE; SILVA; AZEVEDO, 2014).

Modelos de repetibilidade e de regressão aleatória conduzem a soluções bastante distintas, sendo que os modelos de regressão aleatória (RRM) visam colocar mais peso a informações individuais do que informações de parentesco, e em modelos curvilíneos as observações extremas também costumam receber mais peso do que as médias, e quanto maior o grau do polinômio, mais essas características são evidenciadas (LUSH, 1994).

Os RRM geralmente selecionam animais mais jovens quando comparados a modelos de repetibilidade, e nesses modelos a estrutura regressiva será aplicada a todos os fatores aleatórios, que são os efeitos aditivos diretos e maternos e os efeitos permanentes. Quando as medidas repetidas possuem uma correlação alta, o risco de se obter matrizes de covariância não positiva definida e de não convergência no processo da estimação dos componentes de variância é aumentado, assim como quanto maior for a ordem da matriz de covariância, também será maior a chance de que essa matriz não seja positiva definida. O RRM é uma opção interessante para solucionar estas questões, principalmente quando com o aumento da distância entre as idades, vá ocorrendo uma diminuição gradativa das correlações (VEERKAMP; THOMPSON, 1999; RESENDE; SILVA; AZEVEDO, 2014).

Quem introduziu os RRM foram Henderson & Charles (1982) e Laird & Ware (1982). Tanto nos modelos de regressão simples quanto múltiplos, uma característica se mantém, as covariáveis ou variáveis independentes são na verdade constantes e não variáveis, pois em uma amostragem com medidas repetidas hipotética, os valores dessas

covariáveis vão permanecer constantes. Também nestes modelos, os mesmos coeficientes de regressão estarão associados a todos os indivíduos da população estudada, ou seja, eles serão fixos. Em alguns modelos a melhor suposição relacionada aos coeficientes de regressão associados às covariáveis é a de que cada membro possui distribuição aleatória, sendo chamados de modelos de regressão aleatória (RESENDE; SILVA; AZEVEDO, 2014; SEEKER et al., 2018).

Na regressão aleatória, são gerados diferentes coeficientes de regressão que são associados a cada indivíduo diferente dentro de uma população, assim, não somente os indivíduos como também os coeficientes de regressão associados a eles serão considerados obrigatoriamente como efeitos aleatórios (HENDERSON, 1984). Os registros que são feitos desse animal assumem diferentes correlações genéticas entre eles, pois é levada em consideração a premissa de que a característica estudada pode estar mudando de maneira contínua e gradual ao longo do tempo. Nos RRM são geradas curvas para as características avaliadas, e a forma dessas curvas será considerada, de modo a permitir que sejam ajustadas curvas aleatórias para cada um dos indivíduos da população, sendo essas curvas expressas como sendo os desvios de uma curva média da dessa população ou grupos de indivíduos (KIRKPATRICK et al., 1994; SCHAEFFER & JAMROZIK, 2008).

Os RRM são utilizados quando se deseja estimar matrizes que são parte das funções de covariância para se estimar uma gama maior de variâncias e covariâncias. Também, quando se deseja modelar a forma ou trajetória de observações que foram realizadas ao longo do tempo (dados longitudinais) ou em situações onde são feitos registros repetidos das observações de uma característica, sendo coletados esses dados várias vezes durante o curso de vida desse animal. Fenotipicamente essa trajetória precisa ser ajustada, ao mesmo tempo em que a variação ao longo da trajetória precisa ser considerada (HILL & BROTHERTSTONE, 1999; SCHAEFFER, 2016).

Grande parte das trajetórias costumam ser suaves e contínuas, podendo ser ajustadas com poucas covariáveis, porém, às vezes ela pode ser desconhecida e indefinida, com altos e baixos ao longo do tempo. Trajetórias para machos e fêmeas, por exemplo, podem ser diferentes, ou para diferentes raças e linhagens, e também, com o avançar dos anos elas podem sofrer mudanças decorrentes da seleção dos animais para alguma característica de interesse, como produção de ovos, por exemplo, assim, todos esses fatores precisam ser considerados na análise (SCHAEFFER, 2016).

A primeira coisa a se fazer é traçar os dados com relação a escala de tempo que seja de interesse do pesquisador, sendo que as observações podem ser separadas por gênero, raça, linhagem, idade ou até mesmo por anos, olhando-se sempre para todos os aspectos importantes dos dados e, antes de tudo escolhendo um modelo de análise que melhor se ajuste ao comportamento dos dados. No momento de escolher um modelo de regressão apropriado, deve-se escolher aquele com o menor número de parâmetros e que se ajuste bem aos dados, pois um bom ajuste para a trajetória é muito importante em um modelo de regressão aleatória (SCHAEFFER, 2016).

O primeiro estudo que aplicou RRM ao melhoramento genético foi feito no trabalho de Schaeffer (1994), visando avaliar geneticamente vacas leiteiras. Para várias espécies de animais de produção os RRM têm sido utilizados com sucesso na avaliação genética de características longitudinais e esses modelos também tem se mostrado bastante úteis quando aplicados no melhoramento genético de galinhas poedeiras para as características produção de ovos e persistência de postura (SCHAEFFER, 2004; WOLC et al., 2011; BEGLI et al., 2018).

O modelo de regressão aleatória possui a vantagem de ter a capacidade de prever os valores genéticos para a produção cumulativa de ovos ao longo de um período de postura e em qualquer dia específico dessa produção, avaliando a persistência da mesma (BIGNARDI et al., 2011).

Nos estudos em melhoramento genético utilizando RRM, os indivíduos são tratados como efeitos aleatórios por meio de seus valores genéticos, e as medições repetidas realizadas para cada um deles são feitas no decorrer do tempo ou idade do animal, assim as variáveis resposta poderão ser explicadas a partir de uma função linear da idade. Os RRM conseguem explicar melhor a variabilidade que está associada a essas medidas repetidas do que os modelos lineares simples, que incluem o efeito de ambiente permanente e o parâmetro repetibilidade (como em modelos de repetibilidade) ou então consideram que as medidas repetidas são caracteres distintos a partir de um modelo multivariado (JAKOBSEN et al., 2002; RESENDE; SILVA; AZEVEDO, 2014).

A partir das medidas nas diferentes idades dos animais, curvas de valores genéticos serão obtidas e serão associadas a cada um dos indivíduos da população, em que, nos RRM, as mudanças nas variâncias genéticas e residuais através do tempo serão consideradas nessas curvas, assim, se as estimativas obtidas para os componentes de variância forem confiáveis, estes modelos permitirão que se faça uma predição dos valores genéticos de cada indivíduo nas diferentes idades com base em avaliações obtidas de apenas uma idade do mesmo (LEWIS; BROTHERSTONE, 2002; RESENDE; SILVA; AZEVEDO, 2014).

O RRM pode ser considerado como um modelo simplificado e reduzido, pois ele permite que se obtenham os mesmos parâmetros que podem ser obtidos pelo modelo multivariado (herdabilidade em cada idade e correlação genética entre todos os pares de idade existentes), porém ele faz isso utilizando uma parametrização menor e também com menos esforço computacional. Graças a essa abordagem, ele define diretamente as funções de covariância contínuas, permitindo que se possam analisar indivíduos com idades heterogêneas (WILSON; KRUIK; COLTMAN, 2005; RESENDE; SILVA; AZEVEDO, 2014).

Para se ajustar um modelo de regressão aleatória, é preciso um intercepto e inclinação, com um termo quadrático adicional ou cúbico ou de maior ordem de acordo com o polinômio de Legendre ajustado. A trajetória dos valores genéticos ao longo do tempo será modelada pelos RRM como desvios de outros efeitos fixos e aleatórios que forem incluídos no modelo, onde polinômios simples serão utilizados para modelar esses desvios (polinômios de terceira ou quarta ordem), porém, esses polinômios não necessariamente serão eficientes, de forma que funções mais flexíveis como polinômios de ordens mais elevadas poderão conduzir melhores estimativas (HUISMAN; VEERKAMP; VAN ARENDONK, 2002; RESENDE; SILVA; AZEVEDO, 2014).

2.4 Análise fatorial exploratória

Em qualquer estudo em que se pretende analisar suas variáveis, estas podem ser estudadas de duas grandes maneiras ou por dois caminhos e métodos distintos, um em que a análise vai considerar as variáveis de forma isolada, o que é chamado de estatística univariada, e outro em que as variáveis serão analisadas de forma conjunta, o que é chamado de estatística multivariada. É comum dentro da ciência as pessoas serem treinadas para analisar variáveis isoladamente, e a partir da análise inferir sobre a realidade, porém em muitos casos um fenômeno é dependente de muitas variáveis, onde nestes casos a inferência feita por meio de uma análise univariada será falha, pois é necessário conhecer a totalidade das informações fornecidas pelo grupo de variáveis que faz parte daquele acontecimento. Isso não acontece em casos onde as variáveis consideradas no estudo são independentes entre si, o que permite com uma boa segurança que seja feita a interpretação do fenômeno a partir das informações estatísticas de poucas variáveis, que geralmente são medidas de tendência central e de dispersão dos dados (MOITA-NETO, 2015).

Com todo desenvolvimento tecnológico e científico que tem ocorrido ao longo dos anos, as informações sobre fenômenos e acontecimentos tem aumentado expressivamente a ordem de suas grandezas e a quantidade de dados a serem processados, que posteriormente servirão como base para estudos que vão gerar algum tipo de conhecimento. Com isso, a necessidade por ferramentas estatísticas mais robustas e que permitissem uma análise mais global de um determinado fenômeno a ser estudado era necessário, pois neste contexto uma abordagem univariada não seria suficiente. Isso fez com que a análise multivariada fosse a melhor forma de obter esse resultado, podendo ser entendida como um conjunto de métodos e técnicas que simultaneamente utiliza todas as variáveis importantes e necessárias de um determinado fenômeno ao interpretar do ponto de vista teórico seu conjunto de dados (HOWARTH; SINDING, 1983; MOITA-NETO, 2015).

Temos diversos métodos de análise multivariada, dependendo do objetivo do pesquisador, como: a análise por agrupamento hierárquico (HCA) e a análise por componentes principais, feitos quando se pretende verificar o relacionamento entre as amostras e o quanto elas se assemelham de acordo com o conjunto de variáveis utilizadas no estudo; regressão linear múltipla e redes neurais para quando se deseja fazer previsões; e em alguns casos, certos métodos como o simplex e o planejamento fatorial, podem ser utilizados em etapas iniciais e/ou na escolha das variáveis que vão melhor explicar um sistema dentro de um estudo, otimizando o processo (DILLON; GOLDSTEIN, 1984; MOITA-NETO, 2015).

A análise fatorial representa um conjunto de técnicas na estatística cujo objetivo é diminuir a quantidade de variáveis iniciais em um estudo de maneira a provocar a menor perda possível de informação, o que permite encontrar ou identificar fatores em um conjunto de medidas realizadas. Ela foi proposta inicialmente por Spearman e Pearson no começo do século 20, quando Spearman (1904) desenvolveu um método que servia para a criação de um índice geral de inteligência, chamado de fator “g”. Ele foi gerado a partir dos resultados obtidos de vários testes (escalas), que tinham como objetivo refletir supostamente essa aptidão. Este foi um primeiro método de análise fatorial, que neste caso era adequado apenas para a estimação de um único fator (BISHOP; FIENBERG; HLLAND, 2007; HOWARD, 2016; CHATFIELD; COLLINS, 2018; WATKINS, 2018).

As várias medidas obtidas inicialmente em um estudo não podem ser reunidas em um fator sem que se realize uma análise fatorial, que vai determinar quais os fatores ortogonais que melhor descrevem de forma aproximada e sucessiva os vetores resposta de n indivíduos de um conjunto de m testes. Em outras palavras faz com que os dados observados possam ser interpretados mais claramente, analisando-se quais os inter-relacionamentos entre as variáveis, de maneira com que elas sejam descritas de forma conveniente por um grupo de categorias básicas, que será em menor número do que as variáveis originais, o que é chamado de fatores (BISHOP; FIENBERG; HLLAND, 2007; CHATFIELD; COLLINS, 2018).

Os fatores podem ser entendidos como sendo um constructo, que poderia ser escalas, itens, uma variável que não foi observada ou uma medida de qualquer espécie, onde eles irão explicar a variância das variáveis observadas, da forma com que elas se mostram pelas correlações existentes entre as variáveis em análise no estudo, e cada fator é independente um do outro. A partir dos fatores é possível observar o que há de comum nas variáveis originais, pois eles vão expressar isso, de modo com que a análise fatorial vai permitir com que se explique a estrutura das covariâncias entre as variáveis, por meio de um modelo estatístico casual e assumindo a pressuposição de que existem as variáveis não observadas e subjacentes aos dados (TIMM, 2002; BISHOP; FIENBERG; HLLAND, 2007; CHATFIELD; COLLINS, 2018).

Esse método de análise estatística vai determinar quais variáveis pertencem a qual fator e também o quanto que cada variável explica cada fator, e, por ser sensível a correlações pobres entre as variáveis, consequência

de cálculos de estimativas de amostras pequenas, quando isso acontecer, a análise será prejudicada, inviabilizando o uso da técnica, portanto, em geral o mínimo é que se tenha pelo menos cinco casos para cada variável observada (BISHOP; FIENBERG; HLLAND, 2007; CHATFIELD; COLLINS, 2018).

Dentro da análise fatorial temos uma técnica conhecida como análise fatorial exploratória (AFE), que tem por objetivo geral identificar as relações subjacentes entre as variáveis medidas de um estudo. Nela serão estudadas quais as correlações que existem entre um grande número de variáveis agrupando-as em fatores, reduzindo os dados e consequentemente identificando quais as variáveis que são mais representativas ou então criando um conjunto de dados novo e menor que o original (HAIR et al., 2009; ALFARO; DÍAZ; GARCÍA, 2014). Na AFE os resultados têm relação direta com as decisões que o pesquisador tomar durante a obtenção de uma estrutura fatorial adequada, de modo que essas decisões precisam ter um embasamento teórico e metodológico bastante claro, para se evitar resultados errôneos e pouco confiáveis (FABRIGAR; WEGENER, 2011; HOWARD, 2016; WATKINS, 2018).

A abordagem multivariada é bastante utilizada em avaliações genéticas, para obter um aproveitamento das correlações que existem entre as características e aumentar a precisão das estimativas dos valores genéticos (EBV) dos animais estudados. A partir dessa abordagem, os EBVs estimados para todas as características consideradas em um estudo podem ser combinados formando um índice de seleção, que será utilizado para selecionar os animais e ao mesmo tempo obter ganhos nas características desejáveis de acordo com o objetivo do programa de melhoramento genético (GARRICK; GOLDEN, 2009; MOMEN et al., 2017).

Com o avanço científico, cada vez mais características de interesse dentro da avicultura de postura têm sido consideradas relevantes, o que faz com que a estatística multivariada tenha certas limitações ao analisar todas essas características, criando modelos muito complexos, de difícil interpretação e também obtendo um ganho genético menor, já que esse ganho é reduzido à medida que se aumenta a quantidade de características em um índice de seleção. Uma alternativa para solucionar esse impasse é a utilização de metodologias multivariadas, como a análise fatorial, que permitam com que se identifique um pequeno grupo de variáveis que possam representar um grande número de características, o que vai promover uma redução do esforço computacional e garantir um ganho genético elevado para todas as características desejadas dentro de um programa de seleção (MACCIOTTA et al., 2012; PAIVA et al., 2020).

Referências

- ALFARO, I. I.; DÍAZ, J. O.; GARCÍA, F. J. A. Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema*, 2014.
- ALTIERI, M. A.; NICHOLLS, C. I. Agroecology: Challenges and opportunities for farming in the Anthropocene. *Ciencia e investigación agraria: revista latinoamericana de ciencias de la agricultura*, v. 47, n. 3, p. 204-215, 2020.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS - ABNT (Brasil). Dossiê Técnico: Avicultura. 2016. Disponível em: <http://abnt.org.br/paginampe/biblioteca/files/upload/anexos/pdf/2f03721c5afcbc2aee77b4b3c6b0ece2.pdf>. Acesso em: 15 maio 2020.
- BAILEY, R. R.; SRINATH, M. Orthogonal moment features for use with parametric and non-parametric classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 18, n. 4, p. 389-399, 1996.
- BEGLI, H. E. et al. Genetic analysis of egg production traits in turkeys (*Meleagris gallopavo*) using a single-step genomic random regression model. *Genetics Selection Evolution*, v. 53, n. 1, p. 1-9, 2021.

- B EGLI, H. E. et al. Genomic dissection and prediction of feed intake and residual feed intake traits using a longitudinal model in F2 chickens. *animal*, v. 12, n. 9, p. 1792-1798, 2018.
- B ERKHOF F, J. et al. Consumer preferences and sensory characteristics of eggs from family farms. *Poultry science*, v. 99, n. 11, p. 6239-6246, 2020.
- B IGNARDI, A. B. et al. Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. *Genetics and Molecular Research*, p. 3565-3575, 2011.
- B ISHOP, Y. M.; FIENBERG, S. E.; HOLLAND, P. W. *Discrete multivariate analysis: theory and practice*. Springer Science & Business Media, 2007.
- B LOKHUIS, H. J. et al. Measuring and monitoring animal welfare: transparency in the food product quality chain. *ANIMAL WELFARE-POTTERS BAR THEN WHEATHAMPSTEAD-*, v. 12, n. 4, p. 445-456, 2003.
- B ÖLLING, D. et al. Genetic improvement of livestock for organic farming systems. *Livestock production science*, v. 80, n. 1-2, p. 79-88, 2003.
- C AVERO, D. et al. Improving hatchability in white egg layer strains through breeding. *Lohmann information*, v. 46, n. 1, p. 44-54, 2011.
- C HATFIELD, C.; COLLINS, A. J. *Introduction to multivariate analysis*. Routledge, 2018.
- C OSTA, A. J. D.; PETTIT, P.; BITTENCOURT, M. V. L. Innovations, technologies and economic development: An evolutionary view of the Brazilian poultry industry and its firms. In: *RESEARCH WORKSHOP ON INSTITUTIONS AND ORGANIZATIONS, III*. São Paulo: USP, IBMEC and FGV-SP, 2008. p. 13-14.
- D EMATTÊ FILHO, L. C.; MARQUES, P. E. M. Dinâmica tecnológica da cadeia industrial da avicultura alternativa: multifuncionalidade, desenvolvimento territorial e sustentabilidade. *Segurança Alimentar e Nutricional*, v. 18, n. 2, p. 1-11, 2011.
- D IKMEN, B. Y. et al. Egg production and welfare of laying hens kept in different housing systems (conventional, enriched cage, and free range). *Poultry science*, v. 95, n. 7, p. 1564-1572, 2016.
- D ILLON, W. R.; GOLDSTEIN, M. *Multivariate analysis: Methods and applications*. New York (NY): Wiley, 1984., 1984.
- F ABRIGAR, L. R.; WEGENER, D. T. *Exploratory factor analysis*. Oxford University Press, 2011.
- F ALCONER, D. S.; MACKAY, T. F. C. *Introduction to quantitative genetics*. 1996. Harlow, Essex, UK: Longmans Green, v. 3, 1996.
- F ALCONER, D. S.; MACKAY, T. F. C.; FRANKHAM, R. *Introduction to quantitative genetics (4th edn)*. *Trends in Genetics*, v. 12, n. 7, p. 280, 1996.
- F ARUQUE, S. et al. Evaluation of the performance of native chicken and estimation of heritability for body weight. *Journal of Bangladesh Academy of Sciences*, v. 37, n. 1, p. 93-101, 2013.
- F OOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAOSTAT). *Livestock Primary*. 2020. Disponível em: <<http://www.fao.org/faostat/en/#data/QL>>. Acesso em: 26 mar. 2022.
- G ARRICK, D. J.; GOLDEN, B. L. Producing and using genetic evaluations in the United States beef industry of today. *Journal of Animal Science*, v. 87, n. suppl_14, p. E11-E18, 2009.
- H AIR, J. F. et al. *Multivariate data analysis*. Bookman editora, 2009.
- H ELLMEISTER FILHO, P. et al. Efeito de genótipo e do sistema de criação sobre o desempenho de frangos tipo caipira. *Revista Brasileira de Zootecnia*, v. 32, p. 1883-1889, 2003.
- H ENDERSON JR, Charles R. Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random regressions. *Biometrics*, p. 623-640, 1982.

- HENDERSON, C. R. Applications of linear models in animal breeding. 423 p. University of Guelph, Guelph, Ontario, 1984.
- HILL, W. G. Applications of population genetics to animal breeding, from Wright, Fisher and Lush to genomic prediction. *Genetics*, v. 196, n. 1, p. 1-16, 2014.
- HILL, W. G.; BROTHERSTONE, S. Advances in methodology for utilizing sequential records. *BSAP Occasional Publication*, v. 24, p. 55-61, 1999.
- HIRVONEN, K.; BRAUW, A.; ABATE, G. T. Food consumption and food security during the COVID-19 pandemic in Addis Ababa. *American journal of agricultural economics*, v. 103, n. 3, p. 772-789, 2021.
- HORST, P.; MATHUR, P. K. Trends in economic values of selection traits for local egg production. In: *Proceedings, 19th World Poultry Congress, Amsterdam, the Netherlands, 20. 1992.* p. 577-583.
- HOWARD, M. C. A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve?. *International Journal of Human-Computer Interaction*, v. 32, n. 1, p. 51-62, 2016.
- HOWARTH, R. J.; SINDING-LARSEN, R. Multivariate analysis. In: *Handbook of exploration geochemistry.* Elsevier Science BV, 1983. p. 207-289.
- HUISMAN, A. E.; VEERKAMP, R. F.; VAN ARENDONK, J. A. M. Genetic parameters for various random regression models to describe the weight data of pigs. *Journal of Animal Science*, v. 80, n. 3, p. 575-582, 2002.
- JAKOBSEN, J. H. et al. Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. *Journal of Dairy Science*, v. 85, n. 6, p. 1607-1616, 2002.
- KETELAERE, B. De et al. Measuring the eggshell strength of 6 different genetic strains of laying hens: techniques and comparisons. *British poultry science*, v. 43, n. 2, p. 238-244, 2002.
- KIRKPATRICK, M.; HILL, W. G.; THOMPSON, R. Estimating the covariance structure of traits during growth and ageing, illustrated with lactation in dairy cattle. *Genetics Research*, v. 64, n. 1, p. 57-69, 1994.
- LAIRD, N. M.; WARE, J. H. Random-effects models for longitudinal data. *Biometrics*, p. 963-974, 1982.
- LEDUR, M. C. et al. Genetic and phenotypic parameters for productive traits in white egg layer stocks. *Pesquisa Agropecuaria Brasileira*, v. 28, n. 9, p. 1031-1037, 1993.
- LEWIS, R. M.; BROTHERSTONE, S. A genetic evaluation of growth in sheep using random regression techniques. *Animal Science*, v. 74, n. 1, p. 63-70, 2002.
- LUSH, J. L. The genetics of populations. Ames: Iowa State University Press, 1994. 900p.
- MACCIOTTA, N. P. P. et al. Use of multivariate factor analysis to define new indicator variables for milk composition and coagulation properties in Brown Swiss cows. *Journal of Dairy Science*, v. 95, n. 12, p. 7346-7354, 2012.
- MAIA, G. A. R. *Avicultura Alternativa: Carne e Ovos pelo Sistema de Pastejo.* Artigos Técnicos da Sociedade Nacional de Agricultura, Brasília, ano, v. 100, p. 1-6, 1997.
- MEYER, K. Random regression analyses using B-splines to model growth of Australian Angus cattle. *Genetics Selection Evolution*, v. 37, n. 6, p. 1-28, 2005.
- MOMEN, M. et al. A predictive assessment of genetic correlations between traits in chickens using markers. *Genetics Selection Evolution*, v. 49, n. 1, p. 1-14, 2017.
- MRODE, R. A. *Linear models for the prediction of animal breeding values.* Wallingford: CAB International, 2005. 2 Edition.

- MUIR, W. M. et al. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proceedings of the National Academy of Sciences*, v. 105, n. 45, p. 17312-17317, 2008.
- MOITA-NETO, J. M. Estatística multivariada: uma visão didática-metodológica. *Revista crítica na rede—Filosofia da ciência*, v. 9, 2015.
- PAIVA, J. T. et al. Genetic evaluation for latent variables derived from factor analysis in broilers. *British Poultry Science*, v. 61, n. 1, p. 3-9, 2020.
- PROVINE, W. B. Role of Mathematical Population Geneticists in the Evolutionary Synthesis of the 1930's and 40's. In: *Mathematical Models in Biological Discovery*. Springer, Berlin, Heidelberg, 1977. p. 2-31.
- PUTMAN, B. et al. A retrospective analysis of the United States poultry industry: 1965 compared with 2010. *Agricultural Systems*, v. 157, p. 107-117, 2017.
- RESENDE, M. D. V. de; SILVA, F. F. e; AZEVEDO, C. F. *Mathematical, Biometric and Computational Statistics: Mixed, Multivariate, Categorical and Generalized Models (REML/BLUP), Bayesian Inference, Random Regression, Genomic Selection, QTL-GWAS, Spatial and Temporal Statistics, Competition, Survival*. 21. ed. Viçosa: Suprema Gráfica e Editora Ltda, 2014. 881 p.
- SAVINO, V. J. M. et al. Avaliação de materiais genéticos visando à produção de frango caipira em diferentes sistemas de alimentação. *Revista Brasileira de Zootecnia*, v. 36, n. 3, p. 578-583, 2007.
- SCHAEFFER, L. R. Random regression models. Available in <http://animalbiosciences.uoguelph.ca/~lrs/BOOKS/rmbook.pdf>, 2016.
- SCHAEFFER, L. R. Random regressions in animal models for test-day production in dairy cattle. In: *World Congress of Genetics Applied Livestock Production*, 1994. 1994. p. 443-446.
- SCHAEFFER, L. R.; JAMROZIK, J. Random regression models: a longitudinal perspective. *Journal of Animal Breeding and Genetics*, v. 125, n. 3, p. 145-146, 2008.
- SEEKER, L. A. et al. Longitudinal changes in telomere length and associated genetic parameters in dairy cattle analysed using random regression models. *PLoS One*, v. 13, n. 2, p. e0192864, 2018.
- SILVA, M. C. da et al. Innate immunity in chicken lines developed by EMBRAPA Suínos e Aves: antimicrobial activity of macrophages and serum. *Revista Brasileira de Zootecnia*, v. 50, 2021.
- SPEARMAN, C. General intelligence, objectively determined and measured. *American Journal of Psychology*, v.15, p.201-293, 1904.
- TIMM, N. H. (Ed.). *Applied multivariate analysis*. New York, NY: Springer New York, 2002.
- VEERKAMP, R. F.; THOMPSON, R. A covariance function for feed intake, live weight, and milk yield estimated using a random regression model. *Journal of Dairy Science*, v. 82, n. 7, p. 1565-1573, 1999.
- WATKINS, M. W. Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, v. 44, n. 3, p. 219-246, 2018.
- WILSON, A. J.; KRUIK, L. E. B; COLTMAN, D. W. Ontogenetic patterns in heritable variation for body size: using random regression models in a wild ungulate population. *The American Naturalist*, v. 166, n. 6, p. E177-E192, 2005.
- WOLC, A. et al. Evaluation of egg production in layers using random regression models. *Poultry science*, v. 90, n. 1, p. 30-34, 2011.
- WOLC, A. et al. Heritability of egg production in laying hens under cumulative, multitrait and repeated measurement animal models. *Czech Journal of Animal Science*, v. 52, n. 8, p. 254, 2007.

Análise genética de características de produção e qualidade de ovos de linhagens de aves caipiras (*GALLUS GALLUS*) utilizando modelos de regressão e multivariados

Resumo

A avicultura de postura alternativa é um dos mercados com maior potencial de crescimento e nela são utilizadas linhagens de galinhas caipiras criadas com acesso a pasto ou piquetes. Uma das principais vantagens deste sistema, é que ao final do ciclo de postura dessas aves, elas podem ser abatidas com peso corporal superior às aves de postura comerciais. Porém, ainda não existem muitos estudos no sentido de selecionar e estudar linhagens brasileiras caipiras para postura dentro dos programas de melhoramento genético de aves. O objetivo deste trabalho foi analisar os dados de produção e qualidade de ovos de 732 aves de linhagens caipiras, por meio de modelos de regressão e multivariados, para estimar componentes de variância e realizar uma avaliação genética das linhagens Caipirinha, Caipirão e Azul da ESALQ, 7P, Carijó Barbada, Baianinha, Perna Longa e Sooho. Medidas repetidas ao longo do tempo ou dados longitudinais são como regra em estudos com melhoramento genético, e dentre algumas das classes de métodos mais adequados para estes estudos, temos os modelos de regressão aleatória (RRM), utilizados quando se deseja modelar a forma ou trajetória de observações que foram realizadas ao longo do tempo (dados longitudinais). Cada vez mais características da avicultura de postura têm sido consideradas relevantes, o que faz com que a estatística multivariada tenha certas limitações ao analisar todas essas características, criando modelos muito complexos. Uma alternativa para solucionar esse impasse é a utilização de metodologias multivariadas, como a análise fatorial exploratória (AFE), que permitam com que se identifique um pequeno grupo de variáveis (fatores) que possam representar um grande número de características. O modelo de regressão aleatória utilizando polinômios de Legendre de 2ª ordem se ajustou bem aos dados de produção de ovos de galinhas caipiras. A herdabilidade apresentou comportamento decrescente ao longo da trajetória das idades, enquanto que as variâncias genética aditiva e de ambiente permanente apresentaram o mesmo comportamento no início mas ao longo do tempo voltaram a crescer. A linhagem Caipirinha da ESALQ foi a que apresentou os melhores valores genéticos estimados. A análise fatorial exploratória (AFE) utilizando dados de valores genéticos estimados (EBV) para as características de produção e qualidade de ovos com dois fatores apresentou uma melhor explicação das porcentagens das variâncias que ocorrem nas 7 variáveis originais deste estudo, correlacionando melhor as características. A variável altura de albúmen não apresentou importância neste estudo na composição de nenhum dos fatores. No final da análise, das 7 variáveis iniciais, apenas 6 foram importantes para compor os dois fatores da AFE, sendo elas produção de ovos, idade ao primeiro ovo, espessura de casca, peso da gema, peso da casca e peso do ovo.

Palavras-chave: Análise fatorial exploratória, Avicultura alternativa, Medidas longitudinais, Polinômios de Legendre, Regressão aleatória

Abstract

Genetic analysis of production traits and egg quality of free-range bird (*Gallus gallus*) strains using regression and multivariate models

Alternative laying poultry is one of the markets with the greatest growth potential and it uses free range chicken lines raised with access to pasture or paddocks. One of the main advantages of this system is that at the end of the laying cycle of these birds, they can be slaughtered with higher body weight than commercial laying birds. However, there are still not

many studies in the sense of selecting and studying Brazilian redneck lines for laying within the poultry genetic improvement programs. The objective of this work was to analyze data on production and egg quality of 732 free-range birds, using regression and multivariate models, to estimate variance components and carry out a genetic evaluation of the Caipirinha, Caipirão and Azul strains from ESALQ, 7P, Carijó Barbada, Baianinha, Long Leg and Sooho. Repeated measurements over time or longitudinal data are as a rule in studies with genetic improvement, and among some of the most suitable classes of methods for these studies, we have the random regression models (RRM), used when you want to model the shape or trajectory of observations that were carried out over time (longitudinal data). More and more characteristics of laying poultry have been considered relevant, which makes multivariate statistics have certain limitations when analyzing all these characteristics, creating very complex models. An alternative to solve this impasse is the use of multivariate methodologies, such as exploratory factor analysis (EFA), which allow the identification of a small group of variables (factors) that may represent a large number of characteristics. The random regression model using 2nd order Legendre polynomials fitted well to egg production data from free-range hens. Heritability showed a decreasing behavior along the age trajectory, while the additive genetic and permanent environment variances showed the same behavior at the beginning but over time they increased again. The Caipirinha strain from ESALQ was the one that presented the best estimated genetic values. Exploratory factor analysis (EFA) using estimated breeding values (EBV) data for two-factor egg production and quality traits presented a better explanation of the percentages of variances that occur in the 7 original variables of this study, better correlating the traits. The albumen height variable was not important in this study in the composition of any of the factors. At the end of the analysis, of the 7 initial variables, only 6 were important to compose the two AFE factors, namely egg production, age at first egg, shell thickness, yolk weight, shell weight and egg weight.

Keywords: Alternative poultry farming, Exploratory factor analysis, Longitudinal measurements, Legendre polynomials, Random regression

2.5 Introdução

O ovo é um dos alimentos mais consumidos e nutricionalmente completos do mundo, com produção mundial em torno de 116 milhões de toneladas, totalizando pouco mais de 2,2 trilhões de ovos vindos de quase 11 bilhões de poedeiras. Dentre os maiores produtores, o Brasil é o sexto no geral e segundo maior do continente Americano, com uma produção de 3,26 milhões de toneladas (57,2 bilhões de ovos) e um plantel de aves de 252 milhões (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS - FAOSTAT, 2020).

As linhagens utilizadas na avicultura comercial de postura são aves que foram selecionadas para terem alto desempenho produtivo em gaiola, onde geralmente toda a alimentação e ambiente são controlados. Na avicultura alternativa de postura, são utilizadas linhagens de galinhas caipiras que vêm sendo criadas com acesso a pasto ou piquetes nas 24 horas do dia desde que foram introduzidas no Brasil na época da colonização, dessa forma, são linhagens adaptadas e resistentes a essas condições (BOLLING et al., 2003; HELLMEISTER FILHO et al., 2003). Uma das principais vantagens deste sistema, é que ao final do ciclo de postura dessas aves, elas podem ser abatidas com peso corporal superior às aves de postura comerciais, o que pode ser um atrativo para os produtores (FARUQUE et al., 2013; DIKMEN et al., 2016). Adicionalmente, a produção alternativa de ovos é inserida na agricultura familiar, na qual é responsável por grande parte dos alimentos que chegam à mesa das pessoas (ALTIERI; NICHOLLS, 2020; BERKHOFF et al., 2020; SILVA et al., 2021).

A característica mais importante na avicultura de postura é a produção de ovos, a variável mais estudada em programas de seleção, tendo seu progresso sido um sucesso, o que fez com que a quantidade de ovos produzidos por uma ave mais do que dobrasse da década de 50 até os dias atuais, chegando a valores acima de 250 ou 300 ovos em todo seu ciclo de postura. Durante esse ciclo, essa característica é medida diariamente ao longo do tempo de vida produtiva do animal, portanto trata-se de uma variável caracterizada como medida longitudinal (WOLC et al., 2007).

Dentre algumas das classes de métodos mais adequados ao estudo de dados longitudinais ou medidas repetidas, temos a classe não paramétrica, que usa funções de covariância baseadas em funções flexíveis, como os polinômios de Legendre, gerando o método de regressão aleatória (BAILEY; SRINATH, 1996; RESENDE; SILVA; AZEVEDO, 2014).

O modelo de regressão aleatória (RRM) possui a vantagem de ter a capacidade de prever os valores genéticos para a produção cumulativa de ovos ao longo de um período de postura e em qualquer dia específico dessa produção, avaliando a persistência da mesma (BIGNARDI et al., 2011). Além disso, visa colocar mais peso a informações individuais do que informações de parentesco, e em modelos curvilíneos as observações extremas também costumam receber mais peso do que as médias, e quanto maior o grau do polinômio, mais essas características são evidenciadas (LUSH, 1994). A estrutura regressiva será aplicada a todos os fatores aleatórios, que são os efeitos aditivos diretos e maternos e os efeitos permanentes, dessa forma os RRM geralmente selecionam animais mais jovens quando comparados a outros modelos de repetibilidade (RESENDE; SILVA; AZEVEDO, 2014).

Na regressão aleatória, são gerados diferentes coeficientes de regressão que são associados a cada indivíduo diferente dentro de uma população, assim, não somente os indivíduos como também os coeficientes de regressão associados a eles serão considerados obrigatoriamente como efeitos aleatórios (HENDERSON, 1984). Os registros que são feitos desse animal assumem diferentes correlações genéticas entre eles, pois é levada em consideração a premissa de que a característica estudada pode estar mudando de maneira contínua e gradual ao longo do tempo. Nos RRM são geradas curvas para as características avaliadas, e a forma dessas curvas será considerada, de modo a permitir que sejam ajustadas curvas aleatórias para cada um dos indivíduos da população, sendo essas curvas expressas como sendo os desvios de uma curva média da dessa população ou grupos de indivíduos (KIRKPATRICK et al., 1994; SCHAEFFER & JAMROZIK, 2008).

Os RRM são utilizados quando se deseja estimar matrizes que são parte das funções de covariância para se estimar uma gama maior de variâncias e covariâncias. Também, quando se deseja modelar a forma ou trajetória de observações que foram realizadas ao longo do tempo (dados longitudinais) ou em situações onde são feitos registros repetidos das observações de uma característica, sendo coletados esses dados várias vezes durante o curso de vida desse animal, os RRM também podem ser utilizados, ressaltando que fenotipicamente essa trajetória precisa ser ajustada, ao mesmo tempo em que a variação ao longo da trajetória precisa ser considerada (HILL & BROTHERSTONE, 1999; SCHAEFFER, 2016).

No momento de escolher um modelo de regressão apropriado, deve-se escolher aquele com o menor número de parâmetros e que se ajuste bem aos dados, pois um bom ajuste para a trajetória é muito importante em um modelo de regressão aleatória (SCHAEFFER, 2016).

Como o avanço científico, cada vez mais características de interesse dentro da avicultura de postura têm sido consideradas relevantes, o que faz com que a estatística multivariada tenha certas limitações ao analisar todas essas características, criando modelos muito complexos, de difícil interpretação e também obtendo um ganho genético menor, já que esse ganho é reduzido à medida que se aumenta a quantidade de características em um índice

de seleção. Uma alternativa para solucionar esse impasse é a utilização de metodologias multivariadas, como a análise fatorial (MACCIOTTA et al., 2012; PAIVA et al., 2020). A análise fatorial representa um conjunto de técnicas na estatística cujo objetivo é diminuir a quantidade de variáveis iniciais em um estudo de maneira a provocar a menor perda possível de informação, o que permite encontrar ou identificar fatores em um conjunto de medidas realizadas (HOWARD, 2016; CHATFIELD; COLLINS, 2018; WATKINS, 2018).

Os fatores podem ser entendidos como sendo um constructo, que poderia ser escalas, itens, uma variável que não foi observada ou uma medida de qualquer espécie, onde eles irão explicar a variância das variáveis observadas, da forma com que elas se mostram pelas correlações existentes entre as variáveis em análise no estudo, e cada fator é independente um do outro. Esse método de análise estatística vai determinar quais variáveis pertencem a quais fatores e também o quanto que cada variável explica cada fator, e, por ser sensível a correlações pobres entre as variáveis, consequência de cálculos de estimativas de amostras pequenas, quando isso acontecer, a análise será prejudicada, inviabilizando o uso da técnica, portanto, em geral o mínimo é que se tenha pelo menos cinco casos para cada variável observada (TIMM, 2002; CHATFIELD; COLLINS, 2018).

Dentro da análise fatorial temos uma técnica conhecida como análise fatorial exploratória (AFE), que tem por objetivo geral identificar as relações subjacentes entre as variáveis medidas de um estudo. Nela será estudado quais as correlações que existem entre um grande número de variáveis agrupando-as em fatores, reduzindo os dados e consequentemente identificando quais as variáveis que são mais representativas ou então criando um conjunto de dados novo e menor que o original (HAIR et al., 2009; ALFARO; DÍAZ; GARCÍA, 2014).

A abordagem multivariada é bastante utilizada em avaliações genéticas, para obter um aproveitamento das correlações que existem entre as características e aumentar a precisão das estimativas dos valores genéticos (EBV) dos animais estudados. A partir dessa abordagem, os EBVs estimados para todas as características consideradas em um estudo podem ser combinados formando um índice de seleção, que será utilizado para selecionar os animais e ao mesmo tempo obter ganhos nas características desejáveis de acordo com o objetivo do programa de melhoramento genético (GARRICK; GOLDEN, 2009; MOMEN et al., 2017).

São poucas as linhagens desenvolvidas no Brasil, e o material genético puro das linhagens de aves importadas são exclusivos dessas empresas internacionais, as únicas capazes de gerar o animal híbrido industrial que comercializam. Considerando essa demanda por mais estudos, especialmente na área de genética, o objetivo deste trabalho foi analisar os dados de produção e qualidade de ovos de linhagens caipiras, por meio de modelos estatísticos multivariados e univariados, para se estimar os componentes de variância e realizar uma avaliação genética dessas linhagens.

2.6 Material e Métodos

O trabalho foi desenvolvido no Setor de Melhoramento Genético da Escola Superior de Agricultura “Luiz de Queiroz” – ESALQ/USP, Piracicaba, Brasil.

A dieta fornecida foi uma ração comercial à base de milho e farelo de soja e a água foi fornecida ad libitum durante todo o período. O programa de luz adotado foi de 16 horas de luz/dia.

2.6.1 Linhagens e dados de produção de ovos

Foram coletados e analisados dados de produção de ovos diários de aves de oito linhagens caipiras do setor de melhoramento genético animal do Grupo de Estudos em Estatísticas e Melhoramento Genético Animal – GEMA da ESALQ/USP, totalizando 26.637 medidas repetidas ao longo desse tempo, observadas de 732 fêmeas e 72 grupos de contemporâneos gerados a partir de data de nascimento, lote, linha de gaiolas e linhagens comuns.

As aves possuem controle completo do pedigree, foram considerados na análise até três gerações de parentesco e são compostas por oito linhagens, todas com características distintas:

- Caipirinha da ESALQ (HA): linhagem de dupla aptidão para corte e postura, fenótipo para plumagem variado, presença de fenótipo para topete e pescoço pelado, peso vivo médio de 2,42kg e produção de até 240 ovos por ciclo de postura.
- Pinto Preto Pesado de Pasto de Pescoço Pelado de Piracicaba (7P): linhagem de corte e postura, com fenótipo para plumagem preta, por volta de 25% da população com fenótipo para pescoço pelado e peso vivo médio de 3,19kg.
- Carijó Barbada (JO): linhagem de dupla aptidão para corte e postura, fenótipo de plumagem do tipo barrada, presença de fenótipo para barba e costeleta, peso vivo médio de 2,20kg e produção de até 240 ovos por ciclo de postura.
- Caipirão da ESALQ (AO): linhagem de corte e postura, com fenótipos para cor de plumagem variados, média de peso vivo de 3,28kg.
- Azul da ESALQ (AZ): linhagem de postura, de crescimento lento e fenótipo de plumagem variado e presença de fenótipo para ovos azuis e nanismo.
- Baianinha (BA): linhagem para ornamental, com fenótipos para cor variados, presença de fenótipos arrepiados ou gene frisado, sem rabo ou gene suro, pescoço pelado, canela verde, pena nos pés, topete, barba e ovos azuis ou esverdeados.
- Perna Longa (PL): linhagem selecionada para ornamental, com fenótipo longilíneo e canela alongada.
- Carne Preta ou SooHo (CP): ave com fenótipo para plumagem preta, pigmentação escura na pele, crista, barbela, bico e patas.

As aves foram inseminadas e reproduzidas na proporção de 1 macho para 3 ou 5 fêmeas e os ovos foram identificados com informações dos progenitores. Esses ovos foram incubados por 21 dias em incubadoras industriais pertencentes ao Departamento de Zootecnia da ESALQ/USP.

Os pintainhos/as foram identificados com anilha, inserida na pele de sua asa e receberam as vacinas contra Boubá, Gumboro, Marek e Newcastle. Eles ficaram alojados em boxes de 3m², com lotação de até 30 aves e receberam ração comercial para cria e recria, crescimento, pré e pós postura assim como água ad libitum. Nas gaiolas, onde as aves permaneceram durante todo o período de postura, a ração era fornecida para cada ave individualmente, numa quantidade média de 120g.

Foi esperado que as aves atingissem idade a primeira postura (IPP), sendo esta data anotada para cada ave, e a partir desse momento, foram computadas as produções de ovos (Prod_ovos). Os ovos foram coletados diariamente para a análise de regressão aleatória, no intervalo de idades de 120 dias até 456 dias, respectivamente a idade da primeira ave a pôr o primeiro ovo até a idade em que as aves estavam com 65 semanas de vida.

2.6.2 Avaliação dos ovos

Para avaliar as características peso do ovo (PEOVO), altura do albúmen (ALT_ALB), Unidades de Haugh (HU), espessura da casca (ESP_CA), peso da gema (PE_GEMA) e peso da casca (PE_CASCA), foram selecionadas apenas as linhagens AZ, HA, BA e AO, totalizando 504 aves. Uma vez por mês durante 5 meses seguidos (dezembro a abril), foram coletados durante três dias dentro de uma mesma semana os ovos de cada ave pertencente a essas linhagens. Esses ovos foram analisados no laboratório do Instituto de Zootecnia em Nova Odessa/SP, com o equipamento Digital Egg Tester DET 6000, da Nabel. A coleta dos ovos foi feita preferencialmente pela manhã e a análise foi feita no mesmo dia à tarde, e, quando a postura ocorria no período da tarde, esse ovo era armazenado em local refrigerado para análise no dia posterior.

Depois de obtidas as medidas, os dados foram processados e analisados com o software R versão 4.1.0 (R Core Team, 2021), para obtenção da estatística descritiva, remoção de outliers e observação dos dados. Os dados de ovos com duas gemas foram excluídos, assim como as aves que não iniciaram a postura durante todo o período de estudo.

2.6.3 Parâmetros genéticos pelo modelo de regressão aleatória

Numa análise inicial considerando a produção de ovos para cada semana de coleta, foi calculada a variância para essa característica dentro desses intervalos de semanas e feito o teste de homocedasticidade a partir do método proposto por Cochran (1950), onde constatou-se que havia homogeneidade das variâncias residuais, portanto os resultados apresentados neste estudo foram obtidos sob uma suposição de homogeneidade de variâncias residuais.

O modelo geral utilizado foi:

$$y = X\beta + Za + Sc + \varepsilon$$

onde y é o vetor de observações fenotípicas; X é a matriz de incidência dos efeitos fixos e β é o vetor de efeitos fixos (grupo de contemporâneos); Z é a matriz de incidência dos efeitos aleatórios genéticos aditivos; a é o vetor de efeitos genéticos aditivos, com $a \sim N(0, A\sigma_a^2)$, S é a matriz de incidência dos efeitos aleatórios de ambiente permanente, c é o vetor dos efeitos de ambiente permanente e ε é o vetor de efeitos residuais, com $\varepsilon \sim NID(0, I\sigma^2)$.

O critério utilizado para a determinação do melhor modelo foi o Critério de Informação de Akaike (AIC), representado como:

$$AIC = -2\log L + 2p$$

em que $-2\log$ é o logaritmo da função de máxima verossimilhança restrita; p é igual ao número de parâmetros do modelo.

A herdabilidade foi calculada segundo o modelo a seguir:

$$var(\hat{h}_t^2) = var\left(\frac{\hat{g}_{t,t}}{\hat{p}_{t,t}}\right) \approx \frac{(\hat{p}_{t,t}^2 var(\hat{g}_{t,t}) + \hat{g}_{t,t}^2 var(\hat{p}_{t,t}) - 2\hat{g}_{t,t}\hat{p}_{t,t} cov(\hat{g}_{t,t}\hat{p}_{t,t}))}{\hat{p}_{t,t}^4}$$

em que, $\hat{g}_{t,t}$ é a variância genética aditiva, $\hat{p}_{t,t}$ é a variância fenotípica e $\hat{g}_{t,t}\hat{p}_{t,t}$ a covariância entre genótipo e fenótipo.

O valor genético (EBV) foi estimado por meio da seguinte equação:

$$EBV_{it} = \sum_{m=0}^{g_{\alpha-1}} \hat{\alpha}_{im} \varphi_m(\chi_t)$$

em que, EBV é o valor genético do animal *i* na idade *t*; $\hat{\alpha}_{im}$ é o coeficiente de regressão aleatória estimado para o efeito genético aditivo; $\varphi_m(\chi_t)$ é o coeficiente do polinômio de Legendre; e $g_{\alpha-1}$ é a ordem do polinômio de Legendre.

Os componentes de variância foram estimados a partir de um modelo de regressão aleatória utilizando polinômios de Legendre, sob modelo animal, usando os softwares Legendre, RENUMF90 e AIREMLF90 que são parte da família de programas do BLUPF90 (MISZTAL et al., 2018).

As funções de covariância foram estimadas para os componentes genético aditivo direto do animal e de ambiente permanente, tendo sido modeladas por meio de polinômios ortogonais de Legendre de primeira ordem (linear) e segunda ordem (quadrático). Os efeitos fixos considerados no modelo foram Grupo de Contemporâneos (GC) e idade do animal foi considerada como covariável, enquanto que os efeitos aleatórios foram os efeitos genéticos aditivo direto e de ambiente permanente.

Nos RRM se pode utilizar todos os dados disponíveis ao mesmo tempo e não é necessário realizar um pré-ajustamentos para idades padrão. Neles os efeitos de tratamento são modelados por $\sum_r^{l-1} \beta_{ir} \Phi(a_{ik}^*) r$, em que o termo β_{ir} denota o conjunto de *l* coeficientes de regressão aleatória para o *i*-ésimo tratamento, $\Phi(a_{ik}^*) r$ é o *r*-ésimo polinômio sobre a idade padronizada (a_{ik}^*) da medida *k*. A matriz *G* estimada para os efeitos de tratamentos é dada por $G = \Phi B \Phi'$, em que Φ é a matriz contendo os efeitos aleatórios dos polinômios para as idades de medição e *B* é a estimada matriz de variância-covariância dos coeficientes de regressão polinomiais. De forma geral, o valor fenotípico de um indivíduo será dado por $y_{ij} = F_{ij} + \sum_r^{l-1} \beta_{ir} \Phi(a_{ik}^*) r + \varepsilon_{ij}$, em que F_{ij} e ε_{ij} referem-se aos efeitos fixos do modelo e erros aleatórios respectivamente (MEYER & HILL, 1997; RESENDE; SILVA; AZEVEDO, 2014).

A produção de ovos foi modelada por regressão aleatória utilizando polinômios ortogonais de Legendre segundo o modelo animal a seguir:

$$y_{ij} = F + \sum_{m=0}^{n-1} \beta_m \phi_m + \sum_{m=0}^{n_G-1} \alpha_{im} \phi_m + \sum_{m=0}^{n_p-1} \rho_{im} \phi_m + \varepsilon_{ij}$$

onde y_{ij} é a produção de ovos na idade *j*, da ave *i*; *F* são os efeitos fixos compostos por GC e a covariável idade; β_m é o coeficiente de regressão fixo *m* da produção de ovos sobre o polinômio de Legendre para modelar a curva média da população; α_{im} e ρ_{im} são os coeficientes de regressão aleatória para o efeito genético aditivo direto e de ambiente permanente do animal, respectivamente, para a ave *i*; $n - 1$, $n_G - 1$ e $n_p - 1$ são as ordens dos polinômios de Legendre; ϕ_m é a função polinomial de Legendre *m*; e ε_{ij} denota o efeito aleatório residual.

Foram modeladas as RRM por polinômios ortogonais de Legendre de ordem 2 (quadrático), tanto para o componente genético aditivo direto quanto para o componente de ambiente permanente de animal.

Os cálculos das estimativas dos componentes de variância feitos a partir dos resultados da saída da análise do software AIREMLF90 foram feitos no software R versão 4.1.0 (R Core Team, 2021).

2.6.4 Análise fatorial exploratória

Os componentes de variância foram estimados a partir de um modelo univariado, sob modelo animal geral conforme apresentado no item anterior, usando os softwares RENUMF90 e AIREMLF90 que são parte da família de programas do BLUPF90 (MISZTAL et al., 2018). Foram estimadas a variância genética aditiva, variância fenotípica, herdabilidade (h^2) e valores genéticos aditivos (EBV), para as características Prod_ovos, IPP, PEOVO, ESP_CA, HU, PE_GEMA, PE_CASCA e ALT_ALB, considerando os 504 animais das linhagens AZ, HA, BA e AO, as quatro linhagens mais representativas. O efeito fixo considerado no modelo foi Grupo de Contemporâneos (GC) e o efeito aleatório foi o do próprio animal. Após estimar os valores de EBV para cada característica, eles foram utilizados na análise fatorial exploratória (AFE).

A análise fatorial exploratória (AFE) é uma técnica de interdependência que foi empregada para definir a estrutura inerente dos valores genéticos preditos da produção, idade e características da qualidade do ovo. O modelo de análise fatorial empregado é apresentado na Equação (1)

$$\begin{aligned} X_1 &= a_{11} \times F_1 + a_{12} \times F_2 + \dots + a_{1m} \times F_m + e_p \\ X_2 &= a_{21} \times F_1 + a_{22} \times F_2 + \dots + a_{2m} \times F_m + e_p \\ &\vdots \\ X_p &= a_{p1} \times F_1 + a_{p2} \times F_2 + \dots + a_{pm} \times F_m + e_p \end{aligned} \quad \text{Equação (1)}$$

Onde: X_p é o score p th da variável padronizada ($p = 1, 2, \dots, m$), F_m é o fator extraído, a_{pm} é a carga fatorial e e_p é o erro.

Os escores fatoriais para cada grupo foram estimados multiplicando as variáveis padronizadas pelo coeficiente do score do fator correspondente, conforme segue a Equação (2)

$$\begin{aligned} F_1 &= d_{11} \times X_1 + d_{12} \times X_2 + \dots + d_{1j} \times X_{jp} \\ F_2 &= d_{21} \times X_1 + d_{22} \times X_2 + \dots + d_{2j} \times X_{jp} \\ &\vdots \\ F_j &= d_{j1} \times X_1 + d_{j2} \times X_2 + \dots + d_{jp} \times X_{jp} \end{aligned} \quad \text{Equação (2)}$$

Onde: F_j é o j -ésimo fator extraído, d_{pj} é o coeficiente de pontuação do fator e p é o número de variáveis (Hair et al., 2009). O número de componentes principais com autovalores superiores a 1 (regra de Kaiser; Kaiser, 1960). O critério de Kaiser–Meyer–Olkin ($KMO = 0,599$ para dados de EBV e $KMO = 0,625$ para os dados de fenótipo) e o teste de esfericidade de Bartlett ($P < 0,001$) foram atendidos de acordo com Hair et al. (2009). A matriz de componentes rotacionados foi desenvolvida usando rotação ortogonal, método varimax.

2.7 Resultados

2.7.1 Estatística descritiva

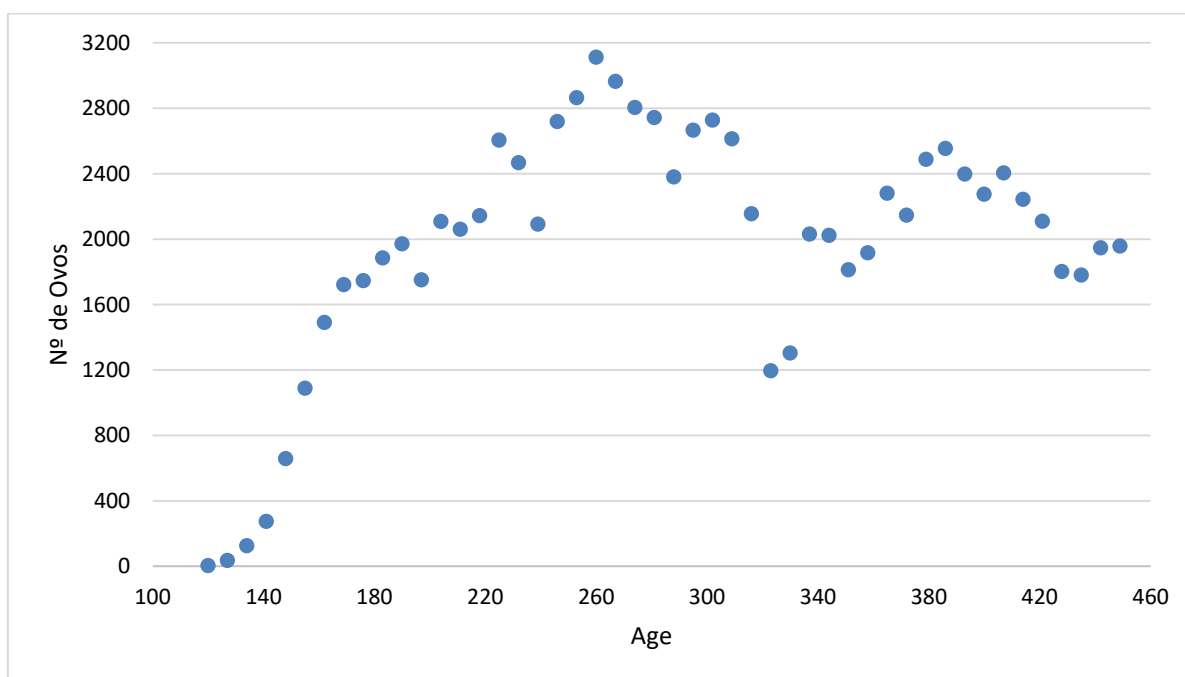
O total de ovos coletados no período deste estudo para as 732 poedeiras foi de 94.624 mil, com média de produção por ave de 3,552 ovos/semana (Tabela 1).

Tabela 1: Resumo descritivo dos valores de produção de ovos das 732 aves.

Min. ⁽¹⁾	Median	Mean	Max. ⁽²⁾	Var. ⁽³⁾	SD ⁽⁴⁾	CV% ⁽⁵⁾	Total de ovos
0	4	3,55	7	3,51	1,87	52,7	94.624

⁽¹⁾Valor mínimo observado; ⁽²⁾valor máximo observado; ⁽³⁾variância; ⁽⁴⁾desvio padrão; ⁽⁵⁾Coefficiente de variação.

Os valores de produção semanal total no período de 120 dias de vida até 456 dias de vida das aves variaram entre quatro (menor valor coletado) e 3112 ovos (maior valor coletado), respectivamente para as idades aos 120 dias e 260 dias, com média de produção semanal de 1971,33 ovos, e o restante dos valores ao longo do tempo também podem ser observados na Figura 1.

**Figura 1:** Produção total por semana de ovos no período de 120 dias a 456 dias de idade das aves.

Considerando os 72 grupos de contemporâneos (GC), o GC I1PL30062017 foi o grupo que teve o maior número de ovos produzidos no período, com 6671 ovos, sendo esse mesmo GC o que tinha o maior número de animais dentro dele (47 aves), seguidos pelos GC D1AZ07072017 (41 aves) e J1BA07072017 (43 aves) com 6193 ovos e 5735 ovos respectivamente. A menor produção registrada foi para os GC X1HA01072017 e X1JO31072017, com 13 ovos e ambos com apenas uma ave em cada. Todos os grupos de contemporâneos com menos de quatro animais foram eliminados da análise de regressão aleatória, assim, dos 72 GC sobraram 43 e das 732 aves ficaram apenas 679 animais. Considerando apenas os grupos que permaneceram na análise RRM, o menor valor de produção registrado foi de 215 ovos para o GC X17P07072017 com apenas quatro aves e ainda considerando esse mesmo tamanho de grupo, o GC B1AZ01082017 conseguiu produzir três vezes mais, totalizando 670 ovos.

Em todos os GC com quatro animais ou mais (43 no total) foram observados valores de produção nula para pelo menos uma das aves durante o período de coleta, e 37 desses GC tiveram ao menos uma observação de sete ovos produzidos em pelo menos uma semana produtiva por alguma ave, enquanto que os outros seis GC o valor máximo produzido em uma semana foi de seis ovos.

As linhagens pertencentes aos GC com maiores produções de ovos foram em primeiro lugar a linhagem PL, seguido por AZ e BA, respectivamente linhagens mais especializadas em ornamental, postura e ornamental,

enquanto que as que as pertencentes aos GC que menos produziram foram JO e HA, linhas especializadas em dupla aptidão, porém podemos destacar a linhagem HA, que também fez parte de GC com boa produção e que foi na sequência a linha que mais produziu ovos depois das 3 primeiras citadas. Os maiores valores de CV foram para os grupos com maiores quantidades de aves e maiores produções, e o contrário também foi verdade neste estudo.

2.7.2 Parâmetros genéticos

Foram testados dois modelos para este estudo num primeiro momento, um modelo mais simples de primeira ordem utilizando RRM com polinômios ortogonais de Legendre linear (Leg1) e outro modelo de segunda ordem utilizando polinômios ortogonais de Legendre quadrático (Leg2), sendo escolhido para o estudo aquele com menor valor de AIC (Critério de Informação de Akaike), pois o menor valor indica o modelo com melhor ajuste aos dados (Tabela 2). O modelo escolhido para este estudo foi o de escala quadrática ou de segunda ordem, com 3 parâmetros, sendo eles o intercepto e os polinômios de primeira e segunda ordem, com valor de AIC= 99.545,03. Assumiu-se a suposição de homogeneidade de variâncias residuais (R) para o modelo.

Tabela 2: Modelos e critérios de seleção para os diferentes graus de polinômios de Legendre.

Modelo	R ⁽³⁾	SE (R) ⁽⁴⁾	-2log(L)	AIC
Leg1 ⁽¹⁾	2,91	0,026	99.932,7	99.946,7
Leg2 ⁽²⁾	2,8	0,026	99.519,03	99.545,03

⁽¹⁾Polinômio de Legendre de 1ª ordem; ⁽²⁾polinômio de Legendre de 2ª ordem; ⁽³⁾variância residual; ⁽⁴⁾erro padrão da variância residual.

Os valores para os componentes de variância foram estimados para 1.157 animais no total, valor obtido pela soma das 679 poedeiras mais os seus parentes presentes no banco de dados, com exceção dos valores para variância de ambiente permanente, que foram estimados apenas para as 679 poedeiras que tiveram seus dados de produção considerados neste estudo.

A variância genética aditiva (AGV) ao longo da trajetória de cada idade no intervalo de 120 a 456 dias pode ser vista na Figura 2, onde podemos observar um comportamento decrescente desses valores nas idades iniciais de postura, indo de 0,9 até aproximadamente 0,1 na idade de 225 dias, permanecendo quase constante, com ligeiro aumento para próximo de 0,2 até a idade de 400 dias, quando começou a crescer um pouco mais até atingir quase 0,3 aos 456 dias. Na mesma figura podemos ver os valores de variância de ambiente permanente (PEV), que apresentou comportamento semelhante ao da AGV, decrescendo de por volta de 1,32 na idade de 120 dias para valores abaixo de 0,2 até a idade de 400 dias, quando voltou a crescer novamente, até atingir um valor de 0,25 na idade de 456 dias. No intervalo de por volta de 288 até 300 dias os valores das duas variâncias se encontraram e foram bastante semelhantes, com diferenças nos decimais entre eles. Como a variância residual foi considerada homogênea (Tabela 4), seu valor permaneceu constante ao longo da trajetória das idades.

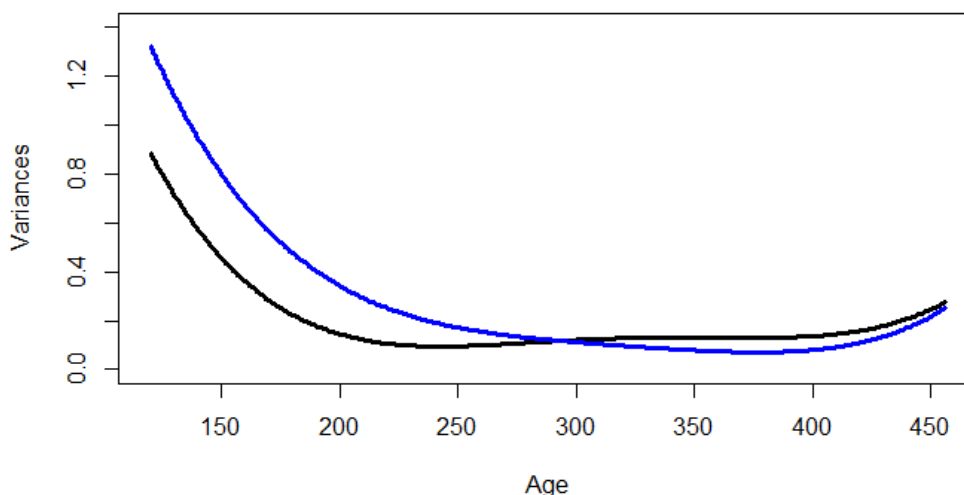


Figura 2: Estimativas da variância genética aditiva (-----) e da variância de ambiente permanente (-----) para a característica de produção de ovos segundo modelo Leg2 no intervalo de idades de 120 a 456 dias de vida das aves caipiras.

Os valores de herdabilidade (h^2) para a característica de produção de ovos ao longo da trajetória das idades de 120 a 456 dias para a população considerada nesta análise variaram de quase 0,2 nas idades iniciais coletadas, caindo para valores próximos de 0,03 por volta dos 200 dias de vida e permanecendo por volta de 0,05 até os 400 dias de vida, quando voltaram a crescer até quase 0,1 aos 456 dias de idade (Figura 3).

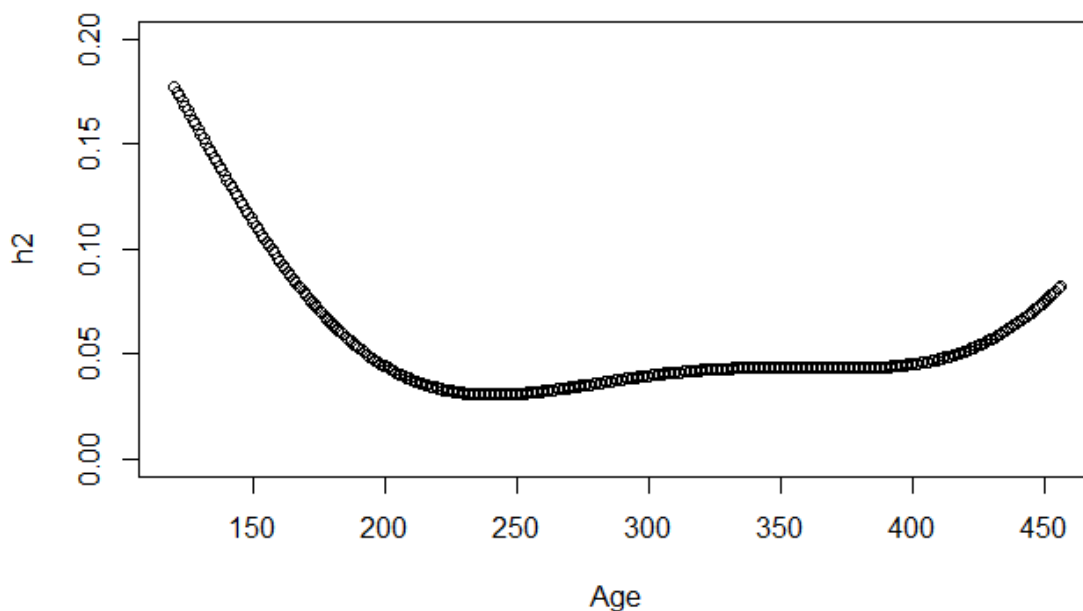


Figura 3: Valores de herdabilidade (h^2) ao longo da trajetória das idades de 120 até 456 dias de vida para a característica produção de ovos.

Os valores genéticos aditivos estimados (EBV) para cada um dos 1.157 animais ao longo da trajetória de idades no intervalo de 120 a 456 dias (Figura 4) variaram de -1,64 a 1,56. A grande maioria dos valores de EBV se concentrou no intervalo de aproximadamente 0,65 a -0,6.

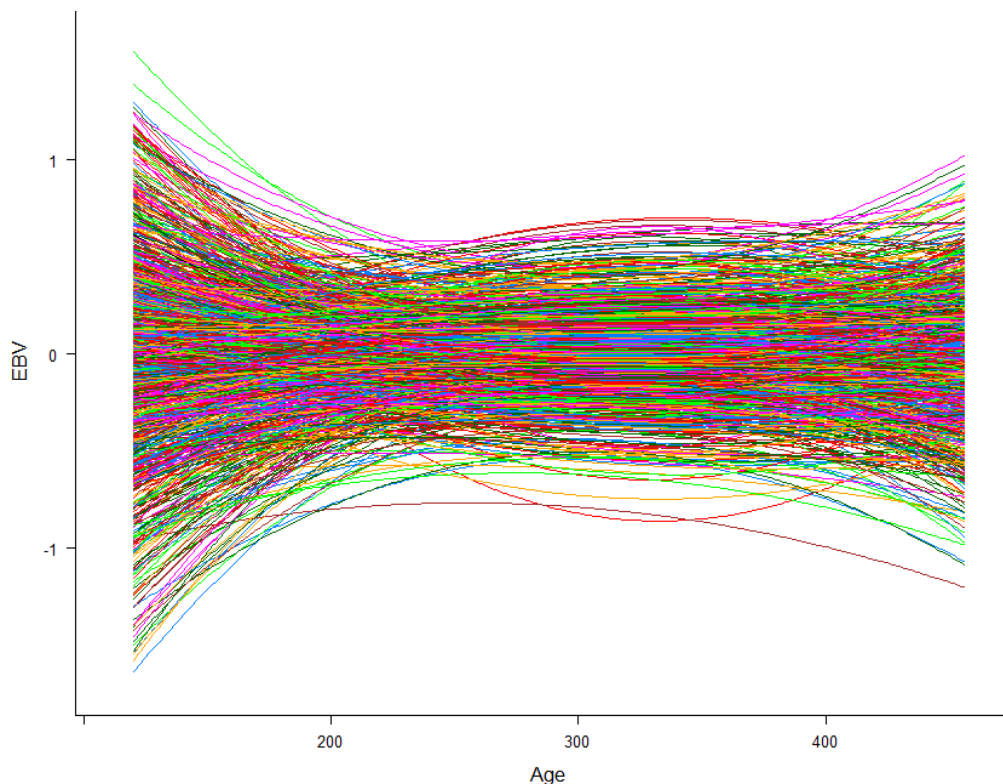


Figura 4: Valores genéticos aditivos estimados (EBV) para cada um dos 1157 animais (representados por cada linha do gráfico) durante a trajetória de idades de 120 a 456 dias para a característica produção de ovos.

O Top 10 das aves com os maiores EBV ao longo da trajetória de 120 a 456 dias pode ser visto na Figura 5. O menor valor de EBV observado dentre os 10 foi para o animal 450 linhagem CP especializada em carne exótica e ornamental, com valor de aproximadamente -0,5 por volta dos 330 dias, e o maior valor observado foi para o animal 160 linhagem HÁ com EBV de aproximadamente 1,55 aos 120 dias. O Top 10 é composto por animais das linhagens HA, que foi a que mais se destacou aparecendo em 4 posições do ranking dos 10 (1ª, 4ª, 5ª e 8ª), 7P que apareceu na 2ª colocação, PL na 3ª, AZ na 6ª, AO nas 9ª e 10ª e CP na 7ª posição. Para todas elas o comportamento para o valor de EBV foi decrescente até por volta de 300 dias, quando então começou a crescer novamente, sendo os valores mais altos de EBV estimados para as idades iniciais de postura.

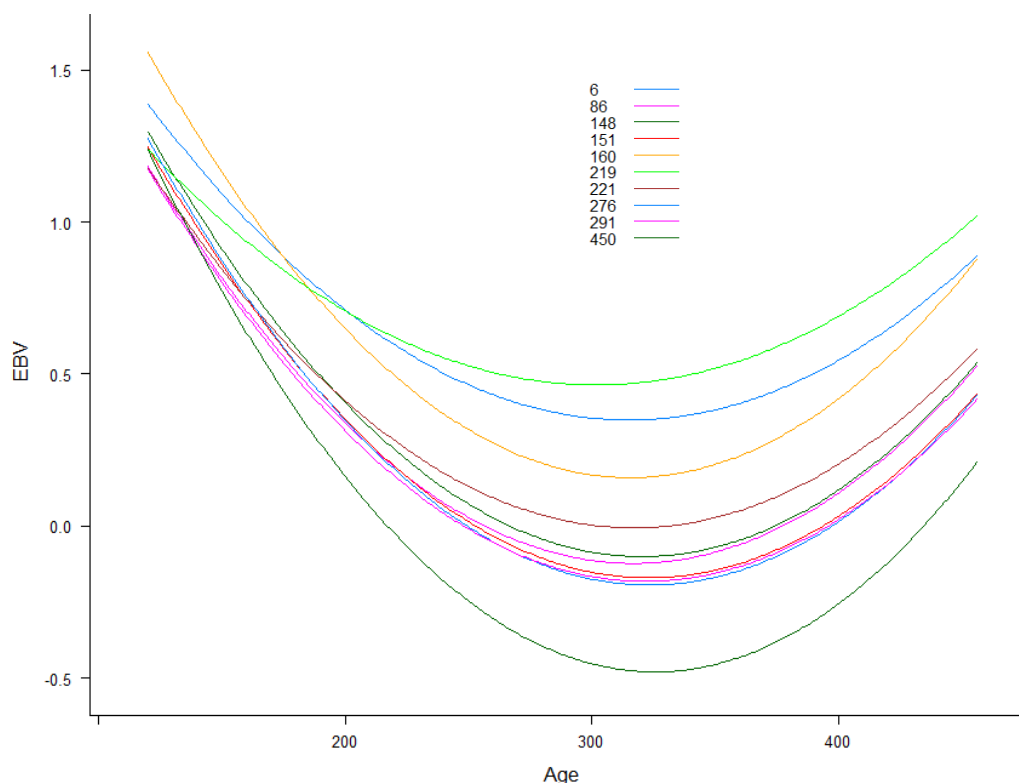


Figura 5: Top 10 melhores animais representados por cada linha do gráfico (6=7P; 86, 151, 160 e 276=HA, 148=PL; 219=AZ; 221 e 291=AO e 450=CP) segundo os valores genéticos aditivos estimados (EBV) durante a trajetória de idades de 120 a 456 dias.

2.7.3 Análise factorial exploratória

Os valores dos componentes de variância da análise univariada para obter os EBVs de cada animal para posteriormente serem usados na AFE são apresentados na Tabela 3. O maior EBV encontrado foi para a característica IPP, seguido por Prod_ovos, respectivamente com 0,12 e 0,05; enquanto que o menor valor foi para a variável PE_GEMA, com -0,0002.

As variáveis Prod_ovos e IPP também foram as que apresentaram as maiores variâncias genética aditiva (171,7 e 58,7) e maiores variâncias residuais (1073,6 e 563,8), assim como os maiores desvios para esses mesmos componentes de variância, porém quanto à herdabilidade, a característica com maior valor foi ALT_ALB com 0,54 o que é considerado uma h^2 alta, e muitos autores concordam que valores abaixo de 0,2 são baixos, entre 0,2 e 0,4 medianos e acima de 0,4 são altos. As h^2 variaram entre todas as características de valores próximos a 0 para a variável PE_GEMA (0,0057) até quase 0,6, mas no geral, três variáveis ficaram no intervalo de baixas h^2 (Prod_ovos, IPP e PE_GEMA), 3 ficaram no intervalo de h^2 medianas (PEOVO, ESP_CA e PE_CASCA) e as outras duas ficaram no intervalo de h^2 altas (ALT_ALB e HU).

O maior valor de CV% foi para a característica Prod_ovos, onde pode observar que teve como valor máximo de produção no período de 120 a 456 dias 243 ovos e como valor mínimo 2 ovos, o que explica neste caso o porquê do maior CV encontrado.

A variável HU, que é uma variável composta a partir de outras duas variáveis (PEOVO e ALT_ALB), é uma medida que indica a qualidade do ovo com relação ao quanto suas proteínas já foram ou não degradadas, pois

quanto mais velho for o ovo, menor será a altura de albúmen e mais leve o ovo fica, devido a degradação pelo tempo da albumina. Podemos observar que na média os valores de HU foram bons (84,3), pois quanto mais próximo de 100, melhor indicativo de que o ovo está novo e bem conservado, e olhando para os valores máximos, algumas medidas chegaram muito próximo do valor máximo de HU.

Tabela 3: Estatística descritiva, componentes de variância e valores fenotípicos para as características de produção e qualidade de ovos das linhagens AZ, HA, AO e BA, resultado da saída da análise univariada.

Resultados	Variáveis							
	Prod_ovos (n°)	IPP (dias)	PE_OVO (g)	Alt_alb (mm)	HU	Esp_ca (mm)	Pe_gema (g)	Pe_casca (g)
Média dos Fenótipos	130,4±51,7	170,2±29,9	49,4±5,2	6,68±0,97	84,3±5,92	0,35±0,03	14,1±1,8	4,87±0,57
Fenótipo Máximo	243	265	70,8	9,3	96,5	0,45	24,8	6,94
Fenótipo Mínimo	2	125	35,6	2,5	39,5	0,25	8,68	2,99
CV (%) do Fenótipo	39,6	17,6	10,5	14,6	7,02	8,36	12,8	11,7
Variância Genética Aditiva (VGA)	171,7	58,7	8,46	0,5	16,3	0,00023	0,014	0,096
Erro Padrão da VGA	143,8	67,8	2,92	0,14	5,07	0,0001	0,22	0,036
Variância Residual (VR)	1073,6	563,8	14,6	0,43	17,8	0,00055	2,45	0,19
Erro Padrão da VR	147,5	71,7	2,68	0,12	4,45	0,00009	0,26	0,034
h²	0,14	0,09	0,37	0,54	0,48	0,3	0,0057	0,34
Médias dos EBVs	0,05	0,12	0,03	0,014	0,086	0,00005	-0,0002	0,0017

Os resultados dos autovalores, soma de rotação de cargas quadradas, variâncias explicadas e acumuladas com dados de EBV são apresentados na Tabela 4. Os dois componentes apresentados, tiveram valores maiores que 1, respectivamente 2,86 e 1,56; o que significa que eles conseguem explicar a variação das variáveis deste estudo, que são produção de ovos (Prod_ovos), idade ao primeiro ovo (IPP), altura do albúmen (ALT_ALB), espessura da casca (ESP_CA), peso da gema (PE_GEMA), peso da casca (PE_CASCA) e peso de ovo (PEOVO). A variável HU não foi considerada nesta análise porque ela apresentou uma alta correlação com as variáveis PEOVO e ALT_ALB, sendo que estas são utilizadas para o cálculo de seus valores, pois HU é uma variável composta pelas informações dessa duas mencionadas, assim, optamos por deixar PEOVO e ALT_ALB e remover HU.

O valor de porcentagem de variâncias para o primeiro componente foi de 40,81% e para o segundo de 22,25%, com porcentagem de variância cumulativa para os dois fatores de 63,06%, que diz que esses dois componentes juntos conseguem explicar esse valor das variações originais dos dados das características obtidas neste estudo. Para o componente 1, as variáveis mais importantes foram ESP_CA, PE_GEMA, PE_CASCA e PEOVO, com valores na matriz de componentes rotacionada de 0,8; 0,71; 0,95 e 0,8 respectivamente, que foram os valores acima de 0,5, enquanto que para o componente 2, as variáveis mais importantes foram Prod_ovos e IPP, com -0,81 e 0,7 respectivamente, representa os valores que foram maiores ou igual a 0,5 ou menos de -0,5. A única variável que não foi considerada relevante em nenhum dos dois fatores foi ALT_ALB, com valores de 0,33 e 0,38, respectivamente para os componentes 1 e 2, pois ficou no intervalo de 0,5 a -0,5.

Quando observamos os valores de comunalidade (Tabela 4), que quanto mais próximos de 1 melhor (RESENDE; SILVA; AZEVEDO, 2014), pois isso indica que os 2 fatores explicam uma grande porcentagem de variação de cada variável, podemos ver que a variável ALT_ALB, que foi a única que não foi importante pra nenhum dos dois componentes, também foi a que apresentou o menor valor de comunalidade (0,26), enquanto que as outras ficaram todas acima de 0,5, com destaque para PE_CASCA que foi a que apresentou o maior valor (0,95).

Na AFE usualmente valores de comunalidade considerados aceitáveis devem estar acima de 0,5 (SCHAWB, 2007), de modo que valores abaixo disso indicam que a variável em questão pode ser eliminada da amostra. Também, valores de comunalidade baixos entre grupos de variáveis mostram que elas estão linearmente não correlacionadas, reforçando o porquê de serem removidas das AFE, como aconteceu com a variável ALT_ALB (HOGARTY et al., 2005; JUNG; LEE, 2011).

Tabela 4. Autovalores iniciais, somas de rotação das cargas quadradas¹, a porcentagem de variância explicada e as cargas fatoriais dos dois primeiros fatores para os valores genéticos preditos da produção e qualidade do ovo de galinhas caipiras criadas em condições tropicais.

EBV						
Componentes ²	Autovalores Iniciais			Somadas de rotação das cargas quadradas		
	Total	% de Variância	Cumulativo %	Total	% de Variância	Cumulativo %
1	3,28	46,91	46,91	2,86	40,81	40,81
2	1,13	16,15	63,06	1,56	22,25	63,06
Variáveis (valores genéticos preditos)	Matriz de Componentes Rotacionados		Comunalidades			
	1	2	Inicial	Extração		
Produção de ovos	0,04	-0,81	1	0,66		
Idade ao primeiro ovo	0,24	0,71	1	0,57		
Altura do albúmen	0,33	0,38	1	0,26		
Espessura da casca	0,8	-0,04	1	0,64		
Peso da gema	0,71	0,25	1	0,57		
Peso da casca	0,95	0,15	1	0,92		
Peso do ovo	0,8	0,39	1	0,79		

¹Pesos dos autovetores $\geq 0,5$ and $\leq -0,5$ foram colocados em negrito.

²O número de componentes foi definido com a regra de Kaiser (1969; autovalor > 1)

Os resultados dos autovalores, soma de rotação de cargas quadradas, variâncias explicadas e acumuladas com dados fenotípicos são apresentados na Tabela 5. Com dados de fenótipo, foram gerados três componentes que conseguiram explicar a variação das variáveis originais deste estudo, com valores acima de um, respectivamente 2,1; 1,5 e 1,03. Assim como na análise com EBVs, nessa com dados de fenótipo também não foi considerado a variável HU, pela mesma razão de ela ser composta por outras duas variáveis já utilizadas na análise (PEOVO e ALT_ALB) e devido à alta correlação entre elas. Para o componente 1, as variáveis mais importantes, com valores na matriz de componentes rotacionados acima de 0,5; foram ALT_ALB, PE_CASCA e PEOVO, para o componente 2 foram PROD_OVOS e IPP, sendo elas correlacionadas inversamente da mesma forma que ocorreu com dados de EBV, e o componente 3 foi composto pelas variáveis ESP_CA, PE_GEMA e PE_CASCA, sendo esta última, comum com o componente 1.

O cumulativo da variância total explicada pelos 3 componentes foi 65,96%, valor um pouco maior que a explicação da variação das variáveis utilizando dados de EBV, que foi de 63,06%. Diferente da análise com EBVs, aqui nesta análise com fenótipos a variável ALT_ALB teve importância e fez parte da composição de um dos componentes.

Apesar de fazer parte do componente 3, a variável ESP_CASCA apresentou um valor de comunalidade baixo (0,36), e sabe-se que quanto mais próximo de um, melhor é esse valor (RESENDE; SILVA; AZEVEDO,

2014); e que valores aceitáveis são aqueles acima de 0,5 (SCHAWB, 2007), como já explicado anteriormente, portanto, essa variável deveria ser removida do respectivo componente, assim como ocorreu com a variável ALT_ALB nos dados de EBV. Isso mostra que quando trabalhamos com dados de EBV comparado com dados de fenótipo, os resultados são diferentes, quanto a relevância das variáveis, número de componentes e variáveis que fazem parte de cada um deles.

Tabela 5. Autovalores iniciais, somas de rotação das cargas quadradas¹, a porcentagem de variância explicada e as cargas fatoriais dos dois primeiros fatores para os valores fenotípicos das variáveis de produção e qualidade do ovo de galinhas caipiras criadas em condições tropicais.

Valores Fenotípicos						
Componentes ²	Autovalores Iniciais			Somadas de rotação das cargas quadradas		
	Total	% de Variância	Cumulativo %	Total	% de Variância	Cumulativo %
1	2,39	34,18	34,18	2,1	29,98	29,98
2	1,20	17,09	51,27	1,5	21,27	51,25
3	1,03	14,69	65,96	1,03	14,71	65,96
Variáveis (valores genéticos preditos)	Matriz de Componentes Rotacionados			Comunalidades		
	1	2	3	Inicial	Extração	
Produção de ovos	0,029	-0,87	-0,041	1	0,75	
Idade ao primeiro ovo	0,28	0,78	-0,019	1	0,68	
Altura do albúmen	0,70	-0,096	0,25	1	0,57	
Espessura da casca	0,075	0,095	-0,59	1	0,36	
Peso da gema	0,12	0,16	0,77	1	0,64	
Peso da casca	0,85	0,18	-0,59	1	0,77	
Peso do ovo	0,89	0,24	-0,067	1	0,86	

¹Pesos dos autovetores $\geq 0,5$ and $\leq -0,5$ foram colocados em negrito.

²O número de componentes foi definido com a regra de Kaiser (1969; autovalor > 1)

Quando comparamos a AFE com dados fenotípicos com a AFE com dados de EBV visualmente (Figura 6), podemos perceber que utilizando dados de fenótipo (Figura 6A), as variáveis ALT_ALB, ESP_CASCA e PE_GEMA foram as menos importantes, por apresentaram os valores menores de comunalidades, respectivamente 0,57; 0,36 e 0,64, porém, baseado nestes valores, ALT_ALB e PE_GEMA ainda assim apresentaram valores aceitáveis para comporem os componentes da análise.

A variável IPP tanto com dados de fenótipo como de EBV ficou mais distante do grupo das outras variáveis, assim como ALT_ALB para os dados de EBV, que ficou entre IPP e as outras, considerando apenas o quadrante positivo do gráfico. Isso mostra de forma visual o que foi apresentado na Tabela 4, onde a característica ALT_ALB não foi importante na composição dos fatores com dados de EBV.

Apesar de a variável IPP apresentar correlação negativa com Prod_ovos, neste caso isso é favorável, pois significa que ao selecionar animais para aumento de produtividade, será selecionado conseqüentemente para diminuição da idade ao primeiro ovo, o que indica uma melhoria na precocidade produtiva dessas aves caipiras, e isso explica de forma visual e prática o porquê dessas duas variáveis ficarem juntas na composição do fator 2 em ambas análises e as outras ficaram juntas nos outros fatores.

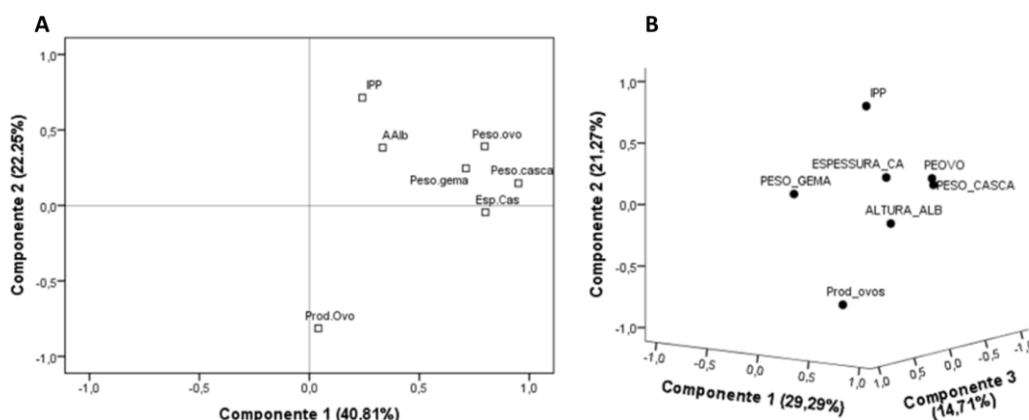


Figure 6. Análise fatorial exploratória utilizando os valores genéticos preditos (EBV) (Imagem A) e os fenótipos (Imagem B) para as variáveis de produção e qualidade do ovo de galinhas caipiras criadas em condições tropicais.

2.8 Discussão

Utilizando o mesmo critério de seleção do melhor modelo (AIC) os autores Wolc & Szwaczkowski (2009) também tiveram melhor ajuste aos dados de produção de ovos com o modelo utilizando polinômios de Legendre de segunda ordem ou quadrático para os efeitos fixos, porém para os efeitos aleatórios de variância genética aditiva e de ambiente permanente, o melhor modelo que encontraram foi o de polinômios de Legendre de terceira ordem ou cúbico. Por outro lado, Wolc et al. (2013), assim como neste estudo, teve o melhor ajuste aos dados com o modelo quadrático para as variâncias genética aditiva e de ambiente permanente, mas para o efeito fixo teve melhor ajuste com polinômios de 5ª ordem.

Assumindo a suposição de homogeneidade de variâncias residuais em seus estudos com aves de postura com dados fenotípicos e utilizando pedigree na sua modelagem, para estimar parâmetros genéticos para a característica produção de ovos utilizando RRM, Wolc & Szwaczkowski (2009) concluíram em seu estudo que a inclusão de variâncias residuais heterogêneas teve pouca influência nos valores das estimativas de parâmetros genéticos, porém relataram que o tempo de computação por interação e a convergência foi mais lento. Em estudos com produção de ovos de codornas japonesas utilizando RRM (KARAMI et al., 2017) e com aves nativas iranianas (TABRIZI et al., 2015), esses autores também modelaram assumindo a suposição de homogeneidade de variâncias residuais em seus estudos, enquanto que utilizando dados genômicos Wolc et al. (2013) assumiu uma heterogeneidade nas variâncias residuais.

As variâncias genéticas aditivas estimadas por Wolc et al. (2013) ficaram torno de 0,1 ao longo do tempo, quando se considerou apenas dados de pedigree por um modelo utilizando polinômios de Legendre de segunda ordem, com dados genômicos essa variância ficou em torno de 0,2; com pico em 0,3 aproximadamente. Esses valores também foram observados em algum momento da curva obtida neste estudo, mas, diferente dos resultados desses autores, a curva que foi observada nesta análise teve um comportamento decrescente ao longo da trajetória

mais pronunciada e valores que ficaram acima de 0,8. Para codornas japonesas Abou Khadiga et al. (2017) obteve valores de variância genética aditiva bem maiores, com observações ao longo do tempo indo de 3,42 até 5,47.

Para variância de ambiente permanente, o menor valor encontrado por Wolc et al. (2013) com dados de pedigree foi de aproximadamente 0,1 e o maior chegando a quase 0,4; e com dados genômicos os valores mínimo e máximo foram iguais, porém ao longo da curva os valores para o modelo com dados de pedigree foram maiores e sempre crescentes, enquanto que para os dados genômicos teve um período de várias semanas que ele ficou em torno de 0,1. Esses comportamentos dos dados de produção de ovos ao longo do tempo diferem dos deste estudo, que foi decrescente e especialmente nas idades iniciais apresentou valores de variância bem mais altos, chegando a mais de 1,3. Todos esses valores ficaram abaixo dos estimados por Abou Khadiga et al. (2017), que obtiveram resultados ao longo da curva de 7,24 a 4,38, apresentando comportamento decrescente no período analisado.

Trabalhando com RRM utilizando polinômios de 4ª ordem como o modelo melhor ajustado para dados de produção de ovos de galinhas, para estimação da variância de ambiente permanente, Luo et al. (2007) obteve todos os valores ao longo do tempo superior ao de 0,5; sendo seu menor valor 0,53 e o maior 3,64; com comportamento crescente ao longo do tempo das medidas. Seus valores para variância genética aditiva, agora utilizando polinômios de Legendre de 2ª ordem, variaram de 0,15 a 2,96 ao longo da trajetória, assim em ambos os modelos eles obtiveram picos maiores que o deste estudo, e mesmo para as h^2 , também para o modelo de 2ª ordem, seu menor valor foi 0,16; próximo da maior h^2 deste trabalho, e com comportamento crescente, as estimativas de h^2 deles chegaram até seu pico de 0,54 quase no final do período, quando teve um leve decréscimo para 0,43; enquanto que o maior valor de h^2 deste estudo não chegou nem a 0,2.

Os valores de h^2 estimados neste estudo foram baixos, não passando de 0,2 durante todo o período de coleta nas idades de 153 a 468 dias, algo que pode ocorrer quando se trata de características reprodutivas, que geralmente sofrem maior influência do ambiente em sua expressão fenotípica, porém, alguns autores encontraram valores de h^2 de 0,3 ou maiores, como Wolc & Szwaczkowski (2009), que estudaram três linhagens de poedeiras com modelos de polinômios de Legendre de até oitava ordem, e no seu estudo, esses valores de h^2 mais altos foram os valores estimados nas coletas iniciais e finais, sendo que o comportamento recorrente em todos os modelos que eles estudaram foi o mesmo, onde a herdabilidade começava alta e ia decrescendo e depois voltava a crescer como também aconteceu neste trabalho. Para o modelo de polinômios de segunda ordem Wolc & Szwaczkowski (2009) obtiveram valores de h^2 variando de próximo de 0 até valores em torno de 0,1 e, valores menores que os das aves caipiras deste estudo, e mesmo para os outros modelos com ordens superiores, eles observaram em algum momento da trajetória das curvas geradas pelas medidas que eles coletaram valores de h^2 abaixo, em torno e acima de 0,1.

Considerando dados de fenótipo e pedigree na análise e modelo de polinômio de 2ª ordem, Wolc et al. (2013) observaram valores de h^2 em torno de 0,1 no início da postura e ao longo do tempo esse valor foi aumentado para 0,2 e um pouco cima em alguns momentos, comportamento crescente diferente do observado na trajetória deste estudo, porém quando os mesmos autores utilizaram dados genômicos na análise, os valores de h^2 em alguns pontos da trajetória chegaram a 0,5; bem acima do maior valor de quase 0,2 deste trabalho, e o comportamento foi crescente no início e decrescente no final do período analisado por eles. A utilização de informações genômicas ajuda a melhorar a precisão da estimativa dos componentes de variância em comparação com abordagens baseadas apenas no uso de informações fenotípicas e de pedigree, porém elas aumentam os custos para sua implementação e demandam maior poder computacional.

Em estudo com produção de ovos de perus utilizando RRM com polinômios de Legendre de 2ª ordem, Kranis et al. (2007) estimaram valores de h^2 variando de 0,07 a 0,12 ao longo do período que eles analisaram as

medidas, porém o comportamento ao longo da curva para os valores de h^2 que eles observaram foi diferente, sendo sempre decrescente ao longo do tempo. Na trajetória das linhagens caipiras desta análise o comportamento voltar a ser crescente depois dos 250 dias de idade e de maneira ainda mais pronunciada a partir dos 400 dias.

Também apresentando comportamento ao longo da curva decrescente para os valores de h^2 , mas agora no estudo de produção de ovos de codornas japonesas, utilizando o modelo com polinômios de Legendre de 2ª ordem para efeitos fixos e de 3ª ordem para efeitos de variância genética aditiva e de ambiente permanente, Abou Khadiga et al. (2017) obteve ao longo da curva de produção de ovos o intervalo de 0,1 a 0,17 de h^2 , sendo que seus valores mais altos foram bem parecidos aos encontrados neste trabalho.

O que pode explicar esse padrão resposta dos valores de herdabilidade sendo maiores no início da postura, decrescendo ao longo da curva e voltando a crescer nas semanas mais tardias de postura, é que tanto no início como nas semanas finais de postura, ou tanto quando as aves são mais jovens e quando são mais velhas, a variabilidade na postura tende a ser maior, enquanto que nas semanas intermediárias tende a ser mais homogênea. Essa maior variabilidade nesses dois períodos, que também pode ser vista nos resultados do padrão resposta das curvas de variância genética aditiva e de ambiente permanente, corroboram com esse resultado, o que pode explicar a trajetória obtida para os valores de herdabilidade deste estudo.

Comparando os valores de EBV obtidos para a característica Prod_ovos na RRM com o valor de EBV para Prod_ovos da AFE, notou-se que na AFE o valor ficou bem próximo de zero (0,0487), um valor mediano dentro do intervalo de valores de EBV da RRM, que variou de -1,635 à 1,558.

Wolc et al. (2011) estimaram EBVs para características de aves de postura baseado em dados de pedigree (PBLUP) e de marcadores (GBLUP) de uma linhagem comercial de galinha poedeira, em dois cenários, um precoce (26-28 semanas de vida) e outro tardio (42-46 semanas de vida). Baseado nos marcadores, eles encontraram que a capacidade preditiva foi maior para se estimar os EBVs do que quando comparado com o método que utilizou relações de pedigree, para ambos os cenários considerados em seu estudo, embora a vantagem de métodos de seleção com marcadores seja maior em idades mais jovens dos animais. Uma das explicações para este resultado é a de que o EBV baseado no GBLUP não é afetado por erros de pedigree, porém, eles não estão isentos de serem afetados por erros de genotipagem ou até mesmo erros na identificação das amostras de DNA. Isso mostra que na possibilidade de se trabalhar com dados de marcadores ao invés de apenas dados de pedigree, a precisão obtida e os valores estimados serão melhores e mais acurados, tanto para uma seleção genômica precoce quanto tardia.

Para todas as características, os EBVs estimados por WOLC et al. (2011) utilizando dados de pedigree foram superiores a 0,15; sendo a variável Prod_ovos a que obteve menor valor de EBV na seleção precoce e Qualidade de Casca na seleção tardia, enquanto que neste estudo o menor valor de EBV foi para a característica ESP_CA. Já a variável com maior EBV que eles estimaram foi ALT_ALB com aproximadamente 0,5 de valor na seleção precoce e 0,6 na tardia, e neste estudo foi a variável IPP com 0,1239.

A maior h^2 encontrada no intervalo de 120 a 456 dias na RRM para Prod_ovos foi de 0,1771 na idade inicial, enquanto que na análise univariada prévia a AFE, o valor de h^2 foi de 0,1378, valor este maior do que qualquer valor de h^2 para qualquer idade da RRM considerando as idades a partir de aproximadamente 150 dias. Venturini et al. (2013) em estudo com poedeiras da linhagem White Leghorn encontrou valores de h^2 no intervalo de 17 a 70 semanas de vida das aves de $0,22 \pm 0,06$ à $0,25 \pm 0,06$; todos maiores que qualquer resultado deste estudo, porém, comparando com as outras características aqui estudadas, tanto PEOVO, como ALT_ALB, HU, ESP_CA e PE_CASCA tiveram valores de h^2 superiores a 0,3; superando os maiores valores da variável Prod_ovos de ambos os estudos.

Na literatura, apesar de se encontrar valores de h^2 baixos (abaixo de 0.2) e moderados para a característica Prod_ovos (entre 0,2 e 0,4), em boa parte dos estudos os valores de h^2 costumam ser de moderados a alto (acima de 0,4), tanto em estudos com produção de ovos por gaiola com mais de um animal (BISCARINI et al., 2010) como em estudos com produção de ovos individual (MISZTAL; BESBES, 2000; NURGIARTININGSH et al., 2004; SINGH et al., 2018).

A IPP também é uma característica bastante relevante em programas de seleção, pois quanto mais cedo uma ave começar a sua postura, por mais tempo ela vai produzir ovos muito provavelmente. Para essa característica o valor de h^2 estimado foi baixo (0,0943), e estudos anteriores obtiveram valores estimados de herdabilidade moderado ($0,22 \pm 0,03$) para perus (BEGLI et al., 2019) e moderados a alto para frangos de corte variando de 0,37 à 0,46 (TONGSIRI et al., 2015). Em seu estudo com perus Begli et al. (2019) registrou uma correlação negativa entre a característica IPP e número de ovos (Prod_ovos), resultado também obtido neste estudo a partir da AFE, pois é natural que aves precoces venham a produzir mais ovos durante sua vida.

A produção de ovos é ambientalmente e geneticamente correlacionada negativamente com a idade na maturidade sexual, porém, apesar de em curto prazo ser interessante selecionar geneticamente aves de postura para idades a primeira postura mais precoce, o que vai levar a uma seleção conjunta para o aumento da produção de ovos, a longo prazo isso pode levar a efeitos fisiológicos prejudiciais segundo Kamali et al. (2007), assim a seleção para essa característica precisa ser melhor investigada antes de ser utilizada em programas de seleção e melhoramento genético de poedeiras. Além da correlação negativa entre IPP e Prod_ovos, Kamali et al. (2007) também obteve correlações negativas entre Prod_ovos e peso corporal da ave, indicando que ao selecionar aves mais pesadas, a característica de produção de ovos ficará prejudicada. Esse resultado mostra a importância de em um próximo estudo, considerar também na AFE a característica peso corporal, pois provavelmente ela terá alguma importância ao se selecionar aves para postura.

A AFE ainda é muito pouco estudada em trabalhos com aves de postura, especialmente quando se utiliza dados de EBV, a maior parte dos estudos tem sido feito para avaliar preferências de pessoas por tipos de alimentos (NRUMMER et al., 2018), a satisfação de pessoas por algo (ISKAMTO et al., 2020), em pesquisas marítimas (MASKEY et al., 2018), na área da psicologia (NORRIS; LECAVALIER, 2010) etc.

Alguns estudos mais recentes começaram a aplicar a AFE em estudos com animais, como Yu et al. (2020) que utilizou a AFE para identificar variáveis latentes para explicar a variância de vários tipos de medidas subjetivas utilizadas para avaliar o temperamento de bezerros em estudos de comportamento animal, onde identificaram dois fatores (difícil e calmo) que dentro de cada um reuniram variáveis relacionadas a temperamentos indesejáveis e desejáveis respectivamente dos animais. Esse resultado vai de encontro ao que foi obtido neste estudo, pois os dois fatores identificados aqui também reuniram variáveis originais com características de mesmo tipo, sendo o primeiro fator constituído por IPP e Prod_Ovos para os dados de EBV, duas variáveis relacionadas à produtividade e podem ser selecionadas em conjunto, enquanto que o segundo fator reuniu as variáveis ESP_CA, PE_GEMA, PE_CASCA e PEOVO, que com exceção de ESP_CA são todas variáveis de peso de alguma parte específica do ovo.

Os autores Paiva et al. (2020) também encontraram um resultado positivo para o uso da AFE em seu estudo com frangos de corte, identificando quatro fatores em que cada um deles reunia um grupo de variáveis correlacionadas, que juntos explicaram 71% da variância original dos dados, valor um pouco maior que os 63,06% encontrados para os dois fatores deste estudo com dados de EBV e 65,96% para os dados de fenótipos. Todas as variáveis reunidas no primeiro fator identificado por esses autores eram características de carcaça e peso corporal, no segundo fator características relacionadas a retenção de água e cisalhamento, no terceiro características do aspecto da

carne quanto a cor e no quarto características de vísceras, mostrando que cada grupo de variáveis dentro de cada fator possui semelhanças e/ou correlação, fazendo com que cada fator seja independente um do outro, o que é esperado em uma AFE.

Comparando a AFE com EBVs e com dados de fenótipo no sentido de qual se adequou melhor ao conjunto de variáveis deste estudo, pode notar que na AFE com valores fenotípicos das características, os fatores 1 e 3 possuem uma variável em comum (PE_CASCA), e isso vai fazer com que exista algum tipo de correlação entre esses fatores, não atendendo a premissa de que cada fator deve ser independente do outro (TIMM, 2002; BISHOP; FIENBERG; HLLAND, 2007; CHATFIELD; COLLINS, 2018). Na AFE com valores de EBV isso não aconteceu, sendo que o fator 1 reuniu as variáveis que compõem os fatores 1 e 3 da AFE com valores fenotípicos (com exceção da característica ALT_ALB que não foi considerada importante a partir dos dados genéticos), ou seja, utilizando EBVs, a análise conseguiu separar e determinar cada fator de modo que cada um seja independente do outro, se adequando melhor a esta análise com este conjunto de variáveis.

Em ambas as análises com fenótipo e EBV, podemos perceber que a característica PROD_OVOS teve uma correlação negativa com as outras variáveis, o que indica que ao selecionar apenas essa característica dentro do programa de melhoramento genético, conseqüentemente estaremos diminuindo ou selecionando inversamente as outras características estudadas, e o contrário também é verdadeiro, de forma que para se obter um ganho genético para todas elas, será necessário selecionar todas elas levando-se em consideração a PROD_OVOS.

Agora quando selecionamos, por exemplo, a característica PE_CASCA na análise com EBVs, podemos perceber que ela tem correlação positiva com outras variáveis como PEOVO, ESP_CA e PE_GEMA, o que mostra que ao selecionar uma dessas características, conseqüentemente estaremos selecionando as outras, algo que facilitaria o programa de melhoramento genético.

2.9 Conclusão

O modelo de regressão aleatória utilizando polinômios de Legendre de 2ª ordem se ajustou bem aos dados de produção de ovos de galinhas caipiras, descrevendo bem a curva de produção ao longo das idades de 120 a 456 dias das aves. Dentre todas as linhagens presentes no estudo, a linha HA foi a que apresentou os melhores valores genéticos estimados. Ter modelado a característica de produção de ovos ao longo das idades de 120 a 456 dias assumindo a homogeneidade da variância residual não afetou os resultados obtidos pelo modelo utilizado neste estudo. A análise fatorial exploratória utilizando dados de valores genéticos estimados para as características de produção de qualidade de ovos com dois fatores apresentou uma melhor explicação das porcentagens das variâncias que ocorrem nas sete variáveis originais deste estudo e separou melhor os fatores e suas respectivas composições. A variável altura de albúmen não apresentou importância neste estudo na composição de nenhum dos fatores com dados de EBV. No final da AFE com EBVs, das 7 variáveis iniciais, apenas 6 foram importantes para compor os dois fatores, sendo elas produção de ovos, idade ao primeiro ovo, espessura de casca, peso da gema, peso da casca e peso do ovo. Como o total da variância explicada por essa AFE foi de 63,06 com os dois fatores, em estudos futuros é interessante que se adicionem mais variáveis e se identifique novamente as variáveis latentes a partir desse novo conjunto de variáveis originais, a fim de se aumentar essa variância total explicada da AFE.

Referências

- ABOU KHADIGA, G. et al. Genetic analysis of partial egg production records in Japanese quail using random regression models. *Poultry Science*, v. 96, n. 8, p. 2569-2575, 2017.
- ALFARO, I. I.; DÍAZ, J. O.; GARCÍA, F. J. A. Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema*, 2014.
- ALTIERI, M. A.; NICHOLLS, C. I. Agroecology: Challenges and opportunities for farming in the Anthropocene. *Ciencia e investigación agraria: revista latinoamericana de ciencias de la agricultura*, v. 47, n. 3, p. 204-215, 2020.
- BAILEY, R. R.; SRINATH, M. Orthogonal moment features for use with parametric and non-parametric classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 18, n. 4, p. 389-399, 1996.
- Begli, H. E. et al. Genetic parameters for clutch and broodiness traits in turkeys (*Meleagris gallopavo*) and their relationship with body weight and egg production. *Poultry Science*, v. 98, n. 12, p. 6263-6269, 2019.
- BERKHOFF, J. et al. Consumer preferences and sensory characteristics of eggs from family farms. *Poultry science*, v. 99, n. 11, p. 6239-6246, 2020.
- BIGNARDI, A. B. et al. Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. *Genetics and Molecular Research*, p. 3565-3575, 2011.
- BISCARINI, F. et al. Estimation of heritability and breeding values for early egg production in laying hens from pooled data. *Poultry science*, v. 89, n. 9, p. 1842-1849, 2010.
- BISHOP, Y. M.; FIENBERG, S. E.; HOLLAND, P. W. *Discrete multivariate analysis: theory and practice*. Springer Science & Business Media, 2007.
- BÖLLING, D. et al. Genetic improvement of livestock for organic farming systems. *Livestock production science*, v. 80, n. 1-2, p. 79-88, 2003.
- BRÜMMER, N.; PETERSEN, W.; CHRISTOPH-SCHULZ, I. Consumer acceptance of dual-purpose chickens a mixed methods approach. *German Journal of Agricultural Economics*, v. 67, n. 670-2021-585, p. 234-245, 2018.
- CHATFIELD, C.; COLLINS, A. J. *Introduction to multivariate analysis*. Routledge, 2018.
- DIKMEN, B. Y. et al. Egg production and welfare of laying hens kept in different housing systems (conventional, enriched cage, and free range). *Poultry science*, v. 95, n. 7, p. 1564-1572, 2016.
- FARUQUE, S. et al. Evaluation of the performance of native chicken and estimation of heritability for body weight. *Journal of Bangladesh Academy of Sciences*, v. 37, n. 1, p. 93-101, 2013.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAOSTAT). *Livestock Primary*. 2020. Disponível em: <<http://www.fao.org/faostat/en/#data/QL>>. Acesso em: 26 mar. 2022.
- GARRICK, D. J.; GOLDEN, B. L. Producing and using genetic evaluations in the United States beef industry of today. *Journal of Animal Science*, v. 87, n. suppl_14, p. E11-E18, 2009.
- HAIR JR., J. F.; WILLIAM, B.; BABIN, B.; ANDERSON, R. E. *Multivariate data analysis*. 6.ed. Porto Alegre: Bookman, 2009.
- HELLMEISTER FILHO, P. et al. Efeito de genótipo e do sistema de criação sobre o desempenho de frangos tipo caipira. *Revista Brasileira de Zootecnia*, v. 32, p. 1883-1889, 2003.
- HENDERSON, C. R. *Applications of linear models in animal breeding*. 423 p. University of Guelph, Guelph, Ontario, 1984.
- HILL, W. G.; BROTHERSTONE, S. Advances in methodology for utilizing sequential records. *BSAP Occasional Publication*, v. 24, p. 55-61, 1999.

- HOGARTY, K. Y. et al. The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and psychological measurement*, v. 65, n. 2, p. 202-226, 2005.
- HOWARD, M. C. A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve?. *International Journal of Human-Computer Interaction*, v. 32, n. 1, p. 51-62, 2016.
- ISKAMTO, D. et al. Exploratory Factor Analysis (EFA) To Measure Entrepreneur Satisfaction. In: *The International Conference on Industrial Engineering and Operations Management*. Detroit, MI, USA. 2020. p. 9.
- JUNG, S.; LEE, S. Exploratory factor analysis for small samples. *Behavior research methods*, v. 43, n. 3, p. 701-709, 2011.
- KAISER, H. F. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, v.20, p.141-151, 1960.
- KAMALI, M. A. et al. Heritabilities and genetic correlations of economic traits in Iranian native fowl and estimated genetic trend and inbreeding coefficients. *British poultry science*, v. 48, n. 4, p. 443-448, 2007.
- KARAMI, K. et al. Genetic evaluation of weekly body weight in Japanese quail using random regression models. *British Poultry Science*, v. 58, n. 1, p. 13-18, 2017.
- KIRKPATRICK, Mark; HILL, William G.; THOMPSON, Robin. Estimating the covariance structure of traits during growth and ageing, illustrated with lactation in dairy cattle. *Genetics Research*, v. 64, n. 1, p. 57-69, 1994.
- KRANIS, A. et al. The application of random regression models in the genetic analysis of monthly egg production in turkeys and a comparison with alternative longitudinal models. *Poultry Science*, v. 86, n. 3, p. 470-475, 2007.
- LUO, P. T.; YANG, R. Q.; YANG, N. Estimation of genetic parameters for cumulative egg numbers in a broiler dam line by using a random regression model. *Poultry Science*, v. 86, n. 1, p. 30-36, 2007.
- LUSH, J. L. *The genetics of populations*. Ames: Iowa State University Press, 1994. 900p.
- MACCIOTTA, N. P. P. et al. Use of multivariate factor analysis to define new indicator variables for milk composition and coagulation properties in Brown Swiss cows. *Journal of Dairy Science*, v. 95, n. 12, p. 7346-7354, 2012.
- MASKEY, R.; FEI, J.; NGUYEN, H. Use of exploratory factor analysis in maritime research. *The Asian journal of shipping and logistics*, v. 34, n. 2, p. 91-111, 2018.
- MEYER, K.; HILL, W. G. Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by restricted maximum likelihood. *Livestock Production Science*, v. 47, n. 3, p. 185-200, 1997.
- MISZTAL, I.; BESBES, B. Estimates of parental-dominance and full-sib permanent environment variances in laying hens. *Animal Science*, v. 71, n. 3, p. 421-426, 2000.
- MISZTAL, I.; TSURUTA, S.; LOURENÇO, D.; MASUDA, Y. *Manual for BLUPF90 family of programs*. University of Georgia, Athens, USA, 2018.
- MOMEN, M. et al. A predictive assessment of genetic correlations between traits in chickens using markers. *Genetics Selection Evolution*, v. 49, n. 1, p. 1-14, 2017.
- NORRIS, M.; LECAVALIER, L. Evaluating the use of exploratory factor analysis in developmental disability psychological research. *Journal of autism and developmental disorders*, v. 40, n. 1, p. 8-20, 2010.
- NURGIARTININGSIH, V. M. A. et al. Estimation of genetic parameters based on individual and group mean records in laying hens. *British poultry science*, v. 45, n. 5, p. 604-610, 2004.
- PAIVA, J. T. et al. Genetic evaluation for latent variables derived from factor analysis in broilers. *British Poultry Science*, v. 61, n. 1, p. 3-9, 2020.

- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available in: <<https://www.R-project.org>> (Accessed on March 10, 2022).
- RESENDE, M. D. V. de; SILVA, F. F. e; AZEVEDO, C. F. Mathematical, Biometric and Computational Statistics: Mixed, Multivariate, Categorical and Generalized Models (REML/BLUP), Bayesian Inference, Random Regression, Genomic Selection, QTL-GWAS, Spatial and Temporal Statistics, Competition, Survival. 21. ed. Viçosa: Suprema Gráfica e Editora Ltda, 2014. 881 p.
- SCHAEFFER, L. R. Random regression models. Available in <http://animalbiosciences.uoguelph.ca/~lrs/BOOKS/rrmbook.pdf>, 2016.
- SCHAEFFER, L. R.; JAMROZIK, J. Random regression models: a longitudinal perspective. *Journal of Animal Breeding and Genetics*, v. 125, n. 3, p. 145-146, 2008.
- SCHWAB, A. J. Data analysis and computers II. Recuperado de http://www.utexas.edu/courses/schwab/sw388r7_spring_2007.
- SILVA, M. C. da et al. Innate immunity in chicken lines developed by EMBRAPA Suínos e Aves: antimicrobial activity of macrophages and serum. *Revista Brasileira de Zootecnia*, v. 50, 2021.
- SINGH, M. K. et al. Heritability estimates of adult body weight and egg production traits in indigenous Uttara chickens. *International Journal of Genetics*, ISSN, p. 0975-2862, 2018.
- TABRIZI, A. E.; TAHMOORESPUR, M.; NEJATI JAVAREMI, A. Exploring the Use of Random Regression Models with Legendre Polynomials to Analyze Clutch Size in Iranian Native Fowl. *Iranian Journal of Applied Animal Science*, v. 5, n. 2, p. 453-458, 2015.
- TIMM, Neil H. (Ed.). *Applied multivariate analysis*. New York, NY: Springer New York, 2002.
- TONGSIRI, S.; JEYARUBAN, M. G.; VAN DER WERF, J. H. J. Genetic parameters for egg production traits in purebred and hybrid chicken in a tropical environment. *British poultry science*, v. 56, n. 6, p. 613-620, 2015.
- VENTURINI, G. C. et al. Genetic parameters and principal component analysis for egg production from White Leghorn hens. *Poultry science*, v. 92, n. 9, p. 2283-2289, 2013.
- WATKINS, M. W. Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, v. 44, n. 3, p. 219-246, 2018.
- WOLC, A. et al. Analysis of egg production in layer chickens using a random regression model with genomic relationships. *Poultry science*, v. 92, n. 6, p. 1486-1491, 2013.
- WOLC, A. et al. Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model. *Genetics Selection Evolution*, v. 43, n. 1, p. 1-9, 2011.
- WOLC, A. et al. Heritability of egg production in laying hens under cumulative, multitrait and repeated measurement animal models. *Czech Journal of Animal Science*, v. 52, n. 8, p. 254, 2007.
- WOLC, A.; SZWACZKOWSKI, T. Estimation of genetic parameters for monthly egg production in laying hens based on random regression models. *Journal of Applied Genetics*, v. 50, n. 1, p. 41-46, 2009.
- YU, H. et al. Deciphering cattle temperament measures derived from a four-platform standing scale using genetic factor analytic modeling. *Frontiers in genetics*, v. 11, p. 599, 2020.