

**ANÁLISE DE COMPONENTES PRINCIPAIS NA
META-ANÁLISE PARA OBTENÇÃO DE EQUAÇÕES
DE PREDIÇÃO DE VALORES ENERGÉTICOS DE
ALIMENTOS PARA AVES**

FLÁVIA CRISTINA MARTINS QUEIROZ

2010

FLÁVIA CRISTINA MARTINS QUEIROZ

**ANÁLISE DE COMPONENTES PRINCIPAIS NA META-ANÁLISE PARA
OBTENÇÃO DE EQUAÇÕES DE PREDIÇÃO DE VALORES
ENERGÉTICOS DE ALIMENTOS PARA AVES**

Dissertação apresentada à Universidade Federal de
Lavras como parte das exigências do Programa
de Pós-Graduação em Estatística e Experimentação
Agropecuária, para a obtenção do título de "Mestre".

Orientador
Prof. Dr. Renato Ribeiro de Lima

LAVRAS
MINAS GERAIS-BRASIL
2010

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Queiroz, Flávia Cristina Martins.

Análise de componentes principais na meta-análise para
obtenção de equações de predição de valores energéticos de
alimentos para aves/ Flávia Cristina Martins Queiroz. -- Lavras:
UFLA, 2010.

71 p. : il.

Dissertação (Mestrado) – Universidade Federal de Lavras, 2010.

Orientador: Renato Ribeiro de Lima.

Bibliografia.

1. CCC. 2. Energia metabolizável. 3. Validação. 4. Grupos
homogêneos. I. Universidade Federal de Lavras. II. Título.

CDD-519.536

FLÁVIA CRISTINA MARTINS QUEIROZ

**ANÁLISE DE COMPONENTES PRINCIPAIS NA META-ANÁLISE PARA
OBTENÇÃO DE EQUAÇÕES DE PREDIÇÃO DE VALORES
ENERGÉTICOS DE ALIMENTOS PARA AVES**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

APROVADA em 23 de Fevereiro de 2010

Prof. Dr. Paulo Borges Rodrigues

UFLA

Prof. Dr. Joel Augusto Muniz

UFLA

Prof. Dr. Denismar Alves Nogueira

UNIFAL-MG

Prof. Dr. Renato Ribeiro de Lima
UFLA
(Orientador)

LAVRAS
MINAS GERAIS-BRASIL

*"Quanto melhor é adquirir sabedoria do que o ouro! E quanto mais excelente
adquirir a prudência do que a prata."*

Provérbios 16:16

"O temor do Senhor é o princípio da sabedoria."

Salmos 111:10

Aos meu pais, Ovídio e Ana Cristina, pelo amor, conselhos, incentivo e apoio em cada instante da minha vida.

Aos meus irmãos, Flaviano e Frederico, pelo carinho, zêlo e motivação em todas as etapas de nossa convivência.

Ao meu noivo Jônatas, pelo amor sincero e por ser sempre meu amigo confidente.

Aos sobrinhos, cunhadas, sogros, demais familiares e amigos, pelo carinho e atenção a mim dedicados.

DEDICO E OFEREÇO

AGRADECIMENTOS

Em primeiro lugar, a Deus, o autor da minha fé, o meu melhor amigo, que sempre iluminou e guiou a minha vida, dando-me forças para vencer e concretizar meus sonhos.

À Universidade Federal de Lavras e ao Departamento de Ciências Exatas, pela oportunidade de cursar o mestrado.

Ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), pela concessão da bolsa de estudos, durante o período de realização deste trabalho.

Ao meu orientador, o professor Dr. Renato Ribeiro de Lima, pelos valiosos ensinamentos, orientação, paciência e prontidão durante o desenvolvimento deste.

Aos professores do programa de pós-graduação em Estatística e Experimentação Agropecuária, pelos ensinamentos.

A todos os funcionários do Departamento de Ciências Exatas, pela educação e bom atendimento dispensados sempre que necessário.

Aos meus pais, imprescindíveis em minha vida, por todo amor, conforto e força que sempre encontro neles.

Ao meu noivo Jônatas, pela companhia e compreensão em cada momento único. Te amo!

Aos meus irmãos Flaviano e Frederico, e às minhas cunhadas, pelo eterno amor fraternal.

Às paixõezinhas da titia, Lury, Ana Laura e Samuel, por me proporcionarem momentos de felicidade.

À melhor amiga Stefani, por me ouvir e me aconselhar sempre.

À nova família Marcon Mariano, pela boa acolhida e carinho a mim de-

monstrados.

Às parceiras de apartamento, Nádia, Fran, Marília e Jessica, pelo companheirismo e compreensão até nos momentos mais críticos.

Às amigas do “Clube da Luluzinha” - Fran, Rosi, Simone e Deyse - pela amizade, parceria e afeto e, também, aos demais colegas de curso, por compartilharmos conhecimentos e experiências acadêmicas.

Ao amigo Fábio, pelo auxílio com a rotina no software R.

Aos professores doutores Paulo Borges Rodrigues, Joel Augusto Muniz, Renato Ribeiro de Lima e Denismar Alves Nogueira, membros da banca avaliadora, pela disponibilidade e interesse em colaborar com o desenvolvimento deste trabalho científico.

Ao tio Adinaldo, pelas palavras de incentivo e preocupação constante com meus estudos.

A todos familiares, amigos e àqueles que, de alguma forma, contribuíram e acreditaram no meu potencial para a realização deste sonho.

Obrigada a todos. Deus os abençoe!

SUMÁRIO

	Página
LISTA DE TABELAS	i
LISTA DE FIGURAS	ii
RESUMO	iii
ABSTRACT	iv
1 INTRODUÇÃO	1
2 REFERENCIAL TEÓRICO	3
2.1 Meta-Análise	3
2.1.1 Surgimento e história da meta-análise	3
2.1.2 Objetivos e consequências da meta-análise	5
2.1.3 Procedimentos em uma meta-análise	6
2.1.4 Importância e Problemas da meta-análise	14
2.2 Análise de Componentes Principais	16
2.3 Regressão linear e estimação de parâmetros	19
2.3.1 Seleção de variáveis para o modelo de regressão	20
2.4 Equações de Predição	21
2.4.1 Equações de Predição dos valores energéticos de alimentos para aves	21
2.4.2 Coeficiente de Determinação	26
2.4.3 Critério de Akaike (<i>AIC</i>)	27
2.5 Técnicas de Validação	28
2.5.1 Uso da regressão na validação	29
2.5.1.1 Coeficiente de Correlação de Condordância (CCC)	32
3 MATERIAL E MÉTODOS	36
3.1 Dados a serem utilizados na obtenção das equações de predição	36
3.2 Obtenção das equações de predição utilizando componentes principais	37

3.3 Seleção e validação de equações de predição da EMAn	40
3.3.1 Equações selecionadas	40
3.3.2 Validação das equações de predição	40
3.4 Rotina utilizada nas análises	41
4 RESULTADOS E DISCUSSÕES	44
4.1 Equações de predição para EMAn selecionadas, obtidas neste estudo .	44
4.2 Comparação de equações de predição via validação com ensaio meta- bólico	52
4.2.1 Alimentos energéticos e concentrados	52
4.2.2 Soja + subprodutos e concentrados	56
5 CONCLUSÕES	59
REFERÊNCIAS BIBLIOGRÁFICAS	60
ANEXOS	65

LISTA DE TABELAS

	Página
1	Números de informações cadastradas no presente estudo, para realização da meta-análise, em função da composição química e valor energéticos dos alimentos de origem vegetal. 37
2	Variáveis explicativas das equações de predição para EMAn de alimentos energéticos encontradas na literatura e seus respectivos R^2 42
3	Variáveis explicativas das equações de predição obtidas para EMAn de soja + subprodutos encontradas na literatura e seus respectivos R^2 43
4	Valores dos critérios de seleção e validação das equações de predição para energéticos obtidas neste trabalho 47
5	Variáveis explicativas das equações de predição obtidas que apresentaram melhores ajustes com o uso de diferentes números de grupos e seus respectivos R^2 48
6	Equações de predição para energia metabolizável (EMAn) de alimentos energéticos 50
7	Equações de predição para energia metabolizável (EMAn) de soja + subprodutos 51
8	Valores dos critérios de seleção e validação das equações de predição para alimentos energéticos 55
9	Valores dos critérios de seleção e validação das equações de predição para soja + subprodutos 58

LISTA DE FIGURAS

	Página
1 Principais etapas de uma meta-análise.	14
2 Reta de concordância perfeita entre as variáveis X e Y.	31
3 Valores dos (a) coeficiente de determinação e (b) critério de Akaike determinados no ajuste das equações de predição para EMAN de alimentos energéticos, obtidas neste trabalho, em função do número de grupos formados	45
4 (a) <i>P-valores</i> de β_0 e β_1 e (b) valores do coeficiente de correlação de concordância determinados na validação das equações de predição para EMAN de alimentos energéticos, obtidas neste trabalho, em função do número de grupos formados	45
5 Valores dos (a) coeficiente de determinação, (b) <i>p-valores</i> de β_0 e β_1 e (c) coeficiente de correlação de concordância determinados na avaliação das equações de predição para EMAN de alimentos energéticos	53
6 Valores dos (a) coeficiente de determinação, (b) <i>p-valores</i> de β_0 e β_1 e (c) coeficiente de correlação de concordância determinados na avaliação das equações de predição para EMAN de soja + subprodutos	57

RESUMO

QUEIROZ, Flávia Cristina Martins. **Análise de componentes principais na meta-análise para obtenção de equações de predição de valores energéticos de alimentos para aves.** 2010. 71 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, MG. *

Neste estudo, foi proposta a utilização da análise de componentes principais, na formação de grupos homogêneos de trabalhos, a serem considerados na meta-análise. Esses grupos foram utilizados na obtenção de equações de predição para energia metabolizável aparente corrigida pelo balanço de nitrogênio (EMAn) de alimentos para aves, como função da composição química desses alimentos (proteína bruta-PB; extrato etéreo-EE; matéria mineral-MM; fibra bruta-FB; fibra em detergente neutro-FDN e fibra em detergente ácido-FDA). Para a aplicação da meta-análise, foram considerados dados de diversos trabalhos já publicados sobre o assunto. Assim, as equações de predição obtidas, juntamente com outras propostas na literatura, foram validadas, visando identificar aquelas que apresentaram melhores ajustes. As equações propostas neste trabalho apresentaram resultados semelhantes às disponíveis na literatura. Porém, a formação de grupos homogêneos de trabalhos, que é um dos grandes problemas na meta-análise, foi facilitada com a utilização da técnica de componentes principais.

Palavras-chave: CCC, energia metabolizável, validação, grupos homogêneos.

* Orientador: Renato Ribeiro de Lima - UFLA

ABSTRACT

QUEIROZ, Flávia Cristina Martins. **Principal components analysis in meta-analysis to obtain prediction equations of energetic values of poultry feedstuffs**. 2010. 71 p. Dissertation (Master in Statistics and Agricultural Experimentation) - Federal University of Lavras, Lavras, MG.*

In this study, it was proposed the use of principal components analysis to obtain homogeneous groups of papers in the development of the meta-analysis. These groups were used to obtain prediction equations of corrected apparent metabolizable energy (AMEn) of poultry feedstuffs, as function of the chemical composition (CP-crude protein; EE-etherial extract; MM-ash, CF-crude fiber; NDF-neutral detergent fiber; ADF-acid detergent fiber). In the process of the meta-analysis, we considered data from several papers published previously. The prediction equations obtained and other proposed equations in the literature were validated in order to identify those that had better adjustments. The equations proposed in this study presented similar results to those available in the literature. However, the formation of homogeneous groups of papers, which is a major problem in meta-analysis, was facilitated with the use of principal components.

Key-words: CCC, metabolizable energy, validation, homogeneous groups.

* Adviser: Renato Ribeiro de Lima - UFLA

1 INTRODUÇÃO

Devido à diversidade de alimentos e seus subprodutos utilizados na formulação de rações, surge a necessidade de se conhecerem bem seus valores nutritivos e energéticos, para melhor aproveitamento e utilização dos mesmos. Por isso, o sucesso na formulação de rações de mínimo custo requer um conhecimento preciso da composição química e dos valores de energia metabolizável (EM) dos alimentos, o que possibilita um fornecimento adequado de energia para os animais. Porém, a determinação do valor da EM não é de fácil execução, além de demandar tempo e ser de custo elevado. No entanto, conforme descrito por Nascimento et al. (2009), existe uma forma rápida de se determinar esse valores, que consiste na utilização de equações de predição, estabelecidas em função da composição química dos mesmos, normalmente de fácil e rápida obtenção.

Existem à disposição na literatura vários trabalhos de pesquisa, onde se estabeleceram equações de predição para os valores energéticos de uma série de alimentos. Mas os resultados obtidos, utilizando-se essas equações não têm sido satisfatórios. Assim, uma maneira interessante de se trabalhar com os vários resultados obtidos é utilizar procedimentos estatísticos, para combinar resultados de estudos distintos, mas relacionados, o que é definido como meta-análise (Kirby, 1993).

Na meta-análise, embora os estudos envolvam um mesmo assunto, é importante que grupos mais homogêneos possíveis sejam formados, visando controlar a heterogeneidade existente entre eles. Essa variabilidade existente pode ser modelada e incorporada na análise. Com isso, essa metodologia permite a combinação de diferentes resultados e determina, inicialmente, conclusões mais confiáveis e informativas. Porém, após a obtenção de tais equações, é imprescindível que

essas equações sejam validadas através de ensaios metabólico e de desempenho, para serem mais concretas e confiáveis perante a comunidade científica (Tedeschi, 2006).

Neste trabalho, foi proposta a utilização da análise de componentes principais, que é uma técnica multivariada, na formação de grupos homogêneos de trabalhos, visto que a formação de grupos é um dos principais problemas encontrados na meta-análise. Estes grupos, que visam controlar a variabilidade existente entre estudos, foram utilizados na obtenção de equações de predição para energia metabolizável aparente corrigida pelo balanço de nitrogênio (EMAn) de alimentos para aves. Para a aplicação da meta-análise, utilizou-se os dados catalogados por Nascimento (2007), sem levar em consideração a maneira com que ele agrupou os dados em seu trabalho. Além disso, efetuou-se a validação das equações de predição obtidas neste estudo e de outras propostas na literatura, visando identificar aquelas que apresentaram melhores ajustes. Para tal, os dados utilizados na validação foram provenientes de ensaios metabólicos conduzidos por Alvarenga (2009).

2 REFERENCIAL TEÓRICO

2.1 Meta-Análise

2.1.1 Surgimento e história da meta-análise

No cenário mundial, nota-se que a produção científica teve um aumento progressivo nas últimas décadas. No entanto, no Brasil, essa evolução veio acontecendo de forma mais lenta, pois nem todas as áreas seguiram esse comportamento mundial. Algumas áreas se desenvolveram mais rapidamente, outras estão se desenvolvendo agora.

Segundo a Produção... (2010), recentemente a produção científica brasileira ultrapassou a da Rússia, antiga potência na área, caminha para superar também a da Índia e se consolidar como a 2ª maior entre os BRICs (Brasil, Rússia, Índia e China). Entre 1990 e 2008, a produção brasileira avançou de 3.665 para 30.021 artigos científicos. Neste mesmo período, a produção russa manteve-se estável. Além disso, levantamentos feitos indicam ainda que a produção científica brasileira é mais forte em áreas como: pesquisas agrícolas e ciências naturais.

Em decorrência desse aumento na produtividade científica, devido à busca de se acompanhar o desenvolvimento tecnológico vigente no mundo, torna-se mais difícil a seleção e análise da qualidade da literatura. É devido a esse problema que a evolução do conhecimento continuou se desenvolvendo e novas técnicas de análises de trabalhos puderam ser desenvolvidas, como é o caso da meta-análise.

O principal interesse na condução de uma meta-análise é considerar os resultados apresentados por um grande número de estudos similares, referentes a uma determinada área de conhecimento, mas conduzidos em situações diferentes. Isso incentiva o desenvolvimento de métodos que combinem informações provenientes desses estudos. Essa síntese de informações de dados coletados sob dife-

rentes condições, e até mesmo diferentes precisões, para produzir conclusões mais abrangentes não é recente.

Um dos primeiros trabalhos que se interessou pelo problema de combinar resultados de vários experimentos independentes, foi desenvolvido por Cochran em 1954. Eysenck, em 1952, concluiu em um estudo que a psicoterapia, que nada mais é que o tratamento de algum distúrbio psíquico, não tinha nenhum efeito benéfico, surgindo com isso uma polêmica entre os profissionais dessa área. Mas só depois de 20 anos de pesquisa e centenas de estudos, tentando resolver esse dilema que, Smith & Glass em 1977, considerando 375 pesquisas em psicoterapia, conseguiram concluir que psicoterapia tem efeitos sim e provaram, então, que Eysenck estava errado (Lovatto et al., 2007). O método utilizado por eles para combinar resultados de diferentes estudos era conhecido como meta-análise.

A meta-análise pode ser definida como um procedimento estatístico que consiste de uma revisão quantitativa e resumida de resultados de estudos distintos, mas relacionados (Glass, 1976), isto é, a meta-análise é a Análise das análises. Conforme Kirby (1993) enunciou, a meta-análise é o uso de métodos estatísticos para combinar ou comparar resultados de estudos distintos, mas relacionados. Esses métodos estatísticos asseguram a obtenção de uma estimativa combinada e precisa, sobretudo em virtude do aumento do número de observações e, conseqüentemente, do poder estatístico e da possibilidade de examinar a variabilidade entre os estudos (Fagard et al., 1996).

Na pesquisa científica é crescente o número de artigos similares conduzidos e publicados em diversas áreas do conhecimento, gerando interesse, muitas vezes, de realizar uma síntese desses resultados (Fagard et al., 1996). Assim, a aplicação da meta-análise tem aumentado em vários campos das ciências, tais como: Agronomia (Martins, 2001), Ciências Sociais (Glass, 1976), Ecologia (Gu-

revitch & Hedges, 1993), Medicina (Martinez, 2007; Meneses-Gaya et al., 2009), Medicina Veterinária (Peters et al., 2000), Melhoramento Animal (Giannotti et al., 2002, 2005) e Nutrição Animal (Rodrigues et al., 2007; Nascimento et al., 2009; Souza, 2009).

A utilização do processo meta-analítico, em trabalhos da área animal tem aumentado nesses últimos anos. Segundo Lovatto et al. (2007), este tema foi considerado como nova metodologia, na 44ª Reunião Anual da Sociedade Brasileira de Zootecnia, em 2007, que é a principal reunião científica da área.

De acordo com Lovatto et al. (2007), a partir de 1980, as publicações científicas internacionais (indexadas PubMed), em se tratando de meta-análise, têm evoluído exponencialmente, passando de 23 em 1980 para 3.700 em 2006. Já no cenário nacional (indexadas Scielo), isso não foi visto, sendo que os dois primeiros artigos sobre o tema surgiram em 1999 e, chegando a apenas 8 publicações em 2006.

Na história da meta-análise, o cenário desde 1980 tem sido marcado pelo interesse crescente pela mesma. Conforme Cooper (1990), isso pode ser comprovado, pois ele disse que: Apesar de críticas e problemas enfrentados pela meta-análise, as evidências indicam que ela é um procedimento estatístico, que está sendo aceito gradativamente pela comunidade científica e sua aplicação tem aumentado em todos os campos da ciência.

2.1.2 Objetivos e consequências da meta-análise

Segundo Lovatto et al. (2007), um dos objetivos da meta-análise é a obtenção de novos resultados, a partir da síntese de *n* observações distintas, porém relacionadas entre si. Em outras palavras, o interesse da meta-análise é encontrar uma resposta única, uma medida resumida, para diversos estudos distintos. No en-

tanto, o reagrupamento de vários tratamentos de n reduzido não garante resultados conclusivos, para tal é desejável que os artigos utilizados tenham um delineamento experimental correto, sobretudo com relação às repetições.

Com o novo resultado obtido pela meta-análise, pode-se evidenciar um efeito de um tratamento que, individualmente, não permite estabelecer conclusões por falta de potência analítica (baixo n). Assim, conseqüentemente, o poder analítico é melhorado, aumentando as chances de identificar diferenças entre os tratamentos, se as mesmas existirem. Além disso, a meta-análise tem por objetivo sintetizar resultados contraditórios, possibilitando a generalização de conclusões para uma amplitude variada de estudos, assim como, resolver incertezas, quando certas conclusões se destoam (Lovatto et al., 2007).

Fagard et al. (1996), também relata que outros propósitos da meta-análise são os de realizar a análise de subgrupos, examinar a variabilidade entre os ensaios ou estudos, identificar a necessidade e planejar ensaios ou estudos maiores, além de responder questões que não foram propostas, de início, nos estudos individuais.

2.1.3 Procedimentos em uma meta-análise

A meta-análise consiste inicialmente de uma revisão de literatura. Porém, pode ser considerada como superior às formas tradicionais de revisão de literatura. Isso se deve ao fato de, além de se efetuar uma revisão, combinam-se resultados de diferentes estudos, ajustando-os para a heterogeneidade experimental, obtendo-se assim, resultados mais precisos e acurados.

Segundo Lovatto et al. (2007), para que o critério da meta-análise seja utilizado, algumas etapas são necessárias:

- a) definição do objetivo do trabalho, do tema da pesquisa;

São os objetivos que determinam todo o processo meta-analítico. Nesta etapa, define-se a maneira com que a meta-análise será conduzida, seja por combinação de estimativas, por estimação de tamanho de efeito, por combinação de testes de hipóteses ou até por estimação de equações de regressão.

b) sistematização das informações;

A meta-análise exige disciplina no processo de sistematizar um conjunto de dados obtidos principalmente da literatura científica. Por isso, é importante definir com que tipo e quais trabalhos devem ser incluídos na base de dados. A busca por esses trabalhos pode ser feita em bases digitais, como, por exemplo, no portal Periódicos Capes, avaliando-se um por um, para que sejam definidos quais serão incluídos no banco de dados a ser trabalhado. Deve-se ter o cuidado em limitar a pesquisa bibliográfica no espaço (amplitude geográfica, como nacionais, internacionais) e no tempo (período de inclusão dos artigos). Também, deve-se tomar o cuidado em não incluir somente trabalhos de um mesmo grupo de pesquisadores, se esse não for o foco do trabalho de meta-análise. Outra importante consideração é a possibilidade ou até mesmo a necessidade de se incluírem resultados não publicados. Sabe-se que é dada preferência para a publicação de resultados significativos, em detrimento aos não significativos, o que gera um vício de publicação. Assim, a inclusão destes pode ser importante. É necessária a definição de critérios pré-determinados, para inclusão e exclusão de artigos. Contudo, esses critérios não devem ser imutáveis ao longo do estudo. Eles devem ser desenvolvidos e adaptados, à medida que se interage com a literatura.

c) codificação dos dados;

Após a sistematização da base de dados, é importante codificar os mesmos

por objetivo experimental anunciado e avaliado em cada publicação. Essas codificações são indispensáveis para interpretar os dados, considerando os vários objetivos experimentais dos dados agrupados na base.

d) filtragem dos dados;

A filtragem dos dados dessa base escolhida e codificada é muito importante e se desenvolve em três fases.

A primeira dessas fases é verificar e ter certeza de que a publicação candidata a entrar na base final tenha coerência com os objetivos do trabalho. Esse tipo de trabalho é feito por pessoas que tenham conhecimento suficiente do tema e que devem explorar a publicação com uma leitura crítica, observando eventuais erros que devem ser anotados. Se a publicação entrar na base, o segundo passo é transferir seus dados para a base de dados. Por fim, se a mesma não apresentar um comportamento anormal, em relação às já transferidas, ela continua na base de dados, caso contrário, ela deve ser deixada de lado, mas sem ser eliminada definitivamente, pois pode ser que, posteriormente, exista a necessidade de utilizá-la no decorrer do trabalho.

e) análise dos dados incluídos na base;

Esta análise deve ser feita tanto gráfica, quanto estatisticamente. Nesta análise gráfica de dados, pode-se fazer uma visualização e identificação rápida das informações e relações importantes. A análise gráfica possibilita o estabelecimento de hipóteses e a escolha de um modelo estatístico, a verificação da existência de linearidade, de interações, localização de zonas com mesmas variações, etc. Já na análise estatística dos dados, diferentes procedimentos devem ser aplicados, para se conhecer a estrutura dos experimentos para cada característica explicativa. Assim, é importante considerarem-se as variações

intra-experimentos e inter-experimentos.

f) formação de grupos

Embora os estudos a serem combinados envolvam o mesmo assunto, muitas vezes, há diferenças entre eles, o que se denomina de heterogeneidade entre os estudos, sendo este um dos problemas correntes em meta-análise. Dando continuidade ao processo meta-analítico é importante agrupar os estudos do banco de dados, os mais homogêneos possíveis, visto que são experimentos independentes. Sendo assim, várias técnicas podem ser utilizadas nessa etapa.

Como são analisados resultados obtidos em trabalhos distintos, mas de uma forma conjunta, nesta análise deve-se considerar o efeito de diferentes fatores que interferem diretamente na variabilidade dos resultados, como: época do ano e ano de realização do trabalho, local de execução do experimento, idade e sexo dos animais utilizados, número de repetições e metodologia utilizada na determinação da variável resposta, etc. Quanto mais homogêneos forem os resultados a serem combinados, mais confiável será a estimativa combinada.

Assim, um método para a formação dos grupos seria considerar esses fatores que influenciam na variabilidade dos resultados e, assim, definir grupos de estudos que tenham em comum os mesmos níveis dos fatores, conforme foi feito por Nascimento et al. (2009), que trabalharam com a predição de valores energéticos de alimentos. Neste caso específico, os resultados são influenciados por alguns fatores como: sexo e idade dos animais utilizados nos experimentos e metodologia, para se determinar a energia metabolizável em ensaios biológicos, dentre outros. Assim, tem-se interesse de que os trabalhos sejam agrupados de acordo com a combinação dos níveis desses fatores.

A técnica multivariada de análise de agrupamento é uma das maneiras de

se obter os grupos homogêneos, pois esta técnica possibilita reunir os estudos em questão, em um determinado número de grupos, de tal forma que exista grande homogeneidade dentro de cada grupo e heterogeneidade entre os grupos (Johnson & Wichern, 2002). Giannotti et al. (2005) utilizaram a análise de agrupamento, o que permitiu a formação de grupos de resultados semelhantes, fornecendo, assim, estimativas combinadas de herdabilidades para grupos específicos.

g) ponderações;

Uma possível solução para o problema da variabilidade existente entre os estudos dos experimentos agrupados seria incorporar essa variabilidade na análise. Para verificar a existência de variabilidade, testes de homogeneidade são executados e, baseando-se nos seus resultados, opta-se por modelar a variância existente entre os estudos, quando há ausência de homogeneidade ou por combiná-los sem considerar a variância entre eles. Assim, é conveniente se utilizar algum fator de ponderação, definindo diferentes pesos a vários grupos de trabalhos e executar a análise de todo o conjunto de estudos. Portanto, a definição dos diferentes grupos, formados por estudos mais homogêneos, é um aspecto importante na meta-análise. Essa ponderação pode ser, no modelo linear geral, por exemplo, o inverso da variância do grupo mais homogêneo de experimentos, podendo se fazer uma normalização em busca de uma mesma escala. No entanto, outros fatores de ponderações podem ser testados e utilizados.

h) escolha de um modelo estatístico;

A etapa da escolha de um modelo estatístico é importante para que a análise dos dados, que envolve a combinação de vários estudos, seja feita de forma

o mais simples possível. Essa análise geralmente envolve a combinação de estimativas (como diferenças médias entre grupos de tratamento e de controle, ou coeficientes de correlação, entre outras), a combinação de resultados de testes de hipóteses (combinando-se os *p-valores*, por exemplo) ou a combinação de equações de regressões lineares estimadas em diferentes estudos.

A análise de regressão pode ser usada para se determinar a relação entre estudos com características contínuas e estimativas do tamanho do efeito (Giannotti, 2000).

De acordo com Hedges & Olkin (1985), considere uma série de n estudos independentes, em que o i -ésimo estudo produz uma estimativa $\hat{\theta}_i$ do tamanho do efeito θ_i , $i = 1, \dots, n$. O termo *tamanho do efeito*, do inglês *effect size*, é bastante utilizado em meta-análise e refere-se, no caso de experimentos, à diferença padronizada entre médias. Para o caso de um modelo linear geral, pode-se assumir que a diferença padronizada entre médias, θ_i , para o i -ésimo experimento é dependente de um vetor de variáveis fixas (variáveis independentes), \mathbf{X}_i , tal que $\mathbf{X}_i = [X_{i1} \ X_{i2} \ \dots \ X_{ip}]$. Assim, considerando n estudos independentes, pode-se definir

$$\begin{aligned}\theta_1 &= \beta_1 X_{11} + \beta_2 X_{12} + \dots + \beta_p X_{1p}, \\ \theta_2 &= \beta_1 X_{21} + \beta_2 X_{22} + \dots + \beta_p X_{2p}, \\ &\vdots \\ \theta_n &= \beta_1 X_{n1} + \beta_2 X_{n2} + \dots + \beta_p X_{np},\end{aligned}$$

em que $\beta_1, \beta_2, \dots, \beta_p$ são coeficientes de regressão desconhecidos.

Portanto, assume-se um modelo de regressão linear,

$$\boldsymbol{\theta} = \mathbf{X}\boldsymbol{\beta}, \quad (1)$$

em que $\boldsymbol{\theta}$ é um vetor coluna, tal que $\boldsymbol{\theta}' = [\theta_1 \ \theta_2 \ \dots \ \theta_n]$, \mathbf{X} é uma matriz de delineamento $n \times p$, assumida como não tendo colunas linearmente dependentes e $\boldsymbol{\beta}$ é um vetor coluna $p \times 1$, com os coeficientes de regressão, isto é, $\boldsymbol{\beta}' = [\beta_1 \ \beta_2 \ \dots \ \beta_p]$.

Partindo-se do modelo apresentado em (1) e, considerando o raciocínio utilizado na análise de regressão por mínimos quadrados ordinários, pode-se definir o modelo

$$\hat{\boldsymbol{\theta}} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2)$$

em que $\boldsymbol{\varepsilon} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}$ é um vetor de erros. Uma vez que $\boldsymbol{\varepsilon} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}$, a distribuição de $\boldsymbol{\varepsilon}$ é aproximadamente normal multivariada, com vetor de médias $\mathbf{0}$ e matriz de covariância diagonal $\boldsymbol{\Sigma}$, dada por

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{\theta_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{\theta_2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{\theta_n}^2 \end{bmatrix}.$$

Assim, os elementos de $\boldsymbol{\varepsilon}$ são independentes, mas não identicamente distribuídos.

De acordo com Hedges & Olkin (1985) não é possível estimar $\boldsymbol{\beta}$, a partir do modelo dado em (2), porque $\boldsymbol{\Sigma}$ depende do vetor de parâmetros desconhecidos $\boldsymbol{\theta}$. Porém, pode-se obter $\hat{\boldsymbol{\beta}}$ utilizando $\hat{\boldsymbol{\Sigma}}$, a qual é obtida substituindo-se $\sigma_{\theta_i}^2$, na matriz $\boldsymbol{\Sigma}$, por $\hat{\sigma}_{\theta_i}^2$.

Assim, o estimador de mínimos quadrados generalizados de θ , $\hat{\theta}$, é dado por

$$\hat{\beta} = \left(\mathbf{X}' \hat{\Sigma}^{-1} \mathbf{X} \right)^{-1} \mathbf{X}' \hat{\Sigma}^{-1} \hat{\theta}. \quad (3)$$

Se o tamanho amostral for suficientemente grande, podem-se aplicar testes de hipóteses ou obter intervalos de confiança para componentes do vetor β , considerando que $\hat{\beta}$ tenha distribuição aproximadamente normal. Porém, é bastante trabalhoso. Uma alternativa mais simples de se estimar β e aplicar testes de hipóteses ou obter intervalos de confiança, consiste em efetuar as análises considerando mínimos quadrados ponderados.

Numa análise por mínimos quadrados ponderados, definindo-se \mathbf{X} como a matriz de delineamento, \mathbf{y} como o vetor de observações e \mathbf{W} como uma matriz diagonal com os pesos, tem-se que o estimador de mínimos quadrados ponderados de β_W , partindo-se do modelo linear $\mathbf{y} = \mathbf{X}\beta$, é dado por

$$\hat{\beta}_W = \left(\mathbf{X}' \mathbf{W} \mathbf{X} \right)^{-1} \mathbf{X}' \mathbf{W} \mathbf{y}. \quad (4)$$

Comparando as equações (3) e (4), observa-se que $\hat{\beta}$ é um caso particular de $\hat{\beta}_W$, apenas substituindo a matrix \mathbf{W} , em (4), por $\hat{\Sigma}^{-1}$ e \mathbf{y} por $\hat{\theta}$.

Para finalizar o processo meta-analítico, é importante que o resultado final seja analisado e, estando o mesmo corretamente desenvolvido, este deve ser disponibilizado à comunidade científica.

Na Figura 1 são apresentadas todas essas etapas descritas anteriormente.

incorporados na análise, a correlação existente entre as estimativas a serem combinadas e o vício de publicação ou viés causado pela dificuldade de acesso a todos os artigos sobre o assunto, incluindo nesse grupo os trabalhos que não foram publicados por algum motivo. Este último, juntamente com a heterogeneidade dos estudos, podem ser considerados como principais preocupações em um estudo de meta-análise. Portanto, quanto mais abrangente for a revisão de literatura sobre o tema em estudo, menor será o vício de publicação. Esse viés de publicação também é decorrente da prática de muitos periódicos aceitarem mais facilmente a publicação de estudos que mostraram efeitos benéficos, que podem acarretar medidas superestimadas do efeito do tratamento em estudo.

Lovatto et al. (2007) descrevem sobre esses e vários outros problemas, sugerindo algumas formas de contorná-los.

Outro problema na meta-análise é que a amostra utilizada não é aleatória e a população não é identificada na maioria das vezes. Costa (1999) considera esse problema como População de Referência. Contudo, Glass et al. (1981) defendem a meta-análise, dizendo que “maçãs e laranjas podem ser combinadas no estudo de frutas”, basta ser coerente na combinação e efetuarla corretamente. Para esse problema, Costa (1999) sugere dois caminhos: no primeiro, que se tenha uma meta-análise com estudos individuais bastante homogêneos, em que as amostras sejam aleatórias e, no segundo, se as amostras não forem aleatórias, a meta-análise tem que incluir, além do tamanho do efeito e dos *p-valores*, também, o tamanho da amostra.

A formação de grupos homogêneos é um outro grande problema da meta-análise. Essa formação pode ser feita por combinação de níveis de fatores, por análises de agrupamento, etc.

2.2 Análise de Componentes Principais

A técnica denominada de análise de componentes principais, popularmente chamada de ACP, foi introduzida por Karl Pearson em 1901 e está fundamentada no artigo de Hotelling (1933). ACP é um método da Estatística Multivariada, que tem como objetivos gerais a redução e a interpretação de dados (Johnson & Wichern, 2002).

Segundo Ferreira (2008), o método consiste na transformação do conjunto original de variáveis em outro, os componentes principais (CP), em que o número de componentes é sempre igual ao número de variáveis. Cada componente retém porcentagem da variância original, sendo os mesmos extraídos na ordem do mais explicativo para o menos explicativo, isto é, o primeiro componente principal (CP1) possui a máxima variância, tendo o segundo componente principal (CP2) a segunda maior variância e, assim, sucessivamente. Entretanto, alguns dos primeiros componentes são responsáveis por grande parte da explicação total e, conseqüentemente, uma grande parte dessa variabilidade pode ser explicada por um número r menor de componentes, $r \leq p$. Apesar da variância explicada por cada fator ir decrescendo ao longo dos componentes, a porcentagem total acumulada de todos é obviamente 100%.

De acordo com Johnson & Wichern (2002), se grande parte da variabilidade total, mais de 70%, pode ser explicada por um, dois ou três componentes principais, então esses componentes poderão substituir as variáveis originais, sem grande perda de informações.

De acordo com Ferreira (2008), essa técnica pode ser aplicada na análise de regressão múltipla, principalmente, nos casos de colinearidade ou de multicolinearidade; aplica-se, também, à análise de agrupamento e como estimadores de fatores nas técnicas multivariadas denominadas de análises fatoriais.

Os CP representam combinações lineares de p variáveis aleatórias X_1, X_2, \dots, X_p , com médias $\mu_1, \mu_2, \dots, \mu_p$ e variâncias $\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2$, respectivamente. Essas variáveis são correlacionadas e, portanto, existe covariância não-nula entre a i -ésima e k -ésima variável definida por σ_{ik} , para $i \neq k = 1, 2, \dots, p$. Assim, o desenvolvimento dos CP depende apenas da matriz de covariância Σ (ou da matriz de correlação ρ) e de X_1, X_2, \dots, X_p . Além disso, não são necessárias pressuposições de normalidade multivariada (Ferreira, 2008).

Seja o vetor aleatório $\mathbf{X}^t = [X_1 \ X_2 \ \dots \ X_p]$ amostrado de uma população com matriz de covariância Σ , cujos autovalores são $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Para considerar a transformação dos elementos da matriz \mathbf{X} em um novo conjunto de variáveis, que seja dois a dois não correlacionados e que o primeiro tenha a maior variância possível, o segundo a segunda maior variância possível e, assim por diante, define-se o i -ésimo componente principal Y_i por:

$$Y_i = \mathbf{a}_i' \mathbf{X} = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p,$$

sendo que o vetor desconhecido \mathbf{a}_i estabelece a i -ésima combinação linear, para $i = 1, 2, \dots, p$.

Conforme Ferreira (2008), Y_i é uma variável latente, isto é, não é mensurável a partir do experimento ou levantamento amostral. O objetivo é determiná-la a partir das p variáveis contidas no vetor \mathbf{X} , a partir da projeção dos pontos coordenados originais em um plano, maximizando a distância entre eles. Isto é equivalente a maximizar a variabilidade da variável latente Y_i . A variância de Y_i é dada por:

$$Var(Y_i) = Var(\mathbf{a}_i' \mathbf{X}) = \mathbf{a}_i' \Sigma \mathbf{a}_i$$

e a covariância entre Y_i e $Y_k (i \neq k)$, por:

$$Cov(Y_i, Y_k) = Cov(\mathbf{a}'_i \mathbf{X}, \mathbf{a}'_k \mathbf{X}) = \mathbf{a}'_i \boldsymbol{\Sigma} \mathbf{a}_k.$$

A definição dos CP baseia-se na maximização de sua variância $Var(Y_i) = \mathbf{a}'_i \boldsymbol{\Sigma} \mathbf{a}_i$ com relação ao vetor \mathbf{a}_i , sujeito à restrição de $\mathbf{a}'_i \mathbf{a}_i = 1$. Esta restrição é imposta porque o máximo da variância do CP não existe devido ao fato que, a medida em que os componentes do vetor de coeficiente \mathbf{a}_i crescem, a variância crescerá para infinito (Ferreira, 2008). As restrições garantem que o sistema tenha solução única e que os CP sejam não correlacionadas e apresentem variância máxima (Johnson & Wichern, 2002).

Assim, a técnica de multiplicadores de Lagrange é utilizada, fazendo

$$\max_{\mathbf{a}_i} [\mathbf{a}'_i \boldsymbol{\Sigma} \mathbf{a}_i - \lambda_i (\mathbf{a}'_i \mathbf{a}_i - 1)],$$

com relação a \mathbf{a}_i , sendo λ_i o multiplicador de Lagrange.

No entanto, tal restrição pode ser imposta dividindo-se a $Var(Y_i)$ por $\mathbf{a}'_i \mathbf{a}_i$, e a maximização é feita da seguinte maneira:

$$\lambda_i = \max_{\mathbf{a}_i} \frac{\mathbf{a}'_i \boldsymbol{\Sigma} \mathbf{a}_i}{\mathbf{a}'_i \mathbf{a}_i}.$$

Em ambos casos acima, o máximo é obtido a partir da derivação da função resultante em relação a \mathbf{a}_i e igualando o sistema de equações a zero, resultando em:

$$(\boldsymbol{\Sigma} - \lambda_i \mathbf{I}) \mathbf{a}_i = \mathbf{0}, \quad (5)$$

sendo λ_1 uma matriz identidade $p \times p$. A demonstração dessa maximização é en-

contrada em Johnson & Wichern (2002) e Ferreira (2008).

A partir do sistema de equações (5), tem-se

$$Var(Y_i) = \mathbf{a}'_i \Sigma \mathbf{a}_i = \mathbf{a}'_i \lambda_i \mathbf{a}_i = \lambda_i \mathbf{a}'_i \mathbf{a}_i = \lambda_i,$$

e também,

$$Cov(Y_i, Y_k) = \mathbf{a}'_i \Sigma \mathbf{a}_k = \mathbf{a}'_i \lambda_k \mathbf{a}_k = \lambda_k \mathbf{a}'_i \mathbf{a}_k = \mathbf{0},$$

com $i \neq k$, sendo que \mathbf{a}_i e \mathbf{a}_k são ortogonais.

Com isso, conclui-se que a definição dos componentes principais é, na verdade, a obtenção dos autovalores e autovetores da matriz Σ , em que os autovetores $\mathbf{a}_i (i = 1, 2, \dots, p)$ representam o sentido de rotação dos eixos coordenados definidos pelas variáveis originais e os autovalores, $\lambda_i (i = 1, 2, \dots, p)$, definem as variâncias dos novos eixos coordenados (Ferreira, 2008).

Então, tomando $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, tem-se definido os componentes principais $Y_1 = \mathbf{a}'_1 \mathbf{X}$, $Y_2 = \mathbf{a}'_2 \mathbf{X}$, ..., $Y_p = \mathbf{a}'_p \mathbf{X}$, respectivamente.

2.3 Regressão linear e estimação de parâmetros

A análise de regressão tem o objetivo de verificar a existência de uma relação funcional entre uma variável com uma ou mais variáveis, obtendo-se uma equação que explique a variação da variável dependente Y , pela variação dos níveis das variáveis independentes X_1, X_2, \dots, X_p .

Esta relação linear pode ser expressa pelo modelo:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i \quad (6)$$

em que os β' s são os parâmetros, e ε_i é o erro aleatório, tal que $\varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$, para todo $i = 1, 2, \dots, n$.

O processo de estimação dos parâmetros das equações de predição pode ser feito por vários métodos. Os dois mais comuns são: método da máxima verossimilhança e o método dos mínimos quadrados ordinários. Este último é sem dúvida o mais amplamente adotado. Caso esteja trabalhando com modelo de regressão linear ponderada, a estimação dos parâmetros pode ser obtida pelo método dos mínimos quadrados ponderados. A descrição detalhada dos métodos de mínimos quadrados, seja o ordinário ou ponderado, pode ser encontrada em Draper & Smith (1998) e Hoffman & Vieira (1998).

2.3.1 Seleção de variáveis para o modelo de regressão

Pelo princípio da parcimônia, o melhor modelo é aquele que apresenta o menor número de variáveis independentes, sendo que explique quase a totalidade da variabilidade da variável dependente. Para avaliar a importância das variáveis independentes sobre o valor da variável dependente é fundamental verificar a significância das variáveis independentes no modelo completo. A verificação desta significância é feita de acordo com algum critério estatístico como, por exemplo, a estatística F, a estatística de *AIC*, *p-valor*, R^2 , entre outros.

Contudo, a escolha do melhor modelo pode ser feita com a aplicação do método *stepwise*. Este método é uma mistura entre os métodos *backward* e *forward*, os quais consistem na retirada e inclusão de variáveis significativas no modelo respectivamente. Considerando um modelo de regressão, em cada passo adiciona-se ou retira-se uma variável preditora X ao modelo. Caso alguma variável independente desse novo modelo não seja significativa, a mesma pode ser retirada. O procedimento chega ao final quando não há mais variável a ser incluída

ou descartada. Maiores detalhes sobre este método pode ser visto em Draper & Smith (1998).

2.4 Equações de Predição

A equação de predição é uma importante ferramenta comumente utilizada para o estabelecimento do valor de uma variável dependente de outras, ou seja, é uma regressão. Ela permite estimar o valor desta variável, num determinado momento do tempo, como uma função de outras variáveis.

2.4.1 Equações de Predição dos valores energéticos de alimentos para aves

A energia, que é o produto gerado pela transformação dos nutrientes durante o metabolismo, é um dos principais fatores a serem considerados na nutrição animal (Nascimento, 2007). Além de ser indispensável para as funções vitais das células, a energia tem participação importante na regulação do consumo, sendo ela que permite que o animal ganhe peso, se desenvolva bem, renda lucros etc. Fischer Júnior et al. (1998) relatam que, entre os nutricionistas, é consenso que a energia é um dos fatores limitantes do consumo, sendo utilizada nos mais diferentes processos.

Visando atender adequadamente às exigências nutricionais dos animais para a expressão do máximo potencial produtivo, é preciso a formulação de rações equilibradas. Sendo assim, é fundamental obter dados precisos dos valores energéticos e da composição química dos alimentos a serem fornecidos aos animais, o que acarreta em sucesso na formulação de rações para aves e, conseqüentemente, no fornecimento adequado de energia (Albino et al., 1992).

Apesar da estimação do valor nutritivo dos alimentos ocorrer inicialmente em meados do século XIX, foi só no século seguinte que os primeiros trabalhos

que estimavam a digestibilidade dos alimentos foram desenvolvidos. Contudo, existem várias formas para expressar a energia nos alimentos, como energia bruta (EB), energia digestível (ED), energia metabolizável (EM) e energia líquida (EL). Dentre elas, a forma mais utilizada para aves é a EM. Segundo Hill & Anderson (1958), foi em 1940 que teve início a estimação do valor energético dos alimentos sob a forma de EM. No entanto, segundo Nascimento (2007) descreve, na determinação da EM é comum fazer a correção dos valores energéticos para o balanço de nitrogênio (BN), visto que o nitrogênio retido no corpo, se catabolizado, é excretado na forma de compostos contendo energia, tal como o ácido úrico. Assim, foi proposto um valor de correção de 8,22 kcal por grama de nitrogênio (Hill & Anderson, 1958). Com isso, alguns autores defendem o uso da energia metabolizável aparente corrigida pelo balanço de nitrogênio (EMAn) (Wolynetz & Sibbald, 1984).

Diante de tantas variações na diversidade de alimentos e seus subprodutos, comumente utilizados na formulação de rações de mínimo custo, não é seguro para as indústrias utilizarem os valores de tabelas. Com isso, muitos métodos têm sido desenvolvidos, visando obter uma melhor estimação do valor energético dos alimentos. Estes podem ser classificados como métodos diretos e indiretos. Nos métodos diretos, utiliza-se o animal e mede-se a diferença entre energia consumida e energia excretada, que é a energia digestível. Já nos métodos indiretos, utilizam-se equações de predição as quais são geradas a partir de análises químicas simples como fibra bruta, extrato etéreo, proteína bruta, etc.

A vantagem de se determinar equações de predição da energia metabolizável, em função da composição química dos alimentos, decorre de que é bem mais fácil realizarem-se análises laboratoriais para determinação das composições químicas, do que efetuar bioensaios com os animais e ficar medindo suas excretas

em calorímetros. Além disso, a maioria dos laboratórios, indústrias e estações de pesquisa do país nem sempre possuem calorímetros e mesmo que possuíssem, a execução desses ensaios biológicos é dispendiosa e demorada. Nesse caso, trabalhar com equações geradas a partir de análises químicas simples pode auxiliar o nutricionista (Albino et al., 1992).

Existem à disposição na literatura nacional alguns trabalhos estabelecendo equações para a predição de valores energéticos de uma série de alimentos. Utilizando os valores de EMAn determinados em ensaios metabólicos, pelo método tradicional de coleta total de excretas com pintos em crescimento, foram obtidas e aplicadas em validações equações para alimentos energéticos (Rodrigues et al., 2001; Nagata et al., 2004) e para alimentos protéicos (Rodrigues et al., 2002; Zonta et al., 2004), com o objetivo de se predizer a energia como função da composição química dos alimentos. Estas foram obtidas sem a proposta de utilização da meta-análise na estimação dos parâmetros das equações de predição para EMAn.

As equações obtidas por Rodrigues et al. (2001) foram estabelecidas para predizer os valores energéticos dos alimentos milho e seus subprodutos, em função de sua composição química. Contudo, as equações de predição ajustadas com as variáveis fibra em detergente neutro (FDN) ou fibra bruta (FB) e matéria mineral (MM) foram as indicadas a serem utilizadas para predizer os valores energéticos do milho e de seus subprodutos, sendo $EMAn = 4281,6 - 39,97FDN - 72,90MM$ e $R^2 = 96\%$ ou $EMAn = 4354,8 - 112,05FB - 151,74MM$, com $R^2 = 95\%$. Em outro trabalho desenvolvido por Rodrigues et al. (2002), foram estabelecidas equações para predizer os valores energéticos dos alimentos soja e seus subprodutos, em função de sua composição química. Neste último, a equação com as variáveis fibra bruta (FB) e extrato etéreo (EE) foi por eles indicada como boa estimadora para valores energéticos da soja e subprodutos foi

$$EMAn = 2822,2 - 90,13FB + 49,96EE, \text{ com } R^2 = 93\%.$$

Nagata et al. (2004) determinaram a EMAn de alguns alimentos energéticos alternativos ao milho (gérmen de milho, quirera de milho, milho moído e em grão e sorgo moído e em grão) e a EMAn de sete híbridos de milho. Entre as equações por eles estudadas, puderam concluir que a equação $EMAn = 4021,8 - 227,55MM$, com $R^2 = 92\%$ é a mais indicada para prever valores de EMAn dos alimentos estudados. Zonta et al. (2004) definiram em seu trabalho, que as equações $EMAn = -822,33 + 69,54PB - 45,26FDA + 90,81EE$, com $R^2 = 92\%$ e $EMAn = 2723,05 - 50,52FDA + 60,40EE$, com $R^2 = 90\%$, foram as mais indicadas para prever valores de EMAn dos alimentos protéicos por eles estudados, sendo cinco amostras de farelos de soja de diferentes marcas comerciais e três amostras de soja integral (extrusada, tostada e micronizada).

Nascimento (2007) estabeleceu equações de predição de valores energéticos de alimentos empregados na formulação de rações para aves, utilizando o princípio da meta-análise. Ele definiu equações de predição da energia metabolizável aparente corrigida pelo balanço de nitrogênio (EMAn), em função da composição química dos alimentos concentrados de origem vegetal (energéticos e protéicos), usualmente utilizados na formulação de rações para aves, considerando diferentes grupos de alimentos. Este autor subdividiu os trabalhos em sete categorias: alimentos energéticos, milho, milho+subprodutos, alimentos protéicos, soja, soja + subprodutos e concentrados de origem vegetal, que consiste de todos os trabalhos utilizados (energéticos + protéicos). No entanto, a formação de grupos mais homogêneos possíveis para o desenvolvimento da meta-análise, foi definida a partir da combinação dos níveis dos três fatores, que influenciavam na variabilidade da EMAn: metodologia empregada no ensaio metabólico, sexo e idade dos animais.

As equações de Nascimento que melhor se ajustaram para prever EMAn,

em função da composição química dos alimentos energéticos comumente utilizados na alimentação de frangos de corte, foram $EMAn = 4371,18 - 26,48PB + 30,65EE - 126,93MM - 52,26FB - 25,14FDN + 24,40FDA$ e $EMAn = 4205,23 + 30,58EE - 130,35MM - 58,29FB - FDN + 16,71FDA$, ambas com $R^2 = 81\%$. Já as equações que melhor se ajustaram para prever a EMAn, em função da composição química dos alimentos protéicos comumente utilizados na alimentação de frangos de corte, foi $EMAn = 2707,71 + 58,63EE - 16,06FDN$, com $R^2 = 81\%$, as quais foram obtidas ao considerarem os alimentos soja + subprodutos. Ao considerar os alimentos concentrados energéticos e protéicos, Nascimento obteve duas equações que melhor se ajustaram na determinação da EMAn, as quais são $EMAn = 4101,33 + 56,28EE - 232,97MM - 24,86FDN + 10,42FDA$, com $R^2 = 84\%$ e $EMAn = 4095,41 + 56,84EE - 225,26MM - 22,24FDN$, com $R^2 = 83\%$.

Em geral, vários trabalhos estabelecendo equações de predição para energia metabolizável de alimentos foram desenvolvidos e apresentaram equações aparentemente bem ajustadas aos dados originais. No entanto, ao serem testados com dados independentes, nem todas as equações responderam satisfatoriamente. Com isso, vem a necessidade de se trabalhar com equações distintas para cada alimento ou para grupos de alimentos semelhantes. Estas equações podem ser coletadas sob diferentes condições e diferentes níveis de precisão. Assim, a combinação das mesmas gera resultados mais coerentes que aquelas disponíveis em cada fonte de informação (Alvarenga, 2009). Por isso, uma boa alternativa é utilizar a meta-análise, para apresentar de uma forma resumida e mais precisa, os vários resultados obtidos.

2.4.2 Coeficiente de Determinação

Para a seleção das melhores equações ajustadas um dos métodos utilizados na verificação da qualidade dos ajustes é o coeficiente de determinação (R^2), o qual é uma razão de soma de quadrados, dado pela expressão:

$$R^2 = \left(\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right) = \left(\frac{\text{SQRegressão}}{\text{SQTotal}} \right) = 1 - \left(\frac{\text{SQResíduo}}{\text{SQTotal}} \right),$$

em que:

y_i = valor real de cada observação;

\hat{y}_i = valor estimado de cada observação;

\bar{y} = média aritmética real;

n = número de observações.

O coeficiente de determinação (R^2) indica o quanto da variabilidade total dos dados é explicada pelo modelo de regressão (Draper & Smith, 1998), ou seja, a quantidade de variação da variável dependente que é explicada pelas variáveis independentes. Este coeficiente pode variar entre $0 \leq R^2 \leq 1$. Se a correlação entre as variáveis dependentes e independentes fosse perfeita, resultaria em $R^2 = 1$. Portanto, na comparação de modelos, aquele que apresentar o maior valor de R^2 é considerado melhor.

Como R^2 é inflacionado pela adição de variáveis explanatórias e, diante da necessidade de comparar equações com diferentes números de parâmetros, utiliza-se o coeficiente de determinação ajustado (R_a^2), o qual pondera o R^2 pelos graus de liberdade da regressão (numerador) e dos graus de liberdade totais (denominador).

O R^2 ajustado é calculado pela seguinte expressão matemática:

$$R_a^2 = 1 - (1 - R^2) \left(\frac{n - 1}{n - p} \right)$$

em que

R_a^2 é o coeficiente de determinação ajustado;

n é o número de observações;

p é o número de coeficientes do modelo;

R^2 é o coeficiente de determinação.

O R_a^2 não tem a mesma interpretação do R^2 e só é utilizado na comparação de modelos, sendo que o modelo que apresenta o maior valor de R_a^2 é considerado o melhor.

2.4.3 Critério de Akaike (*AIC*)

O critério de Akaike utilizado em seleção de modelos é designado, em inglês, pela sigla *AIC* - “Akaike’s Information Criterion” (Akaike, 1973), foi desenvolvido a partir da distância ou informação de Kulback & Leibler (1951). Esta distância é uma medida de discrepância entre as linhas do modelo verdadeiro e o modelo aproximado. Akaike (1973) relacionou a distância de Kulback & Leibler (1951) com a máxima verossimilhança, surgindo o *AIC*. O *AIC* é dado pela seguinte expressão:

$$AIC = -2\ln L + 2K,$$

onde:

L é o valor máximo da função de verossimilhança do modelo candidato;

K é o número de parâmetros deste modelo (incluindo o intercepto).

No critério de Akaike (1983), o modelo com menor valor de AIC é aquele que minimiza a distância de Kulbach-Leibler, por consequência, pode ser considerado como aquele mais próximo do modelo real (hipotético). Assim sendo, a partir de uma série de modelos concorrentes escolhe-se aquele que apresentar o menor valor de AIC .

2.5 Técnicas de Validação

Modelos são representações da realidade e, portanto, eles nunca podem descrever em sua totalidade a realidade em todas as condições propostas. Porém, independentemente do seu objetivo, um modelo funcional é uma aproximação do sistema real (Tedeschi, 2006). Sendo a validação efetuada para avaliar em que medida um modelo é racional e cumpre seus propósitos.

Assim, após o desenvolvimento teórico, com a obtenção das equações de predição de uma forma mais precisa, faz-se necessária a validação dos resultados obtidos através de ensaios metabólicos e de desempenho. Com isso, verifica-se se a metodologia empregada para a obtenção das equações é, realmente, mais precisa do que aquelas utilizadas até então. Portanto, é imprescindível que as mesmas sejam validadas, para serem mais concretas e confiáveis perante a comunidade científica (Nascimento, 2007). Afinal, nenhum estudo compara as equações estimadas com valores determinados posteriormente, ou seja, não há validação delas em condições diferentes daquelas em que foram desenvolvidas.

Segundo Tedeschi (2006), a validação de modelos só é possível através da combinação de vários critérios estatísticos e de uma adequada investigação sobre os fins para o qual ele foi inicialmente concebido e desenvolvido. A utilização de apenas um critério pode acarretar na seleção inadequada em um determinado

cenário. Decorre então que a escolha da técnica é importante, pois algumas contêm problemas e inconsistências. Adicionalmente, a avaliação subjetiva pode ser útil como um guia. Contudo, com a obtenção de vários modelos para um mesmo tema, a validação e seleção destes dá-se através de análises gráficas e de critérios quantitativos. Alguns desses critérios serão descritos a seguir.

2.5.1 Uso da regressão na validação

Um modelo linear utiliza a relação entre duas ou mais variáveis quantitativas (ou qualitativas) de tal forma que uma variável pode ser predita a partir da outra ou outras. Pode ser definido como um modelo de regressão linear simples ou múltipla.

Conforme visto em Hoffman & Vieira (1998), o modelo de regressão linear simples consiste de uma função linear

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad (7)$$

com $i = 1, 2, \dots, n$, em que:

Y_i é o i -ésimo valor da variável resposta;

β_0 e β_1 são os parâmetros (coeficientes de regressão);

X_i é o i -ésimo valor da variável preditora (é uma constante conhecida, fixo);

ε_i é o termo do erro aleatório com $E(\varepsilon_i) = 0$ e $Var(\varepsilon_i) = \sigma^2$.

No entanto, para se fazer inferência sobre os parâmetros do modelo e verificar a qualidade do ajuste, Hoffmann & Vieira (1998) citam quais os pressupostos devem ser considerados ao executar uma regressão linear:

a) A relação entre X e Y é linear;

b) Os valores de X são fixos, isto é, X não é uma variável aleatória;

c) $\varepsilon \sim N(0, \sigma^2)$.

Sendo assim, para se testar a normalidade, o teste de Shapiro-Wilk é frequentemente usado para garantir uma distribuição normal dos dados, no caso de pequenas amostras, caso contrário, recomenda-se utilizar o teste de Kolmogoroff-Smirnov (Tedeschi, 2006).

Para estimar os parâmetros do modelo de regressão linear simples, pode-se utilizar o método dos mínimos quadrados. Contudo, o interesse nesta análise está na relação entre duas variáveis como, por exemplo, X e Y . Essa regressão linear entre os valores observados (Y) e estimados (X) é comumente usada como método de validação do modelo. A hipótese é que a regressão passe pela origem e tenha uma inclinação de unidade (Tedeschi, 2006), ou seja, que ela coincida com a reta de concordância perfeita (bissetriz), conforme Figura 2. Por isso, as estimativas dos parâmetros β_0 e β_1 são testadas sob as hipóteses:

$$\left\{ \begin{array}{l} H_0 : \beta_0 = 0 \\ H_1 : \beta_0 \neq 0 \end{array} \right. \quad \text{e} \quad \left\{ \begin{array}{l} H_0 : \beta_1 = 1 \\ H_1 : \beta_1 \neq 1 \end{array} \right.$$

Os testes de hipóteses sobre os dois parâmetros podem ser feitos de maneira usual pela distribuição t de Student (Draper & Smith, 1998).

Segundo Bussab & Morettin (2003), o p -valor é a probabilidade de ocorrer valores da estatística mais extremos do que o observado, sob a hipótese de H_0 ser verdadeira, isto é,

$$p\text{-valor} = 2 \times P(t > |t^*|)$$

em que t^* é o valor calculado da estatística do teste t .

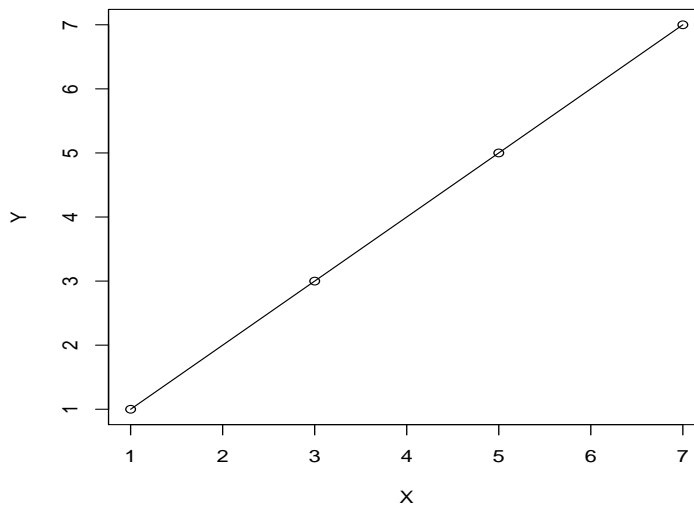


FIGURA 2 Reta de concordância perfeita entre as variáveis X e Y.

Em outras palavras, o *p-valor* é o menor nível de significância para não rejeitar a hipótese nula. Normalmente, considera-se um *p-valor* de 0,05 como base para avaliar a hipótese nula. Se o *p-valor* for inferior ao nível de significância adotado rejeita-se a hipótese nula. Por exemplo, se o pesquisador fixar $\alpha = 0,05$, então, para um *p-valor* $< 0,05$ rejeita-se H_0 , caso contrário, não se rejeita H_0 .

Procurando disponibilizar um resultado preciso e confiável perante a comunidade e demonstrar a importância do processo de validação, Alvarenga (2009) realizou a validação dos resultados obtidos pelas equações de predição obtidas por Nascimento (2007), comparando suas estimativas de EMAn com os resultados de EMAn obtidos por meio de ensaios de digestibilidade.

No entanto, existem várias preocupações na utilização de um modelo de regressão linear. Segundo Mitchell (1997), a regressão não é sensível o suficiente para quantificar o quão bom é a linha, pois os pressupostos raramente são satis-

feitos; uma reta linear ajustada é simplesmente o melhor resumo de uma relação linear entre a amostra de pontos previstos por pares de predições e observações, e não, necessariamente, a confirmação da acurácia dos dados; testes de hipótese de nulidade dão resultados ambíguos, dependendo da dispersão dos dados. Infelizmente, o teste t para a inclinação tem uma interpretação ambígua, porque quanto maior a dispersão dos dados, maior é o erro padrão da inclinação, menor é o valor calculado para o teste estatístico e, portanto, mais difícil é de rejeitar a hipótese nula, que indica o declive é igual à unidade. Portanto, o teste pode deixar de rejeitar a hipótese nula, quer porque a inclinação não é realmente diferente da unidade ou há muita dispersão em torno da linha (Harrison, 1990).

Por isso, é imprescindível que seja considerado mais de um método no processo de validação de equações, pois conforme dito anteriormente, a utilização de apenas alguma técnica pode selecionar um modelo inadequado em um determinado assunto. Neste trabalho, recomenda-se o uso concomitante da análise de regressão linear simples, com o coeficiente de correlação de concordância.

2.5.1.1 Coeficiente de Correlação de Concordância (CCC)

Quando técnicas estatísticas são aplicadas em determinadas circunstâncias, muitas vezes técnicas similares podem conduzir a resultados bem divergentes. Por isso, é preciso utilizar aquelas mais apropriadas para a situação em questão. Alguns procedimentos são utilizados para medir a qualidade da informação em questão, a partir da técnica de validação. Dentre estes, destaca-se o coeficiente de correlação de concordância (CCC), o qual é adequado para variáveis contínuas e avalia a concordância entre duas variáveis, como por exemplo, a concordância entre a EMAn de alimentos predita por uma equação e a EMAn deste mesmo alimento observada em ensaios metabólicos.

Um ponto importante a ser abordado na validação do modelo é se os modelos de previsão de valores são precisos e exatos (acurácia) ao mesmo tempo. Com o coeficiente de correlação de concordância (CCC) proposto por Lin (1989), é possível verificar simultaneamente a acurácia e a precisão de um modelo, isto é, se as variáveis independentes de um modelo estão correlacionadas e concordantes entre si.

O CCC é uma medida do quanto a linha de regressão dos dados observados, dados como função dos preditos, desvia da linha de concordância perfeita, isto é, a linha de 45 graus em um gráfico de dispersão. Para isso, ele combina dois componentes: uma medida de precisão, o coeficiente de correlação de Pearson, com uma medida de acurácia (C_b). Assim, se o valor do CCC é igual a um, significa que a reta de regressão está exatamente sobre a reta de concordância perfeita. O CCC é obtido pela expressão:

$$\rho_c = \frac{2\sigma_{xy}}{\sigma_x^2 + \sigma_y^2 + (\mu_x - \mu_y)^2} = \rho C_b, \quad (8)$$

$$\text{com } C_b = \frac{2}{\nu + \frac{1}{\nu} + u^2}, \quad \nu = \frac{\sigma_x}{\sigma_y} \quad e \quad u = \frac{(\mu_x - \mu_y)}{\sqrt{\sigma_x \sigma_y}},$$

em que:

ρ_c é o coeficiente de correlação de concordância;

ρ é a medida de precisão, expressa pelo coeficiente de correlação linear de Pearson;

σ_{xy} é a covariância entre x e y;

σ_x é o desvio padrão da variável x;

σ_y é o desvio padrão da variável y;

C_b é a medida de acurácia, que expressa a proximidade entre a melhor reta ajustada

tada aos dados e a bissetriz do primeiro quadrante: $0 < C_b \leq 1$;

ν é a medida de afastamento entre a melhor reta ajustada aos dados e a bissetriz;

u é a medida de locação relativa à medida de afastamento;

μ_x e μ_y representam as médias populacionais, para as variáveis X e Y respectivamente.

O fator C_b varia de 0 a 1. Quando C_b é igual a 1, indica que não ocorreu nenhum desvio a partir da linha $Y = X$ e como o fator C_b não inclui a estimativa de ρ como uma componente, tem-se a mesma acurácia, independentemente do grau de precisão, quando a média e a variância são iguais.

As propriedades de ρ_c , como descrito por Lin (1989) são:

- a) $-1 \leq -|\rho| \leq \rho_c \leq |\rho| \leq 1$,
- b) $\rho_c = 0$ se, e somente se, $\rho = 0$,
- c) $\rho_c = \rho$ se, e somente se $\sigma_1 = \sigma_2$ e $\mu_1 = \mu_2$ e
- d) $\rho_c = \pm 1$ se, e somente se $(\mu_1 - \mu_2)^2 + (\sigma_1 - \sigma_2)^2 + 2\sigma_1\sigma_2(1 \pm \rho) = 0$, ou equivalentemente $\rho = \pm 1$, $\sigma_1 = \sigma_2$ e $\mu_1 = \mu_2$, ou equivalentemente, cada par estiver em concordância perfeita (1,1), (2,2), ..., (n,n) ou em reverso perfeito (n,1), (n-1,2), ..., (1,n).

Segundo Lin (1989), o coeficiente de correlação linear representa a medida de precisão o qual, no contexto da concordância, é corrigido pela medida de acurácia. A medida de precisão (ρ) avalia os desvios dos dados em relação à melhor reta ajustada aos dados e à medida de acurácia (C_b) avalia o quanto a melhor reta ajustada aos dados diverge da bissetriz do primeiro quadrante. Portanto, quando ρ tende a 1 (ou a -1), os dados encontram-se próximos à melhor reta ajustada; quando C_b tende a 1, significa que a melhor reta está próxima da bissetriz. Por-

tanto, para existir concordância, é necessário que os dados tenham perfeita relação linear e os valores sejam semelhantes, em todo seu conjunto.

Filgueira (2005) concluiu que o coeficiente de correlação de concordância, apresenta características que o faz ser uma técnica bastante útil no estudo de reprodutibilidade de medidas. Portanto, o coeficiente de correlação de concordância pode ser utilizado como alternativa para avaliar, por exemplo, modelos de equações de predição, visto que o CCC avalia os dados em todo o seu conjunto, sendo, deste modo, bastante confiável.

3 MATERIAL E MÉTODOS

3.1 Dados a serem utilizados na obtenção das equações de predição

Os dados a serem utilizados na obtenção das equações de predição de valores energéticos de alimentos utilizados para aves de corte foram catalogados por Nascimento (2007). Esses dados referem-se a valores de energia metabolizável aparente corrigida pelo balanço de nitrogênio (EMAn) e composição química dos alimentos concentrados energéticos e protéicos, de origem vegetal, usualmente utilizados na formulação das dietas avícolas.

Nessa catalogação foi feita uma ampla e minuciosa revisão bibliográfica de modo a incluir o máximo possível de estudos realizados no Brasil sobre o assunto, nos últimos 40 anos que antecederam 2007. De acordo com Nascimento (2007), foram pesquisados anais de congressos e simpósios, trabalhos não indexados, bibliotecas e base de dados catalogados nos periódicos CAPES, à exemplo do CAB Abstracts, dentre outros. A revisão de literatura foi conduzida, procurando-se obter o máximo de informações e minimizar os erros que pudessem influenciar na análise dos dados, o que é uma necessidade da meta-análise.

Para catalogação das informações, foram considerados neste estudo, os trabalhos que apresentavam na composição química dos alimentos: proteína bruta (PB), extrato etéreo (EE), matéria mineral (MM), fibra bruta (FB), fibra em detergente neutro (FDN) e fibra em detergente ácido (FDA). Esses trabalhos foram classificados como trabalhos com dados completos.

Conforme visto anteriormente, Nascimento (2007) obteve equações de predição, considerando diferentes tipos de alimentos, ou seja, os trabalhos foram subdivididos nas sete categorias: alimentos energéticos, milho, milho+subprodutos, alimentos protéicos, soja, soja + subprodutos e concentrados de origem vegetal,

que consiste de todos os trabalhos utilizados (energéticos + protéicos). Alvarenga (2009) optou por verificar a validação das equações obtidas por Nascimento que apresentaram melhores ajustes, ou seja, quando foram considerados alimentos energéticos, soja + subprodutos e todos alimentos concentrados de origem vegetal. Assim como Alvarenga (2009), optou-se neste estudo por obter e validar apenas equações de predição para EMAn para essas mesmas três categorias de alimentos.

Com isso, para a obtenção das equações de predição para estimar a EMAn dos alimentos foram utilizados diferentes números de trabalhos nas três categorias de alimentos consideradas neste estudo, o que pode ser visto na Tabela 1.

TABELA 1 Números de informações cadastradas no presente estudo, para realização da meta-análise, em função da composição química e valor energéticos dos alimentos de origem vegetal.

	Nº de informações
ENERGÉTICOS ¹	197
SOJA + SUBPRODUTOS ²	67
CONCENTRADOS ³	293

¹ Alimentos energéticos usualmente utilizados nas rações avícolas.

² Soja e seus subprodutos usualmente utilizados nas rações avícolas.

³ Alimentos concentrados usualmente utilizados nas rações avícolas.

3.2 Obtenção das equações de predição utilizando componentes principais

Como na meta-análise é de fundamental importância que a variabilidade entre os estudos seja considerada e incorporada nas análises, devem ser definidos diferentes grupos, formados por estudos mais homogêneos. Diferentes maneiras de formação de grupos podem ser encontradas na literatura, porém neste trabalho foi proposta a formação de grupos utilizando-se a técnica multivariada de componentes principais. Assim, procurando-se controlar a heterogeneidade entre estudos, foram conduzidas várias e diferentes meta-análises, diferenciando-se entre si

pelo número de trabalhos contidos nos grupos.

Em cada uma das meta-análises os grupos continham uma porcentagem dos trabalhos, em que essa porcentagem foi definida com base em uma análise de componentes principais. Estes componentes foram obtidos de acordo com a variabilidade em relação às variáveis EMAn, PB, EE, MM, FB, FDN e FDA.

Para cada um dos trabalhos científicos selecionados de acordo com o tipo de alimento em análise, os valores correspondentes ao primeiro componente principal (PC1) foram calculados e utilizados na formação dos grupos. Neste estudo, foram desenvolvidas várias meta-análises para cada uma das três categorias de alimento considerado, utilizando: 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 25, 30, 40 e 50 grupos homogêneos de trabalhos em cada uma. Além das equações obtidas em cada uma dessas 14 situações, foi considerada, também, a situação sem a formação de grupos. Nesta última situação, não foi utilizado um fator de ponderação e esta foi executada apenas para efeito de comparação com as demais. Assim, foram desenvolvidas 15 meta-análises para alimentos energéticos, 15 para soja + subprodutos e outras 15 para concentrados.

Os passos desde a formação de grupos até a obtenção das equações de predição, em cada meta-análise, estão descritos a seguir:

- (i) efetuar a análise de componentes principais, obtendo-se o valor do PC1 para cada um dos trabalhos a serem considerados na meta-análise;
- (ii) ordenar os valores referentes aos PC1;
- (iii) definir o número de grupos;
- (iv) estabelecer os quantis. Por exemplo: definindo-se 5 grupos, tem-se os quantis 0,20; 0,40; 0,60 e 0,80. Abaixo do quantil 0,20 estão os 20% dos trabalhos com os menores escores de PC1, que compõem o primeiro grupo. Entre o

quantil 0,20 e 0,40 estão os próximos 20% dos trabalhos com menores valores de PC1 superiores ao do primeiro grupo, compondo, assim, o segundo grupo. Esse mesmo procedimento repete-se para a definição dos demais grupos.

- (v) identificar os trabalhos dentro de cada grupo estabelecido;
- (vi) calcular a variância da EMAn observada nos trabalhos de cada grupo;
- (vii) obter o fator de ponderação, que neste trabalho foi definido como o inverso da variância dentro de cada grupo;
- (viii) obter a equação de regressão estimada, utilizando o método dos mínimos quadrados ponderados, sendo os fatores de ponderação os pesos.

Para a predição da EMAn dos alimentos foi proposto um modelo no qual a EMAn é dada como função das variáveis de composição química (PB, EE, MM, FB, FDN e FDA), considerando seus efeitos lineares. Nesse modelo (6), Y_i refere-se ao valor da EMAn do alimento, determinado em ensaio metabólico, no i -ésimo estudo; X_{i1}, \dots, X_{i6} , representam as variáveis de composição química do alimento, no i -ésimo estudo, sendo respectivamente PB, MM, EE, FB, FDN e FDA e ε_i é o erro associado na i -ésima observação, assumido normal e independentemente distribuído, com média 0 e variância σ_i^2 .

Na obtenção das equações de predição utilizou-se o método *stepwise*, sendo o critério de Akaike (*AIC*) utilizado na inclusão e exclusão de variáveis no modelo. Após a aplicação do *stepwise*, se alguma variável foi não significativa a 10%, pelo teste t, esta foi retirada do modelo.

Para efeito de seleção das equações obtidas que apresentaram melhores ajustes, utilizou-se, simultaneamente, o valor de *AIC* com os valores dos critérios de validação de cada meta-análise efetuada.

3.3 Seleção e validação de equações de predição da EMAn

3.3.1 Equações selecionadas

Para o processo de validação foram selecionadas equações de predição da EMAn de alimentos para aves recomendadas pela literatura, as quais encontram-se em Rodrigues et al. (2001, 2002), Nagata et al. (2004), Zonta et al. (2004) e Nascimento (2007). As equações selecionadas, juntamente com seus respectivos valores de R^2 , estão apresentadas nas Tabelas 2 e 3. Além da validação das equações escolhidas na literatura, também validaram-se as equações de predição que apresentaram melhores ajustes neste estudo, que serão apresentadas posteriormente.

Assim, para efeito de comparação, foram confrontados os resultados das equações selecionadas, neste estudo e na literatura, considerando os valores dos critérios de validação de cada uma.

3.3.2 Validação das equações de predição

O processo de validação das equações selecionadas, consistiu basicamente na comparação dos valores de EMAn observados em ensaios metabólicos, conduzidos por Alvarenga (2009), com os valores preditos por tais equações.

A verificação da existência de diferença significativa entre os valores observados e preditos, foi realizada, ajustando-se o modelo de regressão linear simples (7), em que Y_i é a EMAn do alimento i obtida no ensaio metabólico, X_i é o valor obtido utilizando-se uma equação de predição e e_i é o erro tal que $e_i \sim N(0, \sigma^2)$. Aplicou-se o teste t, ao nível de significância de 5%, para verificar se as hipóteses $H_0 : \beta_0 = 0$ e $H_0 : \beta_1 = 1$ são verdadeiras. Assim, nos casos de não-rejeição de ambas as hipóteses de nulidades, têm-se a similaridade entre valores preditos e observados.

Outro critério de validação considerado foi o coeficiente de correlação de concordância entre os valores observados e os valores preditos de EMAn, visando verificar quais modelos apresentavam resultados mais acurados e precisos. Para isso, a expressão (8) foi utilizada.

3.4 Rotina utilizada nas análises

A rotina computacional utilizada neste trabalho foi desenvolvida no software livre R (R Development Core Team, 2009) e está apresentada no Anexo A.

TABELA 2 Variáveis explicativas das equações de predição para EMAn de alimentos energéticos encontradas na literatura e seus respectivos R^2

Equação(*)	Fonte	Variáveis explicativas	R^2 (%)
1	Nascimento (2007)	PB, EE, MM, FB, FDN e FDA	81,00
2		EE, MM, FB, FDN e FDA	81,00
3		EE, MM, FDN e FDA	84,00
4		EE, MM e FDN	83,00
5	Rodrigues et al. (2001)	FDN e MM	96,00
6		FB e MM	95,00
7	Nagata et al. (2004)	MM	92,00

42

(*) Equações obtidas:

Eq. 1: $4371, 18 - 26, 48PB + 30, 65EE - 126, 93MM - 52, 26FB - 25, 14FDN + 24, 40FDA$

Eq. 2: $4205, 23 + 30, 58EE - 130, 35MM - 58, 29FB - 28, 31FDN + 16, 71FDA$

Eq. 3: $4101, 33 + 56, 28EE - 232, 97MM - 24, 86FDN + 10, 42FDA$

Eq. 4: $4095, 41 + 56, 84EE - 225, 26MM - 22, 24FDN$

Eq. 5: $4281, 6 - 39, 97FDN - 72, 90MM$

Eq. 6: $4354, 8 - 112, 05FB - 151, 74MM$

Eq. 7: $4021, 8 - 227, 55MM$

As equações 1, 2, 5, 6 e 7 foram obtidas, considerando como base de dados os alimentos energéticos.

As equações 3 e 4 foram obtidas, considerando como base de dados os alimentos concentrados.

TABELA 3 Variáveis explicativas das equações de predição obtidas para EMAn de soja + subprodutos encontradas na literatura e seus respectivos R^2

Equação(*)	Fonte	Variáveis explicativas	R^2 (%)
1	Nascimento et al. (2009)	EE e FDN	81,00
2		EE, MM, FDN e FDA	84,00
3		EE, MM e FDN	83,00
4	Rodrigues et al. (2002)	FB e EE	93,00
5	Zonta et al. (2004)	PB, FDA e EE	92,00
6		FDA e EE	90,00

(*) Equações obtidas:

$$Eq. 1: 2707,71 + 58,63EE - 16,06FDN$$

$$Eq. 2: 4101,33 + 56,28EE - 232,97MM - 24,86FDN + 10,42FDA$$

$$Eq. 3: 4095,41 + 56,84EE - 225,26MM - 22,24FDN$$

$$Eq. 4: 2822,2 - 90,13FB + 49,96EE$$

$$Eq. 5: -822,33 + 69,54PB - 45,26FDA + 90,81EE$$

$$Eq. 6: 2723,05 - 50,52FDA + 60,40EE$$

As equações 1, 4, 5 e 6 foram obtidas, considerando como base de dados os alimentos soja + subprodutos.

As equações 2 e 3 foram obtidas, considerando como base de dados os alimentos concentrados.

4 RESULTADOS E DISCUSSÕES

4.1 Equações de predição para EMAn selecionadas, obtidas neste estudo

Neste trabalho, notou-se que o R^2 e o R_a^2 apresentaram comportamentos semelhantes, seguindo a mesma estrutura de crescimento ou decréscimo de uma equação para outra. Assim, optou-se por utilizar os valores de R^2 ao invés de R_a^2 , para efeito de comparações com os R^2 apresentados nos trabalhos disponíveis na literatura.

Para os alimentos energéticos, os valores dos critérios utilizados na seleção e validação das equações de predição, estão representados nas Figuras 3 e 4, respectivamente. Na Figura 3b vê-se que com o aumento no número de grupos, os valores de AIC aumentam rapidamente até próximo de 5 grupos. A partir daí, observa-se um aumento menos expressivo com uma tendência de se estabilizar. Conforme era de se esperar, o coeficiente de determinação (Figura 3a) tem comportamento inverso ao AIC , pois quanto menor o AIC e maior o R^2 , melhor o ajuste da equação.

Porém, optou-se por utilizar o AIC , pois apesar do critério do coeficiente de determinação (R^2) ser largamente utilizado para a escolha do melhor modelo, ele não pondera pelo número de parâmetros do modelo estatístico considerado. Para tanto, um compromisso satisfatório entre o bom ajuste e o princípio de parcimônia pode ser alcançado, aplicando-se o chamado critério de informação de Akaike (AIC), como um procedimento para identificação de modelo estatístico.

Analisando os valores de CCC, na Figura 4b, vê-se que para um número maior de grupos, principalmente a partir de 20 grupos, o coeficiente de correlação de concordância diminui. Isso não é desejado, pois, se a partir de uma equação ob-

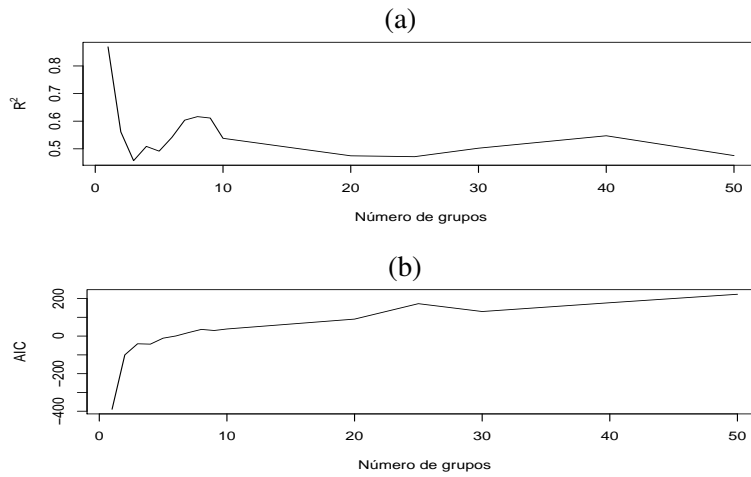


FIGURA 3 Valores dos (a) coeficiente de determinação e (b) critério de Akaike determinados no ajuste das equações de predição para EMAN de alimentos energéticos, obtidas neste trabalho, em função do número de grupos formados

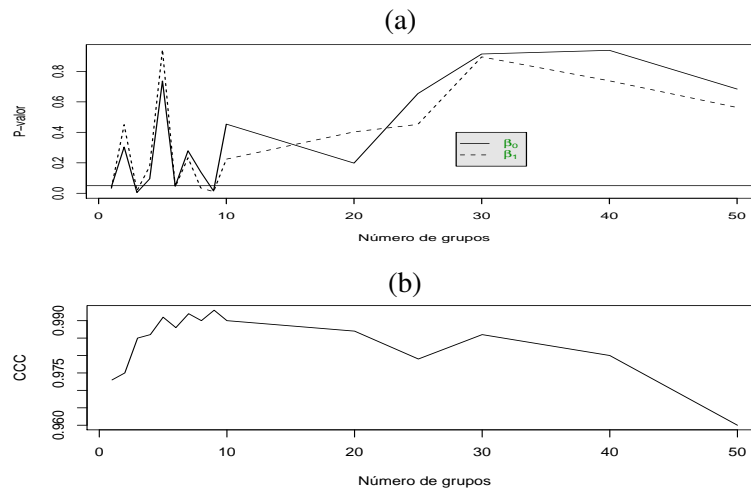


FIGURA 4 (a) P -valores de β_0 e β_1 e (b) valores do coeficiente de correlação de concordância determinados na validação das equações de predição para EMAN de alimentos energéticos, obtidas neste trabalho, em função do número de grupos formados

têm-se valores preditos mais próximos dos valores observados de EMAn, o CCC tende a um. Na Figura 4a, têm-se os *p-valores* do teste t, considerando-se as hipóteses $H_0 : \beta_0 = 0$ e $H_0 : \beta_1 = 1$, os quais correspondem aos resultados da validação utilizando o ajuste da equação linear simples. Apesar de não estar claro no gráfico, em algumas das situações até 9 grupos essas hipóteses foram rejeitadas, o que pode ser verificado na Tabela 4.

Assim, analisando-se concomitantemente o *AIC*, o CCC e os *p-valores*, conclui-se que as equações de predição da EMAn que tiveram os melhores ajustes para os concentrados energéticos, foram aquelas obtidas com a definição de 10 e 20 grupos. Pois, com esses números de grupos, para o conjunto de dados utilizados, têm-se melhores resultados em relação aos critérios utilizados.

No entanto, se fosse analisado apenas o gráfico do R^2 , em função do número de grupos, da Figura 3a, o mesmo não refletiria o que foi concluído, visto que a equação obtida considerando 20 grupos apresentou um baixo valor de R^2 . Com isso, fica comprovado que o R^2 não é indicado a ser usado como critério único de seleção. Portanto, diante destes resultados recomenda-se, sempre que possível, a utilização conjunta dos critérios de seleção e validação de modelos.

Nas outras duas situações, nas quais considera-se os 293 trabalhos de alimentos concentrados e os 67 trabalhos de soja + subprodutos, foram realizados procedimentos semelhantes a dos alimentos energéticos. Coincidentemente, as equações de predição que consideraram 10 e 20 grupos no processo de estimação, para os dados utilizados, foram consideradas as melhores.

Assim, as equações que apresentaram melhores ajustes, segundo o critério *AIC* e os resultados da validação, para os dados utilizados, estão apresentadas na Tabela 5.

TABELA 4 Valores dos critérios de seleção e validação das equações de predição para energéticos obtidas neste trabalho

Número de grupos	R^2 (%)	AIC	p -valor (β_0)	p -valor (β_1)	CCC
1	86,92	-389,737	0,043	0,033	0,973
2	56,14	-100,121	0,304	0,450	0,975
3	45,70	-40,896	0,005	0,016	0,985
4	50,87	-42,913	0,095	0,178	0,986
5	49,15	-11,073	0,737	0,943	0,991
6	54,14	0,695	0,047	0,042	0,988
7	60,37	18,845	0,279	0,233	0,992
8	61,65	35,630	0,140	0,035	0,990
9	61,14	29,774	0,015	0,011	0,993
10	53,80	37,870	0,454	0,224	0,990
20	47,46	90,441	0,198	0,403	0,987
25	47,15	172,569	0,654	0,452	0,979
30	50,22	130,805	0,914	0,895	0,986
40	54,72	177,610	0,938	0,738	0,980
50	47,54	222,657	0,684	0,563	0,960

TABELA 5 Variáveis explicativas das equações de predição obtidas que apresentaram melhores ajustes com o uso de diferentes números de grupos e seus respectivos R^2

Alimentos	Equações(*)	Número de grupos	Variáveis explicativas	R^2 (%)
Energéticos	Eq. 1	10	EE, MM e FDN	53, 80
	Eq. 2	20	PB, EE, MM, FB, FDN e FDA	47, 46
Soja+sub	Eq. 3	10	EE, MM e FDA	86, 90
	Eq. 4	20	PB, EE, MM, FB e FDN	84, 06
Concentrados	Eq. 5	10	EE, MM, FB e FDN	70, 99
	Eq. 6	20	EE, MM, FB e FDN	74, 49

(*) Equações obtidas:

$$Eq. 1: 4215, 773 + 51, 211EE - 193, 436MM - 28, 034FDN$$

$$Eq. 2: 4235, 378 - 11, 487PB + 40, 232EE - 138, 642MM - 42, 872FB - 21, 738FDN + 11, 743FDA$$

$$Eq. 3: 3383, 589 + 51, 412EE - 139, 617MM - 15, 474FDA$$

$$Eq. 4: 1322, 267 + 33, 894PB + 73, 606EE - 84, 217MM + 47, 200FB - 29, 936FDN$$

$$Eq. 5: 4144, 914 + 53, 137EE - 204, 644MM - 26, 214FB - 20, 260FDN$$

$$Eq. 6: 4164, 187 + 51, 006EE - 197, 663MM - 35, 689FB - 20, 593FDN$$

Na Tabela 6 encontram-se equações de predição para alimentos energéticos, sendo as equações 1 a 4 obtidas neste trabalho e, as demais, obtidas na literatura (Rodrigues et al., 2001; Nagata et al., 2004; Nascimento, 2007). No entanto, Nagata et al. (2004) não apresentaram o valor de R^2 em seu trabalho. Porém, observa-se que todas as 11 equações de predição apresentaram teor de MM em sua composição química, sendo esta a única variável presente na equação de predição proposta por Nagata et al. (2004). Considerando-se as oito primeiras equações, nota-se a presença dos teores de EE e FDN, além do teor de MM. As equações 3 e 4 apresentam as mesmas variáveis de composição química, EE, MM, FB e FDN, com valores próximos de R^2 . E as equações 1 e 8, também apresentam variáveis de composição química semelhantes, sendo EE, MM e FDN, porém com valores distantes de R^2 . Apenas nas equações 2 e 5 o teor de PB é encontrado, sendo essas equações semelhantes entre si e, apresentando as mesmas variáveis de composição química: PB, EE, MM, FB, FDN e FDA. No entanto, seus valores de R^2 são muito distantes, sendo $R^2 = 47,46$ e $R^2 = 81,00$ para as equações 2 e 5, respectivamente. Essa diferença nos valores de R^2 se deu devido ao uso de diferentes maneiras de formação de grupos, visto que foi utilizado o mesmo conjunto de dados para obtenção destas equações 2 e 5.

Na Tabela 7 estão apresentadas equações de predição para soja + subprodutos, sendo as equações 1 a 4 obtidas neste trabalho e, as demais, obtidas na literatura (Rodrigues et al., 2002; Zonta et al., 2004; Nascimento, 2007). Porém, Zonta et al. (2004) não apresentaram os valores de R^2 em seu trabalho. Na composição química de todas 10, nota-se a presença do teor EE. Verifica-se, também, que as equações 3 e 4, que são as mesmas 3 e 4 consideradas para os alimentos energéticos, apresentam variáveis de composição química semelhantes, sendo EE, MM, FB e FDN.

TABELA 6 Equações de predição para energia metabolizável (EMAn) de alimentos energéticos

Equação(*)	Fonte	Variáveis explicativas	R ² (%)
1	Energéticos com 10 grupos	EE, MM e FDN	53,80
2	Energéticos com 20 grupos	PB, EE, MM, FB, FDN e FDA	47,46
3	Concentrados com 10 grupos	EE, MM, FB e FDN	71,00
4	Concentrados com 20 grupos	EE, MM, FB e FDN	74,50
5	Nascimento et al. (2009)	PB, EE, MM, FB, FDN e FDA	81,00
6		EE, MM, FB, FDN e FDA	81,00
7		EE, MM, FDN e FDA	84,00
8		EE, MM e FDN	83,00
9	Rodrigues et al. (2001)	FDN e MM	96,00
10		FB e MM	95,00
11	Nagata et al. (2004)	MM	92,00

(*) Equações obtidas:

Eq. 1: $4215,773 + 51,211EE - 193,436MM - 28,034FDN$

Eq. 2: $4235,378 - 11,487PB + 40,232EE - 138,642MM - 42,872FB - 21,738FDN + 11,743FDA$

Eq. 3: $4144,914 + 53,137EE - 204,644MM - 26,214FB - 20,260FDN$

Eq. 4: $4164,187 + 51,006EE - 197,663MM - 35,689FB - 20,593FDN$

Eq. 5: $4371,18 - 26,48PB + 30,65EE - 126,93MM - 52,26FB - 25,14FDN + 24,40FDA$

Eq. 6: $4205,23 + 30,58EE - 130,35MM - 58,29FB - 28,31FDN + 16,71FDA$

Eq. 7: $4101,33 + 56,28EE - 232,97MM - 24,86FDN + 10,42FDA$

Eq. 8: $4095,41 + 56,84EE - 225,26MM - 22,24FDN$

Eq. 9: $4281,6 - 39,97FDN - 72,90MM$

Eq. 10: $4354,8 - 112,05FB - 151,74MM$

Eq. 11: $4021,8 - 227,55MM$

As equações 1 a 4 foram obtidas neste trabalho, utilizando-se a aplicação de meta-análise e componentes principais.

As equações 3, 4, 7 e 8 foram obtidas, considerando como base de dados os alimentos concentrados e as demais equações, considerando alimentos energéticos.

TABELA 7 Equações de predição para energia metabolizável (EM/An) de soja + subprodutos

Equação(*)	Fonte	Variáveis explicativas	R ² (%)
1	Soja + subprodutos com 10 grupos	EE, MM e FDA	86,90
2	Soja + subprodutos com 20 grupos	PB, EE, MM, FB e FDN	84,10
3	Concentrados com 10 grupos	EE, MM, FB e FDN	71,00
4	Concentrados com 20 grupos	EE, MM, FB e FDN	74,50
5	Nascimento et al. (2009)	EE e FDN	81,00
6		EE, MM, FDN e FDA	84,00
7		EE, MM e FDN	83,00
8	Rodrigues et al. (2002)	FB e EE	92,00
9	Zonta et al. (2004)	PB, FDA e EE	92,00
10		FDA e EE	90,00

(*) Equações obtidas:

Eq. 1: 3383,589 + 51,412EE - 139,617MM - 15,474FDA

Eq. 2: 1322,267 + 33,894PB + 73,606EE - 84,217MM + 47,200FB - 29,936FDN

Eq. 3: 4144,914 + 53,137EE - 204,644MM - 26,214FB - 20,260FDN

Eq. 4: 4164,187 + 51,006EE - 197,663MM - 35,689FB - 20,593FDN

Eq. 5: 2707,71 + 58,63EE - 16,06FDN

Eq. 6: 4101,33 + 56,28EE - 232,97MM - 24,86FDN + 10,42FDA

Eq. 7: 4095,41 + 56,84EE - 225,26MM - 22,24FDN

Eq. 8: 2822,2 - 90,13FB + 49,96EE

Eq. 9: -822,33 + 69,54PB - 45,26FDA + 90,81EE

Eq. 10: 2723,05 - 50,52FDA + 60,40EE

As equações 1 a 4 foram obtidas neste trabalho, utilizando-se a aplicação de meta-análise e componentes principais.

As equações 3, 4, 6 e 7 foram obtidas, considerando como base de dados os alimentos concentrados e as demais equações, considerando alimentos soja + subprodutos.

À exceção de uma equação, descrita por Nagata (2004), grande parte das equações de predições, tanto para alimentos energéticos quanto para soja + subprodutos (Tabelas 6 e 7) apresentam de duas a quatro variáveis em sua composição química.

4.2 Comparação de equações de predição via validação com ensaio metabólico

Para verificar a qualidade do ajuste das equações selecionadas utilizando-se meta-análise e componentes principais, estas foram comparadas com outras equações de predição encontradas na literatura. Para isto, foi efetuado a validação das mesmas, utilizando-se os dados provenientes de ensaios metabólicos conduzidos por Alvarenga (2009).

4.2.1 Alimentos energéticos e concentrados

As equações de predição para EMAN de alimentos energéticos, utilizadas na comparação, foram aquelas descritas na Tabela 6. Observa-se que estas equações são obtidas considerando-se diferentes bases de dados. Isto é, as equações 1, 2, 5, 6, 9, 10 e 11 foram estimadas considerando-se apenas alimentos energéticos, enquanto que na obtenção das equações 3, 4, 7 e 8 considerou-se todos os alimentos concentrados catalogados (energéticos e protéicos). Portanto, têm-se equações obtidas a partir de diferentes bases de dados, mas que foram utilizadas na validação da EMAN de alimentos energéticos.

Na Figura 5b, observa-se que a menos das equações 5 e 6, as demais equações de predição para EMAN para alimentos energéticos, as hipóteses $H_0 : \beta_0 = 0$ e $H_0 : \beta_1 = 1$ não foram rejeitadas durante a validação. Isto é, para a maioria das equações consideradas para alimentos energéticos, não foram encontradas diferen-

ças significativas entre os valores ajustados e os observados no ensaio metabólico.

Isso pode ser verificado na Tabela 8.

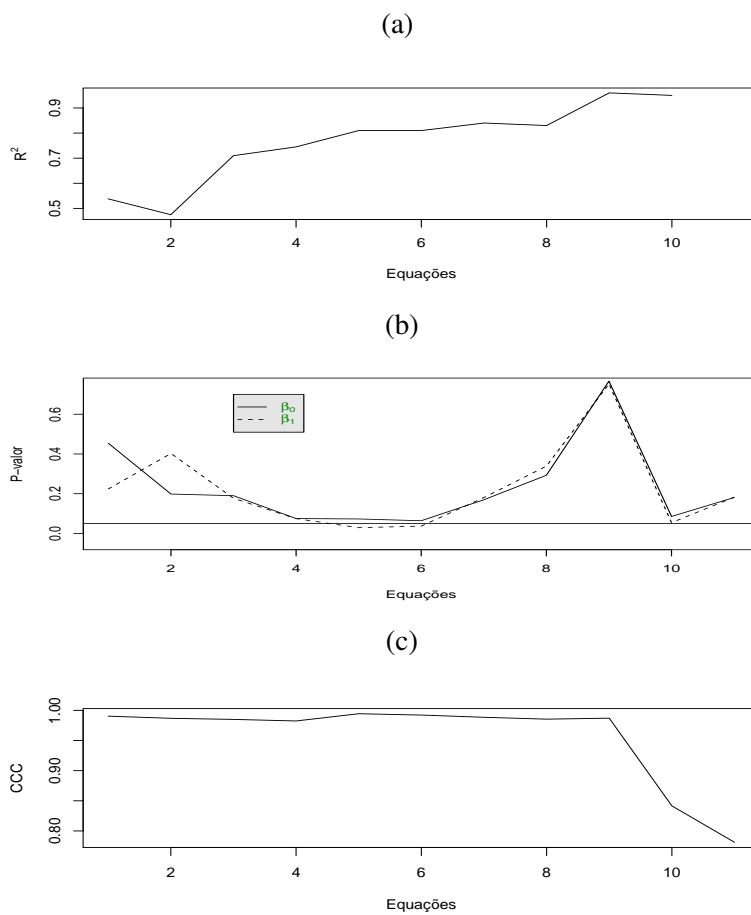


FIGURA 5 Valores dos (a) coeficiente de determinação, (b) *p-valores* de β_0 e β_1 e (c) coeficiente de correlação de concordância determinados na avaliação das equações de predição para EMAn de alimentos energéticos

Ao analisar o coeficiente de correlação de concordância (CCC), na Figura 5c, nota-se que até a equação 9 utilizada nessa comparação, todas possuem valores semelhantes de CCC, sendo todos próximos de um. No entanto, os valores de R^2 não são os maiores apresentados (Figura 5a). Nota-se que, apesar das equações 1, 2, 3, 8 e 9 não apresentarem os maiores coeficientes de determinação (R^2), em sua maioria, estas possuem os maiores *p-valores* para os parâmetros da regressão linear simples. Então, com a análise simultânea destes critérios, conclui-se que as equações 1, 2, 3, 4, 7, 8 e 9 descreveram com acurácia e precisão o valor da EMAn obtidas no ensaio metabólico para alimentos energéticos.

Portanto, as equações obtidas neste trabalho apresentaram resultados semelhantes às outras disponíveis na literatura (Rodrigues, 2001; Nascimento, 2007). Nascimento (2007), assim como este estudo, utilizaram formas menos onerosas, pois não executaram experimentos específicos na obtenção de tais equações, como foi o caso de Rodrigues (2001). Porém, a formação de grupos homogêneos de trabalhos, que é um dos grandes problemas na meta-análise, foi facilitada neste estudo comparando-se com a formação de grupos conforme Nascimento (2007). Isto é, utilizando-se análise de componentes principais não há a necessidade de se definirem fatores mais coerentes na definição dos grupos. Além do que, a definição dos componentes principais pode ser feita em softwares computacionais, como por exemplo no R (R Development Core Team, 2009).

No entanto, se a escolha do melhor modelo fosse baseada unicamente nos valores de R^2 , as equações 9 e 10 seriam consideradas como aquelas que apresentaram melhores ajustes para a predição da EMAn dos alimentos em questão, o que não ocorreu com a equação 10.

TABELA 8 Valores dos critérios de seleção e validação das equações de predição para alimentos energéticos

Equação	Fonte	R^2 (%)	AIC	p-valor (β_0)	p-valor (β_1)	CCC
1	Energéticos com 10 grupos	53,80	37,870	0,454	0,224	0,990
2	Energéticos com 20 grupos	47,46	90,441	0,198	0,403	0,987
3	Concentrados com 10 grupos	71,00	-42,219	0,190	0,177	0,985
4	Concentrados com 20 grupos	74,50	-3,054	0,075	0,075	0,982
5	Nascimento (2007)	81,00	-291,527	0,073	0,029	0,994
6		81,00	-285,706	0,064	0,037	0,992
7		84,00	-512,111	0,169	0,180	0,988
8		83,00	-508,325	0,293	0,339	0,985
9	Rodrigues et al. (2001)	96,00	-	0,767	0,753	0,987
10		95,00	-	0,085	0,054	0,841
11	Nagata et al. (2004)	92,00	-	0,181	0,182	0,781

(*) Equações obtidas:

Eq. 1: 4215, 773 + 51, 211EE - 193, 436MM - 28, 034FDN

Eq. 2: 4235, 378 - 11, 487PB + 40, 232EE - 138, 642MM - 42, 872FB - 21, 738FDN + 11, 743FDA

Eq. 3: 4144, 914 + 53, 137EE - 204, 644MM - 26, 214FB - 20, 260FDN

Eq. 4: 4164, 187 + 51, 006EE - 197, 663MM - 35, 689FB - 20, 593FDN

Eq. 5: 4371, 18 - 26, 48PB + 30, 65EE - 126, 93MM - 52, 26FB - 25, 14FDN + 24, 40FDA

Eq. 6: 4205, 23 + 30, 58EE - 130, 35MM - 58, 29FB - 28, 31FDN + 16, 71FDA

Eq. 7: 4101, 33 + 56, 28EE - 232, 97MM - 24, 86FDN + 10, 42FDA

Eq. 8: 4095, 41 + 56, 84EE - 225, 26MM - 22, 24FDN

Eq. 9: 4281, 6 - 39, 97FDN - 72, 90MM

Eq. 10: 4354, 8 - 112, 05FB - 151, 74MM

Eq. 11: 4021, 8 - 227, 55MM

As equações 1 a 4 foram obtidas neste trabalho, utilizando-se a aplicação de meta-análise e componentes principais.

As equações 3, 4, 7 e 8 foram obtidas, considerando como base de dados os alimentos concentrados e as demais equações, considerando alimentos energéticos.

4.2.2 Soja + subprodutos e concentrados

As equações de predição para EMAn de soja + subprodutos, utilizadas na comparação estão apresentadas na Tabela 7. A comparação foi desenvolvida entre as diferentes equações de predição obtidas e aquelas adquiridas em trabalhos de outros autores, análogo ao feito para os alimentos energéticos e concentrados. Ou seja, considerou-se os resultados obtidos após a aplicação dos critérios de validação de modelos: *p-valor* dos parâmetros da equação linear simples e CCC, sendo que as equações 1, 2, 5, 8, 9 e 10 foram estimadas considerando-se apenas soja + subprodutos, enquanto que na obtenção das equações 3, 4, 6 e 7 foram considerados todos os alimentos concentrados catalogados. Contudo, todas as dez equações foram utilizadas na validação da EMAn de soja + subprodutos.

Na Figura 6b, observa-se que as únicas equações de predição para EMAn para soja + subprodutos, em que as hipóteses estabelecidas foram rejeitadas durante a validação ($H_0 : \beta_0 \neq 0$ e $H_0 : \beta_1 \neq 1$), foram as equações 2 e 9. As demais não apresentaram diferenças significativas entre os valores ajustados e os observados no ensaio metabólico, podendo isso ser verificado na Tabela 9.

Analisando-se a Figura 6c, verifica-se que por possuírem valores próximos de CCC tendendo a um, as equações 2, 3, 4, 6 e 7 apresentam os melhores ajustes dentre as equações comparadas. No entanto, as equações 3, 4, 6 e 7 não apresentaram os maiores coeficientes de determinação (R^2), porém possuem maiores *p-valores* (Figura 6a e 6b). Contudo, se fosse considerado só o coeficiente de determinação como critério único de seleção, o melhor ajuste considerado seria o da equação 8, a qual possui valor de CCC igual a 0,21 (Tabela 9).

Assim, fazendo a análise simultânea destes critérios, conclui-se que as equações 3, 4, 6 e 7 descreveram com maior acurácia e precisão o valor da EMAn obtidas no ensaio metabólico para soja+subprodutos.

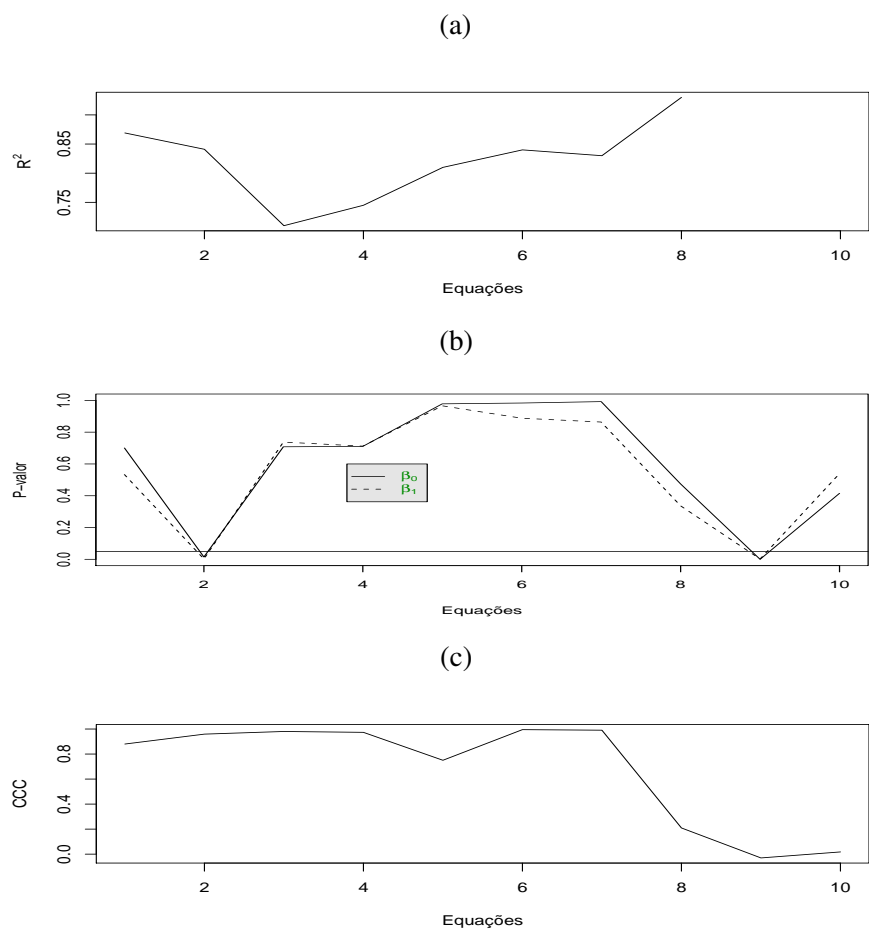


FIGURA 6 Valores dos (a) coeficiente de determinação, (b) *p*-valores de β_0 e β_1 e (c) coeficiente de correlação de concordância determinados na avaliação das equações de predição para EMAn de soja + subprodutos

Portanto, as equações obtidas neste trabalho apresentaram resultados semelhantes às disponíveis na literatura (Nascimento, 2007). Assim, análogo ao discutido na comparação para alimentos energéticos, tem-se que na obtenção de equações de predição para EMAn de soja + subprodutos, pela aplicação de meta-análise, a formação de grupos homogêneos de trabalhos foi facilitada através da técnica multivariada de componentes principais.

TABELA 9 Valores dos critérios de seleção e validação das equações de predição para soja + subprodutos

Equação	Fonte	R^2 (%)	AIC	p-valor (β_0)	p-valor (β_1)	CCC
1	Soja + subprodutos com 10 grupos	86,90	39,685	0,700	0,533	0,880
2	Soja + subprodutos com 20 grupos	84,10	98,979	0,014	0,002	0,959
3	Concentrados com 10 grupos	71,00	-42,219	0,708	0,737	0,980
4	Concentrados com 20 grupos	74,50	-3,054	0,710	0,712	0,973
5	Nascimento (2007)	81,00	-94,118	0,978	0,966	0,750
6		84,00	-512,111	0,983	0,888	0,995
7		83,00	-508,325	0,992	0,864	0,991
8	Rodrigues et al. (2002)	93,00	-	0,474	0,336	0,209
9	Zonta et al. (2004)	92,00	-	0,000	0,002	-0,030
10		90,00	-	0,4152	0,539	0,017

(*) Equações obtidas:

Eq. 1: 3383, 589 + 51, 412EE - 139, 617MM - 15, 474FDA

Eq. 2: 1322, 267 + 33, 894PB + 73, 606EE - 84, 217MM + 47, 200FB - 29, 936FDN

Eq. 3: 4144, 914 + 53, 137EE - 204, 644MM - 26, 214FB - 20, 260FDN

Eq. 4: 4164, 187 + 51, 006EE - 197, 663MM - 35, 689FB - 20, 593FDN

Eq. 5: 2707, 71 + 58, 63EE - 16, 06FDN

Eq. 6: 4101, 33 + 56, 28EE - 232, 97MM - 24, 86FDN + 10, 42FDA

Eq. 7: 4095, 41 + 56, 84EE - 225, 26MM - 22, 24FDN

Eq. 8: 2822, 2 - 90, 13FB + 49, 96EE

Eq. 9: -822, 33 + 69, 54PB - 45, 26FDA + 90, 81EE

Eq. 10: 2723, 05 - 50, 52FDA + 60, 40EE

As equações 1 a 4 foram obtidas neste trabalho, utilizando-se a aplicação de meta-análise e componentes principais.

As equações 3, 4, 6 e 7 foram obtidas, considerando como base de dados os alimentos concentrados e as demais equações, considerando alimentos soja + subprodutos.

5 CONCLUSÕES

As equações de predição propostas neste trabalho utilizaram na meta - análise grupos homogêneos obtidos pela aplicação da técnica de componentes principais. Estas equações, juntamente com outras propostas na literatura, foram validadas e concluiu-se que elas apresentaram resultados semelhantes às disponíveis na literatura. Portanto, a formação de grupos homogêneos de trabalhos, que é um dos grandes problemas na meta-análise, foi facilitada com a utilização da técnica de componentes principais. Isto é, não houve a necessidade da determinação de fatores coerentes para esse agrupamento, como foi proposto em outros trabalhos.

REFERÊNCIAS BIBLIOGRÁFICAS

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, 2., 1973, Budapest. **Proceedings...** Budapest: Academiai Kiado, 1973. p. 267-281.

AKAIKE, H. Information measures and model selection. **International Statistical Institute**, Voorburg, v. 44, n. 1, p. 277-291, 1983.

ALBINO, L. F. T.; ROSTAGNO, H. S.; FONSECA, J. B.; TORRES, R. A. Utilização de diferentes sistemas de avaliação energética dos alimentos na formulação de rações para frangos de corte. **Revista da Sociedade Brasileira de Zootecnia**, Viçosa, MG, v. 21, n. 6, p. 1037-1046, nov./dez. 1992.

ALVARENGA, R. R. **Valores energéticos de alimentos concentrados determinados com frangos de corte e por equações de predição**. 2009. 66 p. Dissertação (Mestrado em Zootecnia) - Universidade Federal de Lavras, Lavras.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. 5. ed. São Paulo: Saraiva, 2003. 526 p.

COOPER, H. M. **Integrating research: a guide for literature reviews**. 2. ed. Newbury Park: Sage, 1990. 157 p.

COSTA, P. A. B. **Um enfoque segundo a teoria de conjuntos difusos para metaanálise**. 1999. 153 f. Tese (Doutorado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Florianópolis.

DRAPER, N. R.; SMITH, H. **Applied regression analysis**. 3. ed. New York: J. Wiley, 1998. 706 p.

FAGARD, R. H.; STAESSEN, J. A.; THIJS, L. Advantages and disadvantages of the meta-analysis approach. **Journal of Hypertension**, London, v. 14, n. 2, p. 9-13, 1996. Supplement.

FERREIRA, D. F. **Estatística multivariada**. Lavras: UFLA, 2008. 661 p.

FILGUEIRA, J. M. Avaliação de satisfação de cliente: proposta de utilização do coeficiente de correlação de concordância. **Holos**, Rio Claro, ano 21, p. 47-56, set. 2005.

FISCHER JÚNIOR, A. A.; ALBINO, L. F. T.; ROSTAGNO, H. S.; GOMES, P. C. Determinação dos valores de energia metabolizável de alguns alimentos usados na alimentação de aves. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 27, n. 2, p. 314-318, mar./abr. 1998.

GIANNOTTI, J. D. G. **Meta-análise de estimativas da correlação genética entre pesos ao nascer e desmama de bovinos**. 2000. 85 p. Dissertação (Mestrado em Zootecnia) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba.

GIANNOTTI, J. D. G.; PACKER, I. U.; MERCADANTE, M. E. Z. Meta-análise para estimativas de correlação genética entre pesos ao nascer e desmama de bovinos. **Scientia Agricola**, Piracicaba, v. 59, n. 3, p. 435-440, jul./set. 2002.

GIANNOTTI, J. D. G.; PACKER, I. U.; MERCADANTE, M. E. Z.; LIMA, C. G. Análise de agrupamento para implementação da meta-análise em estimativas de herdabilidade para características de crescimento em bovinos de corte. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 34, n. 4, p. 1165-1172, jul./ago. 2005.

GLASS, G. V. Primary, secondary, and meta-analysis of research. **Educational Researcher**, Washington, v. 6, n. 1, p. 3-8, Jan. 1976.

GLASS, G. V.; McGRAW, B.; SMITH, M. L. **Meta-analysis in social research**. Beverly Hills: Sage, 1981. 280 p.

GUREVITCH, J.; HEDGES, L.V. Meta-analysis: combining the results of independent experiments. In: SCHEINER, S. M.; GUREVITCH, J. **Design and analysis of ecological experiments**. New York: Chapman & Hall, 1993. chap. 17, p. 378-398.

HARRISON, S. R. Regression of a model on real-system output: an invalid test of model validity. **Agriculture Systems**, Amsterdam, v. 34, n. 3, p. 183-190, 1990.

HEDGES, L.V.; OLKIN, I. **Statistical methods for meta-analysis**. London: Academic, 1985. 369 p.

HILL, F. W.; ANDERSON, D. L. Comparison of metabolizable energy and productive energy determinations with growing chicks. **Journal Nutrition**, Bethesda, v. 64, n. 4, p. 587-603, Apr. 1958.

HOFFMAN, R.; VIEIRA, S. **Análise de regressão: uma introdução à econometria**. 3. ed. São Paulo: Hucitec, 1998. 379 p.

HOTELLING, H. Review of the triumph of mediocrity in business. **Journal of the American Statistical Association**, New York, v. 28, n. 184, p. 463-465, Dec. 1933.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 5. ed. New York: Prentice Hall, 2002. 512 p.

KIRBY, K. N. **Advanced data analysis with SYSTAT**. New York: V.N. Reinhold, 1993. 475 p.

KULBACK, S.; LEIBLER, R. A. On information and sufficiency. **Annals of Mathematical Statistics**, Ann Arbor, v. 22, n. 1, p. 79-86, Mar. 1951.

LIN, L. A concordance correlation coefficient to evaluate reproducibility. **Biometrics**, Washington, v. 45, n. 1, p. 255-268, Mar. 1989.

LOVATTO, P. A.; LEHNEN, C. R.; ANDRETTA, I.; CARVALHO, A. D.; HAUSCHILD, L. Meta-análise em pesquisas científicas: enfoque em metodologias. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 36, p. 285-294, jul. 2007. Suplemento.

MARTINEZ, E. Z. Metanálise de ensaios clínicos controlados aleatorizados: aspectos quantitativos. **Revista de Medicina**, Ribeirão Preto, v. 40, n. 2, p. 223-235, jul. 2007.

MARTINS, R. M. **Estudo do crescimento da cana-de-açúcar através da metaanálise**. 2001. 68 p. Dissertação (Mestrado em Energia na Agricultura) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Botucatu.

MENESES-GAYA, I. C.; ZUARDI, A. W.; LOUREIRO, S. R.; CRIPPA, J. A. S. As propriedades psicométricas do teste de Fagerström para dependência de nicotina. **Journal Brasileiro de Pneumologia**, Brasília, v. 35, n. 1, p. 73-82, jan. 2009.

MITCHELL, P. L. Misuse of regression for empirical validation of models. **Agricultural Systems**, Amsterdam, v. 54, n. 3, p. 313-326, June 1997.

NAGATA, A. K.; RODRIGUES, P. B.; FREITAS, R. T. F.; BERTECHINI, A. G.; FIALHO, E. T. Energia metabolizável de alguns alimentos energéticos para frangos de corte, determinada por ensaios metabólicos e por equações de predição. **Ciência e Agrotecnologia**, Lavras, v. 28, n. 3, p. 668-677, maio/jun. 2004.

NASCIMENTO, G. A. J. **Equações de predição dos valores energéticos de alimentos para aves utilizando o princípio da meta-análise**. 2007. 199 p. Tese (Doutorado em Zootecnia) - Universidade Federal de Lavras, Lavras.

NASCIMENTO, G. A. J.; RODRIGUES, P. B.; FREITAS, R. T. F.; BERTECHINI, A. G.; LIMA, R. R.; PUCCI, L. E. A. Equações de predição para estimar os valores energéticos de alimentos concentrados de origem vegetal para aves utilizando a metanálise. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 38, n. 7, p. 1265-1271, dez. 2009.

PETERS, A. R.; MARTINEZ, T. A.; COOK, A. J. C. A meta-analysis of studies of the effect of GNRH 11-14 days after insemination on pregnancy rates in cattle. **Theriogenology**, Woburn, v. 54, n. 8, p. 1317-1326, Nov. 2000.

PRODUÇÃO científica do Brasil ultrapassa a da Rússia, indica levantamento. Disponível em: <http://www.bbc.co.uk/portuguese/ciencia/2010/01/100127_brasil_russia_ciencia_rw.shtml>. Acesso em: 10 jan. 2010.

R DEVELOPMENT CORE TEAM. **R**: a language and environment for statistical computing. Viena: R Foundation for Statistical Computing, 2009. Software.

RODRIGUES, P. B.; NASCIMENTO, G. A. J.; FREITAS, R. T. F.; BERTECHINI, A. G.; LIMA, R. R. Equações de predição dos valores energéticos de alimentos concentrados para aves utilizando o princípio da meta-análise. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 44., 2007, Jaboticabal. **Anais...** Jaboticabal: UNESP, 2007. 1 CD-ROM.

RODRIGUES, P. B.; ROSTAGNO, H. S.; ALBINO, L. F. T.; GOMES, P. C.; BARBOZA, W. A.; SANTANA, R. T. Valores energéticos do milho, do milho e subprodutos do milho, determinados com frangos de corte e galos adultos. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 30, n. 6, p. 1767-1778, nov./dez. 2001.

RODRIGUES, P. B.; ROSTAGNO, H. S.; ALBINO, L. F. T.; GOMES, P. C.; NUNES, R. V.; TOLEDO, R. S. Valores energéticos da soja e subprodutos da soja, determinados com frangos de corte e galos adultos. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 31, n. 4, p. 1771-1782, jul./ago. 2002.

SOUZA, R. M. **Equações de predição dos valores energéticos de alimentos para aves**. 2009. 123 p. Tese (Doutorado em Zootecnia) - Universidade Federal de Viçosa, Viçosa, MG.

TEDESCHI, L. O. Assessment of the adequacy of mathematical models. **Agricultural Systems**, Amsterdam, v. 89, n. 2/3, p. 225-247, Sept. 2006.

WOLYNETZ, M. N.; SIBBALD, I. R. Relationships between apparent and true metabolizable energy and the effects of a nitrogen correction. **Poultry Science**, Champaign, v. 65, n. 7, p. 1386-1399, July 1984.

ZONTA, M. C. M.; RODRIGUES, P. B.; ZONTA, A.; FREITAS, R. T. F.; BERTECHINI, A. G.; FIALHO, E. T.; PEREIRA, C. R. Energia metabolizável de ingredientes protéicos determinada pelo método de coleta total e por equações de predição. **Ciência e Agrotecnologia**, Lavras, v. 28, n. 6, p. 1400-1407, nov./dez. 2004.

ANEXO

ANEXO A Rotina R para obter as equações de predição, utilizando a técnica de componentes principais e meta-análise e, para calcular os valores dos critérios de seleção e validação das equações de predição

```
# Limpa a memória do software, removendo todos os objetos
rm(list=ls(all=TRUE))

##### 1ª PARTE #####
# Rotina para estimar a equação de predição, utilizando componen-
# tes principais na formação dos grupos para meta-análise. Neste
# caso, tomou-se 10 grupos para alimentos energéticos.
#####

#Função para definir o número de grupos

ngrupos <- 10
pmin<- (100/ngrupos)/100
pmax<- 1 - pmin
each<-pmin

# Função para fazer a leitura dos dados de trabalhos, considerando
# alimentos energéticos arquivo = read.table("energ_refin_c.txt",
header=T) attach(arquivo)

# Definindo a quantidade de quantis
quantil<-seq(pmin,pmax,each)

# Definindo os 1º componentes principais de cada trabalho, obtidos
# de acordo com a variabilidade em relação a EMAn, PB, EE, MM, FB,
# FDN e FDA.

resp<-princomp(~ EMAn + PB + EE + MM + FB + FDN + FDA, data =
arquivo, cor = TRUE)
summary(resp)
cp <- resp$scores[,1]

# Definindo o valor de cada quantil e colocando-os como vetor
quantis<-quantile(resp$scores[,1],probs=quantil)

vetor_Q <- as.vector(quantis)
detach(arquivo)
```

```

# Função que define cada 1° componente principal em seu respectivo
# percentil (grupo)
quati <- function(cp, grupo)
{
  x <- cp
  g <- sort (grupo)
  quantil <- 0
  comp    <- length(grupo)
  m <- new.env()
  for (i in (1:comp)) {
    quantil <- x[x<=g[i]]
    assign(paste("Percentil", i, sep=""), quantil, env=m)
    x <- x[x>g[i]]
  }
  quantil <- cp[cp>g[i]]
  assign(paste("Percentil", length(grupo)+1, sep=""), quantil, env=m)
  eval(expression(as.list(m)))
}

rt <- quati(cp, vetor_Q)

# Função que define o fator de ponderação utilizada em cada grupo
# formado, ou seja, a função que define o inverso da variância dos
# dados de EMAn originais de cada grupo.

namesperc<-as.vector(names(rt))
cnp<-length(namesperc)
grupo <- 0
cont <- 1/each for(i in 1:nrow(arquivo)){
  grupo[as.numeric(names(rt[[cont]])]<-names(rt[i])
  cont <- cont-1
}

# Criando uma lista para identificar as posições nos dados origi-
# nais

cont <- 1/each
posição <- as.list(matrix(0, cont, 1))
for(i in 1:cont){
  posição[[i]]<- as.numeric(names(rt[[i]]))
}
posição          # Posição no conjunto de dados original

invar <- rep(0, cont)
tam    <- 0

for(i in 1:cont){

```

```

    tam <- length(posição[[i]])
    invar[i] <-1/var(arquivo$EMAn[posição[[i]][1:tam]])
  }
  invar                                     # Inverso da variância de cada grupo

for(j in 1:cont){
  tam <- length(posição[[j]])
  arquivo$EMAn[posição[[j]][1:tam]]<-invar[j] }

arquivo$EMAn ->ivar      # Posições dos dados originais com seus
                        # respectivos valores dos fatores de
                        # ponderação

# Ajuste do modelo de regressão linear ponderada, utilizando
# os fatores de ponderação calculados e, aplicando o procedi-
# mento stepwise para seleção das variáveis significativas
# para o modelo

arquivo = read.table("energ_refin_c.txt", header=T)
conjunto = cbind(arquivo,ivar)
names(conjunto)
attach(conjunto)

library(MASS)
reg.aj<-lm(EMAn ~ PB + EE + MM + FB + FDN + FDA, weights=ivar)
summary(reg.aj)
reg <- stepAIC(reg.aj, scope = list(upper = ~ PB + EE + MM + FB +
FDN + FDA, lower = ~1), direction = "both")
summary(reg)

##### 2ª PARTE #####

# De posse da equação de predição, a seguir têm-se os passos
# utilizados para cálculo dos critérios de adequabilidade de
# modelos propostos neste trabalho.

#####
# Validação da EMAn obtido em ensaio metabólico x Estimativa da
# equação de predição, a partir dos valores de composição química
# calculadas no ensaio metabólico

#####
### PARTE VARIÁVEL - ver OBSERVAÇÃO no final desta rotina #####
#####

# Fazendo leitura dos dados obtidos em ensaio metabólico, neste

```



```

# caso, para alimentos energéticos

dados = read.table("Planilha_geral_predict.txt", header=T)
attach(dados)

# Obtenção das estimativas obtidas a partir das equações de predi-
# ção estimadas, utilizando os valores de composição química cal-
# culadas no ensaio metabólico

resp<-summary(reg)
coef<-resp$coefficients[,1]
dados$y.est<-coef[1]
if (length(coef)>1)
  for (i in 2:length(coef))
    dados$y.est<-dados$y.est+coef[i]*dados[,names(coef[i])]
eq <- round(dados$y.est) # Valores das estimativas obtidas
# pela equação de predição estimada eq

#####
##### FIM DA PARTE VARIÁVEL #####
#####

# Ajustando o modelo de regressão linear simples para validação da
# equação estimada

energ.ren<-lm(EMAn ~ eq, data=dados)
energ.ren.sum<-summary(energ.ren)
Pvalue_b0=energ.ren.sum$coefficients[1,4] # p-valor para H0: b0=0

# Teste t para Ho:b1=1
energ.ren.t<-(energ.ren.sum$coefficients[2,1]
              -1)/energ.ren.sum$coefficients[2,2]
energ.ren.p<-2*(1-pt(abs(energ.ren.t),energ.ren$df.residual))
Pvalue_b1=energ.ren.p # p-valor para H0: b1=1
data.frame(t.cal=energ.ren.t,Pvalue=round(energ.ren.p,6))

# Cálculo do Coeficiente de Correlação de Concordância - CCC de Lin
library(epiR)
tmp.ccc <- epi.ccc(eq, EMAn, ci = "z-transform",
                  conf.level = 0.95)
tmp.ccc
CCC <- tmp.ccc$rho.c$est
CCC

# Cálculo do Critério de Akaike (AIC) da Predição
AIC_pred <-as.matrix(extractAIC(reg))[2,]

```

```

#### Saída dos valores dos critérios para interpretações #####

round(data.frame(R_2=summary(reg)$r.squared, AIC_pred,
                 Pvalue_b0, Pvalue_b1, CCC),6)

#####
##### OBSERVAÇÃO #####
#####

# Na seção que se faz a comparação das equações de predição, ao
# calcular os valores dos critérios de adequabilidade de modelos
# para as equações tomadas na literatura, faz-se a aplicação ape-
# nas da 2ª PARTE desta rotina. No entanto, é necessário escre-
# ver a equação observada para se obter as estimativas da EMAn
# das mesmas, a partir dos valores de composição química obtidos
# no ensaio metabólico e, também, retirar o comando que calcula o
# valor do AIC da predição. Assim, a parte variável definida no
# início da 2ª PARTE desta rotina, ao utilizar uma das equações
# obtidas por Nascimento (2009), fica da seguinte forma:

dados = read.table("Planilha_geral_predict.txt", header=T)
attach(dados)
eq <- 4205.23 + 30.58*EE - 130.35*MM - 58.29*FB -
28.31*FDN + 16.71*FDA          # equação proposta por Nascimento

# Ao considerar outros trabalhos de outros autores, a equação
# obtida pelo mesmo, deve ser colocada no lugar desta.

eq <- round(eq)              # Arredondando casas decimais da eq

# Após essa alteração, a rotina segue conforme passos anteriores.

```

ANEXO B Energia metabolizável aparente corrigida para nitrogênio retido (kcal/kg de matéria seca) de alimentos energéticos, obtida por ensaio de metabolismo ou por equações de predição para alimentos energéticos

FONTE	ALIMENTO									
	Farelo de trigo	Farelo de arroz	Sorgo 1	Sorgo 2	Milho pré-cozido	Quirera de milho	Milho 1	Milho 2	Quirera de arroz	
EMAn observada (Alvarenga, 2009)	1941	2682	3529	3598	3624	3676	3699	3813	3862	
Energéticos com 10 grupos	2001	2611	3605	3639	3762	3801	3774	3808	3960	
Energéticos com 20 grupos	2122	2765	3610	3646	3718	3767	3793	3844	3981	
Concentrados com 10 grupos	1990	2394	3569	3591	3718	3743	3768	3795	3896	
Concentrados com 20 grupos	1935	2352	3563	3585	3718	3730	3766	3797	3916	
Nascimento (2007)	1919	2670	3566	3621	3667	3737	3767	3833	4023	
	1917	2596	3579	3610	3680	3709	3747	3832	4058	
	1961	2420	3554	3578	3663	3754	3741	3776	3872	
	1991	2393	3537	3562	3684	3733	3725	3742	3831	
Rodrigues et al. (2001)	1971	2668	3527	3585	3727	3650	3597	3662	4063	
	2393	1776	3776	3765	3941	3779	3977	3991	4150	
Nagata et al. (2004)	2670	1648	3619	3608	3756	3667	3746	3703	3778	

(*) Equações obtidas:

Eq. 1: $4215,773 + 51,211EE - 193,436MM - 28,034FDN$

Eq. 2: $4235,378 - 11,487PB + 40,232EE - 138,642MM - 42,872FB - 21,738FDN + 11,743FDA$

Eq. 3: $4144,914 + 53,137EE - 204,644MM - 26,214FB - 20,260FDN$

Eq. 4: $4164,187 + 51,006EE - 197,663MM - 35,689FB - 20,593FDN$

Eq. 5: $4371,18 - 26,48PB + 30,65EE - 126,93MM - 52,26FB - 25,14FDN + 24,40FDA$

Eq. 6: $4205,23 + 30,58EE - 130,35MM - 58,29FB - 28,31FDN + 16,71FDA$

Eq. 7: $4101,33 + 56,28EE - 232,97MM - 24,86FDN + 10,42FDA$

Eq. 8: $4095,41 + 56,84EE - 225,26MM - 22,24FDN$

Eq. 9: $4281,6 - 39,97FDN - 72,90MM$

Eq. 10: $4354,8 - 112,05FB - 151,74MM$

Eq. 11: $4021,8 - 227,55MM$

As equações 1 a 4 foram obtidas neste trabalho, utilizando-se a aplicação de meta-análise e componentes principais.

As equações 3, 4, 7 e 8 foram obtidas, considerando como base de dados os alimentos concentrados e as demais equações, considerando alimentos energéticos.

ANEXOC Energia metabolizável aparente corrigida para nitrogênio retido (kcal/kg de matéria seca) de alimentos soja + subprodutos, obtida por ensaio de metabolismo ou por equações de predição para alimentos soja + subprodutos

	ALIMENTO									
	Far. soja 1	Far. soja 2	Far. soja 3	Far. soja 4	Proteína texturizada	Far. soja semi-integral	Soja int. micronizada	Soja int. extrusada	Protenose	
EMAn observada (Alvarenga, 2009)	2326	2355	2396	2478	2809	3159	3772	3779	3934	
Soja + subprodutos com 10 grupos	2380	2416	2438	2475	2580	2952	3880	3556	3169	
Soja + subprodutos com 20 grupos	2330	2359	2318	2392	2630	2992	3596	3514	3634	
Concentrados com 10 grupos	2333	2436	2490	2533	2891	2962	3985	3651	3828	
Concentrados com 20 grupos	2318	2443	2498	2533	2937	2910	3967	3589	3837	
Nascimento (2007)	2524	2576	2582	2584	2685	3057	3904	3700	2804	
	2331	2365	2428	2498	2724	3057	3874	3743	3865	
Rodrigues et al. (2002)	2318	2360	2416	2482	2748	3037	3918	3725	3780	
	2207	2455	2467	2411	2754	2525	3989	3235	2873	
Rodrigues et al. (2004)	2141	2387	2237	2264	2842	2614	4045	3143	3786	
Zontia et al. (2004)	2314	2397	2376	2366	2645	2831	4003	3487	2395	

(*) Equações obtidas:

$$Eq. 1: 338,3 \cdot 589 + 51,412EE - 139,617MM - 15,474FDA$$

$$Eq. 2: 1322,267 + 33,894PB + 73,606EE - 84,217MM + 47,200FB - 29,936FDN$$

$$Eq. 3: 4144,914 + 53,137EE - 204,644MM - 26,214FB - 20,260FDN$$

$$Eq. 4: 4164,187 + 51,006EE - 197,663MM - 35,689FB - 20,593FDN$$

$$Eq. 5: 2707,71 + 58,63EE - 16,06FDN$$

$$Eq. 6: 4101,33 + 56,28EE - 232,97MM - 24,86FDN + 10,42FDA$$

$$Eq. 7: 4095,41 + 56,84EE - 225,26MM - 22,24FDN$$

$$Eq. 8: 2822,2 - 90,13FB + 49,96EE$$

$$Eq. 9: -822,33 + 69,54PB - 45,26FDA + 90,81EE$$

$$Eq. 10: 2723,05 - 50,52FDA + 60,40EE$$

As equações 1 a 4 foram obtidas neste trabalho, utilizando-se a aplicação de meta-análise e componentes principais.

As equações 3, 4, 6 e 7 foram obtidas, considerando como base de dados os alimentos concentrados e as demais equações, considerando alimentos soja + subprodutos.