



**UNIVERSIDADE FEDERAL DE ALAGOAS
UNIDADE ACADÊMICA CENTRO DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA**



Análise multivariada para avaliar produção de leite, duração da lactação e intervalo de partos em búfalas

JAILTON DA SILVA BEZERRA JÚNIOR

Rio Largo

2016

JAILTON DA SILVA BEZERRA JÚNIOR

Análise multivariada para avaliar produção de leite, duração da lactação e intervalo de partos em búfalas

Dissertação apresentada ao Programa de Pós- Graduação em Zootecnia do Centro de Ciências Agrárias, da Universidade Federal de Alagoas, como requisito parcial para obtenção do título de Mestre em Zootecnia.

Orientadora: Prof^ª. Dr^ª. Angelina Bossi Fraga

Coorientadoras: Prof^ª. Dr^ª. Aline Zampar

Prof^ª. Dr^ª Fabiane de Lima Silva

Rio Largo

2016

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecário Responsável: Valter dos Santos Andrade

B574a Bezerra Júnior, Jailton da Silva.
Análise multivariada para avaliar produção de leite, duração da lactação e intervalo de partos em búfalas / Jailton da Silva Bezerra Júnior. – 2016.
36 f. : il., tabs.

Orientadora: Angelina Bossi Fraga.
Coorientadora: Aline Zampar.
Coorientadora: Fabiane de Lima Silva.
Dissertação (Mestrado em Zootecnia) – Universidade Federal de Alagoas. Centro de Ciências Agrárias. Programa de Pós-Graduação em Zootecnia, Rio Largo, 2016.

Bibliografia: f. 31-36.

1. Análise por agrupamento. 2. autovalores. 3. Análise de componentes principais. 4. Distância euclidiana. 5. Leite – Produção. 6. Búfalo. I. Título.

CDU:636.293.2

TERMO DE APROVAÇÃO

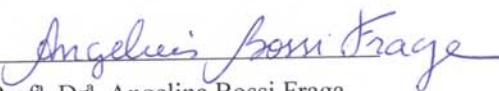
JAILTON DA SILVA BEZERRA

ANÁLISE MULTIVARIADA PARA AVALIAR PRODUÇÃO DE LEITE, DURAÇÃO DA LACTAÇÃO E INTERVALO DE PARTOS EM BÚFALAS"

Esta dissertação foi submetida a julgamento como parte dos requisitos necessários à obtenção do título de Mestre em Zootecnia, outorgado pela Universidade Federal de Alagoas.


A citação de qualquer trecho desta dissertação é permitida, desde que seja feita de conformidade com as normas da ética científica.

Aprovado em 16/02/2016




Prof^ª. Dr^ª. Angelina Bossi Fraga

Orientadora (UFAL)



Prof. Dr. Kuang Hongyu

Membro (UFMT)



Prof^ª. Dr^ª. Chiara Rodrigues Lopes

Membro UFAL-ARAPIRACA)

Rio Largo – AL

2016

A minha amada tia Solange Bezerra

AGRADECIMENTOS

Agradeço primeiramente a Deus por toda a força que tem me dado para que eu pudesse alcançar meus objetivos.

À minha orientadora Prof^a. Dr^a. Angelina Bossi Fraga, pelos ensinamentos, conselhos, paciência e dedicação para o meu desenvolvimento profissional e pessoal.

À Professora Aline Zampar, uma pessoa muito querida que deu todo seu apoio, incentivo e que sempre se colocou a disposição para me ajudar.

Aos Professores Dr. Kuang Hongyu e Dr^a Fabiane de Lima Silva, pela co-orientação durante a elaboração deste trabalho.

Ao Sr. Alberto Couto pela oportunidade de trabalhar com os dados do rebanho de búfalos da Fazenda Castanha Grande.

À CAPES pela concessão da bolsa de estudos.

À minha querida vó Maria Ester, por toda sua dedicação, carinho e amor.

Aos meus pais, Jailton e Margarete que apesar de todas as dificuldades sempre me apoiaram com bastante esforço para que nada me faltasse.

Às minhas irmãs, Morgana, Mayara, Ana Carolina e Natália pela torcida para que tudo desse certo nessa caminhada. Muito obrigado!

Ao meu padrasto Flávio e madrasta Verônica, por todo o apoio.

À minha família, tios (James, Emmanuel e Pedro Júlio - in memória), tias (Solange, Liduina, Marileide, Verônica e Adalgisa), primos (Handson Levi, Hudson, Jorge, Rodrigo e Luis Antônio) e primas (Christianne, Julianne, Lizianne, Luanda e Poliana), que de alguma forma contribuíram neste processo.

À minha esposa, Karine Couto, por toda ajuda, força e paciência. Obrigado meu amor!

À minha sogra (in memoria) Zailde Couto, por todo carinho.

À Dr^a. Rosália de Barros Nascimento de Medeiros, pela ajuda durante a realização deste trabalho. Muito obrigado por todo apoio!

Ao Professor Marcelo Melo, pelo acolhimento em sua sala.

A todos os meus amigos e em especial Filipe Chagas, Otto Cabral, Flávio Baracho, Luciano Gomes, Victor Sales, Renato Prado e Júnior Fragoso, por contribuírem de alguma forma em minha caminhada. Obrigado a todos!

A todos meus sinceros agradecimentos!

Procure descobrir o seu caminho na vida.
Ninguém é responsável por nosso destino, a não ser nós mesmos.

Chico Xavier

RESUMO

O objetivo dessa pesquisa foi avaliar de forma conjunta o potencial da produção de leite por lactação em kg (PL), a duração da lactação em dias (DL) e o intervalo de partos em dias (IDP) de um rebanho de búfalas mestiças Murrah, por meio da análise de componentes principais (ACP) e de agrupamento. Para este estudo foram analisadas 543 lactações provenientes de 105 búfalas coletadas entre os anos de 2002 e 2014. Com a aplicação da ACP, os primeiros componentes principais (CP_s) foram apontados como responsáveis por explicar 90,2% da variação total, sendo 59,1% e 31,1% da variação resultantes do primeiro (CP_1) e segundo (CP_2) componentes, respectivamente. Por meio da análise de agrupamento, foi possível dividir os animais em 3 grupos, de acordo com os desempenhos de PL, DL e IDP. Esses grupos reuniram búfalas semelhantes de acordo com o potencial das características estudadas, os quais foram recomendados para formação de lotes de manejo e reprodução, descarte de animais, além da indicação de animais superiores.

Palavras-chave: Análise de agrupamento. Autovalores. Componentes principais. Distância euclidiana

ABSTRACT

The aim of this research was to evaluate jointly the milk production potential per lactation in kg (MP), lactation length in days (LL) and calving of interval in days (CI) of a crossbred Murrah buffaloes herd through principal component analysis (PCA) and cluster analysis. For this study were analyzed 543 lactations collected from 105 buffaloes between 2002 and 2014. After the implementation of PCA, the first components (PCs) were responsible for 90.2% of the total variation, of which 59.1% and 31.1% were explained by the first (PC₁) and second (PC₂) components, respectively. By means of cluster analysis, it was possible to divide animals into 3 groups according to the performances of PL, DL and CI. These groups which classified according to the trait's potential were recommended for assemble of management and breeding groups, animals disposal and indication of superior animals.

Keywords: Cluster analysis. Eigenvalue. Euclidean distance. Principal component

LISTA DE ILUSTRAÇÕES

Figura 1 -	Box plot das características produção de leite (kg), duração da lactação (dias) e intervalo de partos (dias) em búfalas leiteiras mestiças Murrah.....	24
Figura 2 -	Scree plot do número de componentes principais e autovetores da tabela 4	26
Figura 3 -	Cluster no mapa de fator pelos dois primeiros componentes principais (A) e variáveis no mapa de fator pelas análises de componentes principais (B)	28
Figura 4 -	(A) Dendograma pelo método de Ward em separação de 3 grupos de animais; (B) Biplot de análise de componentes principais baseado em análise de agrupamento	28

LISTA DE TABELAS

Tabela 1 – Técnicas de análise multivariadas de acordo com o caráter de dependência ou independência das variáveis.....	17
Tabela 2- Aplicações de Análise de Componentes Principais (ACP) na produção animal de acordo com os autores, a espécie, as características em estudo, o número de componentes principais retidos (CP retido) e a sua contribuição para a variação total (% VCP).....	19
Tabela 3 - Estatísticas descritivas da produção de leite por lactação em kg (PL), da duração da lactação em dias (DL) e do intervalo de partos em dias (IDP), em búfalas leiteiras mestiças Murrah.....	23
Tabela 4 - Valores dos percentis das características produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP), em búfalas leiteiras mestiças Murrah.....	24
Tabela 5 - Coeficientes de correlação simples entre as características produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP), em búfalas leiteiras mestiças Murrah.....	25
Tabela 6 - Componentes principais (CP), autovalores (λ_i), porcentagem da variância explicada pelos componentes (% VCP) e porcentagem acumulada das características produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP), em búfalas leiteiras mestiças Murrah	25
Tabela 7 - Coeficientes de correlação entre as características produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP) em búfalas leiteiras mestiças Murrah, com os componentes principais e seus coeficientes de ponderação.....	27
Tabela 8 - Resumo das estatísticas produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP), em búfalas leiteiras mestiças Murrah, agrupados nos clusters preto, vermelho e verde, de acordo com a Figura 4A.....	29

SUMÁRIO

1	INTRODUÇÃO	12
2	CONSIDERAÇÕES PRELIMINARES SOBRE A BUBALINOCULTURA	14
3	CARACTERÍSTICAS ESTUDADAS	15
	3.1 Produção de Leite	15
	3.2 Duração da lactação	15
	3.3 Intervalo de partos	16
4	ESTATÍSTICA MULTIVARIADA.....	17
	4.1 Análise de Componentes Principais – ACP.....	18
	4.2 Análise de Agrupamento	19
5	MATERIAL E MÉTODOS.....	21
	5.1 Informação dos animais	21
	5.2 Análise estatística	21
	5.2.1 Análise de Componentes Principal (ACP).....	21
	5.2.2 Formação dos grupos usando Análise de Agrupamento ou Cluster	22
	5.2.3 A escolha do número de clusters a partir de um dendrograma	22
6	RESULTADOS E DISCUSSÃO	23
7	CONSIDERAÇÕES FINAIS	30
	REFERÊNCIAS	31

1 INTRODUÇÃO

O rebanho mundial bubalino é constituído por aproximadamente 195 milhões de cabeças, sendo a Índia (110 milhões de cabeças) o país que detém o maior rebanho (FAO, 2014). Estima-se que o Brasil possui um rebanho bubalino com 1.319.478,00 cabeças, dos quais, 66,5% se encontram na região norte (IBGE, 2014).

A criação de búfalos destinados à produção de leite no Brasil tem apresentado crescimento expressivo desde o surgimento das unidades industriais dedicadas à produção de derivados de leite de búfalas. Além disso, o leite de búfalas, quando comparado com o leite de vacas apresenta maior teor dos principais constituintes, os quais são fundamentais na produção dos derivados (ROSATI ; VLECK, 2002; BERNARDES, 2007; ASPILCUETA et al., 2010).

No Brasil, existem rebanhos que trabalham com o controle da produção de leite, quantitativa e qualitativamente, permitindo a avaliação do potencial dos animais, a prática da seleção e a obtenção de ganhos genéticos (ASPILCUETA et al., 2013). Entretanto, em sua maioria, os rebanhos não possuem controle zootécnico adequado para estimar os valores genéticos dos animais de forma acurada, dificultando a implementação dos programas de seleção. Os criadores possuem dificuldade por ocasião da classificação do potencial produtivo e reprodutivo dos animais. Portanto, surge a demanda por um método que possa identificar animais superiores e inferiores para auxiliar na seleção e descarte, além de permitir o ajuste no manejo, por meio de agrupamento de animais mais similares.

Em geral, na produção animal, os estudos das características de importância econômica são realizados utilizando-se análises univariadas. Essas análises são limitadas, uma vez que não consideram as possíveis correlações existentes entre os grupos e as características em estudo. Essa limitação pode ser contornada com a aplicação das técnicas multivariadas, as quais possuem abordagens que permitem analisar conjuntamente várias características. Essas técnicas têm sido bastante utilizadas na produção e melhoramento genético dos animais. Em geral, objetivam diminuir espaço amostral das características, formar grupos homogêneos de acordo com critérios estabelecidos previamente, além de outros propósitos (KIRKPATRICK ; MEYER, 2004; FRAGA et al., 2015; KADEGOWDA et al., 2008; GONZÁLEZ et al., 2011; VENTURA et al., 2012)

A técnica de análise de componentes principais (ACP), um dos procedimentos multivariados, foi descrita originalmente por Pearson (1901) e em seguida aplicada por Hotelling, (1933). É aplicada em diversas áreas da ciência e consiste em reduzir a

dimensionalidade dos dados por transformação de um conjunto de variáveis correlacionadas em um novo conjunto de variáveis não correlacionadas (CRUZ et al., 2004). A análise de agrupamento é outro tipo de técnica que objetiva reunir indivíduos em grupos específicos, de forma que haja homogeneidade dentro e heterogeneidade entre eles (VENTURA et al., 2012).

O objetivo dessa pesquisa foi avaliar de forma conjunta o potencial da produção de leite, duração da lactação e intervalo de parto de um rebanho de búfalas mestiças Murrah por meio da análise multivariada.

2 CONSIDERAÇÕES PRELIMINARES SOBRE A BUBALINOCULTURA

Os búfalos (*Bubalus bubalis*) são originários da Ásia. Os primeiros búfalos foram introduzidos no Brasil pela Amazônia, entre 1890 e 1895, os quais foram trazidos por condenados foragidos da Guiana Francesa em um barco que ancorou no norte da Ilha do Marajó. Em seguida, foram realizadas várias importações (MARQUES, 2000; SANTIAGO, 2000). Desde 1965, a Associação Brasileira dos Criadores de Búfalos - ABCB efetua os registros genealógicos dos reprodutores bubalinos machos e fêmeas com a finalidade de garantir o controle do rebanho nacional, comprovar a filiação, linhagem e grau de sangue, promover o melhoramento genético, incrementar e aprimorar a seleção dos reprodutores, através de Provas Zootécnicas (ANDRADE ; GARCIA, 2005). Atualmente, quatro raças bubalinas são reconhecidas no Brasil: Murrah, Mediterrâneo, Jafarabadi, e Carabao (ABCB, 2015).

O rebanho mundial bubalino é constituído por aproximadamente 195 milhões de cabeças, sendo a Índia (110 milhões de cabeças) o país que detém o maior rebanho (FAO, 2014). A produção de leite de búfalas mundial é de 102.041.460,00 toneladas, com rendimento/produtividade de 1.666,30 kg/animal/lactação, sendo o continente asiático o responsável por produzir mais de 97% (FAO, 2013). No Brasil o búfalo foi, entre os animais domésticos, àquele que teve o crescimento mais expressivo no Brasil entre 2010 e 2011, com aumento de 7,8%. Estima-se que o Brasil possui um rebanho bubalino com 1.319.478,00 milhões de cabeças, dos quais, 66,5% se encontram na região norte (IBGE, 2014). Embora, não exista informação oficial quanto ao volume do leite de búfala produzido no Brasil, acredita-se que a produção venha acompanhando a tendência mundial (AMARAL ; ESCRIVÃO, 2005).

A criação de búfalos destinados à produção de leite no Brasil tem apresentado crescimento expressivo desde o surgimento das unidades industriais dedicadas à produção de derivados de leite de búfalas. Este fato pode ser justificado devido a qualidade ou sabor singular que os produtos derivados de leite de búfalas apresentam, principalmente, a mussarela (BERNARDES, 2007).

No Brasil, existem rebanhos que trabalham com o controle da produção de leite, quantitativa e qualitativamente, permitindo a avaliação do potencial dos animais, a prática da seleção e a obtenção de ganhos genéticos (ASPILCUETA et al., 2010). Entretanto, em sua maioria, os rebanhos não possuem controle zootécnico adequado para estimar os valores genéticos dos animais de forma acurada, dificultando a implementação dos programas de

seleção. Os criadores possuem dificuldade por ocasião da classificação do potencial produtivo e reprodutivo dos animais, havendo, portanto, a demanda por um método que possa classificar os animais superiores e inferiores para auxiliar na seleção e descarte, além de permitir o ajuste no manejo, por meio de agrupamento de animais mais similares.

3 CARACTERÍSTICAS ESTUDADAS

3.1 Produção de Leite

A produção de leite (PL) é a característica produtiva mais importante no rebanho leiteiro, pois influencia diretamente a lucratividade da atividade. Vários fatores podem interferir nessa produção, dentre eles, podem-se destacar: estágio da lactação, estação de parto, ordem de lactação, raças, entre outros. Embora, a produção de leite das búfalas seja menor que a produção das vacas, o leite bubalino apresenta maior valor nutritivo e rendimento industrial, desta forma, garantindo ao produtor uma remuneração diferenciada pelo litro de leite.

A média estimada de produção de leite dos rebanhos bubalinos no Brasil está entre 1.600 a 2.200 kg de leite por lactação. Autores como Malhado et al. (2013a), trabalhando com búfalas da raça Murrah e Mediterrâneo pertencentes a 9 rebanhos do programa brasileiro de melhoramento genético obtiveram uma média de PL de 1.628,43 kg. Rodrigues et al. (2010), estudando 1.182 registros de fêmeas bubalinas da raça Murrah e seus mestiços em Belém-PA, obtiveram média de produção de leite de 1.663,84 kg. Aspilcueta et al. (2011), obtiveram uma média de 1.714,30 kg para PL. Bezerra Júnior et al. (2014), analisando 446 registros de produção de leite búfalas mestiças Murrah no litoral Norte do estado de Alagoas, obtiveram uma média de 2.218,03 kg. A variação verificada em diferentes estudos, provavelmente, pode ser explicada pelas diversas condições ambientais, além das diferentes constituições genóticas dos rebanhos em estudo.

3.2 Duração da lactação

A duração da lactação (DL) compreende o período de lactação da fêmea. Em geral, quanto maior a duração da lactação, maior será a produção de leite. Em bovinos leiteiros, principalmente de origem européia, preconiza-se um intervalo de partos de 12 meses, com produção de uma cria por ano. Nesse caso, longos períodos de lactação não são recomendados, pois poderão acarretar aumento no intervalo de partos. Por ocasião do final de gestação da fêmea, 60 dias antes do parto, recomenda-se a secagem do leite para descanso do

tecido mamário para o próximo ciclo produtivo, além de destinar maior quantidade de nutrientes para término da formação do feto.

No caso das búfalas, por questões fisiológicas a persistência da lactação é mais curta do que em vacas, o que pode ser explicado pela ausência de seleção da espécie bubalina para esta característica (CATILLO et al., 2002). No Brasil, a média de duração da lactação das búfalas varia entre 252 e 301 dias (SAMPAIO NETO et al., 2001; RAMOS et al., 2006; MALHADO, et al., 2009; RODRIGUES et al., 2010). Nos países asiáticos, essa média varia entre 200 e 305 dias (FLORES et al., 2013; GUPTA et al., 2014; BASHIR et al., 2015; SAHOO et al., 2015). Na Itália, essa média é de 270 dias (PARLATO ; VAN VLECK, 2012; DE ROSA et al., 2015; ZICARELLI, 2015).

Com relação ao ajuste da produção de leite, nos artigos científicos encontram-se vários métodos. Alguns artigos relataram ajustes realizadas aos 305 dias da produção de leite, enquanto que outros fazem o ajuste para 270 dias. Acredita-se que essa discrepância no ajuste da DL ocorre em virtude da variação do potencial de PL por lactação nos rebanhos dos diferentes locais.

Nos países asiáticos, normalmente a produção de leite é ajustada para 305 dias (SANE et al., 1972; MOURAD ; MOHAMED, 1995). Entretanto, na Itália a produção de leite é ajustada para 270 dias, sendo utilizada como critério de seleção (ROSATI ; VAN VLECK, 2002; ICAR, 2014). Já no Brasil, existem pesquisas com o ajuste para 270 dias (TONHATI et al., 2004) e aquelas que trabalham com ajustes de 305 dias (ASPILCUETA et al., 2013).

3.3 Intervalo de partos

O intervalo de partos (IDP) é uma característica reprodutiva de suma-importância no sistema de criação. Sua alteração interfere no ciclo reprodutivo do animal e, conseqüentemente, sobre a lucratividade do produtor. Menores IDP proporcionam maiores números de progênies por fêmeas ao longo de sua vida produtiva e, no caso do sistema de produção de leite, cada parto gera um período de lactação. Fêmeas que recebem nutrição balanceada e que tiveram período de descanso no ciclo reprodutivo anterior, geralmente, são aquelas que vão apresentar os menores períodos de serviço (intervalo entre o parto e o primeiro cio fértil). A produção de dois bezerros a cada três anos é considerada aceitável (PEREIRA et al., 2008).

Essa característica é um dos parâmetros mais importantes para medir a eficiência reprodutiva do rebanho. Encontram-se na literatura alguns relatos sobre os valores médios de

IDP em rebanhos bubalinos. Analisando informações de 12 rebanhos Murrah, Ramos et al. (2006), obtiveram uma média de IDP de 432 dias. Sampaio Neto et al. (2001), trabalhando com búfalas da raça Murrah no Ceará, obtiveram uma média de IDP de 430,79 dias. Malhado et al., (2013b), em 4 rebanhos brasileiro de búfalas Murrah relataram uma média de IDP de 411 dias, porém Téllez et al. (2005), relataram que é possível um intervalo de partos de 365 dias sendo fisiologicamente possível e economicamente vantajoso.

4 ESTATÍSTICA MULTIVARIADA

As técnicas multivariadas, inicialmente, eram utilizadas apenas em algumas áreas, como na área da psicologia, para descrever o comportamento de pessoas, e na área mercadológica, para definição e caracterização de nichos de mercados, além de serem empregadas em outros seguimentos. Com os avanços dos recursos computacionais essas técnicas passaram a ser aplicadas nas diversas áreas de pesquisa. Na produção animal, as técnicas de análises multivariadas têm sido amplamente empregadas em estudos de caracterização racial, qualidade de produtos, critérios de seleção, além de outros (ABREU et al., 1999; BARBOSA et al., 2006; VENTURA et al., 2012; VENTURINI et al., 2013; BIAGIOTTI et al., 2014; REZENDE et al., 2015; ARANDAS et al., 2016). A análise multivariada consiste em um conjunto de técnicas estatísticas que tem por finalidade analisar relações entre 3 ou mais variáveis aleatórias e inter-relacionadas que são medidas simultaneamente em cada unidade experimental (MINGOTI, 2005). Existem várias técnicas multivariadas (Tabela 1) e a adoção da técnica mais apropriada deve levar em consideração alguns critérios, dentre eles, pode-se destacar o caráter de dependência ou independência das variáveis e a escala de medida dos dados coletados, os quais podem ser métricos ou não métricos.

Tabela 1 – Técnicas de análise multivariadas de acordo com o caráter de dependência ou interdependência das variáveis

Dependências	Interdependência
Análise Conjunta	Análise de Agrupamentos
Análise multivariada de variância e covariância	Análise Fatorial Confirmatória e Modelagem de Equações Estruturais
Análise Discriminante Múltipla e Regressão Logística	Análise de Componentes Principais
Análise de Regressão e Correlação Múltipla	Escalonamento Multidimensional
Análise de Correlação Canônica	Análise de Correspondência

Fonte: Adaptado de Hair et al.(2009)

Nas técnicas de dependência existe uma relação de dependência entre as variáveis. Essas técnicas podem ser classificadas como possuindo uma única variável dependente, várias

variáveis dependentes, ou até mesmo, diversas relações de dependência ou independência. Por outro lado, nas técnicas de interdependência as variáveis não podem ser classificadas como dependentes ou independentes. Nesses casos, devem ser realizadas análises simultâneas de todas as variáveis com objetivo de encontrar uma estrutura subjacente a todo conjunto de variáveis (HAIR et al., 2009).

É importante salientar, que as técnicas multivariadas não se aplicam em qualquer conjunto de dados. Para sua aplicação, há necessidade de existência de uma estrutura de correlação entre as variáveis, pois se as mesmas não estiverem relacionadas entre si, as análises univariadas são mais indicadas.

As análises multivariadas que serão utilizadas neste estudo com mais aprofundamento serão análises de componentes principais e análise de agrupamentos.

4.1 Análise de Componentes Principais – ACP

A análise de componentes principais foi originalmente descrita por Pearson (1901) e, posteriormente, aplicada com Hotelling (1933) em diversas áreas da ciência. Essa é uma técnica da estatística multivariada que tem dentre os vários objetivos reduzir a dimensionalidade dos dados, a partir de um conjunto de variáveis correlacionadas em um novo conjunto de variáveis não correlacionadas. Esse novo conjunto é capaz de explicar a variância total observada entre os dados de modo a facilitar a interpretação da interdependência entre elas (CRUZ et al., 2004; KIRKPATRICK ; MEYER, 2004). Outro aspecto é que os componentes principais são combinações lineares de variáveis, as quais possuem a propriedade de reter, nos primeiros componentes estimados, as maiores frações de variabilidade total disponível. Portanto, a escolha com base nos componentes de alto poder discriminatório possibilita a obtenção de resultados satisfatórios (ABREU et al., 1999).

Na maioria dos estudos com ACP o descarte de variáveis é realizado de acordo com a metodologia descrita por Jolliffe (1972, 1973), na qual o número de variáveis descartadas deve ser igual ao número de componentes cuja variância (autovalor) é inferior a 0,7. Cada autovalor representa a variância de cada componente.

Na literatura especializada são encontradas várias pesquisas que aplicam a técnica de ACP na produção animal em várias espécies, Tabela 2.

Tabela 2 - Aplicações de Análise de Componentes Principais (ACP) na produção animal de acordo com os autores, a espécie e características em estudo, número de componentes principais retidos (CP retido) e sua contribuição para a variação total (% VCP)

Autores	Espécie	Características	CP retido	% VCP
Barbosa et al., (2006)	Suínos	TLN, TLD, PN, P21, P42, P63, P77, GPD, CD, CA	5	85,15
Leite et al., (2009)	Codornas	PV, PP, PC, PGA, PVC e PCE	4	75,00
Paiva et al., (2010)	Aves de postura	TP, P34, P42, P63 e P77	3	77,00
Meira et al., (2013)	Equinos	AC, AG, CCa, CC, CD, CG, CE, CP, LC, LA, PT, PC e PM	6	78,57
Vohra et al., (2015)	Búfalos	AC, CC, PT, CO, CCau, CF, CChi, CCB, DPO, DOQ, LF, PA	5	77,70

TLN = tamanho da leitegada; TLD = tamanho da leitegada ao desmame; PN = peso ao nascer; P21, P42, P63 e P77 = peso aos 21, 42, 63 e 77 dias de idade, respectivamente; GPD = ganho de peso diário; CD = consumo diário; CA = conversão alimentar; PV = peso vivo; PP = peso do peito; PC = peso das coxas; PGA = peso da gordura abdominal; PVC = peso das vísceras comestíveis; PCE = peso da carcaça eviscerada; TP = taxa de postura; P34, P42, P50, P58 e P66 = peso médio de idade e do ovo na 34^o, 42^o, 63^o e 77^o semanas, respectivamente; AC = altura da cernelha; AG = altura na garupa; CCa = comprimento da cabeça; CD = comprimento do dorso; CG = comprimento da garupa; CE = comprimento da espádua; CP = comprimento do pescoço; LC = largura da cabeça; LA = largura das ancas; PT = perímetro tórax; PC = perímetro da canela; PM = pontuação de marcha; CC = comprimento do corpo; CO = comprimento da orelha; CCau = comprimento da cauda; CF = comprimento da face; LF = largura da face; CChi = comprimento do chifre; CCB = circunferência do chifre a base; DPO = distância entre o pino ósseo; DOQ = distância entre o osso do quadril; PA = perímetro abdominal.

4.2 Análise de Agrupamento

A Análise de Agrupamento (AA) é uma técnica exploratória da estatística multivariada que tem como propósito a formação de grupos, por algum critério de classificação, de modo que se obtenha homogeneidade dentro do grupo e heterogeneidade entre eles. Para o emprego da AA faz-se necessário a estimativa de uma medida de similaridade ou dissimilaridade entre os dados e a escolha de um método de agrupamento ideal para formação dos grupos (CRUZ et al. 2004).

Visando medir a distância entre dois objetos ou até mesmo saber o quanto eles se parecem, várias metodologias podem ser utilizadas para verificar a distância entre os elementos de uma matriz de dados. Dentre elas, podem ser citadas a distância Euclidiana, Mahalanobis, Absoluta, Minkowski, Chebychev, entre outras. No entanto, a adoção da medida mais apropriada vai depender de cada situação. Afinal, cada metodologia disponibiliza um coeficiente de parença, ou seja, a formação de diferentes agrupamentos (CRUZ ; CARNEIRO, 2003). As medidas mais utilizadas são a distância Euclidiana e a distância de Mahalanobis generalizada (D^2). A distância Euclidiana padroniza as variáveis em estudo utilizando-se a média e o desvio-padrão, de forma a eliminar o viés causados pelas diferentes escalas. A distância de Mahalanobis, além de realizar a padronização das variáveis,

leva em consideração a existência de correlações entre as características avaliadas por meio da matriz de variâncias e covariâncias residuais a ponto de evitar que grupos altamente correlacionados possam superponderar um conjunto de variáveis nos procedimentos de agrupamento (HAIR et al., 2009).

Cruz et al. (2004), relatam que os métodos de agrupamentos são divididos em método hierárquico e método não hierárquico. Nos métodos hierárquicos os indivíduos são agrupados por um processo que se repete em vários níveis ou critérios para então formar uma estrutura de dendograma. Não importa quantos grupos sejam formados, mas sim a estrutura de ligação entre os indivíduos. Esses métodos podem ser classificados como aglomerativos e divisivos. No hierárquico aglomerativo estão os métodos: vizinho mais próximo, vizinho mais distante, ligação simples, ligação média, ligação completa, ponderado ou não, centroide, Ward e outros. No método hierárquico divisivo, o método de Edwards e Cavalli-Sforza é o mais utilizado. É importante ressaltar que todos esses métodos utilizam a matriz de distâncias, dispensando recorrer aos dados originais.

Nos métodos não hierárquicos são realizadas subdivisões do conjunto maior de indivíduos em subconjuntos não vazios e mutuamente exclusivos, de acordo com algum critério. O método não hierárquico de Tocher é o mais utilizado no melhoramento genético.

5 MATERIAL E MÉTODOS

5.1 Informação dos animais

A pesquisa foi realizada com informações de produção de leite em um rebanho de búfalas mestiças de Murrah pertencente à Fazenda Castanha Grande, localizada no município de São Luiz do Quitunde, litoral Norte de Alagoas. Apresenta altitude aproximada de 10 metros, coordenadas geográficas de 9° 10'06''S e 35° 33'40''W e clima tropical chuvoso, de acordo com a classificação de KÖPPEN. Foram utilizados 543 registros da produção de leite provenientes da lactação de 105 búfalas, coletados entre os anos 2000 e 2014. As características estudadas foram: produção de leite por lactação (PL), duração da lactação (DL) e intervalo de partos (IDP).

As búfalas são ordenhadas duas vezes ao dia (manhã e tarde) em ordenhadeira mecânica, com balde ao pé, linha de vácuo canalizada e contenção feita em tandem com capacidade para 24 animais. Antes da ordenha, as búfalas são submetidas ao banho de chuveiro e, posteriormente, banho de imersão, momento em que aproveitam para urinar e defecar. Além de contribuir para a melhor qualidade do leite, essa prática objetiva diminuir o estresse dos animais e, conseqüentemente, contribui para o aumento da produção.

Durante a lactação, os animais são criados em sistema semi-intensivo. São mantidos em pasto de *Braquiária humidícula* e recebem no cocho, como volumoso, a parte aérea da cana-de-açúcar (*Saccharum officinarum*) ou cana-de-açúcar inteira, ambas, enriquecidas com uréia, tanto *in natura*, quanto na forma de silagem. Além disso, é fornecida a suplementação mineral e ração balanceada. A ração é fornecida na seguinte proporção: as búfalas com produção entre 5 a 7,5 kg de leite/dia recebem 1kg de ração/dia; as búfalas com produção entre 8 a 10,5 kg de leite/dia recebem 2 kg de ração/dia.

5.2 Análise estatística

5.2.1 Análise de Componentes Principal (ACP)

Os valores fenotípicos padronizadas de 105 búfalas de todas as características foram utilizados na ACP. Os valores fenotípicos foram padronizados utilizando a fórmula:

$$z = \frac{x - \bar{x}}{s}$$

Onde z é o valor padronizado de x , \bar{x} é a media da característica, e s é o respectivo desvio padrão.

Os componentes principais são calculados por meio de combinações lineares das variáveis originais com autovetores. O valor absoluto de um autovetor determina a importância das características de um componente principal. Cada autovetor é calculado a partir dos autovalores de uma matriz de correlação dos dados e os autovalores são relacionados com a variância em cada um dos componentes principais (RENCHEER, 2002).

O primeiro componente principal (CP₁) explica o maior percentual da variância total. O segundo componente principal (CP₂) explica a segunda maior percentagem e assim por diante, até que toda a variância é explicada. Em um conjunto de dados com p variáveis, o vetor aleatório $\mathbf{x}' = [x_1, x_2, \dots, x_p]$ que tem matriz de correlação \mathbf{R} com os pares autovalor-autovetor $(\lambda_i, \mathbf{e}_i)$, para $i = 1, 2, \dots, p$, em que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ (RENCHEER, 2002).

O i -ésimo componente principal é dado por

$$PC_i = \mathbf{e}_i^t \mathbf{x} = e_{i1}x_1 + e_{i2}x_2 + \dots + e_{ip}x_p$$

onde \mathbf{e}_i^t é o i -ésimo autovetor e x_p é o p -ésimo valor da variável original.

O critério de Jolliffe (1972,1973), foi usado para selecionar os componentes principais que explicam a maior parte da variação no conjunto de dados. Este critério consiste em descartar o CP que apresentar autovalor menor que 0,7, sendo este componente passível de descarte, porque pouco explica a variabilidade total dos dados.

5.2.2 Formação dos grupos usando Análise de Agrupamento ou Cluster

Análise de agrupamento hierárquico pelo critério de Ward foi realizada para encontrar um número adequado de grupos dentro da população. Este critério é baseado no teorema de Huygens que permite a decomposição da variação total entre e dentro de cada grupo. O método de Ward consiste na formação de grupos pela maximização da homogeneidade dentro dos grupos (variância mínima dentro dos grupos). A soma de quadrados dentro dos grupos é usada como medida de homogeneidade. Isto é, esse método tenta minimizar a soma de quadrados dentro do grupo. Os grupos formados em cada passo são resultantes do grupo solução com a menor soma de quadrados (LÊ et al., 2008).

5.2.3 A escolha do número de clusters a partir de um dendrograma

A escolha do número de clusters foi realizada utilizando o pacote “FactorMineR” (HUSSON et al., 2016). Essa escolha é uma questão fundamental e várias abordagens têm sido propostas. O número de grupos pode então ser escolhido olhando para a aparência geral

do dendrograma, o gráfico de barra, a variância dentro dos grupos e etc. Essas regras são muitas vezes baseadas, implicitamente ou não, sobre o crescimento de inércia. Elas sugerem uma divisão em grupos Q quando o aumento de variância entre $Q - 1$ e Q cluster é muito maior do que aquele entre Q e $Q + 1$ cluster.

Todas as análises deste trabalho foram feitas por meio de rotinas computacionais implementadas no software R 3.0.1 (R CORE TEAM, 2015).

6 RESULTADOS E DISCUSSÃO

De modo geral, levando-se em consideração todos os fatores envolvidos no sistema de produção, o rebanho em estudo apresenta superioridade na produção de leite por lactação e duração da lactação, além de menor intervalo de parto quando comparado com outros rebanhos citados na literatura (Tabela 3). Rodrigues et al. (2010), estudando 1.182 registros de fêmeas bubalinas da raça Murrah e seus mestiços em Belém-PA, obtiveram médias de produção de leite por lactação e duração da lactação de 1.663,84 kg e 269,89 dias, respectivamente. Aspilcueta et al. (2010), estimaram média de 1.814,00 kg para PL por lactação. Pereira et al. (2008), analisando 420 informações de intervalos de partos de fêmeas bubalinas Murrah x Mediterrâneo em Porto Velho – RO encontraram uma média de 451 dias.

Tabela 3 - Estatísticas descritivas da produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP) em búfalas leiteiras mestiças Murrah

Variável	Número de informações	Média	Desvio Padrão
PL	543	2232,86	612,91
DL	543	275,44	48,85
IDP	543	420,42	92,30

O box plot das características PL, DL e IDP é apresentado na Figura 1, com a finalidade de apresentar o comportamento de simetria e assimetria das variáveis estudadas. Pelo diagrama do box plot, Figura 1, é possível verificar que a PL e a DL possuem melhor distribuição do que IDP. Nesse caso, a posição da mediana não ocupa a região central do box. Além disso, a média e a mediana da PL e da DL são bem próximas, enquanto que para IDP essas duas medidas de posição são distantes, mostrando que a distribuição dos dados é assimétrica (Tabela 4).

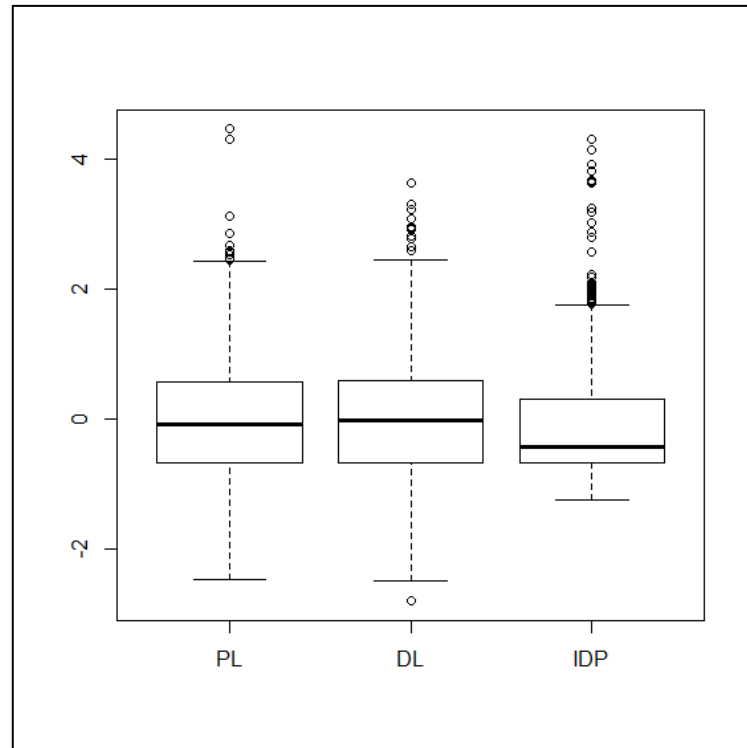


Figura 1 – Box plot das características produção de leite (kg), duração da lactação (dias) e intervalo de partos (dias) em búfalas leiteiras mestiças Murrah

Tabela 4 - Valores dos percentis das características produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP) em búfalas leiteiras mestiças Murrah

	Valor Mínimo	Primeiro Quantil	Mediana	Média	Terceiro Quantil	Valor Máximo
PL	730,2	1.824,7	2.184,0	2.232,9	2.585,3	4.965,4
DL	139,0	243,0	275,0	275,4	305,0	453,0
IDP	306,0	358,0	382,0	420,4	448,5	817,0

A matriz de correlações fenotípicas entre as características e seus respectivos *p-values* são apresentados na Tabela 5. Observaram-se que as variáveis com alta correlação positiva foram PL e DL (0,70), mostrando valor importante entre a produção de leite e a duração a lactação. Por outro lado, a característica IDP apresentou baixa correlação entre demais: IDP/PL (0,11) e IDP/DL (0,23).

Tabela 5 - Coeficientes de correlação simples entre as características produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP) em búfalas leiteiras mestiças Murrah

Características	PL	DL	IDP
PL	1,00 0,00**		
DL	0,70 0,00**	1,00 0,00**	
IDP	0,11 0,00**	0,23 0,00**	1,00 0,00**

** Significativo a 1% (p -valor $<0,01$)

Com base nos resultados obtidos para os componentes principais, seus respectivos autovalores e porcentagens da variância explicada por cada um (Tabela 6), os dois primeiros componentes principais, foram responsáveis por mais de 90,2% da variação total dos dados. O primeiro CP foi responsável por aproximadamente 59,1% e o segundo, por 31,1% dessa variação.

Tabela 6 - Componentes principais (CP), autovalores (λ_i), porcentagem da variância explicada pelos componentes (% VCP) e porcentagem acumulada das características produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP) em em búfalas leiteiras mestiças Murrah

Componente	λ_i	% VCP	% VCP (acumulada)
CP ₁	1,77	59,1	59,1
CP ₂	0,93	31,1	90,2
CP ₃	0,29	9,8	100

A escolha dos dois primeiros componentes principais foi baseada na variância total dos dados, os quais explicaram 90,2% (Tabela 6), além da análise visual do gráfico de *scree plot* (Figura 2). Portanto, os dois primeiros componentes principais resumem efetivamente a variância amostral total e podem ser utilizados para o estudo do conjunto de dados. Cabe ressaltar que, de acordo com Jolliffe (1972, 1973) apenas um componente (CP₃) apresentou autovalor menor que 0,7, sendo este passível de descarte, pois apresentou pouca importância relativa, explicando pouco da variabilidade dos dados.

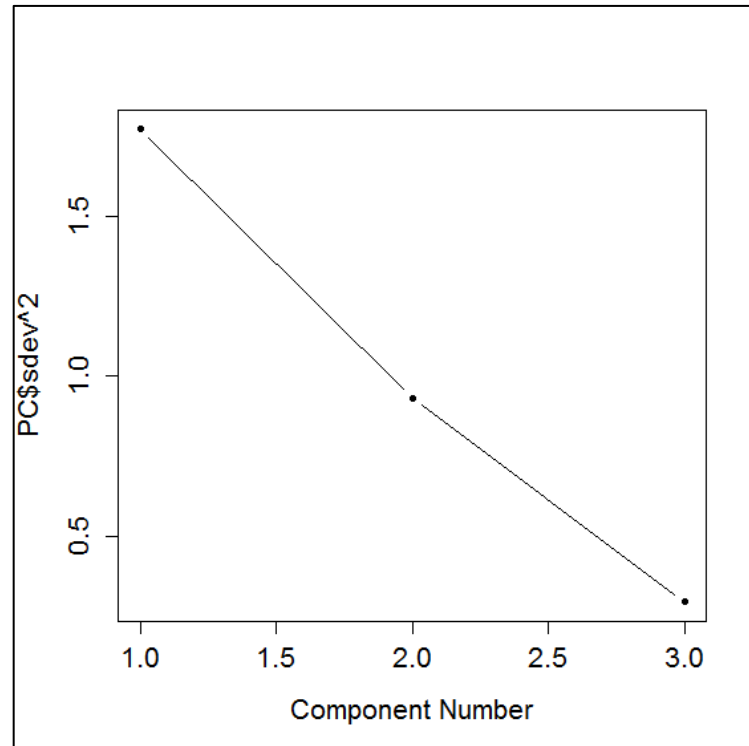


Figura 2 - Scree plot do número de componentes principais e autovetores (λ_i) da Tabela 6

Oliveira et al. (2014) aplicando a técnica de ACP para avaliação genética de 9 características de importância econômica em Búfalos Murrah no Brasil, verificaram que 4 CP pode ser ideal para explicar a estrutura de covariância das características. Por outro lado, Agudelo-Gómez et al. (2015), estudando valores de produção de 8 características produtivas e reprodutivas em búfalos com dupla-aptidão na Colômbia, relataram que com a aplicação da ACP os três primeiros CP apresentaram autovalores maior que um, explicando 65,78% da variação total. Com 13 diferentes medidas corporal de búfalas adultas da raça Gojri, Vohra et al. (2015), concluíram que 4 componentes foram suficientes para explicar mais de 70,9% da variação total, sendo o primeiro componente responsável por explicar 31,5% da variação total.

De posse dos coeficientes de correlação e das ponderações das variáveis originais e os componentes principais foram constituídas as equações 1 e 2, as quais utilizaram, unicamente, os dois primeiros componentes principais (Tabela 7).

O componente principal na equação 1, é o resultado da combinação linear de todas as variáveis estudadas com os coeficientes de ponderação, sendo a PL e a DL, aquelas que apresentaram maiores coeficientes e, portanto, denominado componente de características produtivas. Por outro lado, a equação 2, no segundo componente principal ficou evidente a importância do IDP para esse componente, podendo ser denominado de componente de característica reprodutiva (Tabela 7).

Tabela 7 - Coeficientes de correlação entre as características produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP) em búfalas leiteiras mestiças Murrah com os componentes principais e seus coeficientes de ponderação

Características	Coeficiente de ponderação		Correlação	
	CP ₁	CP ₂	CP ₁	CP ₂
PL	0,66	-0,29	0,88	-0,28
DL	0,69	-0,13	0,91	-0,13
IDP	0,30	0,95	0,40	0,91

$$Y_1 = 0,66 \text{ PL} + 0,69 \text{ DL} + 0,30 \text{ IDP} \quad (1)$$

$$Y_2 = -0,29 \text{ PL} - 0,13 \text{ DL} + 0,95 \text{ IDP} \quad (2)$$

Dois componentes principais explicaram 90,2 % da variação total das características (Figura 3A). De acordo com Rencher (2002), no mínimo 70% da variação total deve ser explicada pelos dois primeiros componentes principais. Portanto, nas condições do nosso estudo, o uso dos dois primeiros componentes principais foi suficiente para a qualidade da análise de agrupamento hierárquico, por meio da dispersão dos escores em gráficos cujos eixos são os referidos componentes (Figura 3B).

A Figura 3B mostra a maior concentração de animais do grupo vermelho situados na mesma região de apresentação dos vetores DL e PL (Figura 3A), os quais exibem elevada incidência sobre o primeiro componente principal. Motivo pelo qual, esses animais são aqueles com maior potencial para DL e PL (Tabela 8). No caso de IDP, animais do grupo verde localizam-se, em sua maioria, no mesmo quadrante do vetor de IDP (Figura 3B), o qual apresenta grande incidência sobre o segundo componente principal (Figura 3A). Entretanto, embora os animais do grupo verde apresentem elevados valores para IDP, não representam o melhor desempenho para a característica. Pois, nesse caso, menores valores para o período entre dois partos refletem os melhores resultados num sistema de produção.

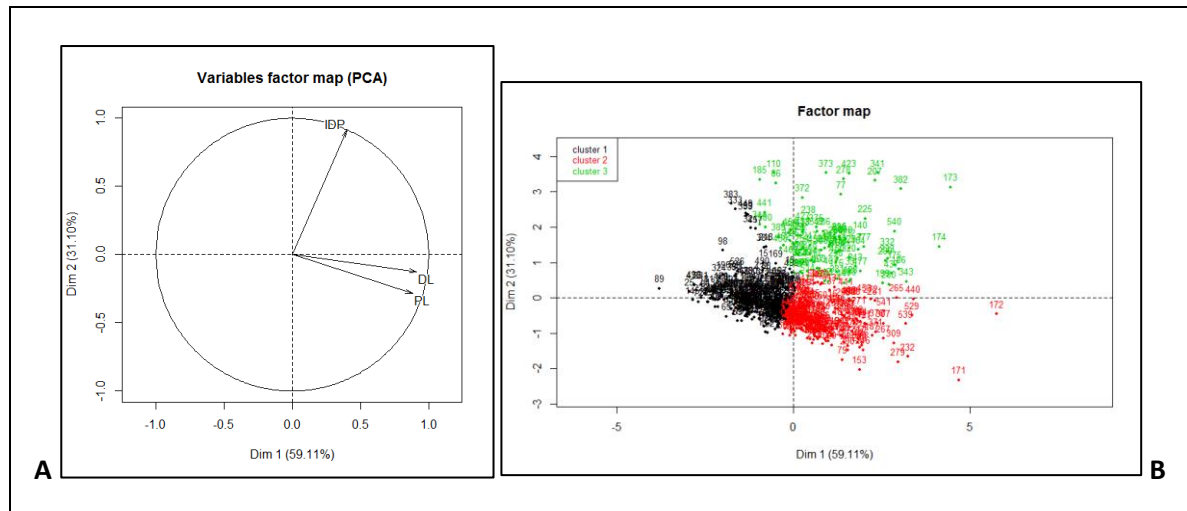


Figura 3: Cluster no mapa de fator pelos dois primeiros componentes principais (A) e variáveis no mapa de fator pelas análises de componentes principais (B)

O dendrograma (Figura 4A) foi obtido pela análise de agrupamento hierárquico com o método de Ward para agrupar os animais baseando-se pela distância Euclidiana, separando-os em 3 grupos conforme as características estudadas. A Figura 4 (B) reúne informações de duas técnicas de análises multivariadas: componentes principais e análise de agrupamento. Pela análise de agrupamento, foi possível dividir os animais em 3 grupos de acordo com os desempenhos de produtividade.

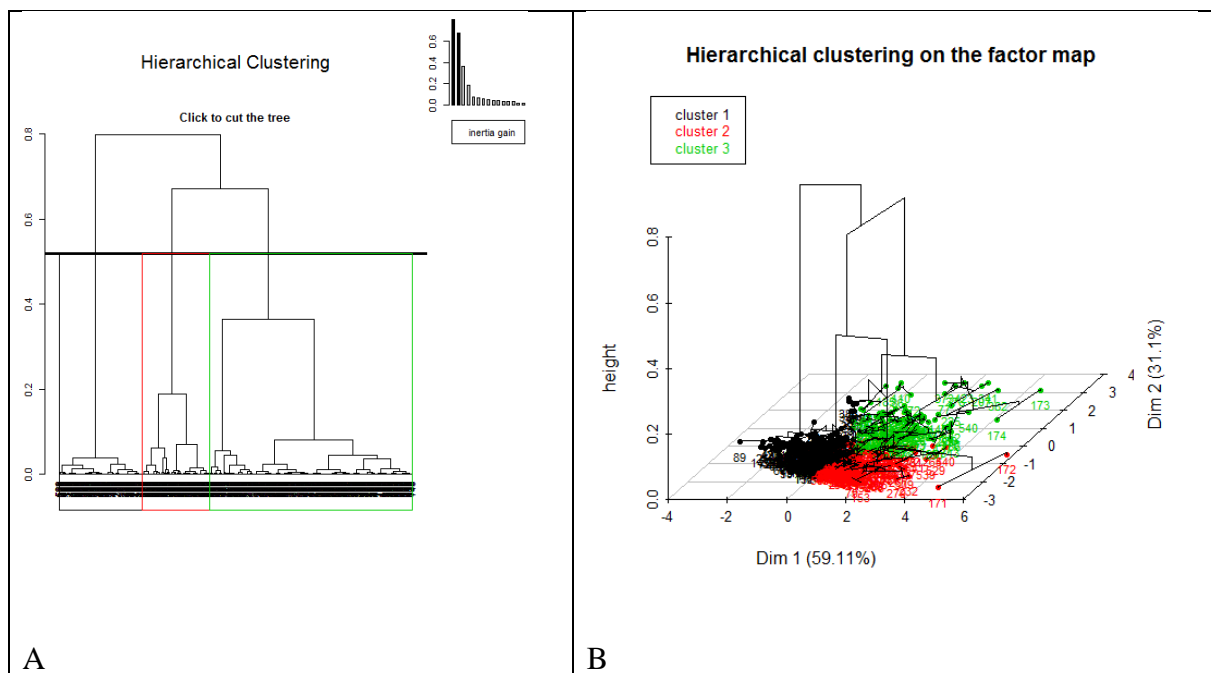


Figura 4- (A) Dendrograma pelo método de Ward em separação de 4 grupos dos animais; (B) Biplot de análise de componentes principais baseado em análise de agrupamento

Na Tabela 8, é possível verificar que no grupo de cor vermelha estão reunidos os animais com maiores lactação e produção de leite por lactação. No grupo verde os animais apresentam maior IDP e, nesse caso, os piores desempenhos para essa característica. Por outro lado, no grupo de cor preta, estão reunidos animais com resultados inferiores para a PL e DL, mas com valores intermediários de IDP.

Tabela 8 - Resumo das estatísticas produção de leite por lactação em kg (PL), duração da lactação em dias (DL) e intervalo de partos em dias (IDP) em em búfalas leiteiras mestiças Murrah agrupados nos clusters preto, vermelho e verde, de acordo com a Figura 4A

Cluster	Preto (n=247)		Vermelho (n=201)		Verde (n=95)	
Variável	<i>Média</i>	<i>SD</i>	<i>Média</i>	<i>SD</i>	<i>Média</i>	<i>SD</i>
PL	1.787,28	346,39	2.704,29	489,60	2.393,94	564,73
DL	238,64	30,73	308,64	36,89	300,89	42,76
IDP	387,77	51,07	383,43	42,73	583,59	76,05

Esses resultados são importantes porque possibilitam a realização de uma classificação dos animais considerando, de forma conjunta, as características de maior importância para o sistema. Em geral, no Brasil, os rebanhos bubalinos não possuem estrutura de informações zootécnicas adequadas para a realização de avaliação genética visando a obtenção de valores genéticos para as características de importância econômica, principalmente, no que diz respeito ao pedigree dos animais.

7 CONSIDERAÇÕES FINAIS

Com a aplicação das técnicas de análises multivariada foi possível observar que os dois primeiros componentes principais foram responsáveis por mais de **90,2%** da variação total dos dados. A análise de agrupamento possibilitou o agrupamento das búfalas em 3 grupos distintos de acordo com os desempenhos da produção de leite por lactação (kg), duração da lactação (dias) e intervalo de parto (dias) avaliados conjuntamente. Com base no agrupamento dos animais, o produtor pode tomar decisões específicas para o manejo nutricional, manejo reprodutivo e descarte de búfalas para cada grupo. Além disso, auxilia no processo da escolha de novos materiais genéticos a serem adquirido para melhorar o potencial do rebanho.

REFERÊNCIAS

- ABREU, V. M. N. et al. Estudo da capacidade de combinação de linhagens de matrizes de frangos de corte, por meio da análise de fatores. Produção de ovos. **Revista Brasileira de Zootecnia**, v. 28, n. 2, p. 258–263, abr. 1999.
- AGUDELO-GÓMEZ, D.; PINEDA-SIERRA, S.; CERÓN-MUÑOZ, M. F. Genetic Evaluation of Dual-Purpose Buffaloes (*Bubalus bubalis*) in Colombia Using Principal Component Analysis. **PloS one**, v. 10, n. 7, p. e0132811, jan. 2015.
- ARANDAS, J. K. G. et al. Multivariate analysis as a tool for phenotypic characterization of an endangered breed. **Journal of Applied Animal Research**, v. 148, n. 3, p. 1–7, fev. 2016.
- AMARAL, F. R.; ESCRIVÃO, S. C. Aspectos relacionados à búfala leiteira. **Revista Brasileira de Reprodução Animal**, v. 29, p. 111–117, 2005.
- ANDRADE, V. J. ; GARCIA, S. K. Padrões raciais e registro de bubalinos. **Revista Brasileira de Reprodução Animal** , Belo Horizonte, v. 29, n. 1, p. 39–45, 2005.
- ASSOCIAÇÃO BRASILEIRA DOS CRIADORES DE BÚFALOS - ABCB. Raças. Disponível em: < <http://www.bufalo.com.br/abcb.html> >. Acesso em: 2 nov. 2015.
- ASPILCUETA, B. R. R. et al. Multiple-trait random regression models for the estimation of genetic parameters for milk , fat , and protein yield in buffaloes. **Journal of Dairy Science**, v. 96, n. 9, p. 5923–5932, 2013.
- ASPILCUETA, B. R. R. et al. Genetic parameters of total milk yield and factors describing the shape of lactation curve in dairy buffaloes **Journal of Dairy Research**, v. 79, p. 60–65, nov 2011.
- ASPILCUETA, B. R. R. et al. Genetic parameters for buffalo milk yield and milk quality traits using Bayesian inference. **Journal of Dairy Science**, v. 93, n. 5, p. 2195–2201, mai 2010.
- BARBOSA, L. et al. Avaliação de características de qualidade da carne de suínos por meio de componentes principais. **Revista Brasileira de Zootecnia**, v. 35, n. 4, p. 1639–1645, ago. 2006.
- BASHIR, K. Environmental Factors Affecting Productive Traits and their Trends in NiliRavi Buffaloes. **Pakistan Journal of Life and Social Sciences**, v. 13, p. 137–144, 2015.

BERNARDES, O. Bubalinocultura no Brasil: situação e importância econômica. **Revista Brasileira de Reprodução Animal**, v. 31, p. 293–298, 2007.

BEZERRA JÚNIOR, J. D. S. et al. Produção de leite, duração da lactação e intervalo de partos em búfalas mestiças murrâh. **Revista Caatinga**, v. 27, n. 2, p. 184–191, 2014.

BIAGIOTTI, D. et al. Uso de estatística multivariada para estudo de caracterização racial em ovinos **Acta Tecnológica**, set. 2014.

CATTILO, G. et al. Effects of age and calving season on lactation curves of milk production traits in Italian Water buffaloes. **Journal Dairy Science**, v.85, p.1298-1306, 2002.

CRUZ, Cosme Damião; CARNEIRO, Pedro Crescêncio Souza. **Modelos biométricos aplicados ao melhoramento genético**. 2. ed. Viçosa: UFV, 2003. 585p.

CRUZ, Cosme Damião; REGAZZI, Adair José.; CARNEIRO, Pedro Crescêncio Souza. **Modelos biométricos aplicados ao melhoramento genético**. 3. ed. Viçosa: UFV, 2004. 480p.

DE ROSA, G. et al. Application of the Welfare Quality protocol to dairy buffalo farms: Prevalence and reliability of selected measures. **Journal of Dairy Science**, v. 98, n. 10, p. 6886–6896, jun. 2015.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS - FAO. Rebanho Mundial. 2014. Disponível em: < http://faostat3.fao.org/browse/Q/*E >. Acesso em: 2 fev. 2015.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS - FAO. Produtividade. 2013. Disponível em: < http://faostat3.fao.org/browse/Q/*E >. Acesso em: 2 fev. 2015.

FLORES, E. B.; KINGHORN, B. P.; VAN DER WERF, J. Predicting lactation yields in dairy buffaloes by interpolation and multiple trait prediction. **Livestock Science**, v. 151, n. 2-3, p. 97–107, 2013.

FRAGA et al. Multivariate analysis to evaluate genetic groups and production traits of crossbred Holstein × Zebu cows. **Tropical Animal Health and Production**, v. 47, p. 1–6, 30 dez. 2015.

GONZÁLEZ, A. et al. Use of Morphometric Variables for Differentiating Spanish Hound Breeds. **International Journal of Morphology**, v. 29, n. 4, p. 1248–1255, 2011.

GUPTA, J. J. et al. diagnostic study on livestock production system in Eastern Region of India. **Indian Journal of Animal Sciences**, v. 84, n. 2, p. 198–203, 2014.

HAIR, J. F. et al. **Análise multivariada de dados - 6ed.** Porto Alegre: Bookman, 2009. 679p.

HOTELLING, H. Analysis of a complex of statistical variables into Principal Components. *Jour. Educ. Psych.*, 24, 417-441, 498-520. **The Journal of Educational Psychology**, v. 24, p. 417–441, 1933.

HUSSON, F. et al. FactoMineR: Multivariate Exploratory Data Analysis and Data Mining. **R package version 1.31.5.** 2016.

INTERNATION COMMITTEE FOR ANIMAL RECORDING - ICAR. **Icar recording guidelines.** Itália, 2014. 619p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. De Geografia E Estatística - Ibge. **Produção da Pecuária Municipal**, v. 42, p. 1 – 39, 2014.

JOLLIFFE. Discarding Variables in a Principal Component Analysis. I: Artificial Data. **Journal of the Royal Statistical Society**, v. 21, n. 1, p. 160–163, 1972.

JOLLIFFE. Discarding Variables in a Principal Component Analysis. II: Real Data. **Journal of the Royal Statistical Society**, v. 22, n. 1, p. 21–31, 1973.

KADEGOWDA, A. K. G.; PIPEROVA, L. S.; ERDMAN, R. A. Principal Component and Multivariate Analysis of Milk Long-Chain Fatty Acid Composition during Diet-Induced Milk Fat Depression. **Journal of Dairy Science**, v. 91, n. 2, p. 749–759, 2 fev. 2008.

KIRKPATRICK, M.; MEYER, K. Direct estimation of genetic principal components: simplified analysis of complex phenotypes. **Genetics**, v. 168, n. 4, p. 2295–2306, 1 dez. 2004.

LÊ, S.; JOSSE, J.; HUSSON, F. FactoMineR : An R Package for Multivariate Analysis. **Journal of Statistical Software**, v. 25, n. 1, p. 1–18, 18 mar. 2008.

LEITE, C. D. S.; TORRES, R. de A. Análise de componentes principais em características de produção de aves de postura. **Revista Brasileira de Zootecnia**, v. 39, n. 2, p. 285–288, 2010.

MALHADO, C. H. M. et al. Inbreeding depression on production and reproduction traits of buffaloes from Brazil. **Animal Science Journal**, v. 84, n. 4, p. 289–295, 2013a.

MALHADO, C. H. M. et al. Genetic parameters for milk yield, lactation length and calving intervals of Murrah buffaloes from Brazil. **Revista Brasileira de Zootecnia**, v. 42, n. 8, p. 565–569, ago. fev. 2013b.

MALHADO, C. H. M. et al. Estimativas de parâmetros genéticos para características reprodutivas e produtivas de búfalas mestiças no Brasil. **Revista Brasileira de Saúde e Produção Animal**, v. 29, n. 4, p. 830–839, out/dez. 2009.

MARQUES, J. R. F.. **Búfalos: o produtor pergunta, a EMBRAPA responde**. Brasília: Embrapa Comunicação para Transferência de Tecnologia, 176p. 2000.

MEIRA, C. T. et al. Seleção de características morfofuncionais de cavalos da raça Mangalarga Marchador por meio da análise de componentes principais. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 65, n. 6, p. 1843–1848, dez. 2013.

MINGOTI, Sueli Aparecida. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2005. 295p.

MOURAND, K. A.; MOHAMED, M. M. Genetic and phenotypic aspects of milk yield traits and reproductive performance of Egyptian buffaloes. **Egypt Journal Animal Production**, v.32, p. 125-137, 1995.

OLIVEIRA, D. P. et al. Principal Components for Reproductive and Productive Traits in Buffaloes from Brazil. in: World Congress of Genetics Applied to Livestock Production, 10th, 2014, Vancouver. p. 3–5.

PAIVA, A. L. C. et al. Análise de componentes principais em características de produção de aves de postura. **Revista Brasileira de Zootecnia**, v. 39, n. 2, p. 285–288, mar. 2010.

PARLATO, E.; VAN VLECK, L. Effect of parentage misidentification on estimates of genetic parameters for milk yield in the Mediterranean Italian buffalo population. **Journal of Dairy Science**, v. 95, n. 7, p. 4059–64, 2012.

PEARSON, K. On lines and planes of closest fit to systems of points in space. **Philosophical Magazine**, v. 2, p. 559–572, 1901.

PEREIRA, R. G. de A. et al. Intervalo de partos em rebanho bubalino no Estado de Rondônia. **EMBRAPA**, Porto Velho - RO. 12p. 2008.

R CORE TEAM. R: A language and environment for statistical computing. Vienna, Austria. **R Foundation for Statistical Computing**, 2015.

- RAMOS, A. et al. Caracterização fenotípica e genética da produção de leite e do intervalo entre partos em bubalinos da raça Murrah. **Pesquisa Agropecuária Brasileira**, v. 41, n. 8, p. 1261–1267, ago. 2006.
- RENCHER, Alvin. **Methods of multivariate analysis**, 2ed. Canadá: Wiley Interscience, 2002.695p.
- REZENDE, M. P. G. et al. Morfometria corporal de equinos utilizados em trabalho, esporte e lazer em três municípios do Mato Grosso do Sul. **Veterinária e Zootecnia**, 6 abr. 2015.
- RODRIGUES, A. E. et al. Estimação de parâmetros genéticos para características produtivas em búfalos na Amazônia Oriental. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 62, n. 3, p. 712–717, jun. 2010.
- ROSATI, A.; VAN VLECK, L. D. Estimation of genetic parameters for milk, fat, protein and mozzarella cheese production for the Italian river buffalo *Bubalus bubalis* population. **Livestock Production Science**, v. 74, n. 2, p. 185–190, mar. 2002.
- SANE, D. D. et al. Studies on Murrah buffalo (*bubalus Bubalis*). **Indian Journal Animal Production**. v.3 p. 61-65, 1972.
- SAHOO, S. K. et al. Comparative evaluation of different lactation curve functions for prediction of bi-monthly test day milk yields in murrah buffaloes. **Animal Science Reporter**. v. 9, n. 3, p. 89–94, jul 2015.
- SAMPAIO NETO, J. C. et al. Avaliação dos desempenhos produtivo e reprodutivo de um rebanho bubalino no Estado do Ceará. **Revista Brasileira de Zootecnia**, v. 30, n. 2, p. 368–373, abr. 2001.
- SANTIAGO, A. A. Búfalos no Brasil. **Associação Brasileira dos Criadores de Búfalos**, p. 13, 2000.
- TÉLLEZ, M. B.; MELÉNDEZ, Y. A.; MARTÍNEZ, A. M. Influencia de la época del parto y región en algunos indicadores reproductivos del búfalo de agua (*bubalus bubalis*) en el territorio oriental de Cuba. **Revista Electrónica de Veterinaria REDVET** *Revista Electrónica de Veterinaria REDVET*, v. 9, 2005.
- TONHATI, H. et al. Estimates of correction factors for lactation length and genetic parameters for milk yield in buffaloes. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 56, n. 2, p. 251–257, abr. 2004.

VENTURA, H. T. et al. Use of multivariate analysis to evaluate genetic groups of pigs for dry-cured ham production. **Livestock Science**, v. 148, n. 3, p. 214–220, out. 2012.

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. 2283–9, 1 set. 2013.

VOHRA, V. et al. Phenotypic Characterization and Multivariate Analysis to Explain Body Conformation in Lesser Known Buffalo (*Bubalus bubalis*) from North India. **Asian-Australasian journal of animal sciences**, v. 28, n. 3, p. 311–7, mar. 2015.

ZICARELLI L, P. E. Genetic Parameters and Sire Ranking for the Adjusted Estimate of Mozzarella Cheese Production in Italian Mediterranean Buffalo. **Advances in Dairy Research**, v. 03, n. 04, out. 2015.