UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA

REDUÇÃO DO TEMPO PARA ESTIMATIVA DO CONSUMO ALIMENTAR RESIDUAL E PREDIÇÃO DO CONSUMO ALIMENTAR EM OVINOS

FERNANDO AMARILHO SILVEIRA
ZOOTECNISTA/UFPEL
MESTRE EM ZOOTECNIA/UFPEL

Porto Alegre (RS), Brasil
abril de 2022

FERNANDO AMARILHO SILVEIRA

REDUÇÃO DO TEMPO PARA ESTIMATIVA DO CONSUMO ALIMENTAR RESIDUAL E PREDIÇÃO DO CONSUMO ALIMENTAR EM OVINOS

Tese apresentada como um dos requisitos à obtenção do grau de Doutor em Zootecnia na área de concentração de Produção Animal, Programa de Pós-Graduação em Zootecnia, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul.

Orientador: Prof. Jaime Araujo Cobuci
Coorientador: Dr. Gabriel Ciappesoni

Porto Alegre (RS), Brasil
abril de 2022

CIP - Catalogação na Publicação

Amarilho Silveira, Fernando
REDUÇÃO DO TEMPO PARA ESTIMATIVA DO CONSUMO ALIMENTAR RESIDUAL E PREDIÇÃO DO CONSUMO ALIMENTAR EM OVINOS / Fernando Amarilho Silveira. -- 2022.
129 f.
Orientador: Jaime Araujo Cobuci.

Coorientador: Gabriel Ciappesoni.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Agronomia, Programa de Pós-Graduação em Zootecnia, Porto Alegre, BR-RS, 2022.

1. ganho de peso médio diário. 2. inteligência artificial. 3. eficiência alimentar. 4. modelo genético. 5. ovinos. I. Araujo Cobuci, Jaime, orient. II. Ciappesoni, Gabriel, coorient. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Fernando Amarilho Silveira
Mestre em Melhoramento Genético de Ovinos

TESE

Submetida como parte dos requisitos
para obtenção do Grau de

DOUTOR EM ZOOTECNIA

Programa de Pós-Graduação em Zootecnia
Faculdade de Agronomia
Universidade Federal do Rio Grande do Sul
Porto Alegre (RS), Brasil

Aprovada em: 27.04.2022
Pela Banca Examinadora

JAIME ARAÚJO COBUCI
PPG Zootecnia/UFRGS
Orientador

Fernando Sebastian Baldi Rey
UNESP - FCAV

Luiz Fernando Brito
Purdue University/EUA

Ricardo Lopes Dias da Costa
Instituto de Zootecnia/Apta/SAA-SP

Homologado em: 05/07/2022
Por

SERGIO LUIZ VIEIRA
Coordenador do Programa de
Pós-Graduação em Zootecnia

CARLOS ALBERTO BISSANI
Diretor da Faculdade de Agronomia

Redução do tempo para estimativa do consumo alimentar residual e predição do consumo alimentar em ovinos[1]

Autor: Fernando Amarilho Silveira
Orientador: Prof. Jaime Araujo Cobuci
Coorientador: Dr. Gabriel Ciappesoni

Resumo: A avaliação do consumo alimentar de ovinos em provas de eficiência alimentar é uma característica de difícil mensuração, associada ao alto custo de mão de obra, alimentação e instalações. Diminuir o período de teste e aplicar abordagens de aprendizado de máquina, considerando potenciais preditores, otimizaria os recursos e forneceria maior número de animais com informações de consumo, e isso reduziria o custo por animal e permitiria avaliar um maior número por ano. Assim, os objetivos desta tese foram explorar modelos de consumo alimentar residual (CAR) para diminuir a duração da prova e testar quatro abordagens para encontrar o melhor método de predição do consumo. Foram realizadas duas pesquisas com o período de teste composto por 56 dias. No primeiro artigo, foram utilizados dados de 286 ovinos Merino Australiano de três provas de consumo. Dois modelos foram usados para calcular o CAR, Modelo 1 (modelo linear) e Modelo 2 (modelo semanal ou de medidas repetidas por semana). O Modelo 1 incluiu o ganho médio diário (GMD) obtido por regressão linear e o consumo médio de todo o período de prova. O Modelo 2 incluiu o consumo médio semanal como medida repetida e o GMD semanal. No segundo estudo, foram utilizados dados de 613 ovinos Merino Australiano, 156 Corriedale e 237 Dohne Merino de dez provas de consumo. Para treinar os modelos, foram utilizados 477, 70 e 93 observações de ovinos Merino Australiano, Corriedale e Merino Dohne, respectivamente. E a validação direta foi realizada em conjuntos de dados de 97 Merino Australiano, 79 Corriedale e 125 Merino. Os modelos de predição foram regressão linear, regressão não linear, regressão de k-vizinhos mais próximos (kNN) e regressão de floresta aleatória (RF). No primeiro estudo, as correlações de Pearson e Spearman entre modelos lineares e semanais, com 42 dias, foram de 0,89 e 0,87, respectivamente, para GMD. As melhores correlações foram detectadas para consumo entre os modelos de 42 e 35 dias, apresentando correlações de Pearson e Spearman de 0,95 e 0,94 para o modelo linear e 0,96 e 0,95 para o modelo semanal. Ao considerar o CAR, as correlações entre os modelos linear e semanal com 42 dias foram de 0,93 e 0,92, respectivamente. Os modelos CAR com 35 dias (linear e semanal) apresentaram correlações de Pearson e Spearman superiores a 0,98 com os modelos de 42 dias. Portanto, os modelos CAR com 35 dias de duração permitiram diminuir sete dias de prova, explicando 75,3% do consumo no modelo linear e 63,6% no modelo semanal. No segundo estudo, o maior R2 foi encontrado no model não linear com valor de 0,76 na validação cruzada no conjunto de dados de treinamento, com correlação de Pearson e Spearman, respectivamente, de 0,87 e 0,85. Na validação direta nos conjuntos de dados do Merino Australiano, Corriedale e Dohne Merino, foram encontrados correlação de Pearson de 0,69 a 0,86 e Spearman de 0,63 a 0,86. A redução

---

[1] Tese em Zootecnia – Faculdade de Agronomia – Departamento de Zootecnia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil. (129 p.), abril, 2022.

de sete dias de teste e o uso de abordagens para previsão de consumo proporcionará uma maior coleta de dados em um ano de avaliação fenotípica, pela otimização de recursos e pela maior quantidade de animais com informação de consumo alimentar.

Palavras-chave: ganho de peso médio diário, inteligência artificial, eficiência alimentar, modelo genético**.**

Reduction of time to estimate residual feed intake and prediction of feed intake in sheep[2]

Author: Fernando Amarilho Silveira
Advisor: Prof. Jaime Araujo Cobuci
Co-advisor: Dr. Gabriel Ciappesoni

Abstract: The evaluation of sheep feed intake (FI) in feed efficiency tests is a trait of difficult measurement, associated with the high cost of labor, feeding, and facilities. Decreasing the test period and applying machine learning approaches, considering traits as potential predictors, could optimize the resources and provide a greater number of animals with feed intake information, reducing the cost of evaluating by animal and allowing to test a greater number per year. For this reason, the objectives of this thesis were to explore residual feed intake (RFI) models to decrease the test duration and to test four approaches to find the best feed intake prediction method. Two research were carried out with the test period consisting of 56 days. In the first paper, the data was collected from 286 Australian Merino sheep of three performed trials. Two models were used to calculate RFI, Model 1 (linear model) and Model 2 (weekly model or repeatability model by week). Model 1 included average daily gain (ADG) estimates by linear regression and average FI of the period. The model 2 included weekly average FI as a repeated measure and the weekly ADG. The second study comes from 613 Australian Merino, 156 Corriedale, and 237 Dohne Merino sheep of ten feed intake trials. To train the models, 477, 70 and 93 observations of Merino Australian, Corriedale and Merino Dohne sheep were used, respectively. And direct validation was performed on 97 Australian Merino, 79 Corriedale and 125 Merino datasets. The prediction models were linear regression, nonlinear regression, k-nearest neighbors regression, and random forest regression. In the first study, the Pearson and Spearman correlations between linear and repeated models with 42-days for ADG were 0.89 and 0.87, respectively. The best correlations were detected for FI between 42 and 35-days models, presenting Pearson and Spearman correlations of 0.95 and 0.94 in the linear model, and 0.96 and 0.95 in the weekly model. When considering RFI, the correlations between linear and weekly 42-days models were from 0.93 and 0.92, respectively. The 35-days RFI length models (linear and weekly) presented Pearson and Spearman correlations greater than 0.98 with the 42-days models. Therefore, the RFI models 35-days of duration allowed to decrease seven days of the FI test while maintaining accuracy and explaining 75.3% of the FI in the linear model and 63.6% of the weekly model. In the second study, the highest R^2 was found in the nonlinear model with a value of 0.76 in the cross-validation of the training data set, with Pearson and Spearman correlation, respectively, of 0.87 and 0.85. In direct validation on the Australian Merino, Corriedale and Dohne Merino datasets, Pearson's correlation from 0.69 to 0.86 and Spearman's correlation from 0.63 to 0.86 were found. Reducing seven days of testing and use approaches for FI prediction would provide a greater data collection into a year of phenotypic evaluation.

[2] Ph.D. thesis in Animal Science – Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil. (129 p.), April, 2022.

Key words: average daily gain, artificial intelligence, feed efficiency, genetic model

SUMÁRIO

LISTA TABELAS

TABELAS CAPÍTULO IV

LISTA FIGURAS

FIGURAS CAPÍTULO II

FIGURAS CAPÍTULO III

FIGURAS CAPÍTULO IV

CAPÍTULO I - INTRODUÇÃO

1. INTRODUÇÃO

O desafio do melhorista é identificar e selecionar animais que, com certa quantidade e qualidade de alimento, e sob certas circunstâncias, produzirão em um menor espaço de tempo maior quantidade de carne, pele, leite e/ou lã com qualidade, proporcionando um maior lucro ao criador e com um menor custo ao consumidor (GOODALE, 1861).

Por muito tempo, o caminho para atingir o lucro dentro de um sistema produtivo era aumentar a eficiência produtiva, que visava melhorar os processos para otimizar a utilização das áreas para pecuária e agricultura. Atualmente, melhorar a produtividade não é mais suficiente frente à demanda mundial, em que os consumidores não apenas querem um alimento barato e saudável, mas também buscam uma garantia de prática amigável ao meio ambiente. Com isso vem à tona um debate, que incide principalmente na pecuária, que são as reduções das emissões de gases de efeito estufa e o aumento da produção de proteína animal, em especial a carne, devido ao constante crescimento da população mundial.

O ACORDO DE PARIS limitou o aquecimento global a 1,5°C com uma redução na emissão de gases de efeito estufa de 16 a 41% até 2050 comparativos ao ano de 2010 (LEAHY; CLARK; REISINGER, 2020). Em paralelo, é projetado que a população mundial aumente em 38% nesse mesmo período (KOCHHAR; OATES, 2014), que irão demandar em 2050 mais do que 70% de carne comparado aos níveis de 2010 (MCLEOD, 2011). Isso representa que o sistema pecuário terá que ser ainda mais eficiente no uso da terra por aumentar os níveis produtivos e ainda proporcionar uma redução nas emissões de gases de efeito estufa.

Em um primeiro momento, está claro quais são as metas a serem atingidas, porém, dentro dos sistemas, alternativas para manter uma produção eficiente e amigável ao meio necessita mais discussões para traçar metas exequíveis e equilibradas, visto que mudanças bruscas podem encontrar produtores e animais inaptos para cumprir tais compromissos externos ao sistema produtivo. O enfoque de tais discussões deveria atacar os desperdícios e estas poderiam começar na busca de animais mais eficientes, que com menos comida produzem mais, visto que o custo com alimentação representa uma grande parcela dos custos de produção. Na Nova Zelândia, com um sistema baseado a pasto, por exemplo, a despesa relativa com alimentação de ovinos e bovinos representou, nos anos de 2019 e 2020, 53% do custo total de produção (BEEF + LAMB NEW ZEALAND, 2021). Já na Irlanda, com um sistema que usa tanto o alimento produzido na própria propriedade como comprado de outras, o National Farm Survey of Ireland, publicado em 2020, destacou que o custo com alimentação pode alcançar 73% do custo total de produção (DILLON; MORAN; DONNELLAN, 2021). Com base nisso, ao selecionar animais que comem menos, além da diminuição dos custos de produção, também terá a diminuição das emissões de gases de efeito estufa, devido à relação linear entre consumo alimentar e emissão de gases (CHARMLEY *et al.*, 2016). Um dos fatores que levam a isso é o fato de que, com o aumento da digestibilidade, o consumo alimentar tende a diminuir e, conjuntamente, ocorre a redução nas emissões de gases (CANTALAPIEDRA-HIJAR *et al.*, 2018), sendo uma das ferramentas para tornar os sistemas produtivamente eficientes sem causar prejuízos ao meio ambiente.

O consumo alimentar possui uma variação natural que ocorre devido ao tamanho e à frequência da refeição, sendo o tamanho influenciado pela taxa e a duração da ingestão e a frequência determinada pelo tempo entre eventos de consumo (ALLEN, 2014), como também a categoria animal e o tipo de alimento (CANTALAPIEDRA-HIJAR *et al.*, 2018). Logo, tanto o tamanho quanto a frequência são diretamente influenciados pela quantidade de nutrientes disponíveis no sangue e pela distensão do rúmen. A quantidade de nutrientes na corrente sanguínea é afetada pelo estágio fisiológico, por exemplo, por diferenças nas taxas de gliconeogênese hepática, velocidade de mobilização e absorção de nutrientes em tecidos extra-hepáticos, secreção e sensitividade dos tecidos a hormônios e citocinas (ALLEN, 2014). Já a distensão ruminal dá-se pelo conteúdo de celulose e lignina que compõe as paredes celulares — principalmente, das forrageiras — e, em uma menor proporção, pelo conteúdo de água dos alimentos (CANTALAPIEDRA-HIJAR *et al.*, 2018). Dessa forma, tanto a saciedade como a fome ao longo de minutos e horas podem ter efeitos a longo prazo na ingestão de alimentos e no balanço energético, caracterizando a variabilidade no consumo diário dos animais.

Entender o papel do consumo alimentar somado ao peso metabólico e à taxa de crescimento dos animais, como uma caracterização de eficiência alimentar, possibilitaria selecionar animais economicamente mais eficientes e amigáveis ao ambiente. Entretanto, o desafio é desenvolver abordagens para a seleção que tenham resultados favoráveis, visto que o aumento nas taxas de crescimento com consequente aumento do peso corporal sem aumentar na mesma proporção o consumo alimentar é geneticamente difícil (CAMMACK *et al.*, 2005). Desse modo, há a necessidade de obter informações de consumo alimentar individual de um maior número de animais, tanto de forma direta como de forma indireta, aproveitando os recursos disponíveis e medidas correlatas para que, quando estes não forem possíveis, possa-se estimar via modelos com tempo reduzido de avaliação, como também por predição via proxy mais fáceis de mensurar.

O sucesso dos programas de melhoramento genético está fortemente relacionado com a quantidade de animais avaliados. No entanto, por muitas vezes, pensando em manter uma quantidade ótima de animais avaliados por geração, poderia-se assumir o risco de obter menor acurácia na tomada de fenótipos em virtude da obtenção de dados de consumo alimentar adicionais aos mensurados (ARCHER; BERGH, 2000), tanto utilizando períodos de testes mais curtos ou por predição com base em informações produtivas e de emissão de gases. Dessa forma, uma agilidade na coleta de dados poderia ser feita para que, no horizonte próximo, fosse possível obter informações suficientes para iniciar as avaliações genéticas para características de consumo e eficiência alimentar.

CAPÍTULO II – REVISÃO BIBLIOGRÁFICA

2. REVISÃO BIBLIOGRÁFICA

2.1. Consumo alimentar

O consumo alimentar é determinado pela integração de sinais centrais e periféricos nos centros de alimentação no cérebro, incluindo sinais de estimulação (orexígeno) que aumentam a fome bem como sinais inibitórios (anorexígeno) que aumentam a saciedade (ALLEN, 2014). Sinais inibitórios são devidos à distensão ruminal, osmolaridade e oxidação dos nutrientes contidos nos alimentos, sendo que ao se afastar do momento da refeição esses eventos inibitórios vão diminuindo e, com o baixo status energético, ocorre um estímulo do consumo, que aumenta do período pré-prandial até o final da refeição (CANTALAPIEDRA-HIJAR *et al.*, 2018). Isso pode ser observado na Figura 1, que mostra a provável dinâmica entre fome e saciedade, em que o início das refeições é determinado quando os sinais inibitórios de uma prévia refeição diminuem e os sinais de estimulação aumentam (a); conforme o consumo aumenta, os sinais de inibição se intensificam e os estimulantes diminuem (b). Isso deixa claro que sinais de fome podem ser sentidos pelos animais como um desconforto e este leva a um evento de consumo; logo, a distensão do trato gastrointestinal, o consumo de substâncias tóxicas e o incremento da produção de calor causado pela metabolização dos nutrientes recém-ingeridos também causarão um desconforto que levará à interrupção do consumo (DADO; ALLEN, 1995). Sendo assim, tanto a saciedade como a fome, ao longo de minutos e horas, podem ter efeitos a longo prazo na ingestão de alimentos e no balanço energético, caracterizando a variabilidade no consumo diário dos animais.

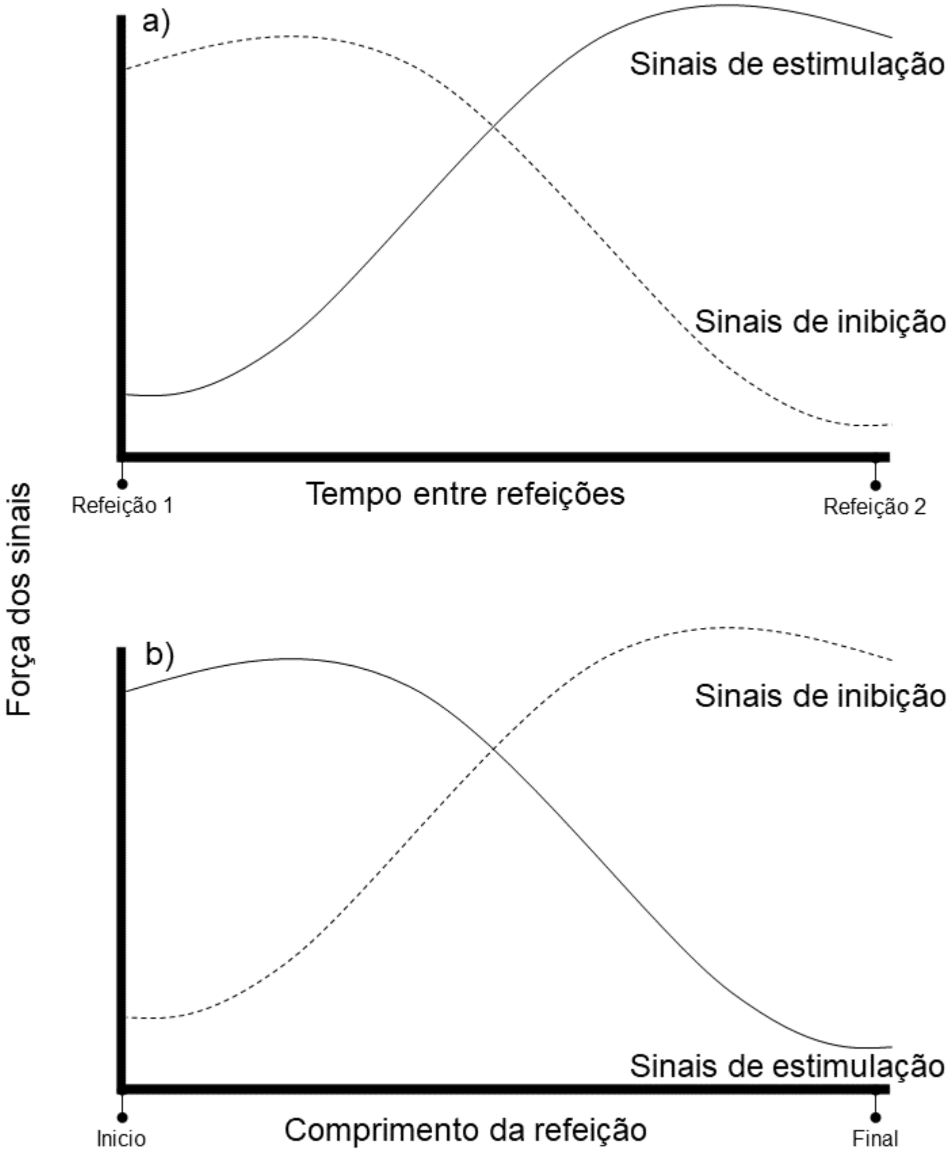


Figure 1. Padrão temporal esperado dos sinais de estimulação (oroxígeno) e de inibição (anoroxígeno) no intervalo (a) entre refeições e (b) durante as refeições (editado de ALLEN, 2014).

Vários estudos utilizam o coeficiente de variação (CV) para descrever a variabilidade do consumo alimentar diário dos animais. Basarab *et al.* (2013) descreveram que valores de CV para o consumo alimentar de bovinos de corte variam de 11 a 20%, sendo considerado que essa variabilidade é um reflexo da variação natural do consumo diário dos animais. Já para ovinos, foram encontrados CV que variaram de 7,3 a 17,1%, com consumos alimentares diários de 0,797 a 3,94 kg.

O consumo médio diário de matéria seca (MS) encontrado no estudo de Ermias *et al.* (2002) em ovinos das raças Mens e Horro a uma idade média de 13 meses, fornecendo feno com 5% de proteína bruta (PB) *ad libitum* e 300 g/dia/ovino de um suplemento concentrado com 18% de PB e 10,5 MJ de energia metabolizável (EM) por kg de MS, variou de 0,797 a 0,838 kg entre duas raças ovinas com um CV de 7.3%.

Redden *et al.* (2011), em ovelhas da raça Targhee alimentadas com feno de gramínea com 8,10% de PB e com pellets com composição de 15,8% de PB, encontraram consumo alimentar variando de 2,08 (cordeiros com nove meses de idade) a 1,77 kg (borregos com 13 meses de idade). Os consumos alimentares encontrados por Johnson et al. (2015), em um composto materno com idade média de 16 meses e com acesso *ad libitum* a uma dieta com pellets de alfada com 10,1 MJ de EM, variam de 2,7 a 3,3 kg.

No estudo de Montelli *et al.* (2019), consumos alimentares de 1,25 a 1,44 kg foram encontrados para ovinos meio sangue Dorper e Santa Inês com idade média de 70 dias, alimentados *ad libitum* com feno de Tifton 85, grão de milho moído e farelo de soja, fechando uma dieta com 20,6% de PB. Já no trabalho de Tortereau *et al.* (2020), em ovinos da raça Romane, com idade média variando de 60 a 90 dias e uma dieta *ad libitum* via pellets concentrado de baixo conteúdo energético com 15,7% de PB, encontraram o consumo alimentar e o CV, respectivamente, de 1,962 kg e 13%.

Nota-se que existe uma grande variação no consumo médio diário em valores absolutos, no entanto, a variação natural no consumo diário tanto para bovinos como para ovinos não extrapola os 20%. Isso se dá em função do tipo de alimento, se é volumoso ou concentrado, bem como sua qualidade, pelo tamanho e pela frequência das refeições, sendo o tamanho determinado pela taxa de consumo e a frequência pelo intervalo entre refeições (ALLEN, 2014).

A taxa de consumo é afetada por fatores dietéticos relacionados com a forma física dos alimentos, a densidade da pastagem, os sinais sensoriais relacionados com o gosto e o olfato, a absorção de nutrientes, bem como interações sociais que causam uma cascata de respostas endócrinas e comportamentais (ALLEN, 2014). Dessa forma, o sistema endócrino serve como um comunicador entre órgãos com a função de regular tanto o consumo alimentar propriamente dito, como a utilização dos nutrientes nas células, o que afeta diretamente a eficiência no uso do alimento ingerido (CANTALAPIEDRA-HIJAR *et al.*, 2018). Na revisão dos mesmos autores, o trato gastrointestinal, com suas glândulas anexas, atua como um primeiro sistema regulador de consumo pela sensibilidade da pressão que o enchimento do trato causa na parede do órgão, principalmente no rúmen (sinal mecânico), e pela disponibilidade de energia. Com base nesta última, hormônios são secretados começando um efeito de regulação química do consumo (por exemplo, insulina e grelina) e a utilização dos nutrientes (por exemplo, polipeptídeo insulinotrópico dependente de glicose, insulina e IGF-1).

Em condições experimentais com ovinos, Texel-Oxford Cammack *et al.* (2005) encontraram correlações genéticas de 0,26 entre o consumo diário e o tempo em alimentação; ou seja, animais que tendem a gastar mais tempo se alimentando também tendem a consumir mais. Nesse mesmo estudo, o número de eventos de consumo apresentou correlação genética de 0,28 com o consumo diário; ou seja, animais que visitam mais o alimentador tendem a comer mais. Já em estudo com bovinos, Basarab *et al.* (2013) mostraram que eventos de consumo de baixa duração estavam associados a um baixo consumo de matéria seca e uma melhor digestibilidade.

É muito importante destacar que, em sistemas a pasto, o comportamento é influenciado por vários fatores. Visto isso, além de fatores inerentes à pastagem (forma física, conteúdo de FDN e digestibilidade), ao comportamento do animal e a interações sociais, em ovinos, existem outros fortes limitantes como o parasitismo por vermes gastrointestinais, que podem levar os animais a fortes eventos de anorexia (COLDITZ, 2008).

Esses múltiplos fatores que afetam o consumo alimentar ainda podem ser mais complexos, visto que existe interação entre o tipo de dieta e o consumo. Vários autores encontraram uma interação entre a eficiência no uso dos alimentos com o tipo de dieta (CAMMACK *et al.*, 2014; CARBERRY *et al.*, 2012; COYLE *et al.*, 2016; DURUNNA *et al.*, 2011; ELLISON *et al.*, 2017; FRANÇOIS *et al.*, 2006; TORAL *et al.*, 2019); sendo que, para ovinos em pastejo, o principal fator limitante do consumo é a distensão ruminal do que propriamente a saturação de nutrientes percebidos pelos tecidos quando o suprimento é maior do que a exigência, como em dietas concentradas

(CANTALAPIEDRA-HIJAR *et al.*, 2018). Forragens com um FDN maior do que 35% podem limitar o consumo devido à sensação de distensão ou resistência da parede ruminal em virtude da forragem consumida (DADO; ALLEN, 1995).

As variáveis comportamentais, como as de capacidade digestiva, possibilitam a identificação do nível de consumo dos animais e, juntamente com a eficiência na utilização dos nutrientes, tornam possível a identificação de ovinos que possuem menores exigências alimentares para produção. Tais animais podem ser identificados por duas métricas: a relação entre a quantidade de alimento consumido e o ganho de peso, chamada de conversão alimentar (CA) e a diferença entre o consumo alimentar observado e a exigência de alimento para mantença e ganho de peso, chamada de consumo alimentar residual (CAR) (CANTALAPIEDRA-HIJAR *et al.*, 2018). A CA é um útil indicador de eficiência alimentar dos animais, no entanto, por se tratar de uma característica composta por uma relação, do ponto de vista genético pode causar desfavoráveis respostas correlacionadas (ZETOUNI *et al.*, 2017). Assim, a diminuição do consumo poderia surtir em uma diminuição no ganho de peso, mantendo ou melhorando a CA, sendo o contrário verdadeiro, em que o aumento no ganho de peso poderia também aumentar o consumo, e o que se deve buscar, no entanto, são animais que comam menos e ganhem mais. Para atender essa demanda, animais com baixo CAR comem menos do que o esperado (conforme suas exigências), com independência do peso corporal e do ganho de peso dos animais, o que possibilita selecionar animais que comem menos sem comprometer o tamanho corporal nem o ganho de peso (KOCH *et al.*, 1963).

2.2. Consumo alimentar residual (CAR)

O conceito de consumo alimentar residual (CAR) pode ser utilizado na constituição de parâmetros referentes à eficiência alimentar, em que diferenças nas exigências de mantença e ganho relacionados à verdadeira capacidade de consumo dos indivíduos podem ser indicativos de animais que comem menos e mantêm os níveis de produção (HERD; ARTHUR, 2009). Esses autores relataram que o CAR pode ser utilizado para identificar animais geneticamente mais eficientes, por se tratar de uma característica herdável. Tal característica apresenta coeficientes de herdabilidade que variam de 0,11 a 0,49 (CAMMACK *et al.*, 2005; FRANÇOIS *et al.*, 2002, 2006; TORTEREAU *et al.*, 2020). Nesse sentido, animais que apresentem valores de CAR negativos são animais que consomem menos do que o estimado (MARIE-ETANCELIN *et al.*, 2019), o que pode ser exemplificado como: consumo observado = 0,900 kg/dia; consumo esperado via uma equação de regressão linear = 1,05 kg/dia; logo CAR = 0,900 – 1,05 → CAR = - 0,15 kg/dia. Isso significa que o animal ilustrado acima consome 150 gramas a menos por dia para cobrir sua exigência de mantença e para ganho de peso.

O CAR é por definição fenotipicamente independente de características produtivas usadas para estimar o consumo individual, ou seja, não há correlações fenotípicas significativas entre o CAR e o ganho de peso, por exemplo (HERD; ARTHUR, 2009). Por sua vez, o consumo individual é estimado dependendo dos diferentes níveis de produção, em que, em termos relativos, animais em mantença consumirão menos do que animais em crescimento. Essa distinção é devido às diferenças nas composições dos ganhos, nas quais animais em mantença apresentam uma exigência basal, apenas para manter suas funções vitais básicas, com ganhos em peso corporal quase que insignificantes. Ao contrário de animais em crescimento, que ainda devem atingir sua maturidade, exigindo uma maior taxa metabólica, consequentemente uma maior exigência relativa (ao peso corporal) no consumo alimentar. Diante disso, é possível atribuir que animais de menores CAR são mais eficientes em crescimento, por comerem menos e apresentarem os mesmos níveis de ganhos.

A seleção por CAR está associada com diferenças no consumo; então, nos animais que comem menos e que atingem o mesmo desempenho, é esperado ter menor quantidade de energia gasta com o incremento de calor com a alimentação (WEBSTER *et al.*, 1975). Isso se dá, principalmente, porque animais com menores CAR, por consumirem menos, produzem menor quantidade de calor devido a uma melhor efetividade em processar o alimento consumido comparativamente a animais com maiores CAR (HERD; ARTHUR, 2009).

Animais com baixo CAR (mais eficientes), além de serem caracterizados por possuírem baixas exigências energéticas para mantença, também são destacados por possuírem uma maior eficiência parcial na utilização da energia metabolizável para crescimento, levando a uma eficiência energética metabólica relativamente alta (CANTALAPIEDRA-HIJAR *et al.*, 2018). Dessa forma, como abordado por esses autores, animais mais eficientes apresentam uma menor taxa metabólica com consequente menor produção de calor (relativa ao peso e à unidade de tempo com similar consumo de energia metabolizável). Essa menor taxa metabólica em animais com baixo CAR pode ter várias origens, visto que cada via metabólica afeta o uso e a produção de ATP (adenosine triphosphate). Como exemplificado por Cantalapiedra-Hijar *et al.* (2018), as mitocôndrias de animais eficientes têm cadeia de transporte de elétrons aprimorada e geram ATP de forma mais eficiente a partir de uma determinada quantidade de substrato.

O metabolismo do nitrogênio microbiano, comparado ao do metabolismo dos carboidratos, está mais relacionado às variações na eficiência alimentar (CAR) em bovinos do que propriamente à maior diversidade na flora microbiana, visto que a inibição da atividade microbiana da flora “não necessária” pode acarretar na melhoria da eficiência alimentar (CANTALAPIEDRA-HIJAR *et al.*, 2018). Em estudo com bovinos de corte, foram encontrados uma relativa abundância do gênero *Prevotella* no rúmen de animais com alto CAR, o que é negativamente associado com isobutirato (r = -0,38), isovalerato (r = -0,43) e digestibilidade da matéria orgânica (r = -0,32) (CARBERRY *et al.*, 2012). Os isoácidos (formas hidrogenadas) apresentam marcada importância na nutrição de ruminantes, tanto o isobutírico (i-C_4) e isovalérico (i-C_5), que são sintetizados no rúmen via uma desaminação oxidativa e uma descarboxilação oxidativa de aminoácidos de cadeia ramificada, como valina e leucina (ANDRIES *et al.*, 1987). Esses autores descreveram que, em cordeiros alimentados com uma dieta purificada contendo 39% de celulose e ureia como fonte de nitrogênio, a digestão da celulose foi melhorada com a adição de uma mistura contendo i-C_4 e i-C_5.

No estudo de Ellison *et al.* (2017), de cinco espécies *Prevotella* que diferiram estatisticamente ($P \leq 0,042$) com o status de eficiência alimentar, quatro (*P. bivia*, *P. paludivivens*, *P. timonensis* e uma espécie de *Prevotella* não conhecida) foram mais abundantes ($P \leq 0,042$) em cordeiros de alto CAR e apenas a *Prevotella nanceiensis* foi mais abundante ($P < 0,001$) em cordeiros com baixo CAR. Nesse mesmo estudo, a espécie degradadora de fibra *Clostridium phytofermentans* foi mais abundante ($P = 0,042$) em cordeiros de alto CAR; a metanogênica *Methanobrevibacter smithii* ($P < 0,001$) e a *Mannheimia haemolytica* ($P \leq 0,031$) foram mais abundantes em capões de baixo

CAR, e em capões de alto CAR, os autores encontraram uma maior abundância ($P < 0{,}001$) das espécies *Sharpea azabuensis*, *Ruminococcus flavefaciens* e *Dialister invisus*.

Ellison *et al.* (2019) foram capazes de predizer com 100% de acurácia animais de baixo, médio e alto CAR, utilizando análise de discriminante múltipla, incluindo cinco *Prevotella genomospecies* (abundante em ovinos de baixo CAR) e duas subespécies de *Pseudobutyrivibrio ruminis* (abundante em baixo CAR), *Schwartzia succinivorans* (abundante em alto CAR) e *Treponema maltophilum* (abundante em baixo CAR).

Não foram encontrados efeitos significativos do CAR ($P > 0{,}05$) em variáveis referente à fermentação ruminal (FITZSIMONS *et al.*, 2013). No entanto, esses autores descrevem uma tendência ($P = 0{,}07$) de animais com alto CAR terem uma maior relação acetato-propionato quando comparados a animais com baixo CAR. Ainda, o CAR não apresentou correlação significativa ($P > 0{,}05$) com nenhuma das variáveis de fermentação ruminal medidas. Concentrações de butirato e valerato foram significativamente ($P < 0{,}01$) maiores em bois com baixo CAR e diferenças não significativas nas concentrações de propionato em relação a bois com altos CAR (GUAN *et al.*, 2008).

No que tange a bovinos, foi constatado que animais com CAR negativo (mais eficientes), além de consumirem menos, apresentaram uma menor produção de calor e uma menor retenção de energia na forma de gordura do que animais com CAR positivo, sem diferenças nas características de crescimento e distribuição de cortes (BASARAB *et al.*, 2003). O mesmo estudo destacou também que animais com CAR positivo apresentaram maior peso dos órgãos digestivos e fígado. Isso mostra que há múltiplos fatores por trás da fisiologia, responsáveis pela variação na eficiência alimentar dos animais, e tal pluralidade passa por uma certa exigência em modelagem do consumo esperado.

O modelo de consumo alimentar de Koch *et al.* (1963) utiliza uma simples equação linear considerando o peso metabólico e o ganho de peso no período de avaliação dos animais. Por exemplo: $Y = \mu + PV^{0,75} + GMD + \varepsilon$, em que o Y é o consumo alimentar diário (consumo alimentar total por dia) expresso em matéria seca; μ é um constante referente à média do consumo alimentar diário; $PV^{0,75}$ é o peso metabólico (peso vivo tomado no meio do período de teste elevado a 0,75); GMD é o ganho de peso médio diário; e ε é o resíduo do modelo referente ao CAR (diferença entre o consumo observado e esperado). Já no estudo de Cammack *et al.* (2005), o consumo alimentar (Y da equação de Kock *et al.* 1963) foi modelado pelo consumo diário total (sempre em matéria seca) = média geral do consumo alimentar + grupo de contemporâneos (ano e curral) + soma dos pesos corporais diários (42 x peso corporal inicial + 21 x ganho de peso médio diário) + ganho de peso total (diferença entre o peso aos 42 dias como peso final e o peso inicial). O CAR no estudo de Durunna *et al.* (2011) foi calculado com base em um modelo de consumo alimentar que possui, além do $PV^{0,75}$, GMD e grupo de contemporâneos, também espessura de gordura subcutânea medida por ultrassonografia sobre o músculo *Longissimus dorsi* entre a 12ª e 13ª costela. Essa variabilidade de modelos é de importante destaque visto que a discussão é pautada no ajuste dos modelos, logo em ovinos foram encontrados R^2 variando de 0,38 a 0,84 nos mais variados modelos.

Assim, Knott *et al.* (2008), usando o modelo proposto por Koch *et al.* (1963), encontraram R^2 de 0,63 e 0,56 para carneiros aos seis e 13 meses de idade, respectivamente. No mesmo trabalho, no entanto, utilizando o modelo composto pelo GMD e peso corporal no meio do teste de Knott *et al.* (2008), estimaram o R^2 em 0,74 e 0,60, respectivamente, para carneiros das mesmas idades como referido acima. Em animais avaliados em duas subsequentes idades, os R^2 foram de 0,41 e 0,38 para ovelhas com 280 e 414 dias de idade, respectivamente (REDDEN *et al.*, 2011). Cockrum *et al.* (2013) encontraram R^2 para vários modelos de predição do consumo alimentar, variando de 0,43 a 0,46. Já Redden *et al.* (2014) encontraram R^2 igual a 0,84. O modelo composto pelo peso corporal e pelo GMD apresentou R^2 igual a 0,79 (JOHNSON *et al.*, 2015). Em Johnson *et al.* (2016), o R^2 encontrado foi de 0,78. Em estudo subsequente, Johnson *et al.* (2017) descreveram que bons ajustes em modelos para predizer o consumo alimentar devem apresentar R^2 maior que 0,70. O que em Zhang *et al.* (2017), Montelli *et al.* (2019) e Tortereau *et al.* (2020), com valores de 0,80, 0,82 e 0,84, respectivamente, conferem bons ajustes com consequentes ótimas predições de consumo alimentar.

Alguns autores atribuem que são necessárias, para uma boa modelagem, acuradas estimativas da taxa de crescimento, sendo que para isso seriam necessárias múltiplas pesagens ao longo do tempo em avaliação. Dessa maneira, a energia necessária para o crescimento seria facilmente distinguida pela computação do consumo e das diferentes taxas de crescimento diário, possibilitando mais acuradamente destacar animais com consumos menores do que o estimado, mantendo níveis produtivos compatíveis com a taxa de crescimento desejada (ARCHER *et al.*, 1997; JOHNSON *et al.*, 2015).

2.3. Período de teste para provas de consumo alimentar

Para conseguir diferenciar os animais por níveis de consumo e taxa de crescimento, normalmente, o período de avaliação em provas de consumo alimentar consiste em 56 dias; sendo 14 dias para adaptação à dieta e a estruturas como comedouros e bebedouros para garantir o acesso *ad libitum* e 42 dias de coleta de dados (CAMMACK *et al.*, 2005; JOHNSON, *et al.*, 2015; JOHNSON, P. L.; MILLER; KNOWLER, 2015; JOHNSON *et al.*, 2017; LEYMASTER *et al.*, 2002; SNOWDER; VAN VLECK, 2002; TORTEREAU *et al.*, 2020). Portanto, existe a possibilidade de capturar a maior parte da variação entre os animais, tanto para consumo alimentar como para o desempenho em ganho de peso, utilizando períodos de avaliação mais curtos, principalmente quando existe a possibilidade de utilizar equipamentos que automaticamente medem o consumo e os pesos diários, de modo que possibilite a diminuição dos custos por animal avaliado com períodos de teste mais curtos sem comprometer a acurácia das mensurações (ARCHER *et al.*, 1997). Alguns trabalhos em ovinos já obtiveram resultados promissores utilizando períodos de avaliação com 35 dias sem comprometimento da acurácia das avaliações (MACLEAY *et al.*, 2016; PAGANONI *et al.*, 2017).

Em bovinos de corte, o estudo de Castilhos *et al.* (2011) foi detalhista em descrever que as variâncias para consumo alimentar diminuíram em torno de 0,18% de 28 a 56 dias, aumentaram 13,67% de 56 a 84 dias e aumentaram novamente em 11,22% de 84 a 112 dias (final do período de teste). Os autores atribuem esse aumento significativo do consumo alimentar após os 56 dias de teste pelo aumento do consumo alimentar dado em resposta ao aumento do peso corporal dos animais. Sendo assim, a variância do consumo de matéria seca não se estabilizou ao longo do período de teste pelas diferenças inerentes aos animais, e as correlações (r) de Pearson e Spearman foram utilizadas para determinar que o período ótimo de teste foi com 84 dias (o que representa

75% do período total), apresentando valores de r > 0,93 para esses autores. Dessa forma, tanto o estudo da variação no consumo alimentar em janelas de dias em teste como o uso de correlações de Pearson e Spearman poderiam ser indicadores sobre o impacto na diminuição dos dias em avaliação. Uma progressão disso, que proporcionará uma validação mais acurada e com garantia de que a redução dos períodos de teste não impactará negativamente na seleção de animais mais eficientes, são estudos a nível genético, em que correlações genéticas entre os diferentes períodos de teste trariam muito mais segurança na identificação de animais no progresso genético para características de consumo alimentar, ganho de peso e eficiência alimentar (GOONEWARDENE *et al.*, 2004).

Outro ponto a ser considerado no momento de propor uma redução no período de avaliação de consumo alimentar é o quanto essa prática impactaria na redução do custo por animal avaliado e qual é o potencial de aumentar o número de animais avaliados em testes de consumo alimentar. Nesse contexto, Archer e Bergh (2000) descrevem que, para programas de melhoramento genético, se recursos são economizados por diminuir o período de teste e com um acréscimo de animais avaliados gerando dados adicionais para as avaliações genéticas, a perda de acurácia por avaliar os animais com um menor tempo de prova de consumo alimentar vai ser parcialmente compensada pela informação extra obtida por avaliar um maior número de indivíduos. Outro ponto que vale destaque, não propriamente reduzindo o tempo de avaliação, é a possibilidade de utilizar características com alto potencial de predição do consumo alimentar, para que em um ciclo de avaliação seja possível mensurar uma amostra de animais e predizer o consumo para outros tantos.

2.4. Metano (CH_4) e dióxido de carbono (CO_2) vs. consumo alimentar

A seleção de animais que comem menos não vai apenas diminuir os custos de produção, mas também indiretamente diminuirá as emissões de gases de efeito estufa devido à linear relação entre o consumo alimentar e a produção de gases (CHARMLEY *et al.*, 2016). Um dos aspectos responsáveis por essa relação é o fato de que uma melhora na digestibilidade causa uma redução no consumo alimentar e, consequentemente, uma redução nas emissões de gases (CANTALAPIEDRA-HIJAR *et* *al.*, 2018). Esse dado é importante, visto que a produção de metano (CH_4) pelos ruminantes constitui cerca de 6 a 10% da energia consumida das pastagens, considerando uma digestibilidade de 8 a 14% (COTTLE; NOLAN; WIEDEMANN, 2011).

A previsão para o ano de 2022, descrito por Alford *et al.* (2006), é de que a produção total de CH_4 pela bovinocultura de corte seja em torno de 1.964.800 ton., taxando ruminantes como grandes fontes emissoras de CH_4 (SCHAEFER *et al.*, 2016). Porém, esse excesso de CH_4 eliminado reflete diretamente na baixa eficiência energética dos ruminantes, que por ser uma característica herdável, poderia ser um critério de seleção para melhorar o uso da energia dos alimentos e, consequentemente, diminuir as emissões ao longo do tempo (JONKER *et al.*, 2018; ROBINSON *et al.*, 2014). Dessa forma, se for obtida uma redução de 25,5 ton. de CH_4 ao longo de 25 anos, isso representaria uma economia média anual equivalente a 20,6 ton. de CO_2 (ALFORD *et al.*, 2006).

Basarab *et al.* (2013) revisaram alguns artigos que descrevem que a produção de CH_4 tem herdabilidade de baixa a moderada (0,13 a 0,38) e positivamente correlacionada com o CAR (0,35 a 0,44). Robinson *et al.* (2014), por sua vez, obtiveram estimativas de herdabilidade, para ovinos, de 0,12. Jonker *et al.* (2018) obtiveram coeficientes de herdabilidade direta de 0,19 a 0,23 e coeficientes de herdabilidade materna de 0,04 a 0,05 para a emissão de CH_4 em ovinos. Esses autores correlacionaram essa variável com o peso corporal, sendo que para cordeiros a correlação genética foi de 0,43 e para ovelhas 0,31. Já correlações genéticas de 0,42 a 0,95 entre o consumo alimentar e CH_4 mensurado via câmaras de acumulação portáteis (CAP) foram obtidas para ovinos (PAGANONI *et al.*, 2017; ROBINSON *et al.*, 2020). Com isso surge a hipótese de que o CH_4 poderia ser usado como proxy para predição do consumo alimentar (ROBINSON; ODDY, 2016), principalmente quando a mensuração de consumo alimentar não é possível ou apresenta um alto custo, as medições de gases utilizando CAPs poderiam ser aplicadas a nível de campo (DOMINIK *et al.*, 2017), facilitando o processo de identificação de animais com menores emissões e, indiretamente, com menores consumos.

Outro gás que supostamente poderia ser considerado na busca de animais com menores consumos alimentares e que apresenta grande impacto ambiental é o dióxido de carbono (CO_2). Visto que correlações fenotípicas e genéticas, em cordeiros e borregos, de 0,71 a 0,74 e 0,86 a 0,96, respectivamente, entre consumo alimentar e medidas de CO_2, foram encontradas por Paganoni *et al.* (2017). Arthur *et al.* (2018) observaram que à medida que o consumo alimentar aumenta a emissão de CO_2 também é aumentada. Renand *et al*. (2019) apontaram que o CO_2 pode ser considerado como uma potencial variável, juntamente com o CH_4, para a seleção indireta com vistas à eficiência alimentar quando medidas de consumo alimentar não estão disponíveis. Isso se dá porque o CO_2 é produzido tanto pela fermentação ruminal quanto pela oxidação dos substratos, sendo proporcional ao gasto energético (HEGARTY, 2013), e a emissão de CH_4 é relatada como um ineficiente processo digestivo, representando uma importante perda de energia bruta dos alimentos (OLIJHOEK *et al.*, 2018). Desse modo, a redução das emissões gasosas de uma maneira geral, devido à já comentada relação com a melhora na digestibilidade com um menor consumo alimentar, ratifica-os como proxy para predição do consumo alimentar dos ruminantes.

2.5. Características produtivas vs. consumo alimentar

Ao contrário das relações existentes entre emissões e consumo alimentar, as características produtivas em ovinos não seguem a mesma linha. As baixas correlações fenotípicas e genéticas entre características produtivas e consumo alimentar encontradas por Safari *et al.* (2007) e Fogarty *et al.* (2009) dificultariam o uso destas como critério de seleção indireta para o consumo alimentar. No entanto, Tortereau *et al.* (2020) descrevem baixa a média de correlações genéticas e fenotípicas entre consumo alimentar e medidas tomadas no período de sua avaliação. Com isso abre-se uma porta para justificar o estudo dessas características como potenciais variáveis com boa capacidade preditiva do consumo alimentar.

Variáveis mensuradas no período de avaliação do consumo alimentar, normalmente, apresentam importante influência na explicação do consumo individual. Desde o proposto por Koch *et al.* (1963), a modelagem do consumo alimentar vem sendo estudada com o objetivo de predizê-lo de forma linear, tendo como base o peso corporal e o ganho de peso como covariáveis. A partir daí vários estudos foram empreendidos para tentar obter os melhores ajustes com uma garantia de melhor predição linear, o que torna as medidas corporais e de desempenho no período de avaliação do consumo alimentar validadas como proxy (JOHNSON; WING; KNOWLER, 2015; KNOTT *et al.*, 2008; MONTELLI *et al.*, 2019; ZHANG *et al.*, 2017).

Com vistas no descrito, não teria evidências suficientes para pensar em empreender estudos de predição do consumo alimentar utilizando características produtivas medidas fora do período de teste de consumo alimentar. Porém, visto que em programas de melhoramento genético essas são tomadas de maneira estruturada e que o conjunto de dados, quando fornecidos para estudos, possuem tais informações, cria-se um anseio de testá-las mesmo que as evidências encontradas na bibliografia digam o contrário. Somado à curiosidade não fundamentada em outros trabalhos e sem um argumento profundo, existe a possibilidade de encontrar resultados promissores com o uso de abordagens diferentes. Nesse contexto, o uso de abordagens de aprendizado de máquina possui vantagens comparadas a modelos lineares comuns em aprender relações de um conjunto de dados utilizado para treinamento e com a possibilidade de generalizar para um conjunto de dados ainda não conhecido pelos algoritmos, com isso possibilitando extrair informações não lineares como a interação entre variáveis na busca da predição (SHAHINFAR; KAHN, 2018), que aplicado na presente proposta seria utilizar o máximo de informações para predizer o consumo alimentar, selecionando as que melhor se relacionam em prol de uma predição mais acurada.

2.6. Abordagens de aprendizado de máquina

Face à importância do consumo alimentar para a eficiência dos sistemas produtivos e por sua dificuldade de mensuração devido ao alto custo com mão-de-obra, alimento e estruturas, a aplicação de abordagens de aprendizado de máquina para identificar variáveis, e a relação entre elas, como potenciais preditoras e desenvolver modelos preditivos, poderia otimizar os recursos e obter informações de consumo alimentar de muitos indivíduos.

Alguns estudos com ovinos vêm sendo conduzidos utilizando abordagens de aprendizado de máquina para predição do crescimento e qualidade de lã em animais adultos (SHAHINFAR; KAHN, 2018), para detectar comportamentos básicos em pastejo (FOGARTY *et al.*, 2020), para classificar a mortalidade de cordeiros dentro de classes de risco (ODEVCI; EMSEN; AYDIN, 2021) e para predizer o escore de condição corporal (SEMAKULA *et al.*, 2021). Contudo, não se tomou conhecimento da aplicação do aprendizado de máquina para predição do consumo alimentar para ovinos, sendo uma fronteira do conhecimento nessa temática.

Em estudo com bovinos leiteiros, foi possível predizer o consumo alimentar com coeficientes de correlação de Pearson variando de 0,71 a 0,76, utilizando modelos mistos e abordagens de aprendizado de máquina (KAMPHUIS *et al.*, 2017). Em outro estudo, também com vacas leiteiras, o modelo Random Forest (modelo que utiliza algoritmos baseado no aprendizado de máquina) mostrou possuir um bom potencial de predição do consumo alimentar de vacas em pastejo (LESO *et al.*, 2019). Isso mostra e reforça a necessidade e o potencial de pesquisa para ovinos.

Dentro das abordagens de aprendizado de máquina, procedimentos de seleção de variáveis são importantes para diminuir o trabalho computacional devido à alta dimensionalidade e à multicolinearidade. Assim, a seleção de variáveis pode ser entendida também como a eliminação de variáveis que não contribuem no entendimento dos dados, com o objetivo de reduzir o requerimento computacional e o efeito da dimensionalidade, e melhorando a performance preditiva (CHANDRASHEKAR; SAHIN, 2014). Quando variáveis possuem alto nível de redundância, diferentes conjuntos de treinamento podem produzir diferentes classificações de variáveis a serem consideradas no modelo de predição (PILES *et al.*, 2021). Dessa forma, a seleção de variáveis que melhorem a performance na predição diminuiria a exigência computacional e proporcionaria uma melhor interpretação dos resultados.

Dentro de uma lógica de ações na aplicação do aprendizado de máquina pode-se apontar: 1ª – edição e análise de consistência dos dados; 2ª – seleção de variáveis e; 3ª – validação cruzada. A validação cruzada é um método estatístico que avalia e compara os algoritmos de aprendizado, dividindo o conjunto de dados, em que uma parte será utilizada para treinar o modelo e a outra para validação (REFAEILZADEH; TANG; LIU, 2016). A validação cruzada K-fold é uma forma básica em que os dados são, primeiramente, particionados em k pedaços de igual tamanho e, a cada interação os algoritmos de aprendizado usam k-1 pedaços do conjunto de dados para ensinar os modelos, então, no pedaço remanescente desta interação, é realizada a validação

(RHYS, 2020). A cada interação são computadas as métricas de performance dos modelos e, após transcorrer todas interações, é calculada a média das métricas de performance de tais, e a partir dessas são verificadas a capacidade de aprendizado dos algoritmos (REFAEILZADEH; TANG; LIU, 2016). Normalmente, em modelos que visam resolver problemas de regressão, as principais métricas utilizadas são o R^2, erro quadrático médio (EQM) e coeficientes de correlação de Pearson e Spearman (RHYS, 2020), proporcionando predizer valores contínuos, como é o caso do consumo alimentar, com base em outras variáveis contínuas ou não.

2.7. Justificativa do trabalho

Identificar animais com baixo consumo alimentar e com bom desempenho produtivo esbarra na aferição do consumo diário, visto que para tal mensuração são necessárias estruturas específicas para este fim. Uma vez encontrado, mesmo de maneira indireta, o nível de consumo alimentar dos indivíduos será possível chegar a valores de eficiência alimentar.

No Uruguai, existe a oportunidade de conduzir abordagens estatísticas e de aprendizado de máquina com base nos bancos de dados de animais com informações de consumo alimentar. Atualmente, nesse país, as raças que possuem informações de consumo alimentar são as Merino Australiano, Corriedale, Merino Dohne e Texel (GIORELLO *et al.*, 2021). No entanto, as raças Merino Australiano, Corriedale e Merino Dohne podem ser agrupadas como raças “laneiras” e representam mais de 91% dos animais com informações de consumo e, na avaliação genética uruguaia, suas características produtivas são praticamente as mesmas (https://www.geneticaovina.com.uy). Frente a isso, o banco de dados da raça Merino Australiano possui um número maior de informações de consumo alimentar e, portanto, pode ser considerado como o principal insumo para estudos, tanto no que se refere a diminuir o número de dias em avaliação quanto na aplicação de abordagens inteligentes, como o aprendizado de máquina para estimar o consumo de animais sem avaliação.

Os resultados esperados são novos horizontes para serem explorados na busca de animais eficientes, sendo possível otimizar a coleta de dados, proporcionando um maior número de indivíduos com informação de consumo alimenta.

3. HIPÓTESES E OBJETIVOS

3.1. Hipóteses

a) O período de teste pode ser encurtado em uma semana ou mais. Isso poderia ser possível aplicando um modelo linear que inclua a medida diária de peso corporal ou cálculo de ganho de peso por semana e utilizando o consumo alimentar como medida repetida. Dessa forma, seria possível diminuir os dias de teste sem comprometer a classificação dos animais e sem perder a precisão das estimativas.
b) O consumo alimentar em animais das raças Corriedale e Merino Dohne pode ser previsto via abordagem de aprendizado de máquina, usando conjunto de dados da raça Merino Australiano como base para treinar os algoritmos de predição.

3.2. Objetivos

a) Testar dois modelos alternativos de CAR para descobrir qual deles permite a redução dos dias de teste de consumo alimentar sem perder a precisão na identificação de animais fenotipicamente eficientes.
b) Testar algoritmos utilizando modelos de regressões linear, não linear, k-Nearest Neighbor e Random Forest para encontrar o melhor método com capacidade de previsão do consumo alimentar no conjunto de dados das raças Merino Australiano, Corriedale e Merino Dohne.

CAPÍTULO III – 4. RESIDUAL FEED INTAKE FOR AUSTRALIAN MERINO SHEEP ESTIMATED IN LESS THAN 42 DAYS OF TRIAL[1]

---

[1] Artigo publicado na revista Livestock Science em março de 2022.
https://doi.org/10.1016/j.livsci.2022.104889

4. RESIDUAL FEED INTAKE FOR AUSTRALIAN MERINO SHEEP ESTIMATED IN LESS THAN 42 DAYS OF TRIAL

Fernando Amarilho-Silveira[ad], Ignacio de Barbieri[c], Jaime Araujo Cobuci[a], Camila Marques Balconi[b], Gracialda Ferreira de Ferreira[c], Gabriel Ciappesoni[b]

[a]Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves, 7712, 91509-900, RS, Brasil.

[b]Instituto Nacional de Investigación Agropecuaria, Ruta 48 km 10, 90200, Canelones, Uruguay.

[c]Instituto Nacional de Investigación Agropecuaria, Ruta 5 km 386, 45000, Tacuarembó, Uruguay.

[d]Corresponding author. e-mail: famarilhosilveira@gmail.com, phone: +55 53 999278297, present address: Avenida Bento Gonçalves, 7712, 91509-900, RS, Brasil.

Abstract

The evaluation of sheep feed intake (FI) in feed efficiency tests is expensive. Decreasing the test period could be a resource-saving tool by reducing the cost of evaluating each animal and allowing to test a greater number of animals per year. For this reason, the objective of this research was to explore residual feed intake (RFI) models and to decreasing the test duration. Data was collected from 286 Australian Merino sheep of three performed trials, the test period consisted of 56 days (14 days of feed and facilities adaptation and 42 days of FI and average daily gain (ADG) evaluation). Two models were used to calculate RFI, Model 1 (based on Koch et al. (1963) linear model) and Model 2 (repeated measures, weekly model). Model 1 included ADG and FI estimates in a linear regression. The second model included weekly average FI as repeated measure and the weekly ADG. The increase in body weight during the test period was not perfectly linear, presenting a marked variance increase in two of the three tests while FI presented a tendency to increase throughout of the evaluation period, however presenting a high variance per day. In the 42-days tests, Pearson and Spearman correlations between models for ADG were of 0.89 and 0.87, respectively. The best correlations were detected for FI between 42 and 35-days models, presenting Pearson and Spearman correlations of 0.95 and 0.94 in the linear model, and 0.96 and 0.95 in the weekly model. When considering RFI, the correlations between linear and weekly 42-days models were from 0.93 and 0.92, respectively. The 35-days RFI length models (linear and weekly) presented a Pearson and Spearman correlations greater than 0.98 with the 42-days models. Therefore, the RFI models 35-days of duration allowed to decrease seven days of the FI test while maintaining accuracy and explaining 75.3% of the FI in the linear model, and 63.6% of the weekly model. Reducing seven days of testing would provide a greater data collection into a year of phenotypic evaluation.

Keywords

feed efficiency, average daily gain, repeated measures, genetic model

Introduction

Understanding the role of feed intake (FI) and growth rate in feed efficiency is indispensable to select economically efficient sheep. Moreover, it is necessary to optimize FI tests without losing precision in order to increase feed efficiency selection intensity. Increasing the number of animals tested without rising trial costs would also provide a faster genetic progress. The development of selection procedures, however, is challenging as increasing growth rate without increasing FI is genetically difficult (Cammack et al., 2005). Daily measurement of body weight (BW) will contribute to decrease the days on trial by monitoring the weight variance over the test and identifying more efficient and responsive animals. Furthermore, it will contribute to achieve fewer errors on data modeling. In this context, Cantalapiedra-Hijar et al. (2018) presented two different metrics for feed efficiency: feed conversion ratio (FCR), a relation between the amount of feed consumed and the animal bodyweight gain, and residual feed intake (RFI), the difference between actual and expected FI, the later based on feed requirements for maintenance and production (bodyweight gain). Although FCR is a useful index to evaluate management practices in production efficiency, it remains a ratio trait, and from a genetic point of view, selecting for FCR could cause unwanted correlated responses (Zetouni et al., 2017). Theoretically, this does not happens with RFI, since it is an independent BW and bodyweight gain measure, allowing to select animals that consume less without compromising the body size, and neither the bodyweight gain (Koch et al., 1963). As RFI is a hereditary trait, the selection for low RFI can be a useful tool to identify more lucrative animals, without affecting BW traits. The RFI heritability coefficients range from 0.11 to 0.49 in sheep (Cammack et al., 2005; François et al., 2006, 2002; Tortereau et al., 2019).

Typically, testing periods for ewes consist of 56 days: 14 days for adaptation to diet and facilities, and 42 days of evaluation period (Cammack et al., 2005; Johnson et al., 2015b, 2016, 2017; Leymaster et al., 2002; Snowder and Van Vleck, 2002; Tortereau et al., 2019). The possibility of reducing the test period with better use of data generated would minimize the costs without compromising accuracy (Archer et al., 1997). Macleay et al. (2016) and Paganoni et al. (2017) showed that it is possible to accurately evaluate sheep FI in 35 days or even less. According to that, our hypothesis is that the test period could be shortened by a week or more. It could be possible by applying a linear model that includes the daily BW measure or weight gain calculation by week and using the FI as a repeated measure, according to the approach used for Cockrum et al. (2013) and Johnson et al. (2016). Besides decreasing the days in the test without compromising the ranking of the animals (phenotypically ordered for RFI) could increase the accuracy estimates. The objective of this research was to test two alternative models of RFI to find out which one allows the reduction of days on feed intake test without losing accuracy in the identification of phenotypically efficient animals.

Methods

Three experiments were conducted with Australian Merino Sheep, the experimental site was located at La Magnolia Experiment Unit (National Agricultural Research Institute of Uruguay), Tacuarembó, Uruguay. Records were collected on 286 Australian Merino Sheep (143 males and 143 females), the offspring of 12 sires. The study dataset comes from three feed intake trials carried out in different periods of the year 2019, with the start dates of the evaluations, respectively, 04-29-2019, 06-17-2019, and 08-12-2019 for the trials 1, 2 and 3. The average age of animals on test per trial was 227±49, 287±48 and 339±48 days old in trial 1, 2 and 3, respectively. The animal's pre-test average BW was of 33.9±5.6, 41.3±5.4 and 36.2±5.0 kg, for trial 1, 2 and 3, respectively (Figure 2). The total precipitation, daily average temperature, and average thermal amplitude were 101.6 mm, 14.5ºC and 8.2ºC, 132.1 mm, 12.3ºC and 9.6ºC, and 210.7 mm, 12.4ºC and 10.5ºC, for trial 1, 2 and 3, respectively (data from: http://www.inia.uy/gras/Clima/Banco-datos-agroclimatico/).

All protocols applied were approved by INIA Animal Ethics Committee (INIA 2018.2), furthermore, according with Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines and following U.K. Animals (Scientific Procedures) Act, 1986 and associated guidelines, EU Directive 2010/63/EU for animal experiments.

Data edition

For BW data, animals with sanitary problems, outliers, and biologically unlikely data were removed from the study, remaining 56,872 BW records (Trial 1= 17,825; trial 2= 18,660; trial 3= 20,387; Figure 3).

For ADG, two calculations were estimated, the first using linear regression and the second by inferring weekly gain calculation. The linear ADG was estimated using PROC REG of software SAS on 12,159 average BW (trial 1 = 3,790; trial 2 = 3,942; and trial 3 = 4,427) considering the average BW per day-animal as dependent variable and the days on evaluation as independent variable. The model corresponded to:

$$Y = \beta_0 + \beta_1 X$$

where Y = daily BW (kg); β_0 = regression intercept; β_1 = ADG (kg/day); and X = experimental day. For the weekly ADG, the calculation was done through considering the difference in average BW of week Y+1 minus the average BW in the week Y, divided by seven. The linear ADG estimate generated only one value per animal, while weekly inferences generated six values.

For feed intake, the average data of 12,225 fresh feed intake was used (3,822; 3,997; and 4,406 average data for trials 1,2 and 3, respectively). The dry matter feed intake was obtained by multiplying the fresh feed intake data by the percentage of dry matter (after drying it in a < 60 ºC air force oven for 72 h) of Festin®. Data considered as biologically unlikely was excluded remaining 1,422,445 fresh feed intake data (447,559; 490,491; and 484,395 feed intake data for trials 1,2 and 3, respectively. The feed bin visits data presented 1,441,475 events, (450,596; 499,490; and 491,389 data for trials 1,2 and 3, respectively).

In summary, the FI and BW were estimated with records considering different length of the test: 42, 35, 28, and 21 days on trial. For the linear models the average values of FI and BW for each respective period were calculated, and for the weekly models, 6, 5, 4, and 3 weekly average values were considered. For the linear ADG the estimates by linear regression on 42, 35, 28, and 21 days on trial were used, while 6, 5, 4 and 3 weekly ADG measures were utilized in the weekly models.

Feed intake trials

After 7 days of acclimatation to the new feed, the animals were allocated to one of five automated feeding systems (pens) in accordance with the BW, sex, type of birth and sire. One day after arriving, the introduction to the new feed started. Animals were fed *ad libitum* with Festin® (Lucerne haylage; DM, 53.73%; crude protein, 21.9%; NDF, 36.1%; ADF, 29.3%). At day 3, electronic radio frequency identification tags (RFID tagged) were applied in the animals' ears. Each pen had five individual automated feeders and two automatic weighing platforms, which were equipped with an electronic tag reader, precision scale, and connected to a central computer (Figure 1), this allowed the control of body weight and feed intake of the animals in a daily basis. At the day 7, the entrance into collective pens was allowed after deworming.

Daily monitoring was applied by a software system that identified the entry of animals into the feeder and the body weighting platform lectures. The equipment and software were provided by Intergado (Belo Horizonte, MG, Brazil). The RFID tagged allows the identification of a specific animal at the feed bin and, consequently, their intake based on the difference in feed weight before and after the visit. The capacity of the feed bin load cells was of 200 kg (±0.025kg of accuracy) and its dimensions were 758.4 mm x 536 mm, x 371 mm. The body weighing platform was set in the water bins where a similar sensor to the feed bins system was present, each time the animal accessed it the BW was automatically recorded. The capacity of the body weighing platform was 400 ± 0.1 kg and its dimensions were 430 mm x 1200 mm x 1200 mm.

Data was continuously registered and transferred to the central computer and Intergado web software data center. To ensure the correct functioning of the equipment daily systematic calibration and on-site and online check were carried out.

Residual feed intake estimates

Two models were used to calculate residual feed intake (RFI). Using a similar methodology as proposed by Koch et al. (1963), average FI and ADG estimates were used for linear regressions of Model 1:

$$\mathrm{Y} = \mu + \mathrm{P} \times \mathrm{T} + \mathrm{BW}^{0.75} + \mathrm{ADG} + e \qquad (1)$$

where Y = observed individual daily average feed intake (total feed intake per day) expressed in dry matter (fresh matter intake × proportion of dry matter – DMI, 53.73%); μ = is an all animals constant referred to average daily feed intake; P × T = is composed for pen per trial (15 levels; fixed effect); $\mathrm{BW}^{0.75}$ = is the metabolic body weight (MBW) (mid-test body weight elevated to 0.75 as covariable); ADG = is the average daily body-weight gain (g/day, covariable); and e = the residual error as RFI (difference between the observed and expected DMI).

For Model 2, the average FI and ADG estimations by week were used:

$$\mathrm{Y}_{rep} = \mu + \mathrm{P} \times \mathrm{T} + \mathrm{PER} + \mathrm{BW}^{0.75}{}_{rep} + \mathrm{ADG}_{rep} + e \ (2)$$

where $\mathrm{Y}_{\mathrm{rep}}$ = observed week average individual daily feed intake (total feed intake per day) expressed in dry matter (fresh matter intake × proportion of dry matter – DMI, 53.73%); μ = is an all animals constant referred to average weekly feed intake; P ×T = is composed for pen per trial (15 levels; fixed effect); PER = is the week as covariable; $\mathrm{BW}^{0.75}{}_{\mathrm{rep}}$ = is the MBW (mid-week body weight elevated to 0.75 as covariable); $\mathrm{ADG}_{\mathrm{rep}}$ = is the daily body-weight gain by week (g/day, covariable); and *e* = the residual error as RFI (difference between the observed and expected DMI).

Statistical analysis

The average and variance of daily BW and FI over the days in the three trials were examined using the R Package ‘ggplot2’ of software R version 4.0.1 (Wickham et al., 2020). To estimate RFI in Model 1, a general linear model (PROC GLM) was used, while to estimate Model 2, the PROC MIXED was performed in SAS program. The R-squares were plotted with radarchart function, using the R Package ‘fmsb’ (Nakazawa, 2019).

The dominance analysis method was used to compare the relative importance of predictors (covariables) in multiple regression model that compose the RFI models, using the R Package ‘dominanceanalysis’ (Navarrete and Soares, 2020), the results were plotted using the R Package ‘ggplot2’ (Wickham et al., 2020).

The FI, ADG and the output of RFI models were submitted of Pearson and Spearman correlation analysis, using the R Package ‘corrplot’ of software R version 3.5.3 (Wei and Simko, 2017).

Results

The average FI range for all RFI models was 1.25 to 1.33 kg dry matter basis (kg DM day^{-1}) and the coefficient of variation (CV) ranged from 17.75% to 19.13%. The ADG ranged from 0.133 to 0.186 kg per day (kg day^{-1}), with a CV from 33.50% to 54.07%. For MBW the values ranged from 15.37 to 16.71 kg, and CV from 11.50% to 11.88%.

On Figures 4A and 4B a BW increase trend throughout the test period and different variabilities can be observed. In Figure 4A, BW per day in trials 1 and 2 had a linear increase, while trial 3 had several BW falls, mostly after day 22, with a new increasing the after 37th day. The variances for BW, as shown in Figure 4B, show the largest values were observed in trial 2 followed by trial 1. In these trials the variance increased, with an increment of 50.2% and 41.4%, respectively, for trials 2 and 1. However, trial 3 kept almost a constant variance throughout the test, with the peak on the 22nd day. The variance increased by just 21% between day one through day 42 in this trial. The standard errors relative per day per trial on average were, respectively, 0.15%, 0.15%, and 0.10% for trials 1, 2, and 3.

The FI behaviour was not constant, with a positive tendency, as shown in Figures 4C and 4D. In Figure 4C the minimum intake point on trial 1 was on the 11th day with dry matter intake decrease of 10.9% compared with the start of the test. Moreover, on the day 22, the decrease was larger in the order of 21.2% compared with the beginning of the feed intake trial. The global increment of intake along the test period was 0.020 kg (+1.8%). In trial 2, which had better constancy in increasing intake along the test, presenting two discrete falls of intakes, one in the seventh and other on the ninth day compared to the beginning of the period of evaluation, respectively, 4.8% and 8.6%

(Figure 4C). Nonetheless, the intake increase over the entire period was 0.236 kg (+19.4%).

Four fall points were noted in trial 3, the first point was the fourth day represented a decrease compared with the first day of 2.5%. The second and greater decrease in FI was on the 12th day that represented 13.5% less dry matter consumed. The third and fourth fall points were on days 16 and 29 with a reduction in the order of 1.4% and 1.2%, respectively (Figure 4C). Like the others, the trial 3 also presented an increase of intake up the test period of 0.175 kg (+16.3%).

As presented in Figure 4D, the variances for FI show the largest values in trial 2. In this trial the variance behaviour presented a large increase of the values, with an increment of 118.4%. In this trial, there were three peak points, in days 2, 12 and 28 with an average increment of 168.4%. On the other hand, trial 1 had an increase between the first day and the 42nd of 26.5%. Despite a small total variation increase, the trial 1 had some peaks in the 13th, 29th and 36th days. The trial 3 kept almost constant the variance until 40th day, with a great increase in the day 41 following an increase in the 42nd day. The variance until day 40 increased by just 1.7% and an extreme increase of 393.5% between day one through day 42. The standard error relative per day per trial on average were, respectively, 0.22%, 0.24%, and 0.20% for trials 1, 2, and 3 (data not shown).

The Pearson and Spearman correlations for the models with a length of 42 days presented values from 0.89 and 0.87, respectively, for the ADG estimate for linear regression and ADG weekly (Figure 5A and Figure 5B).

In the ADG estimate by the linear model, the Pearson and Spearman correlations between the model of 42 days with the models of 35, 28, and 21 days were 0.95 and 0.94, 0.77 and 0.75, and 0.67 and 0.69, respectively. For the weekly ADG model approach the

Pearson and Spearman correlations were from 0.96 and 0.95, 0.83 and 0.81, and 0.87 and 0.88, for the 42-day trial, considering 35, 28, and 21 days length models, respectively.

The FI correlations, both Pearson as Spearman presented values >0.93 not only between different length tests for linear and weekly models but also between linear and weekly models for any given length test (Figures 5C and 5D).

The determinations coefficients (R-square) for the RFI models presented values ranging from 0.610 to 0.753, being the lowest value for the Model 2 (weekly FI and ADG) with 21 days on trial and the greatest value for the Model 1 (linear FI and ADG) with 35 days (Figure 6). Linear models (model 1) with 42, 28 and 21 days presented R-square from 0.752, 0.734 and 0.693, respectively. For the weekly models (model 2), the values of R-square found were 0.634, 0.636 and 0.629 for 42, 35 and 28 days.

The greater contribution in the dominance analysis for all models was the MBW, ranging from 0.29 to 0.33 (Figure 7). The models that MBW had an importance smaller than 0.30 were the Model 1 with 42 days and the Model 2 with 21 days, presented an R-square of 0.29 for both. Sequentially, the Model 1 with 35, 28 and 21 days, presented the R-square of 0.30, 0.31 and 0.31, respectively. For the Model 2 with 42, 35 and 28 days the R-squares, respectively, were 0.33, 0.33 and 0.31.

The ADG contribution ranged from 0.04 (Models 2 with 42 days) to 0.19 (Model 1 with 42 days). The other Models 1 presented, respectively, an ADG contribution from 0.18, 0.13, and 0.12, for 35, 28, and 21 days. The other Models 2 with 35, 28, and 21 days on trial presented R-square for ADG from 0.05, 0.08 and 0.1, respectively (Figure 7).

The fixed effect Pen per trial (Pen ×Trial) in all models presented a contribution greater than 0.19, and the covariable week for Models 2, presented a contribution smaller than 0.05 (data not showed).

The RFI models with 42 days of trials presented a Pearson correlation of 0.93 between the Model 1 and Model 2 (Figure 8A). While the Spearman correlation value between Model 1 and Model 2 was 0.92 (Figure 8B).

In Model 1, comparing the 42 days with the 35, 28, and 21 days, the Pearson and Spearman correlations, respectively, were 0.98 and 0.98, 0.94 and 0.94, and 0.88 and 0.90 (Figure 8A and 8B). Finally, in Model 2, the Pearson and Spearman correlations presented values from 0.98, 0.95, and 0.88, respectively, between 42 days with 35, 28, and 21 days (Figure 8A and 8B).

Discussion

Decreasing the trials period in one week or more would allow test more animals with the same resources without accuracy losses in the phenotypic measurements. The use of methods for weight gain estimation with a linear model or calculations by week, the use of BWs measures, as well as considering the FI as a repeated measure, would allow to reduce the test time without loss of precision.

Average daily gain

A BW evolution along the trials is noted in the Figures 4A and 4B, characterized by a linear increase of BW, along with a variation not so linear. This can be explained by the fluctuations across all animals occurred between weightings, even on consecutive days. Fluctuations on BW are due to possible variations in gut fill, hour of the day, water consumption, FI or rain. However, those fluctuations were corrected when estimating the ADGs by linear regression. The studies of Johnson et al. (2015a, 2016 and 2017) presented ADG ranging from 0.231 to 0.332 kg live weight per day for 9 months-old ewe lambs from maternal breeds (synthetic breed as Coopworth and different crossbreeding by industry sires of different breeds). Thus, our results have not achieved the values described by these studies. This fact may be due the difference in the evaluate animals age, the adult BW inherent the animals' racial composition in the revised studies, and productive fitness of the breeds, compared at Australian Merino sheep raised in Uruguay.

The ADG correlations showed that there is a different classification in the different complete models, however, to maintain accuracy, the linear and weekly models with 35 days presented good Pearson and Spearman correlation (> 0.90) with the respective linear and weekly models with 42 days. According to Waldron et al. (1990), if the duration of an ADG performance test is too short, animals may not be properly ranked for genetic merit, resulting in an extremely low relationship between the sire’s performance and the performance of its progeny. Waldron et al. (1990) suggest that test performance measured on 63 days has an insufficient period as an indicator of breeding value for growth. Results from the single-trait analyses clearly imply that variances for ADG can be influenced by the duration of the adjustment period and of the performance test (Snowder and Van Vleck, 2002). The higher heritability for ADG was found in tests that used an adaptation period of two weeks, and the performance trial duration from eight weeks or longer. The results of Snowder and Van Vleck (2002) point out for the largest amount of genetic variation in ADG must be accounted using a test period of 42 to 98 days for accurately identifying genetic differences between animals for ADG. Archer et al. (1997) obtained optimal periods different of records for beef cattle with 70 days of evaluation. Is important highlight that these studies did not collect daily BW, this practice provides the possibility of decreasing of the days of ADG without loss accuracy in the growth test as shown in the Figure 5A and Figure 5B. Thanks to the fact that in our study the BW was measured daily, it was possible to adjust all daily measurements along of the tests to a straight line with a greater or lesser inclination, which allowed an ADG estimation that reflects the real profile of the BW progression over the days.

Feed intake

As shown in Figures 4C and 4D, the FI is extremely variable, this could be explained due to several mechanisms that act in the regulation of consumption. Allen (2014) describes that the response of feeding behavior to the diet is influenced by the energy supply in blood which is affected by physiological state, such as differences in liver gluconeogenesis, mobilization, extrahepatic tissue energy uptake, tissue secretion and sensitivity to hormones and cytokines. This can affect satiety and hunger over the span of minutes and hours and can have long-term effects on FI and energy balance, explaining the greater variability in FI, as shown in Figure 4D. As the animals in the present study were evaluated outdoors, in addition to internal factors, the effects of the environment such as rain, heat, cold and wind could also influence the FI regulation. Therefore, in the present study the total precipitation, daily average temperature, and average thermal amplitude, not do seem to be challenging enough to influence average FI (data not considered directly in the model).

As there are several factors that affect FI and considering the difference in diets, the comparison of absolute FI values found in this study with other works is not indicated. This fact is very important to highlight since several authors found an interaction between feed efficiency and diet type (Cammack et al., 2014; Carberry et al., 2012; Coyle et al., 2016; Durunna et al., 2011; Ellison et al., 2017; François et al., 2006; Toral et al., 2019). When the diet is based on roughage, the mechanism that controls the intake probably is similar to the mentioned by Cantalapiedra-Hijar et al. (2018). These authors suggest that when a pasture is offered, ruminal distension is more likely to dominate FI control than tissue energy detection. In our study, with an NDF of 36,1%, the intake control was dictated by rumen filling, in which according to Dado and Allen (1995) the forage offered with an NDF greater than 35% would already limit feed intake.

As observed in Figure 4C and 4D, the FI over the test period presents a non-linear behavior and high variability. The FI coefficient of variation (CV) is a good descriptive variable. Basarab et al. (2013) describe that CV for FI in beef cattle is commonly found in values ranging from 11% to 20%, and this is due to the fact that intake is a reflection of natural daily between-animal variation between animals. This is due that FI is a function of meal size and frequency with meal size determined for the rate of eating and meal length and meal frequency determined for the length of time between meals (Allen, 2014). In our study the CV of FI ranged from 17.75% to 19.13%, similarly to the described by Basarab et al. (2013), however higher than results of 7.3%, 13%, 17.1% and 13% reported by Ermias et al. (2002), François et al. (2002), Cammack et al. (2005) and Tortereau et al. (2019), respectively.

In sheep studies, the FI varied from 0.838 at 3.94 kg of dry matter per day, a large variation among studies. In Ermias et al. (2002) study, the average daily total dry matter intake was of 0.84 kg. However, François et al. (2002), Cammack et al. (2005) and Redden et al. (2011) found FI values of 1.79, 1.69 and 1.77 kg, respectively. The FI values found by Cockrum et al. (2013) ranged from 2.85 to 3.94 kg, Johnson et al. (2015a) found feed intake ranges from 2.7 to 3.3 kg. Feed intakes ranging from 1.09 to 1.33 kg were published by Zhang et al. (2017). In the study of Montelli et al. (2019) FI from 1.25 to 1.44 kg were found. Tortereau et al. (2019) found FI from 1.96 kg. Thus, the feed intakes from

1.25 to 1.32 kg found in our study on the different models and times of measurements, is within what is reported in the bibliography, which not was a limiting factor of animals' performance.

The very high Pearson and Spearman FI correlations for linear and weekly models 42-days with respective 35-, 28- and 21-days models showed that is possible to reduce the days on test without losing accuracy in animal's classification. Castilhos et al. (2011) describe that the variances for FI decreased by 0.18% from 28 to 56 days, increased 13.67% from 56 to 84 days, and increased by 11.22% from 84 to 112 days. This might be attributed to the increase of dry matter intake due to the increase in BW gain over the test period. As the variance for dry matter intake did not stabilize over the test period, Pearson and Spearman correlations were used to determine the optimum test period, presented values higher than 0.93 for these authors. Archer et al. (1997) comparing different periods of records in beef cattle found that a 35 days was sufficient for feed intake records. Considering only the FI, the high Pearson and Spearman correlations could make it possible to safely reduce from one to three weeks of experimental period in our study. The increase in FI after 28 tends to decelerate (Figure 4C), tending to maintain the proportion between animals that have a higher and lower FI, reflecting in correlations greater than 0.93.

Metabolic body weight

The values of MBW ranged from 15.67 to 16.14 kg, and CV from 11.63% to 11.75%, showing low variability. Redden et al. (2011) found MBW ranging from 19.6 to 21.1 kg. Moreover, a MBW range from 12.97 to 13.18 kg were published by Zhang et al. (2017). In the study of Montelli et al. (2019) the average MBW was 13.7 kg. Our MBW values do not match with the studies described above. Therefore, this could be explained by the fact that MBW is dependent on several factors such as age of dam, type of birth, weaning and standard/normal adult weight of the breed.

Residual feed intake

The R-square in the FI prediction model is a criterion used to indicate how suitable is the RFI model. In this sense, as showed in Figure 6, our models were able to explain a portion of 61.0% to 75.3% (R-square from 0.61 to 0.753). According to the R-square, the best RFI model was the linear RFI model with 35 days on trial. Using a methodology similar to Model 1, Knott et al. (2008) found an R-square of 0.63 and 0.56 for rams at 6 and 13 months old, respectively. In the model composed for ADG and live weight in the mid-test, Knott et al. (2008) reached R-square equal to 0.74 and 0.60 for 6 and 13 months old rams, respectively. For animals tested in two subsequent ages, the R-square were 0.41 and 0.38 for ewes with 280 and 414 days old, respectively, (Redden et al., 2011). Cockrum et al. (2013) found coefficients of determination for several residual feed intake models estimates range 0.43 to 0.46. In the study of Johnson et al. (2015a), the model fitted with live weight, ADG, previous feed, and feeder in the current trial and the intake of the animals had an R-square of 0.79. Redden et al. (2014) found the R-square equal to 0.84. Johnson et al. (2016) found the coefficient of determination for the RFI model from 0.78. Johnson et al. (2017) present a R-square of the RFI model has been greater than 0.70. For Zhang et al. (2017) the R-square for the RFI model was 0.80. The RFI presented a coefficient of determination of 0.82 in the study of Montelli et al. (2019). For RFI model of Tortereau et al., (2019) the R-square values ranging from 0.63 to 0.84. Our results fall within a wide range of R-square, where the smallest found in the bibliography was 0.38

and the largest was 0.84. Therefore, is a consensus in those several studies reviewed that of goodness of fit values of R-square do not be smaller than 0.70. Thus, the Model 1 with 42, 35 and 28 days of our study are into this range and represents that this model was able to predict FI with good accuracy. That would mean that the covariates of the model would be good predictors of FI (Figure 7 and the next discussion section).

The Figure 8 confirms our results, showing the Pearson and Spearman correlations (P<0.05) of all RFI models, confirming with values greater than 0.98 of Pearson and Spearman correlations (P<0.001) that the RFI models (Models 1 and 2) with 35 days is the best reduced model. The use of this model, in practice, not will cause animal significant reranked in the FI, ADG, and neither on RFI classification.

The validation of the recommendation for a decrease of the days on trial, it will only be conclusive when beyond of phenotypic correlation, the genetic correlations are known (Goonewardene et al., 2004). If the testing time can be shortened while maintaining the same degree of accuracy, by correctly classifying animals into their respective (positive and negative) feed efficiency categories with minimal rank changes, then on-test feeding costs may be reduced. Goonewardene et al. (2004) found Pearson correlations phenotypic in the comparisons of 0-84 days with 0-105 days ranging from 0.85 to 0.93. Knott et al. (2008) describe that highly significant rank correlation values between each of the RFI models, are of key importance as they indicate that animals maintain the same relative rankings in each model at each time of measurement. However, the phenotypic correlations among the models of the Knott et al. (2008) study, ranged from 0.43 to 0.76. The correlations found for Cockrum et al. (2013), between weekly RFI estimates and between weekly RFI rankings were performed to determine the applicability of RFI estimates throughout the testing period, ranged from 0.30 to 0.82 for Pearson correlations and 0.28 to 0.80 for Spearman correlations. With a basis on these studies, our work presented safe results for deciding to decrease the time on trial in one week maintaining the rank of the animals satisfactorily.

Average contribution

For a better understanding of the models, an analysis of dominance as described for Navarrete and Soares (2020) was proposed. The importance of MBW was observed on all models (Figure 7). The portion of the contribution of ADG and MBW of the RFI Models 1 in 42 and 35 days on our study was lower compared with the work of Knott et al. (2008). The models evaluated for Knott et al. (2008) clearly indicate that the MBW and ADG accounted a substantial proportion of the variation in FI in a group of animals at both 6 months old (R-square varied from 0.48-0.73), and at 13 months old (R-square varied from 0.48-0.49). In our study the average contribution from MBW and ADG were of the 0.29-0.31 and 0.12-0.19, respectively, for the Model 1, and 0.29 – 0.33 and 0.04 – 0.10, respectively, for the Model 2. This means that the FI variation among the animals is explained by up to 31% (Model 1) and 33% (Model 2) for the MBW. That is represent the percentage of cases in that the FI increase, or decrease is due to one unit increased or decreased of the MBW. This show that it is possible to find animals with high MBW and low FI and vice versa. The ADG average contribution is lower, by up to 0.19 (Model 1) and 0.10 (Model 2). Using the same logic and comparing with MBW, is possible to find more animals with high ADG and low FI and vice versa.

Shorter duration trials

We found only the works of Redden et al. (2011), Macleay et al. (2016) and Paganoni et al. (2017) where less than 42 days were used in the trial period. Experiment 1 in the study of Redden et al. (2011), worked with a period of 38 days of evaluation, however, the original trial length is 49 days, because the last 11 days not was recorded due to a computer system failure. Already in the Macleay et al. (2016) study, the feed intake in sheep was measured until 35 days. In Paganoni et al. (2017) the RFI was estimated with a trial of 35 days. In others studies the evaluation period was from 42 days in Leymaster et al. (2002), Cammack et al. (2005), Johnson et al. (2015b, 2017, 2016), and in the second approach of Tortereau et al. (2019). With 49 days can mention experiment 1 of Knott et al. (2008), studies of Cammack et al. (2014), Redden et al. (2014), and Ellison et al. (2017). With 50 days, the study of Zhang et al. (2017), 56 days in the work of Cameron (1988), François et al. (François et al., 2007, 2006, 2002), Montelli et al. (2019) and in the first approach of Tortereau et al. (2019). Evaluations periods with more than 62 days, we found in the Waldron et al. (1990), in experiment 2 of Knott et al. (2008), Cockrum et al. (2013), Paula et al. (2013), and in Carneiro et al. (2019). Those evidence, together with correlations presented in Figure 8, reinforce our arguments that it is possible to shorten the tests of feed efficiency, these being promising findings. However, it would be suitable to evaluate the genetic correlations of the different traits with different test periods and a greater number of data to have a validation of the proposal.

Conclusion

Reducing seven days of testing would provide great impacts. In a year where usually six batches would be tested, an additional batch could be included at the same cost. If resources saved by shortening test duration are used to test related animals and the data from relatives are used in genetic evaluations, the loss of accuracy of using a shorter test duration will be partly compensated by the extra information obtained by measuring an additional related individual. In this way, data collection could be accelerated, and selection intensity increased.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. Thank the Instituto Nacional de Investigación Agropecuaria (INIA – UY) for the conduction of the field work in the data collection, which received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement n°772787 (Smarter) and from the Instituto Nacional de Investigación (INIA_CL_38: Rumiar).

References

Allen, M.S., 2014. Drives and limits to feed intake in ruminants. Anim. Prod. Sci. 54, 1513–1524. https://doi.org/10.1071/AN14478

Archer, J.A., Arthur, P.F., Herd, R.M., Parnell, P.F., Pitchford, W.S., 1997. Optimum Postweaning Test for Measurement of Growth Rate, Feed Intake, and Feed Efficiency in British Breed Cattle. J. Anim. Sci. 75, 2024–2032. https://doi.org/10.2527/1997.7582024x

Basarab, J.A., Beauchemin, K.A., Baron, V.S., Ominski, K.H., Guan, L.L., Miller, S.P., Crowley, J.J., 2013. Reducing GHG emissions through genetic improvement for feed efficiency: effects on economically important traits and enteric methane production. Animal 7 Suppl 2, 303–315. https://doi.org/10.1017/S1751731113000888

Cameron, N.D., 1988. Genetic relationships between growth and food intake in performance tested ram lambs: An investigation of variance component estimation procedures. J. Anim. Breed. Genet. 105, 431–440. https://doi.org/10.1111/j.1439-0388.1988.tb00316.x

Cammack, K.M., Ellison, M.J., Conant, G., Lamberson, W.R., Austin, K.J., 2014. Rumen Microbial Taxa and By-products Associated with Feed Efficiency in Sheep, in: Proceedings, 14th World Congress of Genetics Applied to Livestock Production. pp. 14–16.

Cammack, K.M., Leymaster, K.A., Jenkins, T.G., Nielsen, M.K., 2005. Estimates of genetic parameters for feed intake, feeding behavior, and daily gain in composite ram lambs. J. Anim. Sci. 83, 777–785. https://doi.org/10.2527/2005.834777x

Cantalapiedra-Hijar, G., Abo-Ismail, M., Carstens, G.E., Guan, L.L., Hegarty, R., Kenny, D.A., Mcgee, M., Plastow, G., Relling, A., Ortigues-Marty, I., 2018. Review: Biological determinants of between-animal variation in feed efficiency of growing beef cattle. Animal. https://doi.org/10.1017/S1751731118001489

Carberry, C.A., Kenny, D.A., Han, S., McCabe, M.S., Waters, S.M., 2012. Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. Appl. Environ. Microbiol. 78, 4949–4958. https://doi.org/10.1128/AEM.07759-11

Carneiro, M.M.Y., Morais, M. da G., Souza, A.R.D.L., Fernandes, H.J., Feijó, G.L.D., Bonin, M. de N., Franco, G.L., Rocha, R.F.A.T., 2019. Residual intake and gain for the evaluation of performance, non-carcass components, and carcass characteristics of confined crossbred Texel lambs. Rev. Bras. Zootec. 48. https://doi.org/10.1590/rbz4820180206

Castilhos, A.M., Branco, R.H., Razook, A.G., Bonilha, S.F.M., Mercadante, M.E.Z., de Figueiredo, L.A., 2011. Duração do período de avaliação pós-desmame para medidas de desempenho, consumo e eficiência alimentar em bovinos da raça Nelore. Rev. Bras. Zootec. 40, 301–307.

Cockrum, R.R., Stobart, R.H., Lake, S.L., Cammack, K.M., 2013. Phenotypic variation in residual feed intake and performance traits in rams. Small Rumin. Res. 113, 313–322. https://doi.org/10.1016/j.smallrumres.2013.05.001

Coyle, S., Fitzsimons, C., Kenny, D.A., Kelly, A.K., McGee, M., 2016. Repeatability of feed efficiency in steers offered a high-concentrate diet. J. Anim. Sci. 94, 719. https://doi.org/10.2527/jam2016-1482

Dado, R.G., Allen, M.S., 1995. Intake Limitations, Feeding Behavior, and Rumen Function of Cows Challenged with Rumen Fill from Dietary Fiber or Inert Bulk. J. Dairy Sci. 78, 118–133. https://doi.org/10.3168/jds.S0022-0302(95)76622-X

Durunna, O.N., Mujibi, F.D.N., Goonewardene, L., Okine, E.K., Basarab, J.A., Wang, Z., Moore, S.S., 2011. Feed efficiency differences and reranking in beef steers fed grower and finisher diets. J. Anim. Sci. 89, 158–167. https://doi.org/10.2527/jas.2009-2514

Ellison, M.J., Conant, G.C., Lamberson, W.R., Cockrum, R.R., Austin, K.J., Rule, D.C., Cammack, K.M., 2017. Diet and feed efficiency status affect rumen microbial profiles of sheep. Small Rumin. Res. 156, 12–19.

https://doi.org/10.1016/j.smallrumres.2017.08.009

Ermias, E., Yami, A., Rege, J.E.O., 2002. Genetic parameters of feed intake and digestive efficiency in tropical sheep. 7th World Congr. Genet. Appl. to Livest. Prod. 7, 19–23.

François, D., Bibé, B., Bouix, J., Brunel, J.C., Weisbecker, J.L., Ricard, E., 2002. Genetic parameters of feeding traits on meat sheep. 7th World Congr. Genet. Appl. to Livest. Prod. 7, 19–23.

François, D., Bouix, J., Brunel, J.C., Ricard, E., Marcon, D., Weisbecker, J.L., Bibé, B., 2006. Genetic parameters of feeding traits in INRA401 sheep estimated under two different diets. 8th World Congr. Genet. Appl. to Livest. Prod. 8, 13–18.

François, D., Brunel, J.C., Ricard, E., Weisbecker, J.L., Bouix, J., Bourdillon, Y., Bibé, B., 2007. Use in selection of the measurements of feed intake and feeding behaviour parameters in sheep. Options Méditerranéennes Ser. A 447, 441–447.

Goonewardene, L.A., Okine, E., Wang, Z., Spaner, D., Mir, P.S., Mir, Z., Marx, T., 2004. Residual metabolizable energy intake and its association with diet and test duration. Can. J. Anim. Sci. 84, 291–295. https://doi.org/10.4141/A03-073

Johnson, P., Miller, S., Knowler, K., 2015a. Towards a data set to investigate feed efficiency in new zealand maternal sheep. Proc. Assoc. Adv. Anim. Breed. Genet. 106–109.

Johnson, P., Miller, S., Knowler, K., Bryson, B., Dodds, K., 2015b. Modelling liveweight change to inform a residual feed intake model in sheep. Proc. New Zeal. Soc. Anim. Prod. 75, 225–227.

Johnson, P.L., Miller, S.P., Knowler, K., 2016. Preliminary investigations into the trait of residual energy intake in sheep. NZ Soc Anim Prod Proc 76, 34–37.

https://doi.org/10.1079/BJN19660078

Johnson, P.L., Wing, J., Knowler, K., 2017. Relationship between measures of residual energy intake made on growing animals and adults. NZ Soc Anim Prod Proc 77, 85–87. https://doi.org/10.1079/BJN19660078

Knott, S.A., Cummins, L.J., Dunshea, F.R., Leury, B.J., 2008. The use of different models for the estimation of residual feed intake (RFI) as a measure of feed efficiency in meat sheep ઙ 143, 242–255. https://doi.org/10.1016/j.anifeedsci.2007.05.013

Koch, R.M., Swiger, L.A., Chambers, D., Gregory, K.E., 1963. Efficincy of feed use in beef cattle. J. Anim. Sci. 22, 486–494.

Leymaster, K., Cammack, K., Nielsen, M.K., Jenkins, T.G., 2002. Estimates of genetic parameters for daily gain, feed intake, and behavior traits in ram lambs of a composite population. 7th World Congr. Genet. Appl. to Livest. Prod. 1986–1989.

Macleay, C., Blumer, S., Hancock, S., Inglis, L., Paganoni, B., Rose, G., Thompson, A.N., 2016. Feed intake for sheep can be measured precisely in less than 35 days. Proc. Aust. Soc. Anim. Prod. Brief comm.

Montelli, N.L.L., Almeida, A.K. de, Ribeiro, C.R. de F., Grobe, M.D., Abrantes, M.A.F., Lemos, G.S., Furusho Garcia, I.F., Pereira, I.G., 2019. Performance, feeding behavior and digestibility of nutrients in lambs with divergent efficiency traits. Small Rumin. Res. 180, 50–56. https://doi.org/10.1016/j.smallrumres.2019.07.016

Nakazawa, M., 2019. Functions for Medical Statistics Book with some Demographic Data.

Navarrete, C.B., Soares, F.C., 2020. Dominance Analysis - CRAN.

Paganoni, B., Rose, G., Macleay, C., Jones, C., Brown, D.J., Kearney, G., Ferguson, M., Thompson, A.N., 2017. More feed efficient sheep produce less methane and carbon dioxide when eating high-quality pellets. J. Anim. Sci. 95, 3839. https://doi.org/10.2527/jas2017.1499

Paula, E.F.E., Monteiro, A.L.G., Souza, D.F., Prado, O.R., Nomura, T.M., Stivari, T.S.S., Silva, C.J.A., Santana, M.H.A., 2013. Consumo alimentar residual e sua relação com medidas de desempenho e eficiência e características in vivo da carcaça de cordeiros. Arq. Bras. Med. Vet. e Zootec. 65, 566–572. https://doi.org/10.1590/S0102-09352013000200037

Redden, R.R., Surber, L.M.M., Grove, A. V., Kott, R.W., 2014. Effects of residual feed intake classification and method of alfalfa processing on ewe intake and growth. J. Anim. Sci. 92, 830–835. https://doi.org/10.2527/jas.2013-6768

Redden, R.R., Surber, L.M.M., Roeder, B.L., Nichols, B.M., Paterson, J.A., Kott, R.W., 2011. Residual feed efficiency established in a post-weaning growth test may not result in more efficient ewes on the range. Small Rumin. Res. 96, 155–159. https://doi.org/10.1016/j.smallrumres.2010.12.007

Snowder, G.D., Van Vleck, L.D., 2002. Effect of duration of performance test on variance component estimates for lamb growth rate. J. Anim. Sci. 80, 2078–2084. https://doi.org/10.2527/2002.8082078x

Toral, P., Belenguer, Á., Hervás, G., Fernández-Díez, C., Yáñez Ruiz, D., Frutos, P., 2019. Residual feed intake and rumen bacterial diversity in lactating sheep: A preliminary study of their potential link, in: 13º International Syposium on Ruminant Physiology. p. 191.

Tortereau, F., Marie-Etancelin, C., Weisbecker, J.L., Marcon, D., Bouvier, F., Moreno-

Romieux, C., François, D., 2019. Genetic parameters for feed efficiency in Romane rams and responses to single-generation selection. Animal. https://doi.org/10.1017/S1751731119002544

Waldron, D.F., Thomas, D.L., Stookey, J.M., Nash, T.G., Keith, F.K., Fernando, R.L., 1990. Central Ram Tests in the Midwestern United States: III. Relationship Between Sire's Central Test Performance and Progeny Performance. J. Anim. Sci. 68, 45–53. https://doi.org/https://doi.org/10.1093/ansci/68.1.45

Wei, T., Simko, V., 2017. Visualization of a Correlation Matrix - CRAN.

Wickham, H., Chang, W., Henry, L., Pedersen, T.L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., 2020. Create Elegant Data Visualisations Using the Grammar of Graphics Description - CRAN.

Zetouni, L., Henryon, M., Kargo, M., Lassen, J., 2017. Direct multitrait selection realizes the highest genetic response for ratio traits. J. Anim. Sci. 95, 1921–1925. https://doi.org/10.2527/jas2016.1324

Zhang, X., Wang, W., Mo, F., La, Y., Li, C., Li, F., 2017. Association of residual feed intake with growth and slaughtering performance, blood metabolism, and body composition in growing lambs. Sci. Rep. 7, 1–11. https://doi.org/10.1038/s41598-017-13042-7

Figures

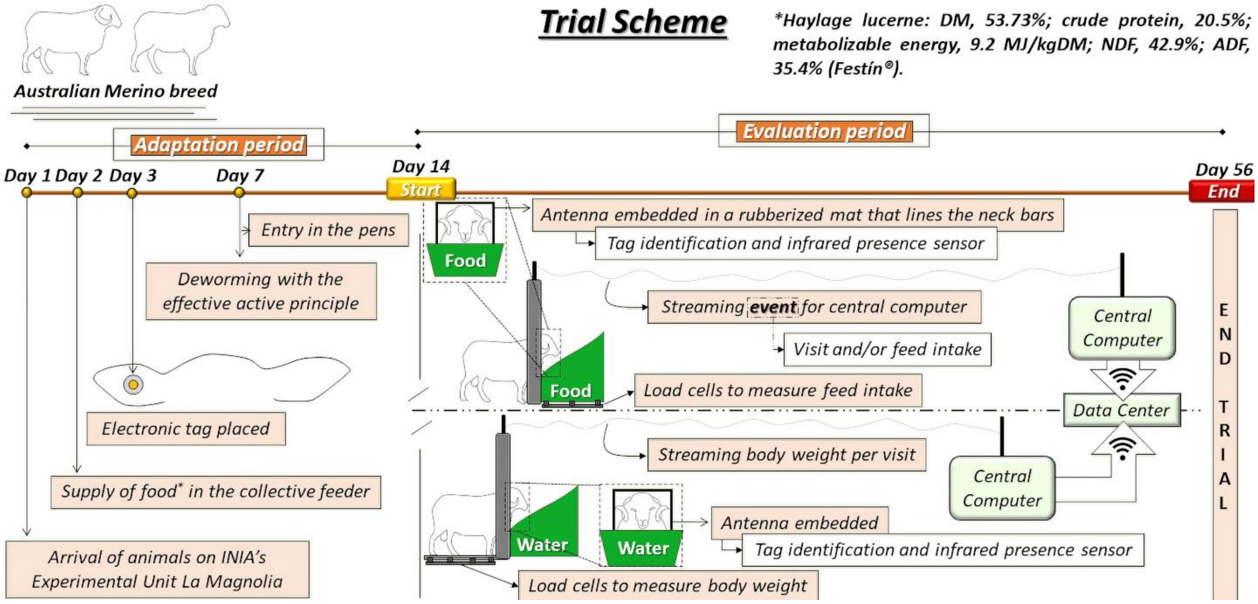
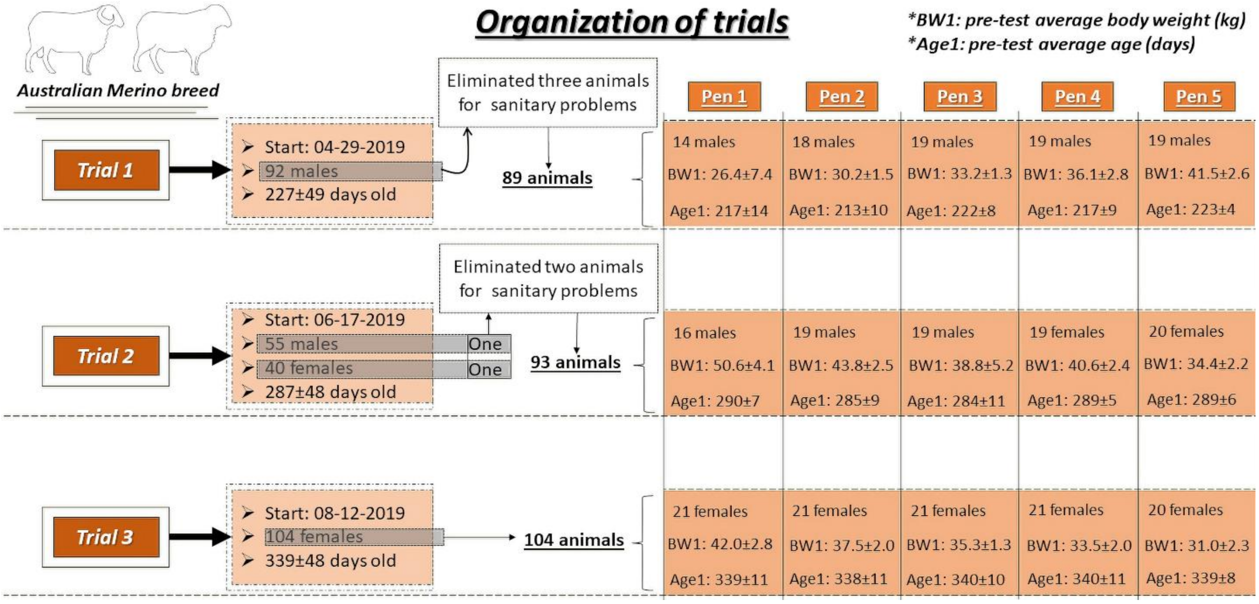


Figure 1. Adaptation period of 14 days, with the first seven days of the feed adaptation and the rest of the feed and facilities adaptation. The evaluation period was from day 14 to day 56, totaling 42 days of feed efficiency trial.



	Pen 1	Pen 2	Pen 3	Pen 4	Pen 5
Trial 1	14 males BW1: 26.4±7.4 Age1: 217±14	18 males BW1: 30.2±1.5 Age1: 213±10	19 males BW1: 33.2±1.3 Age1: 222±8	19 males BW1: 36.1±2.8 Age1: 217±9	19 males BW1: 41.5±2.6 Age1: 223±4
Trial 2	16 males BW1: 50.6±4.1 Age1: 290±7	19 males BW1: 43.8±2.5 Age1: 285±9	19 males BW1: 38.8±5.2 Age1: 284±11	19 females BW1: 40.6±2.4 Age1: 289±5	20 females BW1: 34.4±2.2 Age1: 289±6
Trial 3	21 females BW1: 42.0±2.8 Age1: 339±11	21 females BW1: 37.5±2.0 Age1: 338±11	21 females BW1: 35.3±1.3 Age1: 340±10	21 females BW1: 33.5±2.0 Age1: 340±11	20 females BW1: 31.0±2.3 Age1: 339±8

Figure *2.* There were three food efficiency tests, and in each test the animals were allocated into five pens, divided considering the sex, type of birth and sire of the sheep, and body weight.

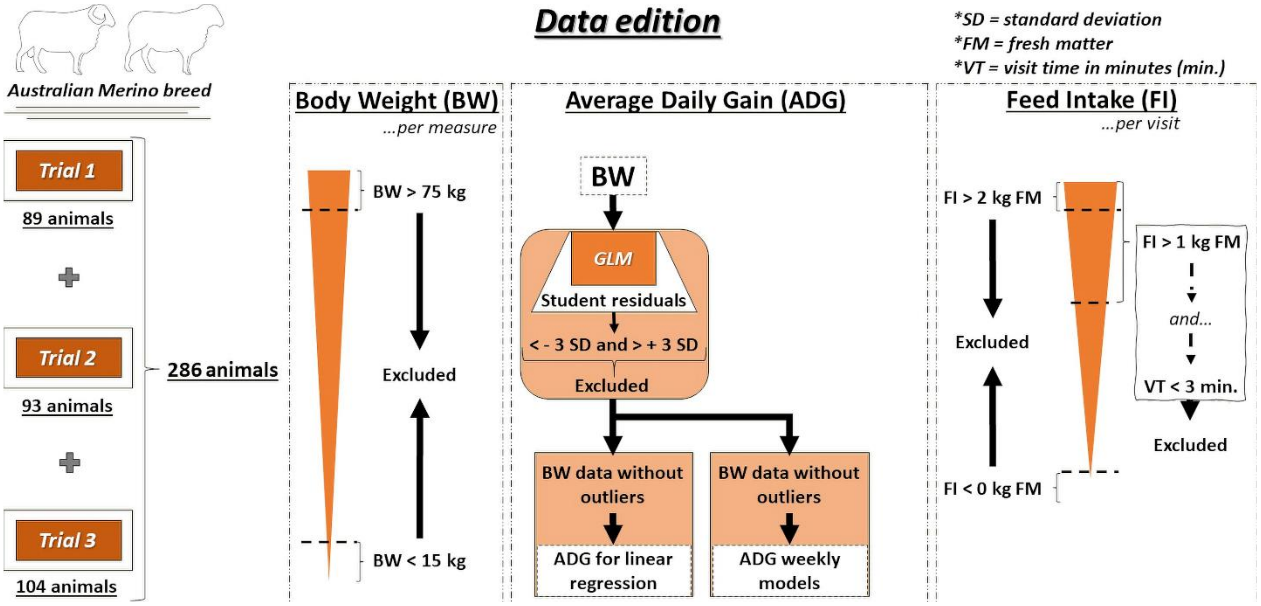


Figure 3. Body weights less than 15 kg or greater than 75 kg per visit were excluded, judging that these weights are biologically improbable. Body weights that presented Student residuals ± 3 SD were also excluded and the remainder were used to calculate the average daily gains by linear and weekly models. Feed intake per visit with values greater than 2 kg or lower than 0 kg, as it were considered biologically improbable, they were excluded. Consumptions greater than 1 kg with a duration of less than 3 minutes were also excluded.

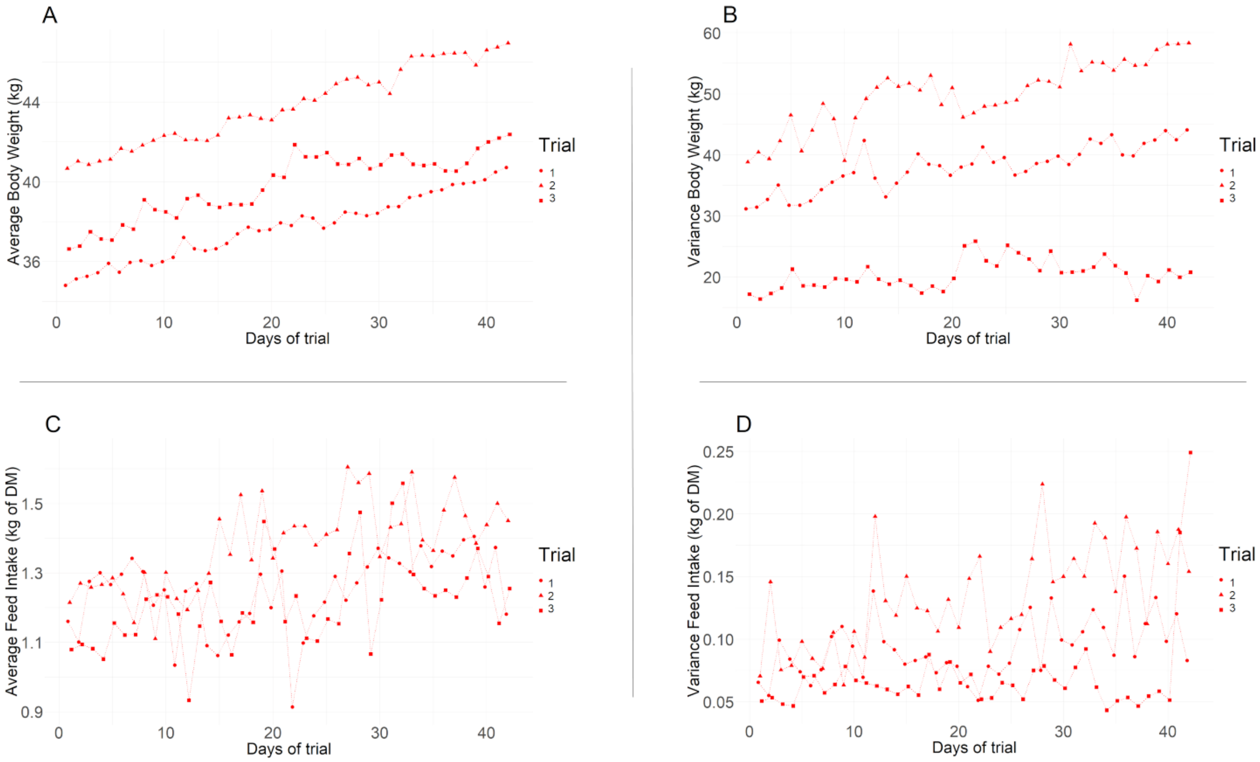


Figure 4. Average and variance of the body weight and feed intake per day in the trials.

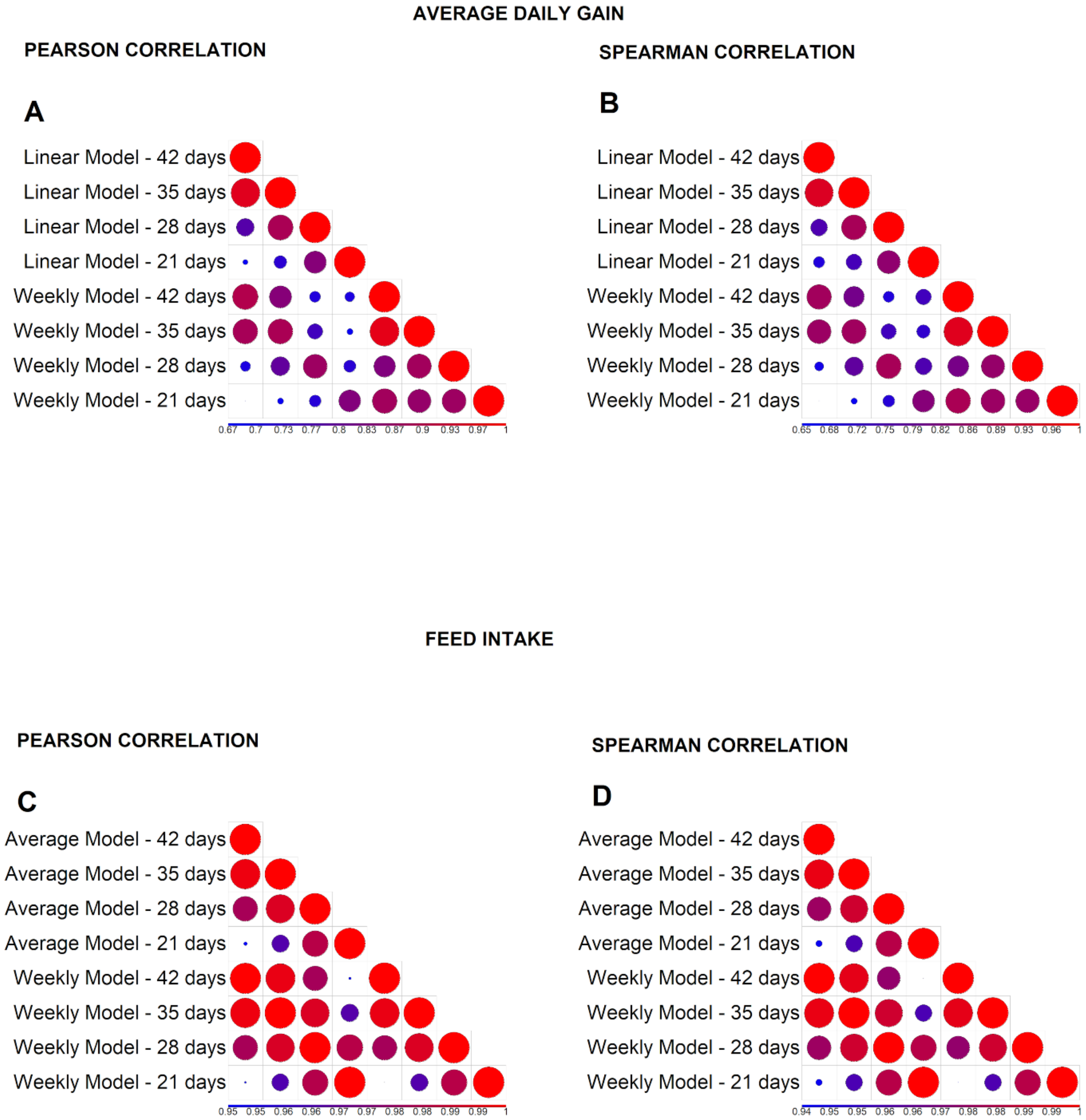


Figure 5. Pearson and Spearman correlation among the ADG estimates (A and B) and FI measurements (C and D), for different models. The bigger and red the circles the closer

to 1 are the correlations and the smaller and blue the circles, the closer to lower values are the correlations.

*p-value < 0.05 for all Pearson and Spearman correlations.

**Linear Model – 42 days – ADG linear model with 42 days on trial; Linear Model - 32 days – ADG linear model with 35 days on trial; Linear model - 28 days – ADG linear model with 28 days on trial; Linear Model - 21 days – ADG linear model with 21 days on trial; Weekly Model - 42 days – ADG weekly model with 42 days on trial; Weekly Model - 35 days – ADG weekly model with 35 days on trial; Weekly Model - 28 days – ADG weekly model with 28 days on trial; Weekly Model - 21 days – ADG weekly model with 21 days on trial; Average Model - 42 days – Average FI model of the 42 days on trial; Average Model - 35 days – Average FI model of the 35 days on trial; Average Model - 28 days – Average FI model of the 28 days on trial; Average Model - 21 days – Average FI model of the 21 days on trial; Weekly Model - 42 days – Average FI model per week into 42 days of trial; Weekly Model - 35 days – Average FI model per week into 35 days of trial; Weekly Model - 28 days – Average FI model per week into 28 days of trial; and Weekly Model - 21 days – Average FI model per week into 21 days of trial.

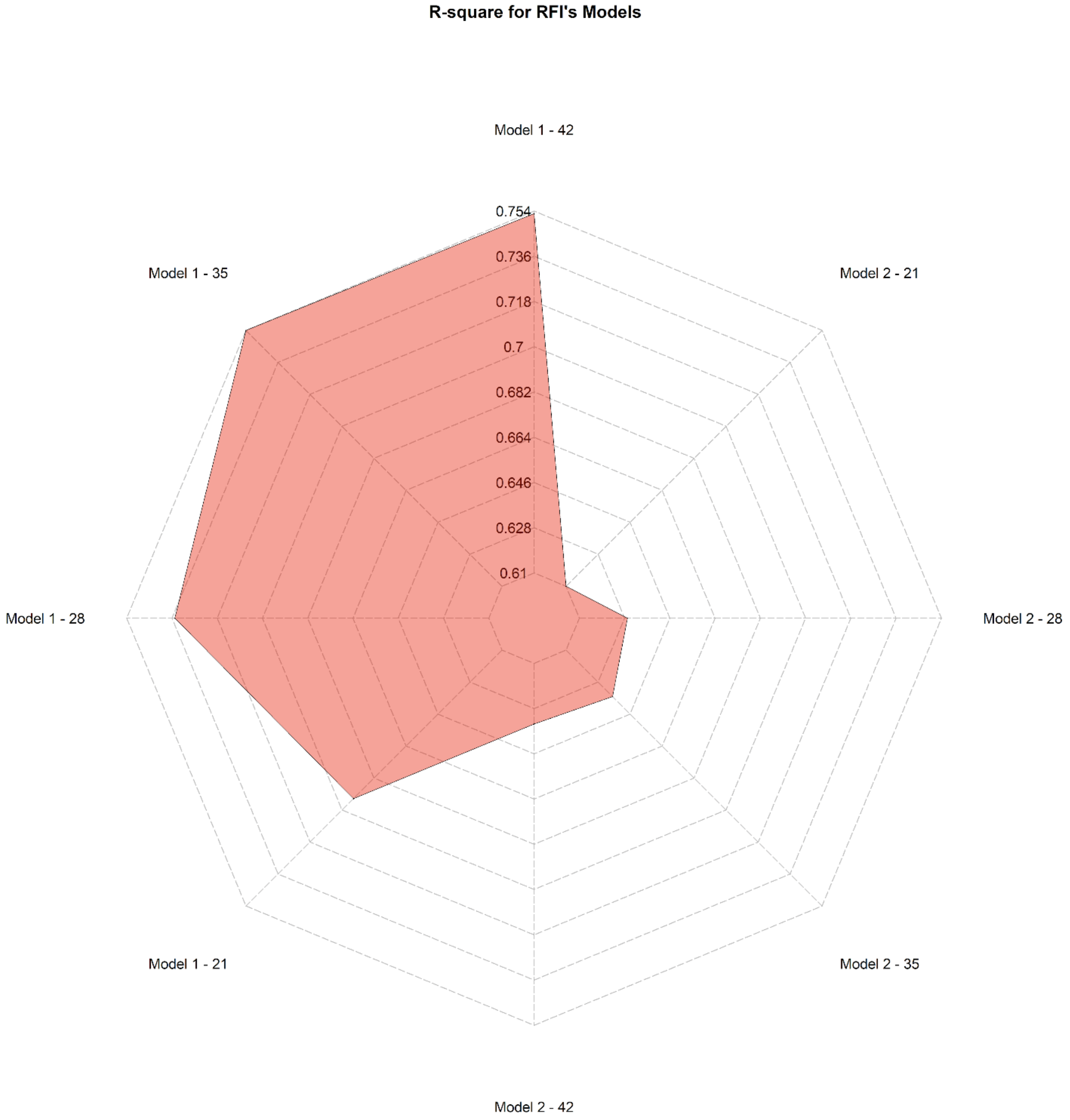


Figure 6. R-square for different RFI models.

*Model 1 - 42 – Linear model with 42 days on trial; Model 1 - 35 – Linear model with 35 days on trial; Model 1 - 28 – Linear model with 28 days on trial; Model 1 - 21 – Linear model with 21 days on trial; Model 2 - 42 – Weekly model with 42 days on trial; Model 2 - 35 – Weekly model with 35 days on trial; Model 2 - 28 – Weekly model with 28 days on trial; and Model 2 - 21 – Weekly model with 21 days on trial.

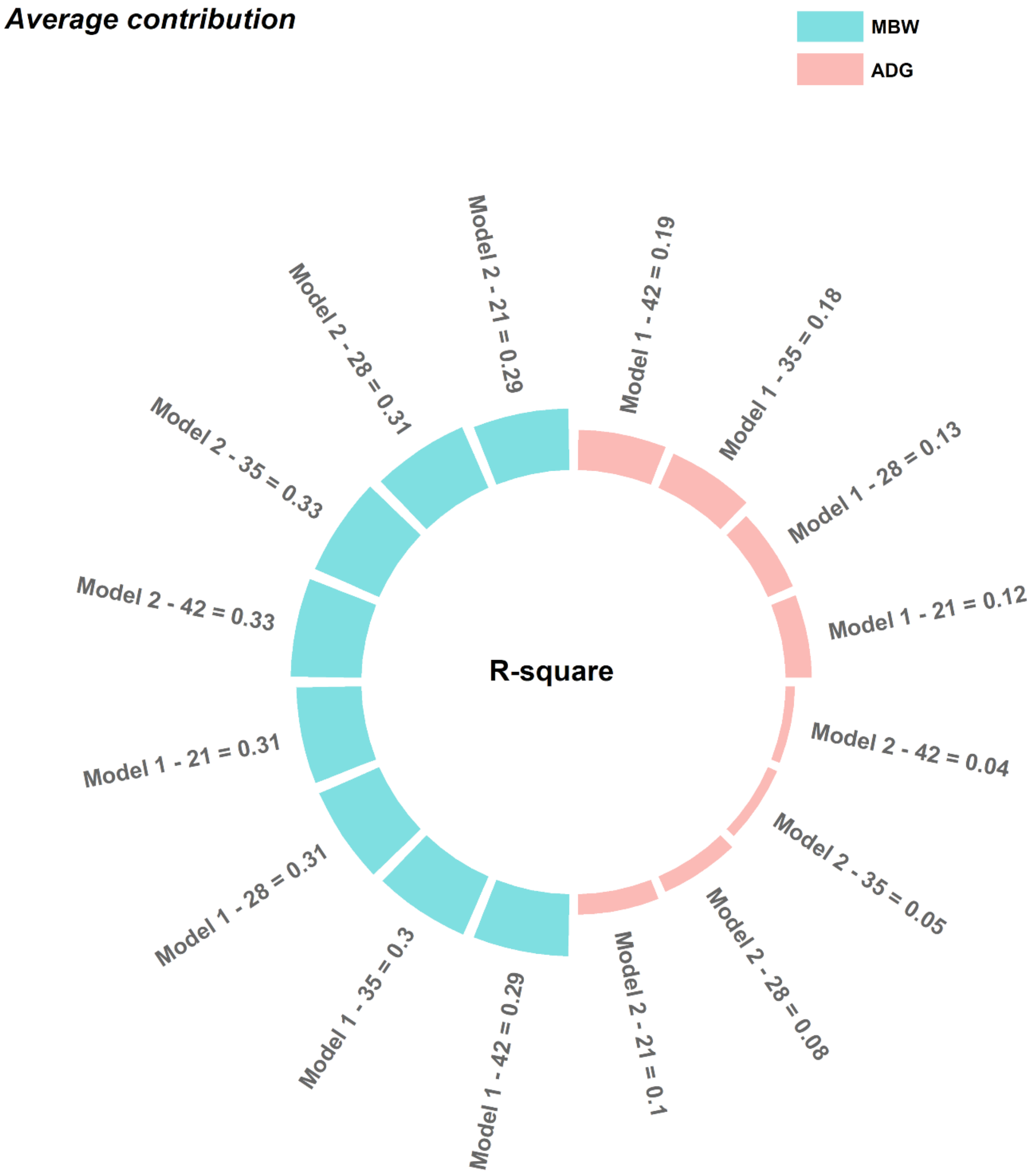
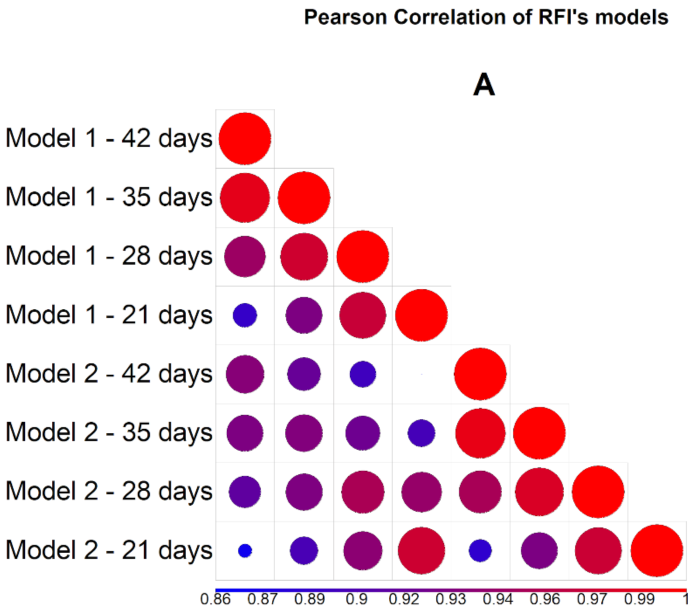


Figure 7. Average contribution (R-square) of the covariables of the RFI models.

*ADG – Covariable of Average Daily Gain and MBW – Covariable of Metabolic Body Weight.

** Model 1 - 42 – RFI linear model with 42 days on trial; Model 1 - 32 – RFI linear model with 35 days on trial; Model 1 - 28 – RFI linear model with 28 days on trial; Model 1 - 21 – RFI linear model with 21 days on trial; Model 2 - 42 – RFI weekly model with 42 days on trial; Model 2 - 35 – RFI weekly model with 35 days on trial; Model 2 - 28 – RFI weekly model with 28 days on trial; and Model 2 - 21 – RFI weekly model with 21 days on trial.


Pearson Correlation of RFI's models

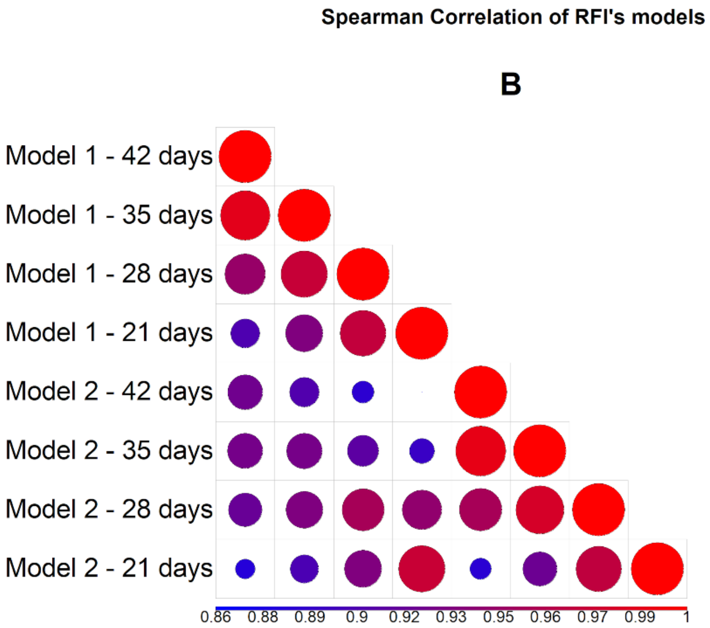

Spearman Correlation of RFI's models

Figure 8. Pearson (A) and Spearman (B) correlation among the different RFI models. The bigger and red the circles the closer to 1 are the correlations and the smaller and blue the circles, the closer to lower values are the correlations.

*p-value < 0.05 for all Pearson and Spearman correlations.

** Model 1 - 42 days – RFI linear model with 42 days on trial; Model 1 - 32 days – RFI linear model with 35 days on trial; Model 1 - 28 days – RFI linear model with 28 days on trial; Model 1 - 21 days – RFI linear model with 21 days on trial; Model 2 - 42 days – RFI weekly model with 42 days on trial; Model 2 - 35 days – RFI weekly model with 35 days on trial; Model 2 - 28 days – RFI weekly model with 28 days on trial; and Model 2 – 21 days – RFI weekly model with 21 days on trial.

CAPÍTULO IV – MACHINE LEARNING FOR FEED INTAKE PREDICTION IN SHEEP[1]

---

[1] Artigo será submetido para The Journal of Agricultural Science.

5. MACHINE LEARNING APPROACHES FOR FEED INTAKE PREDICTION IN AUSTRALIAN MERINO, CORRIEDALE AND DOHNE MERINO SHEEP IN URUGUAY

F. Amarilho-Silveira[1*], I. De Barbieri[2], J. A. Cobuci[1], C. B. Marques[3] and G. Ciappesoni[3]

[1]Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves, 7712, 91509-900, RS, Brasil.

[2]Instituto Nacional de Investigación Agropecuaria, Ruta 5 km 386, 45000, Tacuarembó, Uruguay.

[3]Instituto Nacional de Investigación Agropecuaria, Ruta 48 km 10, 90200, Canelones, Uruguay.

* **Author for correspondence:** F. Amarilho-Silveira, E-mail: famarilhosilveira@gmail.com

Abstract

Feed intake is a trait of difficult measurement, associated with the high cost of labour, feeding, and facilities. Applying machine learning approaches, considering traits as potential predictors, could optimize the resources and provide a greater number of animals with feed intake information. For this reason, the objective of this research was to test four approaches to find the best feed intake prediction method in Australian Merino, Corriedale, and Dohne Merino dataset, using the Australian Merino as the training dataset. The study dataset comes from 613 Australian Merino, 156 Corriedale, and 237 Dohne Merino sheep of ten feed intake trials carried out in different periods of the years 2019 and 2020. The Australian Merino dataset was used to train the algorithms and directly validated in the Corriedale and Dohne Merino datasets. The prediction models were the linear regression, nonlinear regression, k-nearest neighbors regression and random forest regression. The greater R^2 was found in the Nonlinear Model with a value of the 0.76 in the cross-validation in training dataset, with Pearson and Spearman correlations coefficients, respectively, from 0.8y and 0.86. In the Australian Merino, Corriedale and Dohne Merino dataset validation, Pearson correlation coefficients ranging 0.69 to 0.86 and Spearman correlation coefficients range from 0.66 to 0.86 were presented. The best model for predict feed intake in the Australian Merino, Corriedale and Dohne Merino datasets were, respectively, the k-Nearest Neighbors, Random Forest and Nonlinear models, respectively.

Key words: artificial intelligence, feed efficiency, k-nearest neighbor, methane, random forest

Introduction

The great debate in the livestock sector is the challenges of mitigating the emission of greenhouse gases and increasing food production for a constantly growing world population. The Paris agreement limits global warming to 1.5ºC with greenhouse gases emission reduction from 16 to 41% in 2050 relative to 2010 (LEAHY; CLARK; REISINGER, 2020) and the global population is projected to increase by 38% from the same period (KOCHHAR; OATES, 2014) with meat consumption projected to increase by more than 70% compared with 2010 levels (MCLEOD, 2011). This represents that the livestock system will have to be more efficient in land use for increasing the production levels and decreasing the emission of greenhouse gases.

The out-farm impact view about the status of environment-friendly production system has well-defined goals; however, the view on-farm needs more discussions in the search of alternatives for economically sustainable systems. This involves approaches to find animals that are more efficient, that can be productive with less exigency of feed intake, principally for the feed costs to represent an important expenditure in the production system. In New Zealand with a system based on pasture, for example, the expenditure relative at sheep and beef feeding represented in the years 2019-2020 approximately 56% of the total direct cost (BEEF + LAMB NEW ZEALAND, 2021).

In Ireland, which uses both produced food and purchased food, the National Farm Survey of Ireland published in 2020 that direct and indirect costs with sheep feeding can reach 73% to total direct cost of production (DILLON; MORAN; DONNELLAN, 2021). These situations show the importance of identifying and selecting animals that eat less and that will have an optimal production level indifference of the type of production system.

Selecting animals that eat less will not only provide a decrease in production costs but also indirectly decrease greenhouse gas emission, due to the linear relationship between feed intake and gas production (CHARMLEY *et al.*, 2016). An of the probable reasons this relation is because when the digestibility is improved, the feed intake is decreased, and together the reduction of gas emissions occurs (CANTALAPIEDRA-HIJAR *et al.*, 2018). This fact can be considered, principally, when the feed intake measurement is not practical or presents high cost of measurement, since the gas measurements using portable accumulation chambers (PAC) can be used in the field (DOMINIK *et al.*, 2017). Thus, the gas emission, with emphasis on methane emissions measure with PACs, can be used as a proxy for feed intake prediction (ROBINSON; ODDY, 2016) presented correlations ranging from 0.86-0.95 between feed intake and methane emissions (ROBINSON *et al.*, 2020).

Unlike the relationship with the methane emission, the productive traits do not present a marked relationship with feed intake. Safari *et al.* (2007b) found low genetic and phenotypic correlations between growth traits and feed intake. The lack of strong genetic and phenotypic relationships between feed intake and the production traits was defined for Fogarty *et al.* (2009) as a precludes in their use as indirect selection criteria for feed intake. However, the genetic and phenotypic low to medium correlations found for Tortereau *et al.* (2020) between feed intake and average daily gain (0.59 and 0.78), weight at the beginning of the test (0.45 and 0.29), weight at the end of the test (0.60 and 0.78), backfat thickness measured at the end of the test (0.31 and 0.28) and muscle depth measured at the end of the test (-0.12 and 0.18) put theses variables as potentially proxy for feed intake prediction.

Individual feed intake is a trait of difficult measurement associated with a high cost with labour, feeding, and facilities that applying the machine learning approaches considering traits (features) as potential predictors, could optimize the resources and provide a greater number of animals with feed intake information. Thus, using the machine learning advantage compared to common linear models to learn relationships from training data and generalise it to the unseen testing set and to overcome non-linearity and interactions among features (SHAHINFAR; KAHN, 2018), the gas emissions and production traits could be proxy for feed intake prediction. Several studies with sheep have been carried out using the machine learning approaches, of those the machine learning was utilized, for example, for accurately and precisely predict adult wool growth and quality traits, based in yearling wool, conformation and health traits along with pasture and climate data (SHAHINFAR; KAHN, 2018). The machine learning algorithms also were used to detect basic behaviours in sheep such as grazing, lying, standing, and walking, activity behaviour or detection of body posture as prostrate or upright (FOGARTY *et al.*, 2020) or used as classifying lamb mortality into risk classes (ODEVCI; EMSEN; AYDIN, 2021). In Romney Ewes study showed that ewe body condition can be predicted with great accuracy using previous liveweight with machine learning algorithms (SEMAKULA *et al.*, 2021). However, these several studies none used the machine learning approaches for feed intake prediction in sheep, having an important field in the research to be explored.

Currently in Uruguay the breeds that have feed intake information are the Australian Merino, Corriedale, Dohne Merino and Texel (GIORELLO *et al.*, 2021). The Australian Merino, Corriedale and Dohne Merino can be grouped as woolly breeds and represent more than 91% of the animals with feed intake information and in the Uruguayan genetic evaluation their productive traits are practically the same (https://www.geneticaovina.com.uy). The Australian Merino breed database has greater feed intake information and could be therefore considered as the main input for machine learning application in this first study using this approach. Thus, we hypothesized that feed intake in animals Australian Merino, Corriedale, and Dohne Merino could be predicted using the majority of the Australian Merino dataset and samples of the Corriedale and Dohne Merino dataset as train data with good accuracy using machine learning approaches. We aimed to test four approaches to find the best method with the capacity to predict feed intake in the Australian Merino, Corriedale, and Dohne Merino dataset using a dataset with Australian Merino, Corriedale, and Dohne Merino information.

Materials and methods

The study dataset comes from ten feed intake trials carried out in different periods of the years 2019 and 2020, six with Australian Merino, two with Corriedale, and two with Dohne Merino breed. The experimental site where the data was recorded is located at La Magnolia Experiment Unit (National Agricultural Research Institute of Uruguay), Tacuarembó, Uruguay. Records were collected on 1,006 Sheep (613 Australian Merino, 156 Corriedale and 237 Dohne Merino) offspring of 37 rams (16 Australian Merino, 11 Corriedale and 10 Dohne Merino). The Australian Merino evaluations started: 04-29-2019, 06-07-2019, 08-12-2019, 4-28-2020, 6-30-2020 and 8-24-2020, in the six experiments respectively. The Corriedale evaluations started: 2-12-2019 and 3-3-2020, in the two experiments respectively. The Dohne Merino evaluations started: 10-07-2019 and 10-19-2020, in the two experiments, respectively.

All protocols applied were approved by INIA Animal Ethics Committee (INIA 2018.2 and 2018.3).

Data edition and description

From 1,006 records within the dataset, animals without its sire information and the following data were excluded: body weight in the start of the trial; yearling greasy fleece weight; yearling wool diameter; yearling body weight; average of the methane emissions; dam age at lambing and shearing age, remaining 943 records (575 Australian Merino, 149 Corriedale and 219 Dohne Merino).

The data description is presented in the Table 1, where the body weight (BW) average at the start and the end of the trial are: 36.40±5.96 kg - 43.90±6.86 kg for Australian Merino, 29.90±4.12 kg - 52.90±6.42 kg for Corriedale, and 45.30±5.43 kg - 36.40±4.77 kg for Dohne Merino, respectively. The individual daily feed intake was gotten by total feed intake per day expressed in dry matter (fresh matter intake × range 50.9 - 78% (proportion of dry matter); DMI). The average feed intake based in dry matter into 42 days on the trials for Australian Merino, Corriedale, and Dohne Merino, respectively, were of 1.41±0.293, 1.28±0.256 and 1.66±0.296 kg. In the Corriedale and Dohne breeds, only females were evaluated (149 and 219, respectively). In Australian Merino 296 males and 279 were evaluated. The average age at the start of trial was 297±47, 189±12 and 420±8 days for Australian Merino, Corriedale, and Dohne Merino breed, respectively.

Feed intake trials

The trial duration period was of 56 days, with a 14 days feed and facilities adaptation period, making 42 days of feed intake evaluation. Animals were fed *ad libitum* with Lucerne haylage (DM 50.6-78.0%; crude protein 19.6-23.5%; NDF 29.8-38.7%; ADF 23.5-30.7%). Each pen had five individual automated feeders and two automatic weighing platforms, which were equipped with an electronic tag reader, precision scale, and connected to a central computer, this allowed the control of body weight and feed intake of the animals in a daily basis. At the day 7, the entrance into collective pens was allowed after deworming and allocated to one of five automated feeding systems (pens) in accordance with the body weight (BW), sex, type of birth and sire.

Daily monitoring was applied by a software system that identified the entry of animals into the feeder and the body weighting platform lectures. The equipment and software were provided by Intergado (Intergado®, Belo Horizonte, MG, Brazil). The RFID tagged allows the identification of a specific animal at the feed bin and, consequently, their feed intake based on the difference in feed weight before and after the visit. The body weighing platform was set in the water bins where a similar sensor to the feed bins system was present, each time the animal accessed it the BW was automatically recorded.

After each feed bin and BW platform visit, the system documented the visit events by recording the animal's identification tag, bin, and platform number. The data collection and the functioning of the equipment can be found in Amarilho-Silveira *et al.* (2022).

Gas measurements

Methane emissions were estimated following the portable accumulation chambers (PAC) protocol described by Goopy *et al.* (2011, 2016), Paganoni *et al.* (2017), and Robisnon *et al.* (2014). In brief, two estimates per animal were performed during the last two weeks of the feed intake test (at least one week between estimates), allowing to know the feed intake and body weight of the animals on the day and previous days of gases emission estimation. The traits evaluated were methane emissions, CO_2 emissions and O_2 consumption. In the week of measure, 1 pen per day was measured in consecutive runs of 10 animals, therefore, 20 animals per day were measured and by the end of the week 100 animals. If the feed intake test considers more than 100 animals, an extra run per day was performed if necessary. In accordance with Robinson *et al.* (2020) animals were on feed until the moment of the measure. After, were allocated to one of 10 sealed chambers of 862 lts of volume. At 20 to 30 and 40 to 50 minutes later, estimates of the concentrations of CH_4, CO_2 and O_2 were performed, in parallel, estimates of temperature, atmospheric pressure and gases concentration in the air were done. Gases measures were performed with an Eagle 2 equipment (RKI instruments, Union City, CA, USA). The Eagle 2 and chambers were checked between measure weeks and the Eagle 2 was calibrated periodically in accordance with the specifications provided by RKI instruments.

In regard to calculations, first the concentration of CH_4, CO_2 and O_2 were calculated in litres per day (l/d), as shown below (edited from Goopy *et al.,* 2011; 2016, Paganoni *et al.*, 2017, and Robisnon *et al.,* 2014):

$$\mathrm{CH_4\ (l/d)} = \frac{\left(\frac{\Delta \mathrm{CH_4}}{\Delta \mathrm{PAC}_{time}}\right) \times (\mathrm{PAC}_v - \mathrm{BW} \times 1.01)}{1.000.000} \quad \text{(Eq. 1)}$$

$$\mathrm{CO_2\ (l/d)} = \frac{\left(\frac{\Delta \mathrm{CO_2}}{\Delta \mathrm{PAC}_{time}}\right) \times (\mathrm{PAC}_v - \mathrm{BW} \times 1.0)}{100} \quad \text{(Eq. 2)}$$

$$\mathrm{O_2\ (l/d)} = \frac{\left(\frac{\Delta \mathrm{O_2}}{\Delta \mathrm{PAC}_{time}}\right) \times (\mathrm{PAC}_v - \mathrm{BW} \times 1.0)}{100} \quad \text{(Eq. 3)}$$

Where the ΔCH_4, ΔCO_2 and ΔO_2 is the difference between end and start gases concentrations and ΔPAC_{time} is the difference between end and start time (hour:minutes:seconds) per each animal measurement period; the PACv is the PAC volume measure between the water line that seals the chamber to PAC top, and BW is the body weight of the animal on the day of the measure.

In a second step, the l/d has been converted to grams per day (g/d) at standard temperature and pressure, as shown (edited from Goopy *et al.,* 2011; 2016, Paganoni *et al.*, 2017, and Robisnon *et al.,* 2014):

$$\mathrm{CH_4\ (g/d)} = \mathrm{CH_4\ l/d} \times \frac{\mathrm{Pressure} \times 0.1}{8.3143 \text{ x } (\mathrm{Temperature} + 273.15)} \times 16 \quad \text{(Eq. 4)}$$

$$\mathrm{CO_2\ (g/d)} = \mathrm{CO_2\ l/d} \times \frac{\mathrm{Pressure} \times 0.1}{8.3143 \text{ x } (\mathrm{Temperature} + 273.15)} \times 44 \quad \text{(Eq. 5)}$$

$$O_2\ (g/d) = O_2\ l/d \times \frac{\text{Pressure} \times 0.1}{8.3143 \text{ x } (\text{Temperature} + 273.15)} \times 32 \quad \text{(Eq. 6)}$$

Where the 0.1 is a multiply to transform hPa for KgPa, the 8.3143 is the universal gas constant (R), 273.15 is equivalent to 1ºC in Kelvin and 16, 44 and 32 are the molar weight, respectively, to CH_4, CO_2 and O_2.

Feature selection

The feature selection was done aiming to find the more explaining variables of the average feed intake based on dry matter in kg (FI). The considered variables were the following: the breed (BREED), dam age at lambing (DA, years); weaning age (WA, days); test start age (TSA, days); body weight at the start of the trial (BWsT, kg); body weight (BWmT, kg) and metabolic body weight (MBW = $BWmT^{0.75}$, kg) at the middle of the trial; rib eye area (REA, cm^2), fat thickness (FT, mm) and body weight (BWeT, kg) at the end of the trial; average of CH_4 emissions (CH_4, g/day); average of CO_2 emissions (CO_2, g/day); average of O_2 consumption (O_2, g/day); body weight difference between the start and end of the trial (BWdif, kg); sex of the animal (SEX); birth type (BT, 1 or 2); yearling greasy fleece weight (YGFW, kg); yearling wool diameter (YFD, microns); yearling body weight (YBW, kg); yearling staple length (YSL, cm) and yearling shearing age (YSA, days).

The REA and FT were measures for backfat ultrasound recorded ate end of the trial, at the level of the last floating rib at 1 cm from the spine (between the 12th and 13th ribs).

Linear regression, nonlinear regression with generalized additive models, k-nearest neighbor regression, and random forest regression were used for feature selection.

Statistics and machine learning algorithms procedures

The training dataset was constituted for 477, 70, and 93 observations of the Australian Merino, Corriedale, and Dohne Merino dataset, respectively, and divided in the training and testing using the k-fold cross-validation and with a direct validation in samples from Australian Merino (n =97), Corriedale (n=79), and Dohne Merino (n=125) dataset not known by the models.

All procedures were used for to solve regression problems using the linear regression, nonlinear regression, k-nearest neighbors regression and random forest regression of the mlr R package . Linear regression uses the equation of a straight line (Linear Model); nonlinear regression uses supervised learns generalized additive models (Nonlinear Model); k-nearest neighbor regression uses the k-nearest neighbors' algorithm for regression (k-Nearest Neighbors Model); and random forest regression uses tree-based algorithms for regression (Random Forest Model; Rhys 2020).

The feature selection was done using the same prediction procedure, that is, for example, when the prediction was performed using Random Forest Model the feature selection was done using the same procedure. Thus, the algorithms of mlr R package were, respectively, regr.lm, regr.gamboost, regr.kknn and regr.randomForest for Linear Model, Nonlinear Model, k-Nearest Neighbor Model, and Random Forest Model.

The sequential floating backward search wrapper method for feature selection was used with the cross-validation using k-fold resample with 10 iterations using the same model to make predictions.

For k-Nearest Neighbor Model was tuned the k hyperparameter, using cross-validation with k-fold resample with 20 iterations, getting 25 k-nearest neighbor. The hyperparameters for Random Forest Model considered were the number of individual trees in the forest between 50 and 500, number of features to randomly sample at each node between 100 and 575, minimum number of cases allowed in a leaf between 1 and 10, maximum number of leaves allowed between 5 and 50, maximum from 500 iterations in the random search method, and cross-validation with k-fold resample with 20 iterations. The hyperparameters tuned were 161 trees, 379 features randomly sampled, 6 minimum number in a leaf and 48 maximum number of leaves.

The models training was done using models with cross-validation performed using k-fold resample with 20 iterations.

The k-fold is cross-validation randomly split the data into approximately equal-size chunks called folds. Then was reserve one of the folds as a test set and use the remaining data as the training set. After pass the test set through the model or learning algorithm is record of the performance metrics: R^2, root mean squared error (*RMSE*) and Pearson and Spearman correlations. Was used different folds of the data as test set and doing the same thing continuing until all the folds have been used once as the test set. The performances metrics were the average of every test set run.

The Pearson and Spearman correlations were used as validations metrics in the Corriedale and Dohne Merino dataset with the cor.test function from stats R package (R CORE TEAM, 2021).

The model proposed by Koch et al. (1963) was used to obtain the estimated feed intake which was correlated with the observed and predicted feed intake by the models proposed.

For the confusion matrix calculation, the observed and predicts feed intake were transformed in class. Animal with feed intake less or equal to the 25th percentile was classified as low, greater than or equal to the 75th percentile as high and between 25th and 75th percentiles as medium. This classification is analogous at classification for residual feed intake (ELLISON *et al.*, 2019). The confusion matrix was plotted using the plot_confusion_matrix function from cvms R package (OLSEN, 2021 and R CORE TEAM, 2021).

Results

The selected variables in the feature selection in the Australian Merino dataset were the MBW,

CO_2 and BWdif for Linear Model, Nonlinear Model and Random Forest Model and BWmT, CO_2 and BWdif for k-Nearest Neighbor Model, all with a mean squared error, respectively of the 0.021, 0.021, 0.024 and 0.022 (Table 2).

On Table 3 is shown the performance metrics for feed intake algorithms cross-validation in the training dataset. The greater R^2 was found in the Nonlinear Model with a value of the 0.76 and for the other models the R^2 were equal from 0.75. The Pearson and Spearman correlations coefficients ranging from 0.84 and 0.87, respectively. The Pearson correlations coefficients in the Random Forest was from 0.86 and for the other models was equal to 0.87. The Spearman correlations coefficients for Linear, Nonlinear, k-Nearest Neighbor and Random Forest models were 0.85, 0.85, 0.87 and 0.84, respectively.

The validation in the Australian Merino, Corriedale and Dohne Merino datasets presented Pearson correlation coefficients ranging 0.69 to 0.86 ($P < 0.001$) and Spearman correlation coefficients range from 0.63 to 0.86 ($P < 0.001$; Table 4). The best methods for feed intake prediction in the Australian Merino was the Nonlinear model (*r*-Pearson = 0.86 and *r*-Spearman = 0.86). The best methods for feed intake prediction in the Corriedale dataset were the Linear and Nonlinear Models (*r*-Pearson = 0.74 and *r*-Spearman = 0.69) and the best method for Dohne Merino dataset were the Nonlinear and Random Forest models (*r*-Pearson = 0.79 and *r*-Spearman = 0.77).

In the Table 5 is show the Pearson and Spearman correlation coefficients among the observed and predict feed intakes and the feed intake expected for Kock model. Between the observed and estimated the Pearson and Spearman correlations presented values raging from 0.75 to 0.78 in the validation's datasets. The Pearson and Spearman correlations between estimated and predicts feed intake ranging from 0.86 to 0.94. The greater Pearson Spearman correlations were found in the Nonlinear Model in Corridale dataset (*r*-Pearson = 0.94 and *r*-Spearman = 0.93) and the lower were found in the k-Nearest Neighbor Model in Australian Merino dataset (*r*-Pearson = 0.86 and *r*-Spearman = 0.86).

In the Australian Merino dataset, the confusion matrix presented accuracy for feed intake classification between observed and predict feed intake of 0.69, 0.69, 0.73 and 0.67, for Linear Model, Nonlinear Model, k-Nearest Neighbor Model and Random Forest Model (Fig. 1), respectively.

The confusion matrix in the Australian Merino dataset not presented animals misclassified as low and high or as high and low feed intake. However, the Corriedale dataset was found this type of misclassified, but in a very low frequency.

The Corriedale dataset (Fig. 2) the accuracy was 0.67, 0.67, 0.61 and 0.70 for Linear Model, Nonlinear Model, k-Nearest Neighbors Model and Random Forest Model respectively. This dataset presented 10% misclassification of low when in really the feed intake was high for all models and 5.2% misclassification of high as low the k-Nearest Neighbors Model.

The Dohne Merino dataset the accuracy presented values from 0.64, 0.66, 0.63 and 0.64 for the Linear Model, Nonlinear Model, k-Nearest Neighbors Model and Random Forest Model, respectively (Fig. 3).

Discussion

In Uruguayan genetic evaluation, the Australian Merino dataset was the one presenting more feed intake information, the development of prediction models could be applied not only in future animals of this breed but also in Corriedale and Dohne Merino, improving the feed intake data even if indirectly. In this mode, machine learning approaches can bring interesting generalization to the use of information available to obtain feed intake data.

The difference between feed intake and gaseous emissions observed in Table 1 is dependent on several factors, as the Corriedale breed showed lower feed intake and gaseous emissions, but also lower growth and final weight because they were animals that were evaluated in a younger age.

Feature selection

Feature selection can be understood as variable choice that best help in understanding data, aiming reducing computation requirement and the dimensionality with consequently improvement in performance of the predictors (CHANDRASHEKAR; SAHIN, 2014). Thus, when the traits presented a high level of redundancy, can produce different feature rank in different training samples and consequently different models will be produced with the same prediction accuracy (PILES *et al.*, 2021). In the present study 20 variables were tested and selected only three for Linear Model, Nonlinear Model and Random Forest Model, and four for k-Nearest Neighbor Model.

The small number of variables selected that had capacity of explaining of the variation in the feed intake for feature selection reinforce the found in the studies of Safari *et al.* (2007b) and Fogarty *et al.* (2009) where no productive variable taken outside the feed intake test period was selected as a potential predictor, and the found for Tortereau *et al.* (2020) about the low to medium relationship among feed intake and the measures traits in the test period. The CH_4 pointed out by Donoghue *et al.* (2015), Robinson & Oddy (2016) and Robinson *et al* (2020) like a great proxy, in the present study it was not selected as a variable with good explanatory power of feed intake. Among the variables referring to gas measurements the selected was the CO_2 emission, that presented a correlation with feed intake from 0.79 to 0.83 in Robinson *et al.* (2016) and phenotypic and genetic correlations range from 0.71 to 0.74 and 0.86 to 0.96 for lambs and hogget, respectively, in Paganoni *et al.* (2017). This can be explain for the founded for Arthur *et al.* (2018)

that as the quantity of feed consumed increases, the amount of CO_2 produced also increases. Renand *et al.* (2019) mention the CO_2 emission as a potential variable for indirect selection of feed efficiency together with CH_4 when no daily feed intake measurement is available. However, as the correlations between CH_4 and CO_2 ranging from moderate to very strong for Robinson *et al.* (2016), moderate to strong for Paganoni *et al.* (2017), and weak to moderate for Jonker *et al.* (2018) could explain why CO_2 was selected; since the selection of one excludes the other due to high redundancy if both are considered in the model. The CO_2 is produced by both ruminal fermentation and oxidation of substrates is proportional to energy expenditure (HEGARTY, 2013) and the CH_4 emission is related with an inefficient digestive process representing an important loss of feed gross energy (OLIJHOEK *et al.*, 2018). Thus, a reduction of gas emissions due to its relationship with digestibility improvement, consequently, also would be observed a feed intake decrease (CANTALAPIEDRA-HIJAR *et al.* 2018) without necessarily compromising the daily supply of nutrients to the animal. Therefore, a low apparent CO_2 emission could be considered as a potential proxy for low feed intake in sheep (HEGARTY, 2013), since in feature selection, together body weight difference between the start and end of the feed trial (BWdif), the CO_2 was selected for all models in the feature selection.

Variables measured in the feed intake test period, normally present important influence in the explaining the animal consumption. Since the proposal for Koch *et al.* (1963) several studies analysed the feed intake records by linear form considering body weight and body weight gain as covariables in order to estimate the residual feed intake. The feed intake was modelling using the average daily gain (ADG) and live weight in the mid-test in Knott *et al.* (2008). In the studies of Johnson *et al.* (2017), Zhang *et al.* (2017), Montelli *et al.* (2019) and Amarilho-Silveira *et al.* (2022) the feed intake was adjusted using the ADG and metabolic body weight. Based in these studies and that for the estimation of residual feed intake it is necessary to predict the feed intake, the variables of body weight (or metabolic body weight) and weight gain (BWdif is an analogue variable to ADG, however, not regressed by test days) selected in the feature selection for all methods were acceptable proxy for feed intake prediction.

Feed intake prediction

Cross-validation is a statistical method of evaluating and comparing learning algorithms, dividing the dataset where one is used to training a model and the other is used for validation. K-fold cross-validation is a basic form in that data is first partitioned into k equally sized folds (RHYS, 2020). Cross-validation is used to evaluate or compare learning algorithms as follows: in each iteration, one or more learning algorithms use k-1 folds of data to learn one or more models, and subsequently, the learned models are asked to make predictions about the data in the validation fold (REFAEILZADEH; TANG; LIU, 2016). In the present study were used model for regression problems in that R^2, *RMSE*, *r*-Pearson, and *r*-Spearman were metrics used as an average of all iterations was used to obtain an aggregate measure from these samples.

The feed intake in the training dataset was predicted using the k-folds cross-validation with metrics that made it possible to observe predictions with relative security. To date, no study has been found using the machine learning approach to predict feed intake in sheep, however, several studies using linear models have described equations for such prediction with R^2 ranging from 0.38 to 0.84. Knott *et al.* (2008) found R^2 ranging from 0.56 to 0.74; Redden *et al.* (2011) found R^2 between 0.38 and 0.41; Cockrum *et al.* (2013) found R^2 range 0.43 to 0.46; Johnson *et al.* (2015) found R^2 equal to 0.79; Redden *et al.* (2014) found R^2 equal to 0.84; Johnson *et al.* (2016) found R^2 equal to 0.78; in Zhang *et al.* (2017) the R^2 was 0.80; Montelli *et al.* (2019) found R^2 equal to 0.82; Tortereau *et al.* (2020) the R^2 values ranging from 0.63 to 0.84 and; Amarilho-Silveira *et al.* (2022) the R^2 values ranging from 0.61 to 0.75. Even in these studies that did not use the machine learning approach, it is possible to visualize the baseline values for prediction models of feed intake. Thus, the results of the present study are between 0.34 and 0.35 points above the lowest R^2 and between 0.11 and 0.12 below the higher value found in the literature. However, the four models in the present study presented R^2 above 0.70 as described Johnson *et al.* (2017) as a model with good explanatory power.

In addition to acceptable R^2*'s* found in the present study, the similar distributions among the predicted and observed feed intakes may explain the strong Pearson and Spearman correlation coefficients obtained. However, these results were expected because the dataset is of the Australian Merino breed, which, by presenting a good number of observations distributed in six tests over two years, provided a great number of contrasts at each iteration of the k-fold cross-validation.

In the validation datasets the greater Pearson and Spearman correlation coefficients were found using the Nonlinear Model in the Australian Merino dataset. This is because the Australian Merino dataset made up 75% of the training set data. Therefore, the other validation datasets the best methods were the Linear and Nonlinear models for Corridale and Nonlinear and Random Forest models for Dohne Merino dataset. In a study on predicting feed intake of cows, Pearson's correlation coefficients using mixed and machine learning models showed values ranging from 0.71 to 0.76 between actual and predicted feed intake in the test dataset, with a decrease in these coefficients in machine learning models applied on the training dataset (KAMPHUIS *et al.*, 2017). Already in present study the correlations coefficients decreased in the validation (test) compared to training dataset. However, the majority models presented Pearson correlations coefficients greater than 0.71.

Confusion matrices

The correlation metrics gave a good indication together with a graphical look at the distributions, however, to visualize how the animals could be classified and thus extrapolate to practice it was proposed to present the results in the light of a confusion matrix. A confusion matrix presents information about how often a certain behaviour (in present study the feed intake) is detected correctly and how often it is classified as another behaviour (RUUSKA *et al.*, 2018).

The confusion matrices for the Australian Merino dataset presented accuracy greater than 0.65 for all models, with an emphasis on the k-Nearest Neighbors Model that presented an accuracy from 0.73. However, in the Corriedale dataset the Random Forest Model performed better with accuracy 0.70. The Random Forest Model has been shown to have good potential for the prediction of feed intake in grazing dairy cattle in the Leso *et al.* (2019). In lamb survival study the Random Forest performed very well in classification of the mothering ability with an accuracy form 0.83 (ODEVCI; EMSEN, 2019). In the Odevci *et al.* (2021) the accuracy for Random Forest Model were 0.83 and 0.93 for mothering ability and lamb survival rate, respectively. In Romney ewes the body condition score was predict with accuracy above of the 0.88 for pre-breeding, pregnancy diagnosis, pre-lambing and weaning using the Random Forest Model (SEMAKULA *et al.*, 2021). Therefore, when exist relatively large variation between the training and validation dataset the Random Forest Model may produce unstable results (LESO *et al.*, 2019). Therefore, the Corriedale dataset presented in all model’s classification errors of the observed low feed intake predicted as high and vice-versa.

In Dohne Merino dataset, in no model was found classification problems between low and high food consumption and vice versa and the best model was the Nonlinear Model. This can be explained by the fact that this dataset presents feed intake, CO_2, and metabolic weight (or body weight) similar to the Australian Merino dataset that is the with more information in the training dataset and, mainly, for the predictive capacity of the Nonlinear Model using generalized additive models. The generalized additive models are sufficiently flexible to capture the trends without strong assumptions about it shapes, beyond using substantially fewer parameters are able to adjust a wide variety of shapes of underlying data trend (BORCHERS *et al.*, 1997).

However, the weaknesses of generalized additive models is that they have a propensity to overfit on the training dataset (not observed in the present study) and the poor prediction when the validation dataset is outside the range of the training dataset values (noticeably in the Corriedale dataset; Rhys 2020). This form the substantial improvements in precision was achieved using Nonlinear Model, for its best accuracy among models apply in the Dohne Merino dataset and the not confusion between low and high feed intake.

Thus, the prediction of the feed intake of the Australian Merino and Dohne Merino animals using the k-Nearest Neighbors and Nonlinear models, respectively, would provide great impacts as increasing the number of animals with evaluation for feed intake as an anticipated decision ensuring that animals with low feed intakes not being classified as high and vice-versa. With such prediction, it will be possible resources saved because the same number of sheep will be evaluated in the feed intake test, but with more animals with feed intake information.

The application of cross-validation procedures, applied both in statistical models and in models based on algorithms that intelligently learn patterns and try to predict feed intake, may also increase the number of animals with daily feed intake information. For example, if structures for daily feed intake and weight collections are available that allow the collection of a specific number of the animals per evaluation period, simple structures with collective feeders and drinkers, in which the animals will have *ad libitum* access to feed and water would be enough to phenotype them. In these, it would only be necessary to obtain information from the animals of initial weight, weight in the middle of the evaluation period, final weight and two measurements of CO_2 emission using portable accumulation chambers. Thus, it would be able to have a feed intake evaluation centre and tests dispersed in several locations where the animals would have access to the same feed as the central test station and their body weight would be controlled at three times and two measurements of CO_2 emissions, being able to obtain the double of the animals with feed intake information in each evaluation period.

Conclusion

In conclusion, the best model for predict feed intake in the Australian Merino and Dohne Merino datasets were, respectively, the k-Nearest Neighbors and Nonlinear models. Therefore, the model with better metrics in the Corriedale was the Random Forest, but this was not able to not confound the feed intake of some animals with low and high classification.

Financial Support

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. The field work and data collection received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement n°772787 (Smarter) and from the Instituto Nacional de Investigación (INIA_CL_38: Rumiar).

Conflicts of Interest

The authors declare there are no conflicts of interest.

Ethical Standards

All protocols applied were approved by INIA Animal Ethics Committee (INIA 2018.2 and 2018.3).

References

AMARILHO-SILVEIRA, F., DE BARBIERI, I., COBUCI, J.A., BALCONI, C.M., DE FERREIRA, G.F. & CIAPPESONI, G. (2022). Residual feed intake for Australian Merino sheep estimated in less than 42 days of trial. *Livestock Science*, **258**, 104889.

ARCHER, J.A. & BERGH, L. (2000). Duration of performance tests for growth rate, feed intake and feed efficiency in four biological types of beef cattle. *Livestock Production Science*, **65**, 47–55.

ARTHUR, P.F., BIRD-GARDINER, T., BARCHIA, I.M., DONOGHUE, K.A. & HERD, R.M. (2018). Relationships among carbon dioxide , feed intake , and feed efficiency traits in ad. *Journal of Animal Science*, **96**, 4859–4867.

BEEF + LAMB NEW ZEALAND. (2021). *Compendium of New Zealand Farm Facts 2021*.

BISCHL, B., LANG, M., KOTTHOFF, L., SCHRATZ, P., SCHIFFNER, J., RICHTER, J., JONES, Z., CASALICCHIO, G., GALLO, M., BOSSEK, J., STUDERUS, E., JUDT, L., KUEHN, T., KERSCHKE, P., FENDT, F., PROBST, P., SUN, X., THOMAS, J., VIEIRA, B., BEGGEL, L., AU, Q., BINDER, M., PFISTERE, F., COORS, S., BRONDER, S., ENGELHARDT, A., MOLNAR, C. & SPOONER, A. (2021). Machine Learning in R. *R CORE TEAM*, 304.

BORCHERS, D.L., BUCKLAND, S.T., PRIEDE, I.G. & AHMADI, S. (1997). Improving the precision of the daily egg production method using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences*, **54**, 2727–2742.

CANTALAPIEDRA-HIJAR, G., ABO-ISMAIL, M., CARSTENS, G.E., GUAN, L.L., HEGARTY, R., KENNY, D.A., MCGEE, M., PLASTOW, G., RELLING, A. & ORTIGUES-MARTY, I. (2018). Review: Biological determinants of between-animal variation in feed efficiency of growing beef cattle. *Animal*, **12**, S321–S335.

CHANDRASHEKAR, G. & SAHIN, F. (2014). A survey on feature selection methods. *Computers and Electrical Engineering*, **40**, 16–28.

CHARMLEY, E., WILLIAMS, S.R.O., MOATE, P.J., HEGARTY, R.S., HERD, R.M., ODDY, V.H., REYENGA, P., STAUNTON, K.M., ANDERSON, A. & HANNAH, M.C. (2016). A universal equation to predict methane production of forage-fed cattle in Australia. *Animal Production Science*, **56**, 169–180.

COCKRUM, R.R., STOBART, R.H., LAKE, S.L. & CAMMACK, K.M. (2013). Phenotypic variation in residual feed intake and performance traits in rams. *Small Ruminant Research*, **113**, 313–322.

DILLON, E., MORAN, B. & DONNELLAN, T. (2021). *Teagasc National Farm Survey 2020: Preliminary Results*.

DOMINIK, S., ROBINSON, D.L., DONALDSON, A.J., CAMERON, M., AUSTIN, K.L. & ODDY, V.H. (2017). Relationship between feed intake, energy expenditure and methane emissions: implications for genetic evaluation. *Proc. Assoc. Advmt. Anim. Breed. Genet.*, **22**, 65–68.

DONOGHUE, K.A., BIRD-GARDINER, T.L., ARTHUR, P.F., HERD, R.M. & HEGARTY, R.F. (2015). Genetic parameters for methane production and relationships with production traits in Australian beef cattle. *Proc. Assoc. Advmt. Breed. Genet.*, **21**, 114–117.

FOGARTY, N.M., SAFARI, E., MORTIMER, S.I., GREEFF, J.C. & HATCHER, S. (2009). Heritability of feed intake in grazing Merino ewes and the genetic relationships with production traits. *Animal Production Science*, **49**, 1080–1085.

FOGARTY, E.S., SWAIN, D.L., CRONIN, G.M., MORAES, L.E. & TROTTER, M. (2020). Behaviour classification of extensively grazed sheep using machine learning. *Computers and Electronics in Agriculture*, **169**, 105175.

GIORELLO, D., DE BARBIERI, I., AGUERRE, J.I., BANCHERO, G., ROVIRA, F., RODRIGUEZ, B. & CIAPPESONI, G. (2021). Feedlot forrajero: Confinamiento estratégico de corderos con forraje conservado. *Revista INIA Uruguay*, **68**, 12–16.

GOOPY, J.P., ROBINSON, D.L., WOODGATE, R.T., DONALDSON, A.J., ODDY, V.H., VERCOE, P.E. &

HEGARTY, R.S. (2016). Estimates of repeatability and heritability of methane production in sheep using portable accumulation chambers. *Animal Production Science*, **56**, 116–122.

GOOPY, J.P., WOODGATE, R., DONALDSON, A., ROBINSON, D.L. & HEGARTY, R.S. (2011). Validation of a short-term methane measurement using portable static chambers to estimate daily methane production in sheep. *Animal Feed Science and Technology*, **166–167**, 219–226.

HEGARTY, R.S. (2013). Applicability of short-term emission measurements for on-farm quantification of enteric methane. *Animal*, **7**, 401–408.

JOHNSON, P., MILLER, S. & KNOWLER, K. (2015). Towards a data set to investigate feed efficiency in new zealand maternal sheep. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics*, 106–109.

JOHNSON, P.L., MILLER, S.P. & KNOWLER, K. (2016). Preliminary investigations into the trait of residual energy intake in sheep. *NZ Soc Anim Prod Proc*, **76**, 34–37.

JOHNSON, P.L., WING, J. & KNOWLER, K. (2017). Relationship between measures of residual energy intake made on growing animals and adults. *NZ Soc Anim Prod Proc*, **77**, 85–87.

JONKER, A., HICKEY, S.M., ROWE, S.J., JANSSEN, P.H., SHACKELL, G.H., ELMES, S., BAIN, W.E., WING, J., GREER, G.J., BRYSON, B., MACLEAN, S., DODDS, K.G., PINARES-PATIÑO, C.S., YOUNG, E.A., KNOWLER, K., PICKERING, N.K. & MCEWAN, J.C. (2018). Genetic parameters of methane emissions determined using portable accumulation chambers in lambs and ewes grazing pasture and genetic correlations with emissions determined in respiration chambers. *Journal of Animal Science*, **96**, 3031–3042.

KAMPHUIS, C., VAN RIEL, J.W., VEERKAMP, R.F. & DE MOL, R.M. (2017). Traditional mixed linear modelling versus modern machine learning to estimate cow individual feed intake. In *Precision Livestock Farming '17* (Eds D. Berckmans & A. Keita), pp. 366–376. Nates: 8th European Conference on Precision Livestock Farming.

KNOTT, S.A., CUMMINS, L.J., DUNSHEA, F.R. & LEURY, B.J. (2008). The use of different models for the estimation of residual feed intake (RFI) as a measure of feed efficiency in meat sheep. *Animal Feed Science and Technology*, **143**, 242–255.

KOCH, R.M., SWIGER, L.A., CHAMBERS, D. & GREGORY, K.E. (1963). Efficiency of feed use in beef cattle. *Journal of Animal Science*, **22**, 486–494.

KOCHHAR, R; OATES, R. (2014). *Attitudes about Aging: A Global Perspective*.

LEAHY, S., CLARK, H. & REISINGER, A. (2020). Challenges and Prospects for Agricultural Greenhouse Gas Mitigation Pathways Consistent With the Paris Agreement. *Frontiers in Sustainable Food Systems*, **4**, 1–8.

LESO, L., WERNER, J., MCSWEENEY, D., KENNEDY, E., GEOGHEGAN, A. & SHALLOO, L. (2019). Random forest regression for estimating dry matter intake of grazing dairy cows. In *Precision Livestock Farming '19* (Eds B. O'Brien, ShallooThe, D. Hennessy, & L. Shalloo), pp. 606–612. Cork: The 9th European Conference on Precision Livestock Farming.

MCLEOD, A. (2011). *World Livestock 2011*. Rome: Food and Agriculture Organization of the United Nations (FAO).

MONTELLI, N.L.L., ALMEIDA, A.K. DE, RIBEIRO, C.R. DE F., GROBE, M.D., ABRANTES, M.A.F., LEMOS, G.S., FURUSHO GARCIA, I.F. & PEREIRA, I.G. (2019). Performance, feeding behavior and digestibility of nutrients in lambs with divergent efficiency traits. *Small Ruminant Research*, **180**, 50–56.

ODEVCI, B.B. & EMSEN, E. (2019). Machine learning model for maternal quality in sheep. In *Precision Livestock Farming '19* (Eds B. O'Brien, ShallooThe, D. Hennessy, & L. Shalloo), pp. 69–73. Cork: The 9th European Conference on Precision Livestock Farming.

ODEVCI, B.B., EMSEN, E. & AYDIN, M.N. (2021). Machine learning algorithms for lamb survival. *Computers and Electronics in Agriculture*, **182**, 105995.

OLIJHOEK, D.W., LØVENDAHL, P., LASSEN, J., HELLWING, A.L.F., HÖGLUND, J.K., WEISBJERG, M.R. & NOEL, S.J. (2018). Methane production , rumen fermentation , and diet digestibility of Holstein and Jersey dairy cows being divergent in residual feed intake and fed at 2 forage-to-concentrate ratios. *Journal of Dairy Science*, **101**, 9926–9940.

OLSEN, L.R. (2021). Cross-Validation for Model Selection - CRAN. *R CORE TEAM*, 100.

PAGANONI, B., ROSE, G., MACLEAY, C., JONES, C., BROWN, D.J., KEARNEY, G., FERGUSON, M. & THOMPSON, A.N. (2017). More feed efficient sheep produce less methane and carbon dioxide when eating high-quality pellets. *Journal of Animal Science*, **95**, 3839.

PILES, M., BERGSMA, R., GIANOLA, D., GILBERT, H. & PILES, M. (2021). Feature Selection Stability and Accuracy of Prediction Models for Genomic Prediction of Residual Feed Intake in Pigs Using Machine Learning. *Frontiers in Genetics*, **12**, 1–14.

R CORE TEAM. (2021). R: A language and environment for statistical computing.

REDDEN, R.R., SURBER, L.M.M., GROVE, A. V. & KOTT, R.W. (2014). Effects of residual feed intake classification and method of alfalfa processing on ewe intake and growth. *Journal of Animal Science*, **92**, 830–835.

REDDEN, R.R., SURBER, L.M.M., ROEDER, B.L., NICHOLS, B.M., PATERSON, J.A. & KOTT, R.W. (2011). Residual feed efficiency established in a post-weaning growth test may not result in more efficient ewes on the range. *Small Ruminant Research*, **96**, 155–159.

REFAEILZADEH, P., TANG, L. & LIU, H. (2016). Cross-Validation. In *Encyclopedia of Database Systems* (Eds L. Liu & M.T. Özsu), pp. 1–7. New York, NY: Springer New York.

RENAND, G., DECRUYENAERE, V., MAUPETIT, D. & DOZIAS, D. (2019). Methane and Carbon Dioxide Emission of Beef Heifers in Relation with Growth and Feed E ffi ciency. *Animals*, **3**, 1–17.

RHYS, H.I. (2020). *Machine Learning with R, the Tidyverse and Mlr*, 1st ed (Eds M. Michaels, D.

Warren, A. Dragosavljevic, L. Weidert, & T. Taylor). Shelter Island: Manning.

ROBINSON, D.L., CAMERON, M., DONALDSON, A.J., DOMINIK, S. & ODDY, V.H. (2016). One-hour portable chamber methane measurements are repeatable and provide useful information on feed intake and efficiency. *Journal of Animal Science*, **94**, 4376–4387.

ROBINSON, D.L., DOMINIK, S., DONALDSON, A.J. & ODDY, V.H. (2020). Repeatabilities , heritabilities and correlations of methane and feed intake of sheep in respiration and portable chambers. *Animal Production Science,* **60**, 880–892.

ROBINSON, D.L., GOOPY, J.P., HEGARTY, R.S., ODDY, V.H., THOMPSON, A.N., TOOVEY, A.F., MACLEAY, C.A., BRIEGAL, J.R., WOODGATE, R.T., DONALDSON, A.J. & VERCOE, P.E. (2014). Genetic and environmental variation in methane emissions of sheep at pasture. *Journal of Animal Science*, **92**, 4349–4363.

ROBINSON, D.L. & ODDY, V.H. (2016). Benefits of including methane measurements in selection strategies. *Journal of Animal Science,* **94**, 3624–3635.

RUUSKA, S., HÄMÄLÄINEN, W., KAJAVA, S., MUGHAL, M. & MATILAINEN, P. (2018). Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. *Behavioural Processes*, **148**, 56–62.

SAFARI, E., FOGARTY, N.M., MORTIMER, S.I., GREEFF, J.C., HATCHER, S., LEE, G.J. & GILMOUR, A.R. (2007). Feed intake and its genetic relationship with growth traits in merino sheep. *Proc. Assoc. Advmt. Anim. Breed. Genet.*, **17**, 199–202.

SEMAKULA, J., CORNER-THOMAS, R.A., MORRIS, S.T., BLAIR, H.T. & KENYON, P.R. (2021). Application of machine learning algorithms to predict body condition score from liveweight records of mature romney ewes. *Agriculture (Switzerland),* **11**, 1–22.

SHAHINFAR, S. & KAHN, L. (2018). Machine learning approaches for early prediction of adult wool growth and quality in Australian Merino sheep. *Computers and Electronics in Agriculture*,

148, 72–81.

TORTEREAU, F., MARIE-ETANCELIN, C., WEISBECKER, J.-L., MARCON, D., BOUVIER, F., MORENO-ROMIEUX, C. & FRANÇOIS, D. (2020). Genetic parameters for feed efficiency in Romane rams and responses to single-generation selection. *Animal*, **14**, 681–687.

WICKHAM, H., CHANG, W., HENRY, L., PEDERSEN, T.L., TAKAHASHI, K., WILKE, C., WOO, K., YUTANI, H. & DUNNINGTON, D. (2020). Create Elegant Data Visualisations Using the Grammar of Graphics Description - CRAN. *R CORE TEAM*, 277.

ZHANG, X., WANG, W., MO, F., LA, Y., LI, C. & LI, F. (2017). Association of residual feed intake with growth and slaughtering performance, blood metabolism, and body composition in growing lambs. *Scientific Reports*, **7**, 1–11.

Table 1. General data description presented as the mean±sd for all traits

Traits[1]	Australian Merino	Corriedale	Dohne Merino
FI (kg DM/day)	1.40±0.293	1.28±0.256	1.65±0.196
DA (years)	3.27±2.100	3.12±2.070	2.76±1.75
WA (days)	137.00±12.300	111.70±7.540	133.30±27.000
TSA (days)	297.00±47.800	188.60±12.700	420.50±8.46,
REA (cm^2)	7.38±1.560	6.29±1.100	9.76±2.140
FT (mm)	2.04±0.733	2.04±0.616	2.76±0.899
BWmT (kg)	40.55±6.290	32.98±4.290	49.76±5.770
MBW (kg)	16.03±1.850	13.74±1.340	18.71±1.64,
BWsT (kg)	36.43±5.960	29.93±4.120	45.28±5.430
BWeT (kg)	43.95±6.860	36.41±6.420	52.92±4.770
CH_4 (g/day)	23.23±5.470	15.25±4.800	27.97±5.970
CO_2 (g/day)	1039.00±228.000	753.20±134.000	1450.60±356.000
O_2 (g/day)	951.40±179.000	782.20±137.000	1292.60±336.000
BWdif (kg)	7.51±2.720	6.47±2.230	7.63±2.59
BT (1 or 2)	1.58±0.539	1.52±0.540	1.63±0.545
YGFW (kg)	4.13±0.696	3.23±0.443	2.63±0.39,
YWD (microns)	14.87±0.918	23.37±1.640	18.28±1.350
YBW (kg)	49.21±11.500	35.90±4.570	42.57±5.190
YSL (cm)	10.92±1.320	12.21±1.580	10.04±1.330
YSA (days)	411.90±8.700	398.10±7.340	396.30±6.680

[1] Average feed intake (FI, kg DM/day); dam age at lambing (DA, years); weaning age (WA, days); test star age (TSA, days); rib eye area in the end of the trial (REA, cm^2); fat thickness in the end of

the trial (FT, mm); body weight in the middle of the trial (BWmT, kg); metabolic body weight in the middle of the trail (MBW = $BWmT^{0.75}$, kg); body weight in the start of the trial (BWsT, kg); body weight in the end of the trial (BWeT, kg); average of CH_4 emission (CH_4, g/day); average of CO_2 emission (CO_2, g/day); average of O_2 consumption (O_2, g/day); body weight difference between the start and end of the trial (BWdif, kg); birth type (BT, 1 or 2); yearling greasy fleece weight (YGFW, kg); yearling wool diameter (YFD, microns); yearling body weight (YBW, kg); yearling staple length (YSL, cm) and yearling shearing age (YSA, days).

Table 2. Selected variables with feature selection using the methods linear regression model (Linear Model), nonlinear regression model (Nonlinear Model), k-nearest neighbor regression (k-Nearest Neighbor Model) and random forest regression (Random Forest Model) with sequential floating backward search control method

Methods[1]	Selected variables[2]	Mean squared error
Linear Model	MBW, CO_2, BWdif	0.021
Nonlinear Model	MBW, CO_2, BWdif	0.021
k-Nearest Neighbor Model	BWmT, CO_2, BWdif	0.024
Random Forest Model	MBW, CO_2, BWdif	0.022

[1] Linear Model: linear regression uses the modelling relationship based in the straight-line equation; Nonlinear Model: nonlinear regression uses the generalized additive models; k-Nearest Neighbor Model: k-nearest neighbor regression uses the k-nearest neighbors' algorithm for regression; Random Forest Model: random forest regression uses tree-based algorithms for regression.

[2] Body weight in the middle of the trial (BWmT, kg); metabolic body weight in the middle of the trail ($MBW = BWmT^{0.75}$, kg); average of CO_2 emission (CO_2, g/day) and body weight difference between the start and end of the trial (BWdif, kg).

Table 3. Performance metrics: R^2, root mean squared error (*RMSE*), Pearson and Spearman correlations coefficients (*r*); for feed intake predict algorithms cross-validation (observed x predicted feed intake) using the nine trials of Australian Merino dataset, one of Corriedale dataset and one of Dohne Merino dataset (n=640)

Method[1]	Algorithm	R^2	*RMSE*	*r*-Pearson	*r*-Spearman
Linear Model	regr.lm	0.75	0.146	0.87	0.85
Nonlinear Model	regr.gamboost	0.76	0.145	0.87	0.85
k-Nearest Neighbor Model	regr.kknn	0.75	0.140	0.87	0.86
Random Forest Model	regr.randomForest	0.75	0.148	0.86	0.84

[1] Linear Model: linear regression uses the modelling relationship based in the straight-line equation; Nonlinear Model: nonlinear regression uses the generalized additive models; k-Nearest Neighbor Model: k-nearest neighbor regression uses the k-nearest neighbors' algorithm for regression; Random Forest Model: random forest regression uses tree-based algorithms for regression.

Table 4. Pearson and Spearman correlations coefficients for feed intake predict algorithms for Australian Merino (AM), Corriedale (Cor) and Dohne Merino (DM) dataset

Method[1]	Algorithm	Dataset	Pearson[2]	Spearman[2]
Linear Model	regr.lm	AM	0.76	0.77
Linear Model	regr.lm	Cor	0.74	0.69
Linear Model	regr.lm	DM	0.78	0.76
Nonlinear Model	regr.gamboost	AM	0.86	0.86
Nonlinear Model	regr.gamboost	Cor	0.74	0.69
Nonlinear Model	regr.gamboost	DM	0.79	0.77
k-Nearest Neighbor Model	regr.kknn	AM	0.81	0.80
k-Nearest Neighbor Model	regr.kknn	Cor	0.69	0.63
k-Nearest Neighbor Model	regr.kknn	DM	0.73	0.73
Random Forest Model	regr.randomForest	AM	0.76	0.78
Random Forest Model	regr.randomForest	Cor	0.70	0.66
Random Forest Model	regr.randomForest	DM	0.79	0.77

[1] Linear Model: linear regression uses the modelling relationship based in the straight-line equation; Nonlinear Model: nonlinear regression uses the generalized additive models; k-Nearest Neighbor Model: k-nearest neighbor regression uses the k-nearest neighbors' algorithm for regression; Random Forest Model: random forest regression uses tree-based algorithms for regression.

[2] Cor: Corriedale dataset and DM: Dohne Merino dataset.

[3] $P < 0.001$.

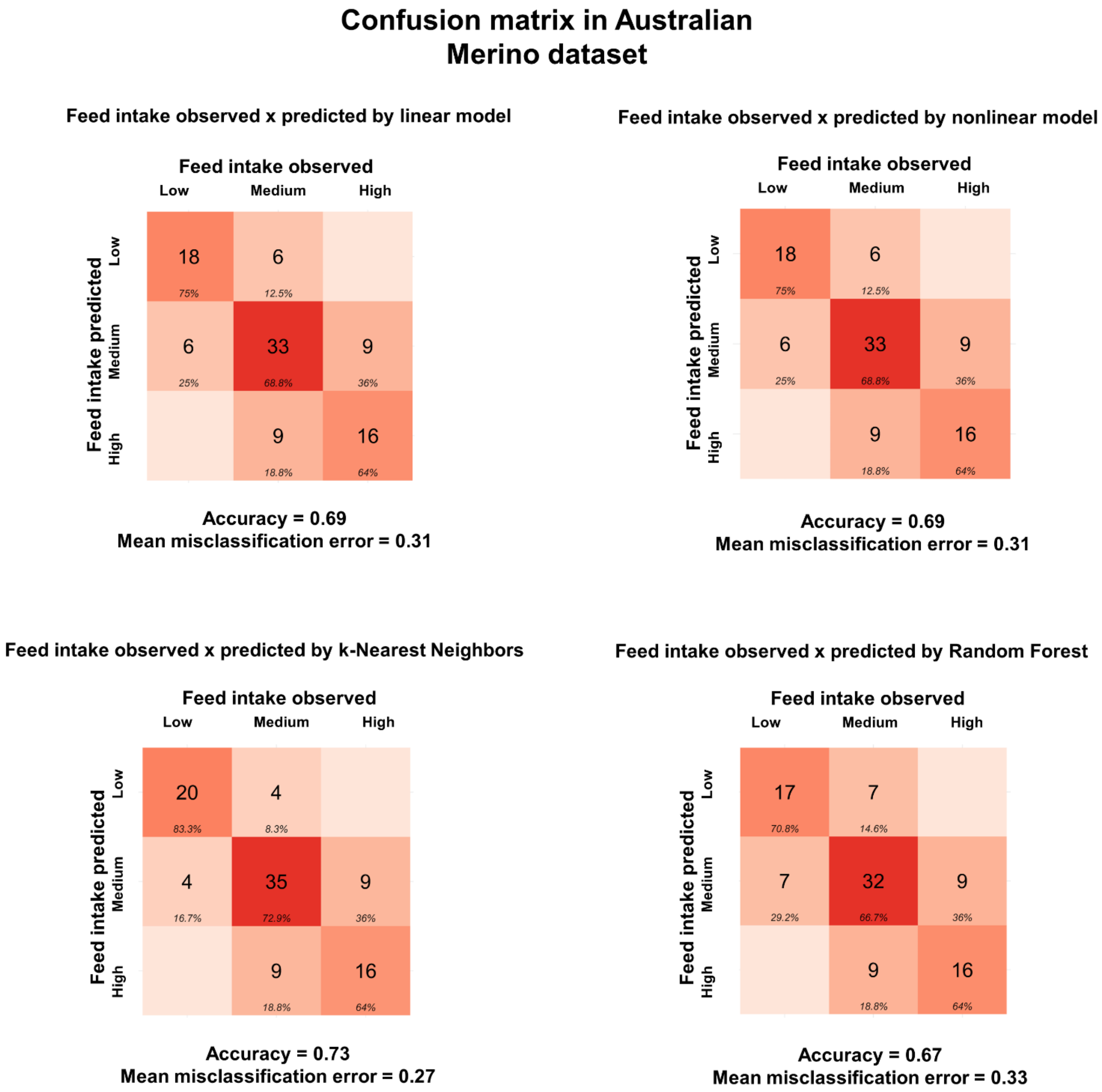
Table 5. Pearson and Spearman correlations coefficients between feed intake expected for Koch model and observed and predict feed intake for Australian Merino (AM), Corriedale (Cor) and Dohne Merino (DM) dataset

Feed intake	Dataset	Pearson[2]	Spearman[2]
Observed	AM	0.75	0.76
Observed	Cor	0.78	0.76
Observed	DM	0.77	0.75
Predicted by Linear Model	AM	0.90	0.91
Predicted by Linear Model	Cor	0.94	0.93
Predicted by Linear Model	DM	0.91	0.90
Predicted by Nonlinear Model	AM	0.86	0.88
Predicted by Nonlinear Model	Cor	0.94	0.93
Predicted by Nonlinear Model	DM	0.92	0.92
Predicted by k-Nearest Neighbor Model	AM	0.86	0.86
Predicted by k-Nearest Neighbor Model	Cor	0.89	0.89
Predicted by k-Nearest Neighbor Model	DM	0.85	0.89
Predicted by Random Forest Model	AM	0.89	0.91
Predicted by Random Forest Model	Cor	0.90	0.89
Predicted by Random Forest Model	DM	0.92	0.92

[1] Linear Model: linear regression uses the modelling relationship based in the straight-line equation; Nonlinear Model: nonlinear regression uses the generalized additive models; k-Nearest Neighbor Model: k-nearest neighbor regression uses the k-nearest neighbors' algorithm for regression; Random Forest Model: random forest regression uses tree-based algorithms for regression.

[2] Cor: Corriedale dataset and DM: Dohne Merino dataset.

[3] $P < 0.001$.


Confusion matrix in Australian Merino dataset

Feed intake observed x predicted by linear model

Accuracy = 0.69
Mean misclassification error = 0.31

Feed intake observed x predicted by nonlinear model

Accuracy = 0.69
Mean misclassification error = 0.31

Feed intake observed x predicted by k-Nearest Neighbors

Accuracy = 0.73
Mean misclassification error = 0.27

Feed intake observed x predicted by Random Forest

Accuracy = 0.67
Mean misclassification error = 0.33

Fig. 1. Confusion matrix between observed and predicted feed intake classifications in Australian Merino dataset (n=97). Linear model: linear regression uses the modelling relationship based in the straight-line equation; nonlinear model: nonlinear regression uses the generalized additive models; k-Nearest Neighbors: k-nearest neighbor regression uses the k-nearest neighbors' algorithm for regression; Random Forest: random forest regression uses tree-based algorithms for

regression.

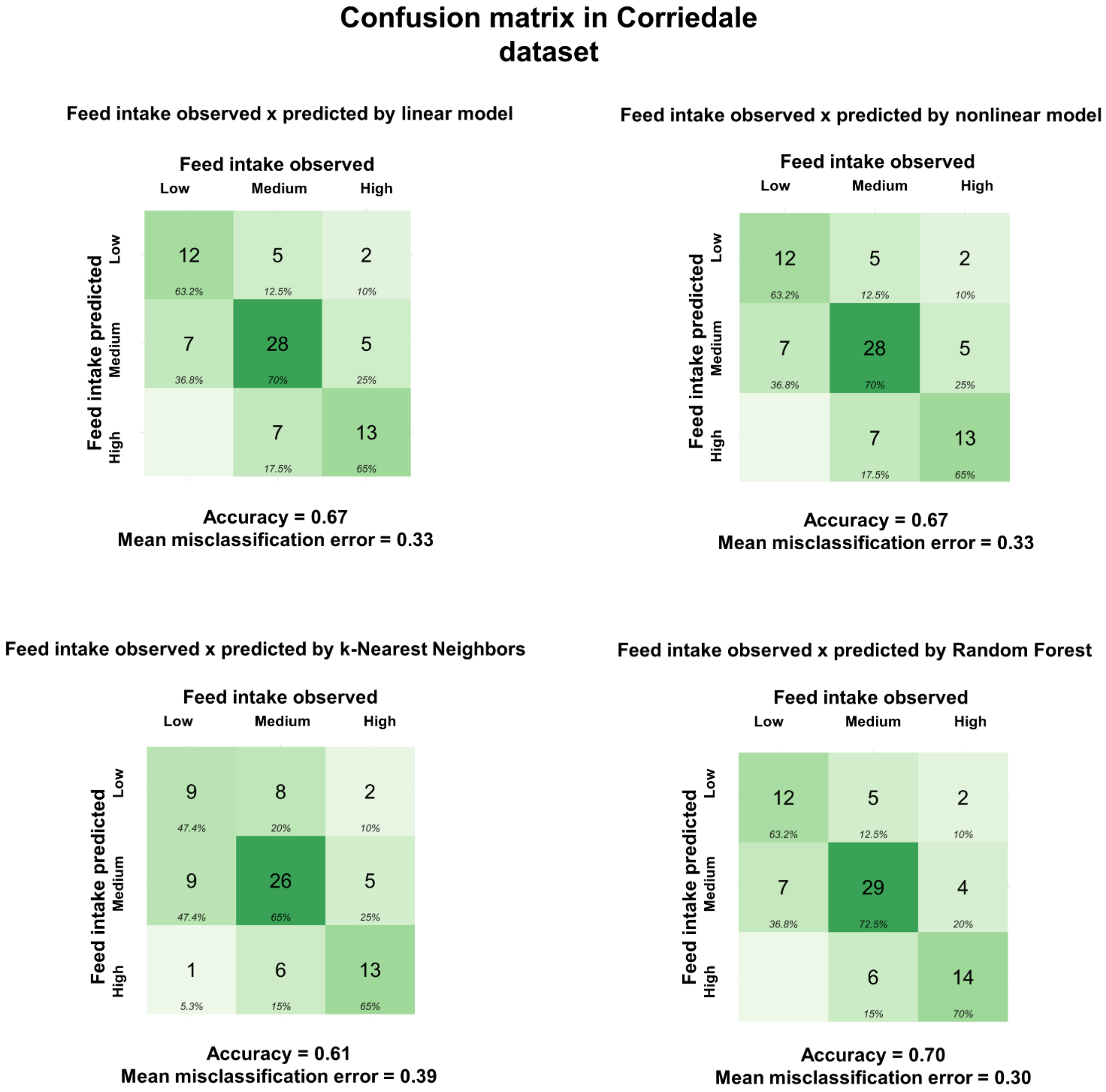
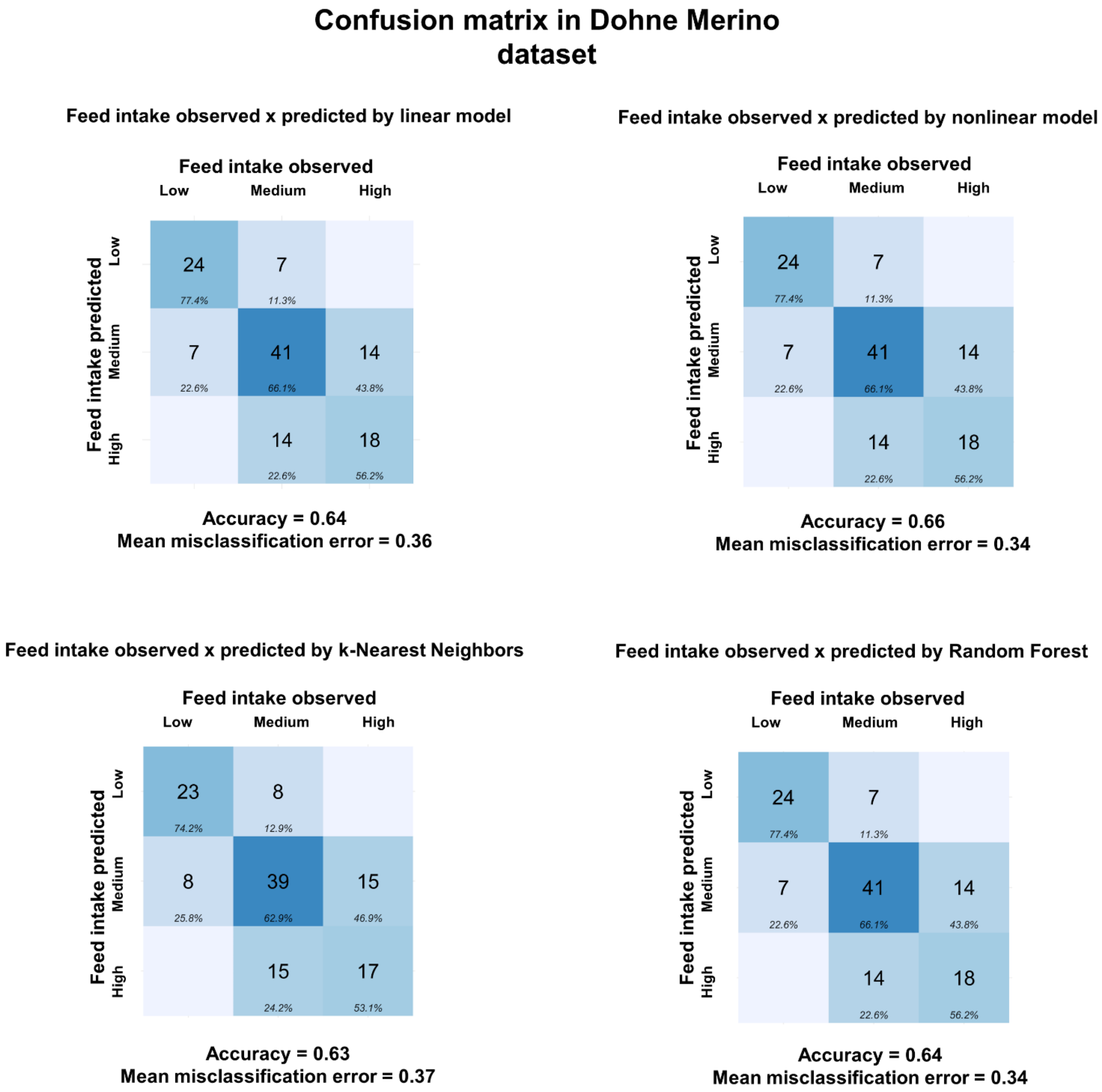


Fig. 2. Confusion matrix between observed and predicted feed intake classifications in Corriedale dataset (n=79). Linear model: linear regression uses the modelling relationship based in the straight-line equation; nonlinear model: nonlinear regression uses the generalized additive models; k-Nearest Neighbors: k-nearest neighbor regression uses the k-nearest neighbors' algorithm for regression; Random Forest: random forest regression uses tree-based algorithms for regression.



Confusion matrix in Dohne Merino dataset

Feed intake observed x predicted by linear model

Feed intake predicted \ Feed intake observed	Low	Medium	High
Low	24 77.4%	7 11.3%	
Medium	7 22.6%	41 66.1%	14 43.8%
High		14 22.6%	18 56.2%

Accuracy = 0.64
Mean misclassification error = 0.36

Feed intake observed x predicted by nonlinear model

Feed intake predicted \ Feed intake observed	Low	Medium	High
Low	24 77.4%	7 11.3%	
Medium	7 22.6%	41 66.1%	14 43.8%
High		14 22.6%	18 56.2%

Accuracy = 0.66
Mean misclassification error = 0.34

Feed intake observed x predicted by k-Nearest Neighbors

Feed intake predicted \ Feed intake observed	Low	Medium	High
Low	23 74.2%	8 12.9%	
Medium	8 25.8%	39 62.9%	15 46.9%
High		15 24.2%	17 53.1%

Accuracy = 0.63
Mean misclassification error = 0.37

Feed intake observed x predicted by Random Forest

Feed intake predicted \ Feed intake observed	Low	Medium	High
Low	24 77.4%	7 11.3%	
Medium	7 22.6%	41 66.1%	14 43.8%
High		14 22.6%	18 56.2%

Accuracy = 0.64
Mean misclassification error = 0.34

Fig. 3. Confusion matrix between observed and predicted feed intake classifications in Dohne Merino dataset (n=125). Linear model: linear regression uses the modelling relationship based in the straight-line equation; nonlinear model: nonlinear regression uses the generalized additive models; k-Nearest Neighbors: k-nearest neighbor regression uses the k-nearest neighbors'

algorithm for regression; Random Forest: random forest regression uses tree-based algorithms for regression.

CAPÍTULO V – CONSIDERAÇÕES FINAIS

6. CONSIDERAÇÕES FINAIS

Tanto a redução do período de avaliação como o aumento de animais com informações de consumo alimentar impactarão na quantidade de animais fenotipados e isso proporcionará uma maior quantidade de animais com avaliação genética, principalmente, para consumo alimentar.

A redução de sete dias no período de testes de consumo alimentar, utilizando alimentadores automáticos, trará grandes impactos, visto que em um ano em que normalmente seis lotes seriam testados nas estruturas em que os dados para os estudos foram coletados, um lote adicional poderia ser incluído com o mesmo custo. Dessa forma, a coleta de dados poderia ser acelerada e a intensidade de seleção aumentada.

Ainda, a aplicação de procedimentos de validação cruzada, aplicados tanto em modelos estatísticos como em modelos baseados em algoritmos que aprendem padrões e inteligentemente tentam predizer o consumo alimentar, poderá também aumentar a quantidade de animais com informações de consumo alimentar. Por exemplo, se são disponibilizadas estruturas de coletas de consumo e pesos diários que proporcionam a coleta de um número específico de animais por período de avaliação, estruturas simples, com comedouros e bebedouros coletivos, em que os animais terão acesso *ad libitum* ao alimento e água, seriam suficientes para obter informações de consumo alimentar. Assim, seria necessário apenas obter informações dos animais de peso inicial (dia 0), peso no meio do período de avaliação (dia 21 ou 17), peso final (dia 42 ou 35) e duas medições de emissão de CO_2 utilizando câmaras de acumulação portáteis. Dessa maneira, será possível ter uma central de avaliação do consumo alimentar e testes dispersos em várias localidades, em que os animais terão acesso ao mesmo alimento do teste central e terão seus pesos corporais pontualmente acompanhados, com duas medidas de emissão de CO_2, podendo obter o dobro de animais com informações de consumo alimentar em cada período de avaliação.

No futuro, com a entrada de mais dados de provas de consumo alimentar, poderão ser empreendidos estudos que utilizem períodos mais curtos e consumo alimentar preditos a nível de abordagens computacionalmente mais profundas, como aplicação de redes neurais artificiais com entradas não só fenotípicas, mas também a nível genético e genômico, podendo-se utilizar valores genéticos dos progenitores e toda informação genealógica disponível para predizer o consumo alimentar dos animais em teste de forma ainda mais acurada.

REFERÊNCIAS

.

ALLEN, M. S. Drives and limits to feed intake in ruminants. **Animal Production Science**, Melbourne, v. 54, n. 10, p. 1513–1524, 2014.

AMARILHO-SILVEIRA, F. *et al.* Residual feed intake for Australian Merino sheep estimated in less than 42 days of trial. **Livestock Science**, Amsterdam, v. 258, [art.] 104889, [p. 1-12], 2022.

ANDRIES, J. I. *et al.* Isoacids in ruminant nutrition: their role in ruminal and intermediary metabolism and possible influences on performances - A review. **Animal Feed Science and Technology**, Amsterdam, v. 18, n. 3, p. 169–180, 1987.

ARCHER, J. A. *et al.* Optimum postweaning test for measurement of growth rate, feed intake, and feed efficiency in British Breed Cattle. **Journal of Animal Science**, Champaign, v. 75, n. 8, p. 2024–2032, 1997.

ARCHER, J. A.; BERGH, L. Duration of performance tests for growth rate, feed intake and feed efficiency in four biological types of beef cattle. **Livestock Production Science**, Amsterdam, v. 65, n. 1/2, p. 47–55, 2000.

ARTHUR, P. F. *et al.* Relationships among carbon dioxide, feed intake, and feed efficiency traits in ad. **Journal of Animal Science**, Champaign, v. 96, p. 4859–4867, 2018.

BASARAB, J. A. *et al.* Reducing GHG emissions through genetic improvement for feed efficiency: effects on economically important traits and enteric methane production. **Animal**, Cambridge, v. 7, p. 303–315, 2013. Suppl. 2.

BASARAB, J. A. *et al.* Residual feed intake and body composition in young growing cattle. **Canadian Journal of Animal Science**, Ottawa, v. 83, n. 2, p. 189–204, 2003.

BEEF + LAMB NEW ZEALAND. **Compendium of New Zealand farm facts 2021**. Wellington: Beef+Lamb, 2021.

BISCHL, B. *et al.* **Machine learning in R-CRAN**. Vienna: CRAN, 2021.

BORCHERS, D. L. *et al.* Improving the precision of the daily egg production method using generalized additive models. **Canadian Journal of Fisheries and Aquatic Sciences**, Ottawa, v. 54, p. 2727–2742, 1997.

CAMERON, N. D. Genetic relationships between growth and food intake in performance tested ram lambs: an investigation of variance component estimation procedures.

Journal of Animal Breeding and Genetics, Berlin, v. 105, n. 1/6, p. 431–440, 1988.

CAMMACK, K. M. *et al.* Estimates of genetic parameters for feed intake, feeding behavior, and daily gain in composite ram lambs. **Journal of Animal Science**, Champaign, v. 83, n. 4, p. 777–785, 2005.

CAMMACK, K. M. *et al.* Rumen microbial taxa and by-products associated with feed efficiency in sheep. *In*: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 10., 2014, Vancouver. **Proceedings of the** [...]. [*S. l.*]: WCGALP, 2014. p. 14–16.

CANTALAPIEDRA-HIJAR, G. *et al.* Review: Biological determinants of between-animal variation in feed efficiency of growing beef cattle. **Animal**, Cambridge, v. 12, n. s2, p. s321-s335, 2018.

CARBERRY, C. A. *et al.* Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. **Applied and Environmental Microbiology**, Michigan, v. 78, n. 14, p. 4949–4958, 2012.

CARNEIRO, M. M. Y. *et al.* Residual intake and gain for the evaluation of performance, non-carcass components, and carcass characteristics of confined crossbred Texel lambs. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 48, [art.] e20180206, [p. 1-13], 2019.

CASTILHOS, A. M. *et al.* Duração do período de avaliação pós-desmame para medidas de desempenho, consumo e eficiência alimentar em bovinos da raça Nelore. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 40, n. 2, p. 301–307, 2011.

CHANDRASHEKAR, G.; SAHIN, F. A survey on feature selection methods. **Computers and Electrical Engineering**, New Jersey, v. 40, n. 1, p. 16–28, 2014.

CHARMLEY, E. *et al.* A universal equation to predict methane production of forage-fed cattle in Australia. **Animal Production Science**, Melbourne, v. 56, n. 3, p. 169–180, 2016.

COCKRUM, R. R. *et al.* Phenotypic variation in residual feed intake and performance traits in rams. **Small Ruminant Research**, Amsterdam, v. 113, n. 2/3, p. 313–322, 2013.

COLDITZ, I. G. Six costs of immunity to gastrointestinal nematode infections. **Parasite Immunology**, Edinburgh, v. 30, n. 2, p. 63–70, 2008.

COYLE, S. *et al.* Repeatability of feed efficiency in beef cattle offered grass silage and zero-grazed grass. **Journal of Animal Science**, Champaign, v. 94, p. 719, 2016. Suppl. 5.

COTTLE, D. J.; NOLAN, J. V.; WIEDEMANN, S. G. Ruminant enteric methane mitigation: a review. **Animal Production Science**, Melbourne, v. 51, n. 6, p. 491-514, 2011.

DADO, R. G.; ALLEN, M. S. Intake limitations, feeding behavior, and rumen function of cows challenged with rumen fill from dietary fiber or inert bulk. **Journal of Dairy Science**, Champaign, v. 78, n. 1, p. 118–133, 1995.

DILLON, E.; MORAN, B.; DONNELLAN, T. **Teagasc National farm survey 2020:** preliminary results. Carlow: Teagasc, 2021.

DOMINIK, S. *et al.* Relationship between feed intake, energy expenditure and methane emissions: implications for genetic evaluation. *In*: CONFERENCE OF THE ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS - AAABG, 22., 2017, Townsville, Queensland, Australia. **Proceedings of the** [...]. Perth: Association for the Advancement of Animal Breeding and Genetics, 2017. p. 65-68.

DONOGHUE, K. A. *et al.* Genetic parameters for methane production and relationships with production traits in Australian beef cattle. *In*: CONFERENCE OF THE ASSOCIATION FOR THE ADVANCEMENT OF ANIMAL BREEDING AND GENETICS – AAABG, 21., 2015, Lorne. **Proceedings of the** [...]. Perth: Association for the Advancement of Animal Breeding and Genetics, 2015. p. 114–117.

DURUNNA, O. N. *et al.* Feed efficiency differences and reranking in beef steers fed grower and finisher diets. **Journal of Animal Science**, Champaign, v. 89, n. 1, p. 158–167, 2011.

ELLISON, M. J. *et al.* Diet and feed efficiency status affect rumen microbial profiles of sheep. **Small Ruminant Research**, Amsterdam, v. 156, p. 12–19, 2017.

ELLISON, M. J. *et al.* Predicting residual feed intake status using rumen microbial profiles in ewe lambs. **Journal of Animal Science**, Champaign, v. 97, p. 2844–2854, 2019.

ERMIAS, E.; YAMI, A.; REGE, J. E. O. Genetic parameters of feed intake and digestive efficiency in tropical sheep. *In*: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings of the** [...]. Perth: Association for the Advancement of Animal Breeding and Genetics, 2002. p. 19–23.

FITZSIMONS, C. *et al.* Methane emissions, body composition, and rumen fermentation traits of beef heifers differing in residual feed intake. **Journal of Animal Science**, Champaign, v. 91, n. 12, p. 5789–5800, 2013.

FOGARTY, E. S. *et al.* Behaviour classification of extensively grazed sheep using machine learning. **Computers and Electronics in Agriculture**, Washington, DC, v. 169, [art.] 105175, [p. 1-10], 2020.

FOGARTY, N. M. *et al.* Heritability of feed intake in grazing Merino ewes and the genetic relationships with production traits. **Animal Production Science**, Melbourne, v. 49, n. 12, p. 1080–1085, 2009.

FRANÇOIS, D. *et al.* Genetic parameters of feeding traits in INRA401 sheep estimated under two different diets. *In*: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 8., 2006, Belo Horizonte, Minas Gerais, Brazil. **Proceedings of the** [...]. Perth: Association for the Advancement of Animal Breeding and Genetics, 2006. p. 13–18.

FRANÇOIS, D. *et al.* Genetic parameters of feeding traits on meat sheep. *In*: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings of the** [...]. Perth: Association for the Advancement of Animal Breeding and Genetics, 2002. p. 19–23.

FRANÇOIS, D. *et al.* Use in selection of the measurements of feed intake and feeding behaviour parameters in sheep. **Options Méditerranéennes: Series A**, Paris, v. 447, n. 74, p. 441–447, 2007.

GIORELLO, D. *et al.* Feedlot forrajero: confinamiento estratégico de corderos con forraje conservado. **Revista INIA Uruguay**, Montevideo, v. 68, p. 12–16, 2021.

GOODALE, S. L. **The principles of breeding**. Boston: A. Williams, 1861.

GOONEWARDENE, L. A. *et al.* Residual metabolizable energy intake and its association with diet and test duration. **Canadian Journal of Animal Science**, Ottawa, v. 84, n. 2, p. 291–295, 2004.

GOOPY, J. P. *et al.* Estimates of repeatability and heritability of methane production in sheep using portable accumulation chambers. **Animal Production Science**, Melbourne, v. 56, p. 116–122, 2016.

GOOPY, J. P. *et al.* Validation of a short-term methane measurement using portable static chambers to estimate daily methane production in sheep. **Animal Feed Science and Technology**, Amsterdam, v. 166/167, p. 219–226, 2011.

GUAN, L. L. *et al.* Linkage of microbial ecology to phenotype: correlation of rumen microbial ecology to cattle’s feed efficiency. **FEMS Microbiology Letters**, Oxford, v. 288, n. 1, p. 85–91, 2008.

HEGARTY, R. S. Applicability of short-term emission measurements for on-farm quantification of enteric methane. **Animal**, Cambrigde, v. 7, n. s2, p. 401–408, 2013.

HERD, R. M.; ARTHUR, P. F. Physiological basis for residual feed intake. **Journal of Animal Science**, Champaign, v. 87, p. 64–71, 2009. Suppl. 14.

JOHNSON, P. L. *et al.* Modelling liveweight change to inform residual feed intake models in growing lambs. *In*: NEW ZEALAND SOCIETY OF ANIMAL PRODUCTION, 75., 2015, Dunedin. **Proceedings of the** [...]. Wellington: NZSAP, 2015. p. 225–227.

JOHNSON, P. L.; MILLER, S. P.; KNOWLER, K. Preliminary investigations into the trait of residual energy intake in sheep. *In*: NEW ZEALAND SOCIETY OF ANIMAL PRODUCTION, 76., 2016, Adelaide. **Proceedings of the** [...]. Wellington: NZSAP, 2016. p. 34–37.

JOHNSON, P. L.; WING, J.; KNOWLER, K. Relationship between measures of residual energy intake made on growing animals and adults. *In*: NEW ZEALAND SOCIETY OF ANIMAL PRODUCTION, 77., 2017, Rotorua. **Proceedings of the** [...]. Wellington: NZSAP, 2017. p. 85–87.

JOHNSON, P.; MILLER, S.; KNOWLER, K. Towards a data set to investigate feed efficiency in new zealand maternal sheep. **Proceedings of the Association for the Advancement of Animal Breeding and Genetics**, Lorne, v. 21, p. 106–109, 2015.

JONKER, A. *et al.* Genetic parameters of methane emissions determined using portable accumulation chambers in lambs and ewes grazing pasture and genetic correlations with emissions determined in respiration chambers. **Journal of Animal Science**, Champaign, v. 96, n. 8, p. 3031–3042, 2018.

KAMPHUIS, C. *et al.* Traditional mixed linear modelling versus modern machine learning to estimate cow individual feed intake. *In*: BERCKMANS, D.; KEITA, A. (org.). **Precision livestock farming'17**: papers presented at the 8th European Conference on Precision Livestock Farming. Nantes: [*s. n.*], 2017. p. 366–376.

KNOTT, S. A. *et al.* The use of different models for the estimation of residual feed intake (RFI) as a measure of feed efficiency in meat sheep. **Animal Feed Science and Technology**, Amsterdam, v. 143, p. 242–255, 2008.

KOCH, R. M. *et al.* Efficiency of feed use in beef cattle. **Journal of Animal Science**, Champaign, v. 22, n. 2, p. 486–494, 1963.

KOCHHAR, R.; OATES, R. **Attitudes about aging:** a global perspective. Washington, DC: Pew Research Center, 2014.

LEAHY, S.; CLARK, H.; REISINGER, A. Challenges and prospects for agricultural greenhouse gas mitigation pathways consistent with the Paris agreement. **Frontiers in Sustainable Food Systems**, Lausanne, v. 4, [art.] 69, [p. 1–8], 2020.

LESO, L. *et al.* Random forest regression for estimating dry matter intake of grazing dairy cows. *In*: O'BRIEN, B.; HENNESSY, D.; SHALLOO, L. (ed.). **Precision livestock farming '19**: papers presented at the 9th European Conference on Precision Livestock Farming Cork, Ireland
26-29 August '19. Cork: Teagasc, 2019. p. 606–612.

LEYMASTER, K. *et al.* Estimates of genetic parameters for daily gain, feed intake, and behavior traits in ram lambs of a composite population. *In*: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings of the** [...]. Perth: Association for the Advancement of Animal Breeding and Genetics, 2002. p. 1986–1989.

MACLEAY, C. *et al.* Feed intake for sheep can be measured precisely in less than 35 days. *In*: BIENNIAL CONFERENCE, 31., 2016, Adelaide, South Australia. **Proceedings of the** [...]. Orange East: Australian Society of Animal Production, 2016. Brief comm.

MARIE-ETANCELIN, C. *et al.* Detailed genetic analysis of feeding behaviour in Romane lambs and links with residual feed intake. **Journal of Animal Breeding and Genetics**, Berlin, v. 136, n. 3, p. 174–182, 2019.

MCLEOD, A. **World livestock 2011**. Rome: FAO, 2011.

MONTELLI, N. L. L. L. *et al.* Performance, feeding behavior and digestibility of nutrients in lambs with divergent efficiency traits. **Small Ruminant Research**, Amsterdam, v. 180, p. 50–56, 2019.

NAKAZAWA, M. **Functions for medical statistics book with some demographic data - CRAN**. Vienna: CRAN, 2019.

NAVARRETE, C. B.; SOARES, F. C. **Dominance analysis - CRAN**. Vienna: CRAN, 2020.

ODEVCI, B. B.; EMSEN, E. Machine learning model for maternal quality in sheep. *In*: O'BRIEN, B.; HENNESSY, D.; SHALLOO, L. (ed.). **Precision livestock farming '19:** papers presented at the 9th European Conference on Precision Livestock Farming Cork, Ireland 26-29 August '19. Cork: Teagasc, 2019. p. 69–73.

ODEVCI, B. B.; EMSEN, E.; AYDIN, M. N. Machine learning algorithms for lamb survival. **Computers and Electronics in Agriculture**, Washington, DC, v. 182, n. February 2020, [art.] 105995, [p. 1-7], 2021.

OLIJHOEK, D. W. *et al.* Methane production , rumen fermentation, and diet digestibility of Holstein and Jersey dairy cows being divergent in residual feed intake and fed at 2 forage-to-concentrate ratios. **Journal of Dairy Science**, Champaign, v. 101, n. 11, p. 9926–9940, 2018.

OLSEN, L. R. **Cross-validation for model selection - CRAN**. Vienna: CRAN, 2021.

PAGANONI, B. *et al.* More feed efficient sheep produce less methane and carbon dioxide when eating high-quality pellets. **Journal of Animal Science**, Champaign, v. 95, n. 9, p. 3839-3850, 2017.

PAULA, E. F. E. *et al.* Consumo alimentar residual e sua relação com medidas de desempenho e eficiência e características in vivo da carcaça de cordeiros. **Arquivo Brasileiro de Medicina Veterinaria e Zootecnia**, Belo Horizonte, v. 65, n. 2, p. 566–572, 2013.

PILES, M. *et al.* Feature selection stability and accuracy of prediction models for genomic prediction of residual feed intake in pigs using machine learning. **Frontiers in Genetics**, Lausanne, v. 12, [art.] 611506, [p. 1–14], 2021.

R CORE TEAM. **R:** a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2021.

REDDEN, R. R. *et al.* Effects of residual feed intake classification and method of alfalfa processing on ewe intake and growth. **Journal of Animal Science**, Champagign, v. 92, n. 2, p. 830–835, 2014.

REDDEN, R. R. *et al.* Residual feed efficiency established in a post-weaning growth test may not result in more efficient ewes on the range. **Small Ruminant Research**, Amsterdam, v. 96, n. 2/3, p. 155–159, 2011.

REFAEILZADEH, P.; TANG, L.; LIU, H. Cross-validation. *In*: LIU, L.; ÖZSU, M. T. (org.). **Encyclopedia of database systems**. New York: Springer New York, 2016. p. 1–7.

RENAND, G. *et al.* Methane and carbon dioxide emission of beef heifers in relation with growth and feed efficiency. **Animals**, Cambridge, v. 9, n. 12, [art.] 1136, [p. 1–17], 2019.

RHYS, H. I. **Machine learning with R, the tidyverse and mlr**. Shelter, Island: Manning, 2020.

ROBINSON, D. L. *et al.* Repeatabilities, heritabilities and correlations of methane and feed intake of sheep in respiration and portable chambers. **Animal Production Science**, Melbourne, v. 60, n. 7, p. 880–892, 2020.

ROBINSON, D. L. *et al.* One-hour portable chamber methane measurements are repeatable and provide useful information on feed intake and efficiency. **Journal of Animal Science**, Champaign, v. 94, p. 4376–4387, 2016.

ROBINSON, D. L.; ODDY, V. H. Benefits of including methane measurements in selection strategies. **Journal of Animal Science**, Champaign, v. 94, p. 3624–3635, 2016.

ROBINSON, D. L. *et al.* Genetic and environmental variation in methane emissions of sheep at pasture. **Journal of Animal Science**, Champaign, v. 92, n. 10, p. 4349–4363, 2014.

RUUSKA, S. *et al.* Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. **Behavioural Processes**, Utrecht, v. 148, n. January, p. 56–62, 2018.

SAFARI, E. *et al.* Across population genetic parameters for wool, growth, and reproduction traits in Australian Merino sheep. 2. Estimates of heritability and variance components. **Australian Journal of Agricultural Research**, Clayton South, v. 58, n. 2, p. 177–184, 2007a.

SAFARI, E. *et al.* Feed intake and its genetic relationship with growth traits in merino sheep. **Proceedings of the Association for the Advancement of Animal Breeding and Genetics**, Armidale, v. 17, p. 199–202, 2007b.

SCHAEFER, H. *et al.* A 21st-century shift from fossil-fuel to biogenic methane emissions indicated by 13CH4. **Science**, Washington, DC, v. 352, n. 6281, p. 80–84, 2016.

SEMAKULA, J. *et al.* Application of machine learning algorithms to predict body condition score from liveweight records of mature romney ewes. **Agriculture**, Basel, v. 11, n. 2, [art.] 162, [p. 1–22], 2021.

SHAHINFAR, S.; KAHN, L. Machine learning approaches for early prediction of adult wool growth and quality in Australian Merino sheep. **Computers and Electronics in Agriculture**, Washington, DC, v. 148, n. March, p. 72–81, 2018.

SNOWDER, G. D.; VAN VLECK, L. D. Effect of duration of performance test on variance component estimates for lamb growth rate. **Journal of Animal Science**, Champaign, v. 80, n. 8, p. 2078–2084, 2002.

TORAL, P. *et al.* Residual feed intake and rumen bacterial diversity in lactating sheep: a preliminary study of their potential link. *In*: INTERNATIONAL SYPOSIUM ON RUMINANT PHYSIOLOGY, 13., 2019, Leipzig, Germany. **Proceedings of the** [...]. Cambridge: Cambridge University Press; British Society of Animal Science, 2019. [Abstract] P 191, p. 641.

TORTEREAU, F. *et al.* Genetic parameters for feed efficiency in Romane rams and responses to single-generation selection. **Animal**, Cambridge, v. 14, n. 4, p. 681–687, 2020.

WALDRON, D. F. *et al.* Central ram tests in the midwestern United States: III. Relationship between sire’s central test performance and progeny performance. **Journal of Animal Science**, Champaign, v. 68, n. 1, p. 45–53, 1990.

WEBSTER, A. J. F. *et al.* The influence of food intake on portal blood flow and heat production in the digestive tract of sheep. **British Journal of Nutrition**, Cambridge, v. 34, n. 1, p. 125–139, 1975.

WEI, T.; SIMKO, V. **Visualization of a correlation matrix - CRAN**. Vienna: CRAN, 2017.

WICKHAM, H. *et al.* **Create elegant data visualisations using the grammar of**

graphics description - CRAN. Vienna: CRAN, 2020.

ZETOUNI, L. *et al.* Direct multitrait selection realizes the highest genetic response for ratio traits. **Journal of Animal Science**, Champaign, v. 95, n. 5, p. 1921–1925, 2017.

ZHANG, X. *et al.* Association of residual feed intake with growth and slaughtering performance, blood metabolism, and body composition in growing lambs. **Scientific Reports**, London, v. 7, [art.] 12681, [p. 1–11], 2017.

VITA

Fernando Amarilho Silveira nasceu em 11 de maio de 1990, no município de Pelotas, no estado do Rio Grande do Sul. É filho de Fernando Carlos Costa Silveira e Luciane Gonçalves Amarilho. No segundo semestre de 2010, ingressou no curso de Zootecnia da Universidade Federal de Pelotas (UFPel). Em agosto de 2015, obteve o título de Bacharel em Zootecnia. No ano de 2016, iniciou o curso de mestrado no Programa de Pós-Graduação em Zootecnia da Faculdade de Agronomia Eliseu Maciel da UFPel, na área de Melhoramento Genético Animal, sob orientação do prof. Dr. José Laurino Dionello, obtendo em 2018 o título de Mestre em Melhoramento Genético Animal. No mesmo ano, ingressou no curso de Doutorado em Zootecnia pela Universidade Federal do Rio Grande do Sul (UFRGS), na área de Melhoramento Genético Animal, sob orientação do prof. Dr. Jaime Araujo Cobuci. Nesse período, foi desenvolvido o trabalho de tese sobre redução do tempo para estimativa do consumo alimentar residual e predição do consumo alimentar em ovinos. Por fim, submeteu-se à banca de defesa de tese em abril de 2022 pela UFRGS em Porto Alegre, RS.