Animal Feed Science and Technology 312 (2024) 115974

Contents lists available at ScienceDirect

Animal Feed Science and Technology

journal homepage: www.elsevier.com/locate/anifeedsci

# Evaluation of equations for predicting ileal nutrient digestibility and digestible nutrient content of broiler diets based on their gross chemical composition

S. Thiruchchenthuran [a,b], N. Lopez-Villalobos [a], F. Zaefarian [a], M.R. Abdollahi [a], T.J. Wester [a], N.B. Pedersen [c], A.C. Storm [c], A.J. Cowieson [d], P.C.H. Morel [a,*]

[a] School of Agriculture and Environment, Massey University, Private Bag 11 222, Palmerston North 4442, New Zealand
[b] Department of Animal Science, Faculty of Agriculture, University of Jaffna, Kilinochchi 44000, Sri Lanka
[c] Novozymes A/S, Animal Health and Nutrition, Lyngby DK-2800, Denmark
[d] DSM-Firmenich, Animal Nutrition and Health, Wurmisweg 576, Kaiseraugst, Switzerland

A R T I C L E I N F O

*Keywords:*
Bootstrapping
Ileal digestibility
Ileal digestible content
Prediction equation
Stepwise multiple regression
Validation

A B S T R A C T

The coefficient of apparent ileal digestibility (CAID) and ileal digestible contents (IDC) of nutrients of 56 diets using 10 feed ingredients were measured in broilers (21–24 d post-hatch). Diets contained varying inclusion levels of traditional and non-traditional ingredients and differed widely in chemical composition. The chemical composition and *in vivo* digestibility values were used to establish prediction equations for CAID and IDC of nutrients using stepwise multiple regression. The strength and accuracy of the developed equations were evaluated by root mean square error (RMSE), coefficient of determination ($R^2$), adjusted $R^2$ (adj. $R^2$), and Akaikie's Information Criteria (AIC). The bootstrap method was used to validate the choice of variables by stepwise selection method in the original equation based on their frequencies of selection. Selection of variables was validated if the variables that appear in the original stepwise model were selected in more than 30% of the 1000 bootstrap samples. A close agreement between the original equations and bootstrap resampling was observed for CAID of nitrogen (N) and energy and IDC of energy, starch, and calcium (Ca). Additionally, the original data was subjected to another run of stepwise regression analysis using the selected variables by bootstrapping. The initial regression showed that the CAID of N and energy was highly dependent on crude fibre (CF) and energy contents of the diets. The CAID of energy can be predicted ($R^2 = 0.89$ and RMSE $= 0.035$) by CF, gross energy (GE), $CF^2$, and starch-to-CF ratio (starch:CF). Calcium content had a positive influence, while phosphorus (P) content had a negative influence on the prediction of CAID of fat. The main variable to predict CAID and IDC of most nutrients was the dietary CF content. Based on the lowest RMSE and AIC, the best predictors for IDC of N were ash, N, fat, CF, $CF^2$, and starch:CF, while the best predictors for IDC of energy were CF, GE, $CF^2$, and starch:CF. The results of the

*Abbreviations:* AA, amino acids; AIC, Akaike's information criteria; Adj. $R^2$, adjusted $R^2$; CAID, coefficient of apparent ileal digestibility; Ca, calcium; CF, crude fibre; $CF^2$, square value of crude fibre; CM, canola meal; CP, crude protein; DCP, dicalcium phosphate; DDGS, distillers dried grains with solubles; DM, dry matter; EE, ether extract; FFSB, full-fat soybean; Fat:CF, fat-to-crude fibre ratio; GE, gross energy; IDC, ileal digestible content; MB, mean bias; MBM, meat and bone meal; ME, metabolisable energy; N, nitrogen; NDF, neutral detergent fibre; NSP, non-starch polysaccharides; P, phosphorus; PKM, palm kernel meal; RMSE, Root mean square error; $R^2$, coefficient of determination of regression; SBM, soybean meal; Starch:CF, starch-to-crude fibre ratio; $TiO_2$, titanium dioxide; WB, wheat bran.

* Corresponding author.
*E-mail address:* P.C.Morel@massey.ac.nz (P.C.H. Morel).

https://doi.org/10.1016/j.anifeedsci.2024.115974
Received 8 January 2024; Received in revised form 3 April 2024; Accepted 18 April 2024
Available online 19 April 2024

original stepwise regression models and the stepwise regression with the selected variables from the bootstrap results for CAID of N, energy, fat, and DM, as well as IDC of energy, starch, and Ca, were the same with no differences in $R^2$, Adj. $R^2$, RMSE, and AIC. This method can be useful for developing stable and reproducible models using stepwise regression. However, an external validation is needed to confirm the use of these equations in commercial settings.

## 1. Introduction

Advances in genetics, health, nutrition, and management practices have contributed to rapid growth of the poultry sector in recent years. These improvements raise the nutritional demands of birds that now require more complex diets. It is essential to accurately determine the chemical composition, digestible nutrients, and energy content of feedstuffs to formulate nutritionally balanced diets to fulfil birds' requirements (Alvarenga et al., 2015). Knowledge of the digestibility coefficients and requirement of digestible contents enables diet formulations closer to the requirements of the bird.

A rapid, inexpensive, and accurate method for estimation of the nutritive value of feedstuffs is a goal for animal production. Direct *in vivo* methods provide greater accuracy in feed evaluation, utilization of nutrients, and better prediction of bird performance. However, *in vivo* analysis is costly, time-consuming, and laborious (Zaefarian et al., 2021). Table values and prediction equations are used to quickly obtain digestibility values of feed ingredients that are used in feed formulation. However, errors in formulations may occur in using data from tables as these values represent an average of several previous studies in poultry (Mateos et al., 2019; Zaefarian et al., 2021).

In recent years, prediction equations have attained much interest and are used by most industries related to animal feed manufacturing (Mateos et al., 2019). Several researchers (Cerrate et al., 2019; Sheikhhasan et al., 2020a; Pedersen et al., 2021) recently proposed equations to predict energy and nutrient digestibility from chemical composition of feedstuffs.

Alvarenga et al. (2013) stated that to obtain the energy values of feed, it may be more appropriate to use prediction equations considering the chemical composition of the feed rather than performing *in vivo* assays for every raw material and utilising table values with numerous variations. In addition, predictions based on chemical composition were more accurate in terms of reflecting *in vivo* results (Yegani et al., 2013; Sheikhhasan et al., 2020b). However, more research and robust validation are warranted to develop more accurate equations. A proper validation procedure is necessary to assume that a prediction equation is effective.

Usually, mathematical models have good predictive power (coefficient of determination ($R^2$), root mean square error (RMSE)) using the original data set (Castilho et al., 2015). External validation using new data by an independent research team is the best approach to validate a model. However, it can be costly and slow, and disappointing results could often be avoided with rigorous internal validation performed earlier in the process. One such internal validation method is the bootstrap resampling technique (Steyerberg and Harrell, 2016).

Bootstrap resampling was first presented by Efron (1979). It is used mainly for estimation of parameters and their variability in a given model. Split-sample is a popular approach for internal validation in which a dataset is split into training (model development) and test (model validation) by a random process (Harrell, 2015). However, bootstrapping procedures are more useful than split-sample method and produce better results in terms of bias and variability (Steyerberg et al., 2001; Harrell, 2015). It is most appropriate in situations where the sample size is small and external validation data are not readily available (Chowdhury and Turin, 2021). The theory behind bootstrapping is that it replicates the process of sample generation from an underlying population by drawing samples with replacements from the original dataset (Steyerberg et al., 2001). This method provides stable results in terms of less variance than other methods with a large number of repetitions. When the same selection procedure as for the original data is used in an ideal validation study, then (nearly) the same variables should be selected. This is sometimes called 'replication stability' (Sauerbrei, 1999).

Each internal model validation strategy has pros and cons, and no one technique is consistently superior to another. Different researchers have different ideas on what approach is best for internal model validation. Before reaching a decision, a number of criteria need to be considered, including sample size, best indicators of a model's performance, and choice of models (Chowdhury and Turin, 2021).

Recently, bootstrap resampling techniques have been promoted to evaluate the degree of stability of models resulting from stepwise procedures (Nunez et al., 2011). In stepwise regression, after a variable has been added to the model at each step of the variable selection process, it is possible to remove variables from the model. For instance, if the significance of a given predictor is above a specific threshold, it will be eliminated from the model. When a prespecified stopping rule has been satisfied, the iterative process will end (Austin and Tu, 2004a). If applied to regression analysis, bootstrapping provides variables that have a high degree of reliability (Brunelli, 2014).

Methods based on statistical models have been continually proposed for prediction of nutrient digestibility and digestible content of nutrients in feed ingredients. However, only a few published reports are available that use bootstrap resampling for validation of regression models (Castilho et al., 2015; Smith et al., 2015; Oliveira et al., 2019). Therefore, the objective of the current study was to formulate prediction equations to estimate ileal digestibility coefficients and digestible content of nutrients in broiler diets using stepwise multiple regression and application of bootstrap resampling as validation.

## 2. Materials and methods

A digestibility study conducted at Massey University to predict digestible nutrient content of poultry diets published by Pedersen et al. (2021). With the permission of authors, their data were used to develop prediction equations and then validate them using a bootstrapping resampling technique. The original experimental design (Pedersen et al., 2021), stepwise regression, and bootstrapping are described below.

**Table 1**
Major ingredients inclusion levels of 56 diets (g/kg) as per Pedersen et al.(2021).

| Diets | Maize | Wheat | Sorghum | SBM | CM | MBM | WB | DDGS | FFSB | PKM |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 820 |
| 2 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 420 | 420 |
| 3 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 820 | 20 |
| 4 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 420 | 20 | 420 |
| 5 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 420 | 420 | 20 |
| 6 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 820 | 20 | 20 |
| 7 | 20 | 20 | 20 | 20 | 20 | 20 | 420 | 20 | 20 | 420 |
| 8 | 20 | 20 | 20 | 20 | 20 | 20 | 420 | 20 | 420 | 20 |
| 9 | 20 | 20 | 20 | 20 | 20 | 20 | 420 | 420 | 20 | 20 |
| 10 | 20 | 20 | 20 | 20 | 20 | 20 | 820 | 20 | 20 | 20 |
| 11 | 20 | 20 | 20 | 20 | 20 | 420 | 20 | 20 | 20 | 420 |
| 12 | 20 | 20 | 20 | 20 | 20 | 420 | 20 | 20 | 420 | 20 |
| 13 | 20 | 20 | 20 | 20 | 20 | 420 | 20 | 420 | 20 | 20 |
| 14 | 20 | 20 | 20 | 20 | 20 | 420 | 420 | 20 | 20 | 20 |
| 15 | 20 | 20 | 20 | 20 | 20 | 820 | 20 | 20 | 20 | 20 |
| 16 | 20 | 20 | 20 | 20 | 420 | 20 | 20 | 20 | 20 | 420 |
| 17 | 20 | 20 | 20 | 20 | 420 | 20 | 20 | 20 | 420 | 20 |
| 18 | 20 | 20 | 20 | 20 | 420 | 20 | 20 | 420 | 20 | 20 |
| 19 | 20 | 20 | 20 | 20 | 420 | 20 | 420 | 20 | 20 | 20 |
| 20 | 20 | 20 | 20 | 20 | 420 | 420 | 20 | 20 | 20 | 20 |
| 21 | 20 | 20 | 20 | 20 | 820 | 20 | 20 | 20 | 20 | 20 |
| 22 | 20 | 20 | 20 | 420 | 20 | 20 | 20 | 20 | 20 | 420 |
| 23 | 20 | 20 | 20 | 420 | 20 | 20 | 20 | 20 | 420 | 20 |
| 24 | 20 | 20 | 20 | 420 | 20 | 20 | 20 | 420 | 20 | 20 |
| 25 | 20 | 20 | 20 | 420 | 20 | 20 | 420 | 20 | 20 | 20 |
| 26 | 20 | 20 | 20 | 420 | 20 | 420 | 20 | 20 | 20 | 20 |
| 27 | 20 | 20 | 20 | 420 | 420 | 20 | 20 | 20 | 20 | 20 |
| 28 | 20 | 20 | 20 | 820 | 20 | 20 | 20 | 20 | 20 | 20 |
| 29 | 20 | 20 | 420 | 20 | 20 | 20 | 20 | 20 | 20 | 420 |
| 30 | 20 | 20 | 420 | 20 | 20 | 20 | 20 | 20 | 420 | 20 |
| 31 | 20 | 20 | 420 | 20 | 20 | 20 | 20 | 420 | 20 | 20 |
| 32 | 20 | 20 | 420 | 20 | 20 | 20 | 420 | 20 | 20 | 20 |
| 33 | 20 | 20 | 420 | 20 | 20 | 420 | 20 | 20 | 20 | 20 |
| 34 | 20 | 20 | 420 | 20 | 420 | 20 | 20 | 20 | 20 | 20 |
| 35 | 20 | 20 | 420 | 420 | 20 | 20 | 20 | 20 | 20 | 20 |
| 36 | 20 | 20 | 820 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 37 | 20 | 420 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 420 |
| 38 | 20 | 420 | 20 | 20 | 20 | 20 | 20 | 20 | 420 | 20 |
| 39 | 20 | 420 | 20 | 20 | 20 | 20 | 20 | 420 | 20 | 20 |
| 40 | 20 | 420 | 20 | 20 | 20 | 20 | 420 | 20 | 20 | 20 |
| 41 | 20 | 420 | 20 | 20 | 20 | 420 | 20 | 20 | 20 | 20 |
| 42 | 20 | 420 | 20 | 20 | 420 | 20 | 20 | 20 | 20 | 20 |
| 43 | 20 | 420 | 20 | 420 | 20 | 20 | 20 | 20 | 20 | 20 |
| 44 | 20 | 420 | 420 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 45 | 20 | 820 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 46 | 420 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 420 |
| 47 | 420 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 420 | 20 |
| 48 | 420 | 20 | 20 | 20 | 20 | 20 | 20 | 420 | 20 | 20 |
| 49 | 420 | 20 | 20 | 20 | 20 | 20 | 420 | 20 | 20 | 20 |
| 50 | 420 | 20 | 20 | 20 | 20 | 420 | 20 | 20 | 20 | 20 |
| 51 | 420 | 20 | 20 | 20 | 420 | 20 | 20 | 20 | 20 | 20 |
| 52 | 420 | 20 | 20 | 420 | 20 | 20 | 20 | 20 | 20 | 20 |
| 53 | 420 | 20 | 420 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 54 | 420 | 420 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 55 | 820 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| 56 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

CM = canola meal; DDGS = wheat distillers dried grains with solubles; FFSB = full-fat soybeans; MBM = meat and bone meal; PKM = palm kernel meal; SBM = soybean meal; WB = wheat bran.

### 2.1. Dietary treatments

A total of 56 experimental diets was formulated based on 10 feed ingredients including maize, wheat, sorghum, soybean meal (SBM), canola meal (CM), palm kernel meal (PKM), full-fat soybeans (FFSB), meat and bone meal (MBM), wheat bran (WB), and wheat distillers dried grains with solubles (DDGS) (Table 1). A geometrically central diet was formulated by mixing equal proportions of each of the above 10 ingredients at the same level (100 g/kg). The inclusion level of each feed ingredient in different feed mixtures was either 20, 420, or 820 g/kg and all ingredients were included in all the diets. In dietary treatments based on cereal source, either 820 g/kg of one cereal source or 420 g/kg of two cereal sources were used. In dietary treatments based on protein source, either 820 g/kg of one protein source or 420 g/kg of two protein sources were used. By-product-based dietary treatments had 820 g/kg of either MBM, WB, or wheat DDGS or 420 g/kg of two by-product sources. All diets contained 5 g/kg of indigestible marker titanium dioxide ($TiO_2$) to determine apparent ileal nutrient digestibility. All diets were steam-conditioned at 60 °C for 30 s and pelleted through a pellet mill (Model Orbit 15; Richard Sizer Ltd., Kingston-Upon-Hull, UK) capable of manufacturing 180 kg of feed/h and equipped with a die ring with a 3 mm hole and 35 mm thickness) (Pedersen et al., 2021).

### 2.2. Experimental design

A total of 2688, day-old male broiler chicks (Ross 308) were obtained from a commercial hatchery in 2 batches (1344 chickens per batch) and were fed a common starter diet from 1 to 21 d of age. These birds were used in two batches (3 replicates for each dietary treatment in every batch due to limits on housing) with a two-week interval between batches. In each batch on d 21, 1344 birds were allocated to 168 cages (8 chicks per cage). The 56 feed mixtures were then randomly assigned to 3 replicate cages each in each batch (6 replicate cages in total). Feed mixtures were fed in pelleted form from d 21–24.

### 2.3. Chemical analysis

The diets and digesta samples were analysed for DM, ash, $TiO_2$, N, starch, CF, fat, Ca, P, and GE. DM was determined using standard procedures (Methods 930.15 and 925.10; AOAC, 2005). Ash was determined by standard procedures (method 942.05; AOAC, 2016) using a muffle furnace at 550 ºC for 16 h. Samples were assayed for titanium (Ti) on a UV spectrophotometer following the method of Short et al. (1996). N was determined by combustion (Method 968.06; AOAC, 2016) using a CNS-200 carbon, N, and sulphur auto analyser (LECO Corporation, St. Joseph, MI, USA). Total starch was determined using the assay procedure (Megazyme Total Starch Assay Procedure; Megazyme International Ireland Ltd., Wicklow, Ireland) based on thermostable α-amylase and amyloglucosidase. CF was determined using standard procedures (Methods 962.09 and 978.10; AOAC, 2005). Fat was determined using the Soxhlet extraction procedure (Method 991.36; AOAC, 2005). Ca and P were determined by colorimetric methods after combustion of the samples at 550 ºC and acid digestion in 6.0 M HCl using standard procedures (Method 968.08D; AOAC, 2005). GE was determined by adiabatic bomb calorimetry (Gallenkamp Autobomb, London, UK) standardised with benzoic acid.

### 2.4. Calculations

The CAID of nutrients and energy was calculated using the following formula:

$$\text{CAID} = [(\text{Nutrient} / \text{TiO}_2)_{\text{diet}} - (\text{Nutrient} / \text{TiO}_2)_{\text{ileal}}] / (\text{Nutrient} / \text{TiO}_2)_{\text{diet}} \quad (1)$$

where,

$(\text{Nutrient} / \text{TiO}_2)_{\text{diet}}$ = ratio of nutrient to $TiO_2$ in the diet, and $(\text{Nutrient} / \text{TiO}_2)_{\text{ileal}}$ = ratio of nutrient to $TiO_2$ in the ileal digesta.

The IDC of nutrients and energy was calculated using the following formula:

$$\text{IDC} = \text{Gross composition of nutrient} \times \text{CAID} \quad (2)$$

### 2.5. Statistical analysis

#### 2.5.1. Initial development of prediction equations for CAID and IDC of nutrients

Predictive multiple regression equations for CAID and IDC of N, energy, fat, starch, Ca, P, and DM were established by stepwise model selection procedure using the chemical composition of the 56 broiler diets (ash, N, fat, starch, CF, Ca, P, and GE) and ratios between chemical composition in the diet (fat:CF, starch:CF) as well as non-linear relationship ($CF^2$). The best model that fits the data was selected using AIC by comparing different models. All variables left in the model were significant at 0.15 level.

#### 2.5.2. Bootstrapped variable selection and validation of stepwise regression equations

The bootstrap method was used to validate the choice of variables by the stepwise selection method in the original equation. The following method was used to determine inclusion or exclusion of variables (Royston and Sauerbrei, 2009). If a total of *B* bootstrap samples is used to explore variations among the possible models for the original dataset with *k* variables,

1. Draw a bootstrap sample of size n.
2. Apply the model selection procedure.
3. For each variable $x_j$ (j =1, ….,*k*), record whether $x_j$ is selected in the model.
4. Repeat the steps above to a larger number, *B* number of times.
5. Summarize the results.

The outcome of this analysis will comprise a matrix consisting of *B* number of rows and *k* columns (Royston and Sauerbrei, 2009). This technique is dependent on the dataset because it relies on resampling observations from the dataset (Bertolini et al., 2022) (Fig. 1).

Sixty percent of the original data set (56 diets) was assigned for variable selection by bootstrap resampling by drawing repeated 1000 bootstrap samples from the original data set. Thus, some cases in the original data set were duplicated in the bootstrap sample while others were not included at all. The remaining 40 % was used to test the obtained results. As was done during the original equation development, stepwise multiple regression was used to select variables within each bootstrap sample. For each candidate variable, the proportion of bootstrap samples in which that variable was identified as an independent predictor of the outcome was determined. Then, variables were ordered according to the proportion of bootstrap samples in which they were selected as predictors of the outcome. A preliminary predictive model will consist of those variables identified as significant in all bootstrap samples (Austin and Tu, 2004b). The selection of variables was validated if the variables that appeared in the original stepwise model were selected in more than 30 % of the bootstrap samples.

Another approach was used to rerun the stepwise regression analysis for original data with selected variables that appeared above 30 % in 1000 bootstrap resamples. The final model was constructed using a stepwise regression procedure where top-ranked variables that were above 30% appearance in bootstrap samples were added to the model. This choice of cut points for percentage inclusion in a model and choice of section levels was arbitrary in this strategy, as explained by Sauerbrei and Schumacher (1992). Selection of prediction models in all instances was made using $R^2$, Adj. $R^2$, AIC, and RMSE values. Bootstrap resampling and stepwise multiple regression analyses were done using SAS software, version 9.4 package with various procedures (SAS, 2016).

## 3. Results

### 3.1. Chemical composition and nutrient digestibility coefficient of the diets

As expected, the analysed chemical composition of diets was quite variable as many traditional and non-traditional ingredients were used in varying amounts (Table 2). The N and GE contents of diets ranged from 22.3 to 78.1 g/kg DM and 17.87–24.03 MJ/kg, respectively. The digestibility coefficient of N varied between 0.466 and 0.806 (mean = 0.683), whereas the digestibility coefficient of energy ranged from 0.346 to 0.820 (mean = 0.611).

### 3.2. Prediction of coefficient of apparent ileal digestibility of N, energy, fat, starch, Ca, P, and DM

Prediction equations were calculated with an intercept and the best equation was selected as having the highest $R^2$ and minimum AIC values (Table 3). Prediction accuracy was improved for both CAID of N and energy when interactions between chemical composition and non-linear relationship in the diet were considered, with the lowest AIC and RMSE values and highest $R^2$ values.

The CAID of N was predicted by dietary content of N, CF, GE, $CF^2$, and starch:CF ratio, where $CF^2$ had a negative relationship and N, CF, and GE content had a positive relationship, with an $R^2$ of 0.78 and AIC of −305 (Eq. 1). For each additional unit of $CF^2$, digestibility of N decreased by 0.00003 units. High $R^2$ (0.87) and low AIC (-305) for prediction of CAID of energy was obtained when CF, GE, $CF^2$,

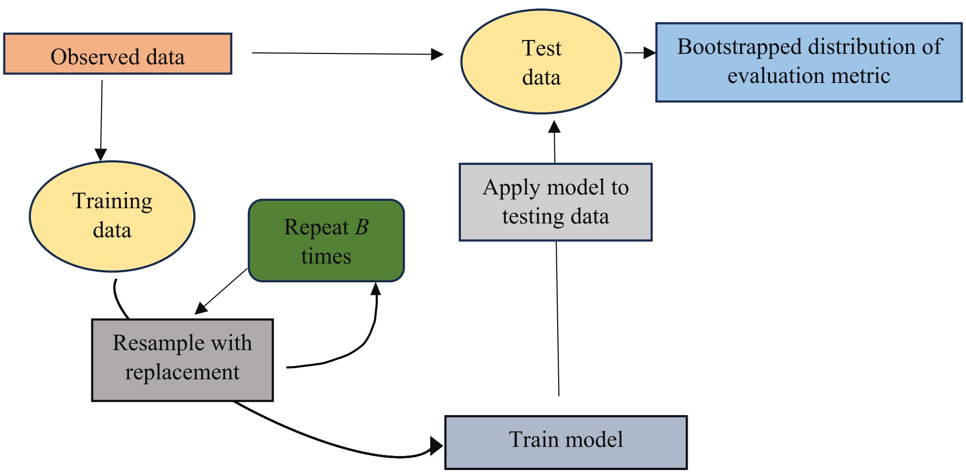


**Fig. 1.** Bootstrap resampling method adapted from Bertolini et al. (2022).

**Table 2**
Analysed chemical composition (g/kg DM), apparent ileal digestibility coefficient of nutrients, and ileal digestible nutrient content (g/kg DM) of the 56 diet mixtures used in the study as per Pedersen et al. (2021).

| Variable (g/kg DM) | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|
| Chemical Composition | | | | |
| Ash | 77.57 | 56.74 | 27.29 | 300.02 |
| N | 44.42 | 14.77 | 22.32 | 78.08 |
| Fat | 94.90 | 40.72 | 51.80 | 243.36 |
| Starch | 230.20 | 170.3 | 44.90 | 636.30 |
| CF | 68.13 | 29.56 | 26.63 | 144.73 |
| Ca | 15.36 | 22.39 | 3.31 | 105.17 |
| P | 11.24 | 9.02 | 4.18 | 46.96 |
| GE (MJ/kg) | 20.10 | 11.82 | 17.87 | 24.03 |
| Digestibility Coefficient | | | | |
| N | 0.683 | 0.076 | 0.466 | 0.806 |
| Fat | 0.752 | 0.081 | 0.506 | 0.879 |
| Starch | 0.929 | 0.040 | 0.810 | 0.982 |
| Ca | 0.332 | 0.137 | 0.122 | 0.633 |
| P | 0.498 | 0.155 | 0.137 | 0.704 |
| GE | 0.611 | 0.100 | 0.346 | 0.820 |
| DM | 0.536 | 0.121 | 0.228 | 0.795 |
| Digestible Content | | | | |
| N | 30.28 | 10.40 | 14.73 | 55.23 |
| Fat | 71.79 | 34.70 | 37.50 | 206.07 |
| Starch | 213.60 | 158.20 | 43.20 | 616.10 |
| Ca | 3.77 | 4.52 | 0.61 | 19.99 |
| P | 4.47 | 1.57 | 2.46 | 9.51 |
| GE | 12.26 | 2.13 | 7.00 | 16.66 |
| DM | 536.20 | 120.70 | 228.10 | 794.90 |

Ca = calcium; CF = crude fibre; DM = dry matter; GE = gross energy; N = nitrogen; P = phosphorus.

**Table 3**
Prediction equations of apparent ileal digestibility coefficients (CAID) of fat, starch, Ca, P, and DM based on chemical composition of the diets (g/kg DM) in broiler chickens.

| No | Equation | RMSE | $R^2$ | Adj.$R^2$ | AIC |
|---|---|---|---|---|---|
| 01 | CAID N = −0.37 +0.002 N +0.005 CF +0.03 GE −0.00003 $CF^2$ +0.02 starch:CF | 0.037 | 0.780 | 0.758 | -304.771 |
| 02 | CAID E = 0.14 +0.002 CF +0.021 GE −0.000021 $CF^2$ +0.013 starch:CF | 0.038 | 0.870 | 0.860 | -304.976 |
| 03 | CAID Fat = −1.78–0.003 fat +0.03 Ca −0.06 P +0.15 GE −0.000003 $CF^2$ | 0.041 | 0.774 | 0.751 | -295.149 |
| 04 | CAID Starch = 1.36 +0.0003 fat +0.0006 CF −0.025 GE | 0.033 | 0.355 | 0.318 | -319.192 |
| 05 | CAID Ca = −0.228–0.007 N −0.0008 starch +0.055 GE −0.00002 $CF^2$ −0.04 fat:CF +0.03 starch:CF | 0.081 | 0.692 | 0.654 | -217.540 |
| 06 | CAID P = 1.32–0.008 N −0.001 starch −0.016 P −0.000023 $CF^2$ +0.012 starch:CF | 0.068 | 0.825 | 0.808 | -237.822 |
| 07 | CAID DM = −0.24–0.0004 fat +0.004 CF +0.033 GE −0.00004 $CF^2$ +0.02 starch:CF | 0.042 | 0.888 | 0.877 | -290.507 |

Adj. R2 = adjusted R2; AIC = Akaikie's Information Criteria; Ca = calcium; CAID = coefficient of apparent ileal digestibility; CF = crude fibre; CF2 = square value of crude fibre; DM = dry matter; Fat:CF = fat-to-crude fibre ratio; GE = gross energy; N = nitrogen; P = phosphorus; R2 = coefficient of determination, RMSE = root mean square error; Starch:CF = starch-to-crude fibre ratio.

and starch:CF comprised the equation (Eq. 2).

The CAID of fat was negatively affected by dietary content of fat, P, and CF. For each increment of P content in the diet, CAID of fat was reduced by 0.06 units. The Ca and GE contents were positively correlated with CAID of fat ($R^2$ = 0.77 and AIC = −295; Equation

**Table 4**
Prediction equations of ileal digestible content (IDC) based on chemical composition (g/kg DM) in broiler chickens.

| No | Equation | RMSE | $R^2$ | Adj.$R^2$ | AIC |
|---|---|---|---|---|---|
| 08 | IDC N = −10.17–0.04 ash +0.8 N +0.03 fat +0.12 CF −0.0008 $CF^2$ +0.4 starch:CF | 1.660 | 0.977 | 0.975 | 121.281 |
| 09 | IDC E = −9.87 +0.03 CF +1.06 GE −0.0004 $CF^2$ +0.24 starch:CF | 0.749 | 0.885 | 0.876 | 30.400 |
| 10 | IDC Fat = −75.76–1.49 ash +0.48 N +0.61 fat −0.053 starch +3.46 Ca +7.23 GE −0.0005 $CF^2$ | 3.506 | 0.991 | 0.990 | 205.868 |
| 11 | IDC Starch = 52.41 +0.92 starch −2.5 GE | 13.086 | 0.993 | 0.993 | 348.929 |
| 12 | IDC Ca = 0.804 +0.193 Ca | 1.295 | 0.919 | 0.918 | 88.927 |
| 13 | IDC P = 5.71 +0.11 ash −0.1 N −0.005 starch −0.218 Ca −0.0002 $CF^2$ | 0.721 | 0.808 | 0.789 | 26.985 |
| 14 | IDC DM = −237.47–0.37 fat +3.89 CF +32.67 GE −0.04 $CF^2$ +18.2 starch:CF | 42.277 | 0.888 | 0.877 | 483.009 |

Adj. R2 = adjusted R2; AIC = Akaikie's Information Criteria; Ca = calcium; CF = crude fibre; CF2 = square value of crude fibre; DM = dry matter; GE = gross energy; IDC = ileal digestible content; N = nitrogen; P = phosphorus; R2 = coefficient of determination, RMSE = root mean square error; Starch:CF = starch-to-crude fibre ratio.

3). The best prediction was achieved for CAID of starch when fat, CF, and GE were included in the equation (Equation 4). For each additional MJ of GE, CAID of starch was reduced by 0.025 units and fat and CF contents were positively correlated with the CAID of starch.

The CAID of Ca (Equation 5) was predicted by N, starch, and GE content, $CF^2$, fat:CF and starch:CF. The CAID of P was influenced by content of N, starch, P, $CF^2$, and starch:CF. There was a negative relationship between CAID of P and dietary content of N, starch, P, and $CF^2$ (Equation 6). The CAID of DM depended on dietary content of fat, CF, and GE, and $CF^2$ and starch:CF. Fat content and $CF^2$ negatively correlated with CAID of DM, and other parameters positively correlated with CAID of DM. Prediction was more accurate ($R^2 = 0.89$) when $CF^2$ and starch:CF were included in the equation with a minimum AIC of −291 (Equation 7).

### 3.3. Prediction of ileal digestible content of nutrients and energy

The IDC of N was positively affected by dietary N, fat, and starch:CF ratio, and negatively related to ash and $CF^2$ (Table 4; Equation 8). The IDC of energy was best predicted by CF, GE, $CF^2$, and starch:CF, with $R^2$ of 0.89 and AIC of 30 (Equation 9). Digestible fat was linearly influenced by dietary fat (Equation 10). Moreover, ash, starch, and $CF^2$ had a negative relationship with the IDC of fat.

The IDC of starch relied on dietary starch and GE content. For each MJ increment of GE, IDC of starch was reduced 0.25 g, and for each 1 g increment in dietary starch, IDC of starch increased by 0.92 g (Equation 11). The starch content of the diets varied from 44.90 to 636.30 g/kg DM while gross energy had the range between 17.87 and 24.03 MJ/kg. The best predictor of IDC of Ca was dietary content of Ca, and it was positively correlated with IDC of Ca (Equation 12).

Ash, N, starch, and Ca content and $CF^2$ were the best variables to predict IDC of P, with $R^2$ of 0.81 and AIC of 27 (Equation 13). Ash content of the diet was positively correlated with IDC of P, whereas N, starch, Ca, and $CF^2$ had a negative relationship with IDC of P. The prediction equation for IDC of DM was composed of fat, CF, GE content, $CF^2$, and starch:CF (Equation 14).

### 3.4. Bootstrap validation of stepwise regression analysis for CAID and IDC of nutrients

Generalizability and reproducibility of the variables in the stepwise regression, as determined using measures of goodness of fit for the 1000 bootstrap resampling of training (60 % of the original data set) and test data (remaining 40 % of the original data set), showed that there was a good agreement between the goodness of fit results of the training and test data (Table 5).

Table 6 shows the frequency of the variables entering the model for stepwise regression on 1000 bootstrap samples. The number of variables entered in the model varied from 2 to 9, 2–10, 2–8, and 1–9 for CAID of N, energy, Ca, and P, respectively. Eight variables were selected more than 30 % (range, 31–85 %), whereas three variables were selected infrequently (range, 9–17 %) for CAID of N. Four variables were selected more than 30 % (range, 43–89 %) and seven variables were between 5 % and 27 % for CAID of energy. Similarly, five variables were selected for CAID of Ca and P whose frequency was above 30 %.

The number of variables entered in the model varied from 3 to 9, 3–9, 1–10, and 1–9 for IDC of N, energy, Ca, and P, with mean number of variables of 4.4, 3.9, 1.5, and 4.2, respectively for the 1000 bootstrap resamples. Variables ash, N, fat, starch, P, and $CF^2$ were selected more than 30 % of the time for IDC of N, whereas CF, GE, $CF^2$, and starch:CF were selected for IDC of energy. The only predictor selected for IDC of Ca was content of Ca, with 98.7 % appearance rate in 1000 bootstrap samples. Starch was appeared in 100 % bootstrap resamples for prediction of IDC of starch. Variables ash, N, Ca, GE, and $CF^2$ were selected more than 30 % in 1000 bootstrap samples for predicting IDC of P. The mean value of parameter estimates of variables selected using bootstrap resampling for CAID and IDC of nutrients are shown in Table 7 at $P < 0.15$.

The comparison between the original equation and the bootstrap variable selection shows that most variables were the same in

**Table 5**
Measures of goodness of fit for 1000 bootstrap resampling of 56 diets (60 % training data and 40 % test data).

| | Training (60 %) | | Test (40 %) | |
|---|---|---|---|---|
| | RMSE | $R^2$ | RMSE | $R^2$ |
| CAID N | 0.04 | 0.76 | 0.04 | 0.62 |
| CAID Energy | 0.04 | 0.87 | 0.04 | 0.82 |
| CAID Fat | 0.05 | 0.70 | 0.04 | 0.59 |
| CAID Starch | 0.03 | 0.37 | 0.02 | 0.24 |
| CAID Ca | 0.09 | 0.64 | 0.08 | 0.51 |
| CAID P | 0.07 | 0.82 | 0.08 | 0.72 |
| CAID DM | 0.04 | 0.89 | 0.05 | 0.84 |
| | | | | |
| IDC N | 1.65 | 0.98 | 1.97 | 0.96 |
| IDC Energy | 0.74 | 0.89 | 0.79 | 0.84 |
| IDC Fat | 3.55 | 0.99 | 4.51 | 0.98 |
| IDC Starch | 12.88 | 0.99 | 12.97 | 0.99 |
| IDC Ca | 1.23 | 0.92 | 1.17 | 0.92 |
| IDC P | 0.74 | 0.79 | 0.78 | 0.68 |
| IDC DM | 42.1 | 0.89 | 45.36 | 0.84 |

Ca = calcium; CAID = coefficient of apparent ileal digestibility; DM = dry matter; N = nitrogen; IDC = ileal digestible content; P = phosphorus; RMSE = root mean square error; R2 = coefficient of determination.

**Table 6**
Summary of appearance of predictors (%) in 1000 bootstrap resampling for coefficient of apparent leal digestibility (CAID) and ileal digestible content (IDC) of nutrients in broiler diets.

| | CAID | | | | | | | IDC | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Predictors | N | Energy | Fat | Starch | Ca | P | DM | N | Energy | Fat | Starch | Ca | P | DM |
| Ash | 8.8 | 26.8 | 17.8 | 5.4 | 15.6 | 24.5 | 51.5* | 56.2* | 5.9 | 41.1* | 2.3 | 2.1 | 93.3* | 49.8* |
| N | 44.2* | 14.8 | 56.7* | 6.8 | 37.4* | 48.0* | 11.0 | 100.0* | 9.4 | 72.2* | 1.6 | 7.1 | 59.1* | 11.0 |
| Fat | 30.7* | 7.6 | 30.2* | 23.5 | 9.2 | 27.0 | 6.8 | 61.6* | 13.9 | 99.9* | 0.7 | 1.9 | 18.0 | 5.9 |
| Starch | 9.8 | 12.9 | 12.0 | 9.6 | 44.1* | 32.7* | 19.0 | 32.4* | 10.7 | 29.1 | 100.0* | 1.5 | 26.0 | 15.3 |
| CF | 33.5* | 42.5* | 8.2 | 49.6* | 12.1 | 24.3 | 53.2* | 20.0 | 38.0* | 12.3 | 4.4 | 5.8 | 13.6 | 50.5* |
| Ca | 38.2* | 24.1 | 78.1* | 13.7 | 9.4 | 10.8 | 18.1 | 15.7 | 11.5 | 96.2* | 3.4 | 98.7* | 56.3* | 16.4 |
| P | 40.7* | 8.7 | 74.3* | 7.2 | 12.6 | 79.0* | 14.6 | 38.1* | 13.5 | 69.1* | 2.2 | 8.6 | 13.2 | 15.8 |
| GE | 39.7* | 77.5* | 49.6* | 87.4* | 82.3* | 20.2 | 39.0* | 18.1 | 93.1* | 66.0* | 37.1* | 5.2 | 38.4* | 37.4* |
| $CF^2$ | 85.3* | 88.7* | 11.7 | 31.2* | 33.7* | 58.9* | 97.8* | 55.4* | 94.0* | 19.2 | 0.6 | 10.4 | 52.9* | 98.4* |
| Fat:CF | 17.0 | 5.3 | 6.0 | 6.1 | 17.7 | 40.9* | 8.2 | 23.6 | 8.7 | 17.6 | 1.7 | 2.9 | 23.6 | 6.4 |
| Starch:CF | 48.8* | 78.4* | 26.3 | 9.5 | 53.3* | 24.2 | 85.2* | 18.6 | 86.2* | 6.9 | 13.7 | 2.4 | 25.3 | 86.6* |
| Predictors ≥ 30 % | 8 | 4 | 5 | 3 | 5 | 5 | 5 | 6 | 4 | 6 | 2 | 1 | 5 | 5 |
| No. of variables entered | | | | | | | | | | | | | | |
| Minimum | 2 | 2 | 1 | 0 | 2 | 1 | 2 | 3 | 3 | 2 | 1 | 1 | 1 | 2 |
| Maximum | 9 | 10 | 11 | 9 | 8 | 9 | 10 | 9 | 9 | 10 | 9 | 10 | 9 | 8 |
| Average | 3.97 | 3.87 | 3.71 | 2.50 | 3.27 | 3.91 | 4.04 | 4.40 | 3.85 | 5.30 | 1.68 | 1.45 | 4.20 | 3.94 |

Ca = calcium; CF2 = square value of crude fibre; DM = dry matter; Fat:CF = fat-to-crude fibre ratio; GE = gross energy; N = nitrogen; P = phosphorus; Starch:CF = starch-to-crude fibre ratio; * (≥ 30 %) selected for applying stepwise regression of the original diet data.

**Table 7**
Mean value of parameter estimates obtained using 1000 bootstrapping resamples of 56 broiler diets data for determining coefficient of apparent ileal digestibility of nutrients (CAID) and ileal digestible content of nutrients (IDC).

| | Ash | N | Fat | Starch | CF | Ca | P | GE | $CF^2$ | Fat:CF | Starch:CF |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CAID N | 0.004 | 0.003 | 0.001 | 0.0002 | 0.004 | -0.003 | -0.008 | 0.026 | -0.00002 | -0.016 | 0.015 |
| CAID Energy | -0.004 | 0.002 | -0.002 | -0.0001 | 0.002 | 0.010 | -0.003 | 0.030 | -0.00002 | -0.011 | 0.013 |
| CAID Fat | -0.006 | 0.004 | -0.003 | -0.0001 | 0.001 | 0.022 | -0.050 | 0.112 | -0.00001 | 0.011 | 0.008 |
| CAID Starch | 0.003 | -0.001 | 0.0004 | -0.0001 | 0.001 | 0.001 | -0.016 | -0.021 | 0.00000 | -0.009 | -0.001 |
| CAID Ca | 0.003 | -0.007 | -0.002 | 0.00004 | -0.001 | 0.015 | -0.052 | 0.067 | -0.00002 | -0.060 | 0.024 |
| CAID P | 0.009 | -0.009 | -0.001 | -0.001 | -0.003 | -0.014 | -0.029 | 0.119 | -0.00002 | -0.007 | 0.018 |
| CAID DM | -0.002 | -0.0005 | -0.001 | -0.0002 | 0.005 | 0.006 | -0.006 | 0.028 | -0.00003 | 0.003 | 0.016 |
| IDC N | 0.007 | 0.781 | 0.044 | 0.019 | 0.022 | -0.103 | -0.636 | 0.191 | -0.0004 | -1.130 | 0.426 |
| IDC Energy | -0.062 | 0.033 | -0.014 | 0.004 | 0.046 | 0.200 | -0.236 | 1.194 | -0.0004 | 0.106 | 0.241 |
| IDC Fat | -1.213 | 0.405 | 0.663 | -0.049 | 0.014 | 2.407 | -4.345 | 9.737 | -0.001 | 1.376 | -0.047 |
| IDC Starch | 1.335 | -0.547 | -0.030 | 0.923 | 0.265 | 0.145 | -7.392 | -3.441 | -0.002 | -3.237 | -0.206 |
| IDC Ca | -0.011 | -0.040 | -0.002 | -0.004 | 0.025 | 0.206 | -0.178 | 0.880 | -0.0002 | -1.411 | -0.063 |
| IDC P | 0.090 | -0.103 | -0.029 | -0.005 | -0.004 | -0.213 | -0.150 | 1.212 | -0.0002 | -0.730 | -0.061 |
| IDC DM | -1.482 | -0.095 | -0.710 | -0.129 | 4.666 | 6.305 | -5.657 | 28.095 | -0.029 | 0.278 | 15.005 |

Ca = calcium; CF2 = square value of crude fibre; DM = dry matter; Fat:CF = fat-to-crude fibre ratio; GE = gross energy; N = nitrogen; P = phosphorus; Starch:CF = starch-to-crude fibre ratio.

both. However, some variables in the original model were omitted or some new variables were selected in bootstrap resampling except the models for CAID of energy and IDC of energy, starch, and Ca. Fat, Ca, and P contents which were selected above 30 % in bootstrap resampling, did not appear in the original equation for CAID of N. Likewise, N content, which appeared above 30 % in bootstrap resampling, was not included in the original equations for CAID of fat. At the same time, $CF^2$, which was selected only in 12 % of the bootstrap resamples was included in the original equation for CAID of fat. For CAID and IDC of DM, ash content was not included in the original equation even though it was selected above 30 % in the 1000 bootstrap resamples.

### 3.5. Stepwise regression models using selected variables by 1000 bootstrap resampling

Equations for prediction of CAID of N, energy, and DM were similar when regressed with the selected variables using bootstrapping (Table 8). The $CF^2$ was removed for CAID of fat and fat was removed for CAID of starch (Equations 17 and 18, respectively) compared to the original equations. The CAID of Ca was predicted by the variables N, GE, $CF^2$, and starch:CF (Equation 19), whereas CAID of P was predicted by content of P and $CF^2$ (Equation 20). Fig. 2 (*a* and *b*) shows a close agreement between the predicted value and the observed value for CAID of N and energy, respectively, where the $R^2$ values obtained were 0.78 and 0.87 for CAID of N and energy, respectively. The calculated mean bias (MB) is 0 for both CAID of N and energy.

Equations for prediction of IDC of energy, starch, and Ca were similar when using all variables and selected variables with the same parameter estimates (Equations 23, 25, and 26, respectively; Table 9). Starch:CF was not selected as a predictor for prediction of IDC of N (Equation 22) and $CF^2$ was not selected for IDC of fat (Equation 24) when compared to the original equation. The equations developed using the selected variables for the IDC of N and energy showed good performance as the $R^2$ obtained was 0.98 and 0.89, respectively, with MB = 0 (Fig. 3a and b).

IDC of P was represented by content of ash, N, Ca, GE, and $CF^2$ where starch was removed, and GE was added (Equation 27) compared to the original equation (Equation 12). For IDC of DM, ash was replaced for the fat in the original equation (Equation 28). The $R^2$ and adj. $R^2$ decreased, while RMSE and AIC increased when using the selected variables from bootstrapping for CAID of fat, starch, Ca, and P. Similarly, $R^2$ and adj. $R^2$ decreased while RMSE and AIC increased for IDC of N, fat, P, and DM.

**Table 8**
Prediction equations for coefficient of apparent ileal digestibility (CAID) of nutrients based on chemical composition (g/kg DM) in broiler chickens using selected predictors by bootstrapping.

| No | Equation | RMSE | $R^2$ | Adj.$R^2$ | AIC |
|---|---|---|---|---|---|
| 15 | CAID N = −0.37 +0.002 N +0.005 CF +0.03 GE −0.00003 $CF^2$ +0.02 starch:CF * | 0.037 | 0.780 | 0.758 | -304.771 |
| 16 | CAID E = 0.14 +0.002 CF +0.021 GE −0.000021 $CF^2$ +0.013 starch:CF * | 0.038 | 0.870 | 0.860 | -304.976 |
| 17 | CAID Fat = −1.74–0.003 fat +0.03 Ca −0.06 P +0.15 GE | 0.042 | 0.757 | 0.738 | -293.130 |
| 18 | CAID Starch= 1.25 +0.0004 CF −0.02 GE | 0.034 | 0.315 | 0.289 | -317.773 |
| 19 | CAID Ca = −0.47–0.004 N +0.05 GE −0.000009 $CF^2$ +0.01 starch:CF | 0.086 | 0.635 | 0.607 | -212.090 |
| 20 | CAID P = 0.71–0.02 P −0.000008 $CF^2$ | 0.081 | 0.739 | 0.729 | -221.071 |
| 21 | CAID DM = −0.24–0.0004 fat +0.004 CF +0.033 GE −0.00004 $CF^2$ +0.02 starch:CF * | 0.042 | 0.888 | 0.877 | -290.507 |

Adj. R2 = adjusted R2; AIC = Akaikie's Information Criteria; Ca = calcium; CAID = coefficient of apparent ileal digestibility; CF = crude fibre; CF2 = square value of crude fibre; DM = dry matter; Fat:CF = fat-to-crude fibre ratio; GE = gross energy; N = nitrogen; P = phosphorus; R2 = coefficient of determination, RMSE = root mean square error; Starch:CF = starch-to-crude fibre ratio; *, same equation as original equation developed using all variables (56 diets).

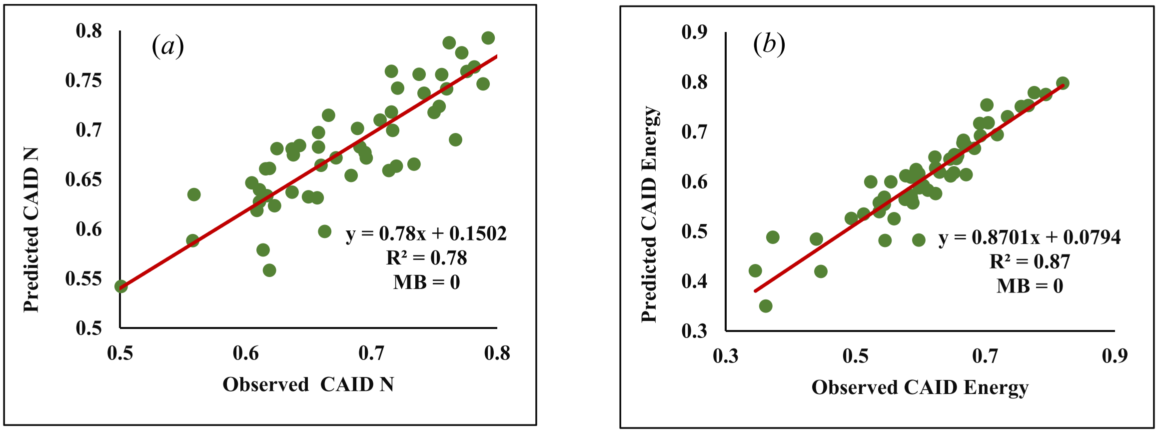


**Fig. 2.** Predicted vs. observed values for coefficient of apparent ileal digestibility (CAID) of N (*a*) and energy (*b*) based on the gross chemical composition using selected predictors from bootstrapping.

**Table 9**
Prediction equations of ileal digestible content (IDC) based on chemical composition (g/kg DM) in broiler chickens using selected predictors by bootstrapping.

| No | Equation | RMSE | $R^2$ | Adj.$R^2$ | AIC |
|---|---|---|---|---|---|
| 22 | IDC N = −4.7–0.04 ash +0.8 N +0.02 fat +0.01 starch −0.0002 $CF^2$ | 1.710 | 0.976 | 0.973 | 123.744 |
| 23 | IDC E = −9.87 +0.03 CF +1.06 GE −0.0004 $CF^2$ +0.24 starch:CF * | 0.749 | 0.885 | 0.876 | 30.400 |
| 24 | IDC Fat= −94.27–1.36 ash +0.57 N +0.64 fat −0.03 starch +3.2 Ca +7.2 GE | 3.565 | 0.991 | 0.989 | 206.896 |
| 25 | IDC Starch = 52.41 +0.92 starch −2.5 GE* | 13.086 | 0.993 | 0.993 | 348.929 |
| 26 | IDC Ca = 0.804 +0.193 Ca* | 1.295 | 0.919 | 0.918 | 88.927 |
| 27 | IDC P = 3.77 +0.15 ash −0.1 N −0.3 Ca +0.3 GE −0.0001 $CF^2$ | 0.723 | 0.806 | 0.787 | 27.377 |
| 28 | IDC DM = 543.7–0.63 ash +2.2 CF +32.67 GE −0.03 $CF^2$ +11.1 starch:CF | 43.831 | 0.878 | 0.868 | 486.161 |

Adj. R2 = adjusted R2; AIC = Akaikie's Information Criteria; Ca = calcium; CF = crude fibre; CF2 = square value of crude fibre; DM = dry matter; GE = gross energy; IDC = ileal digestible content; N = nitrogen; P = phosphorus; R2 = coefficient of determination, RMSE = root mean square error; Starch:CF = starch-to-crude fibre ratio; *, same equation as original equation developed using all variables (56 diets).

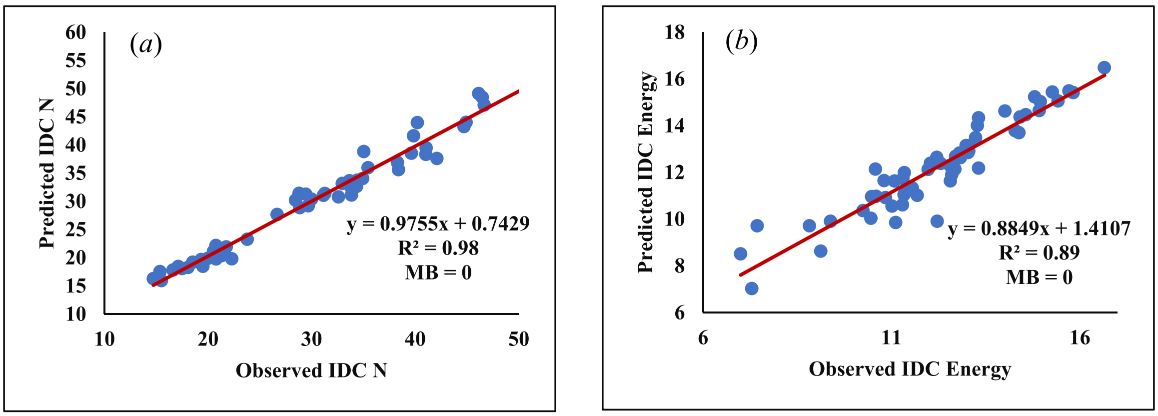


**Fig. 3.** Predicted vs. observed values for ileal digestible content (IDC) of N (*a*) and energy (*b*) based on the gross chemical composition using selected predictors from bootstrapping.

## 4. Discussion

In the current study, the bootstrap approach was used primarily to select variables and construct prediction models for CAID and IDC of dietary nutrients for broilers. The frequency of the variables entering 1000 bootstrap models is used as a guide to validate variables selected in original stepwise regression. Selection based on a particular cut-off (i.e., > 30 % frequency) is arbitrary. The data set used in the present study as per Pedersen et al. (2021) was very robust because it had a wide range of digestibility and chemical composition needed to develop the predictions with more complex diets. It is noteworthy that some variables selected were only responses to non-traditional ingredients viz. diets consisting of high levels of MBM had high amounts of Ca and P. This would not be applicable under practical situations as commercial diets consist of Ca and P in low concentrations and they are less variable (Pedersen et al., 2021). Therefore, the equations with Ca and P could be only considered as a general guide.

The variations in the stepwise regression models for sub-samples may have been caused by correlations between variables (Scalon et al., 1998; Steyerberg et al., 2001). The stepwise method can also be biased by an observation that has a significant impact on the inference process. The perfect model should have good stability that is the variables selected and predictive ability of the model should be the same among different data sets from the same population (Scalon et al., 1998). With the bootstrapping, we estimated the whole distribution of important variables under consideration. Therefore, from a whole set of variables we can select only the important variables which are more stable.

In the current study, CF has been included as a variable along with its interactions to develop equations as it has been traditionally used in feed analysis because of its simplicity, feasibility, and historical usage. Prediction equations including CF as an important variable were reported in previous broiler studies (Campbell et al., 1986; Alvarenga et al., 2015; Pedersen et al., 2021). Moreover, variable components of an equation should come from simple analytical procedures and generally CP, ash, fat, starch, and sometimes a fibre crieteria like CF are important parameters (Zaefarian et al., 2021).

According to Pedersen et al. (2021), ileal digestibility of protein in broiler diets can be predicted by dietary starch, CF, and fat contents. Starch and fat contents had a positive relationship, while CF had a negative relationship with the ileal digestibility of protein ($R^2 = 0.42$; $P < 0.05$). In the present study, CAID of N was predicted by Eq. 1, with $R^2 = 0.78$ ($P < 0.15$). However, CF had a positive relationship and $CF^2$ had a negative relationship with CAID of N. Cerrate et al. (2019) reported that digestibility of protein in pigs and poultry reduced with increasing dietary fibre content and the negative effect was dependent upon the level and type of fibre. On the other hand, in the current study, fat content of the diet did not have an impact on digestibility of protein. This agrees with previous results by Honda et al. (2009) who found no effects on protein digestibility with chickens fed 3–10 % crude fat.

Results of the present study showed that CAID of energy was best explained by dietary CF and GE content, $CF^2$, and starch:CF (Eq. 2 and 16). Cerrate et al. (2019) stated that dietary ME can be predicted from digestible nutrients and crude dietary nutrients, where the interaction between dietary protein and fat was considered. A 0.02 unit increment in the CAID of energy was observed with a unit increase in dietary content of energy. Pedersen et al. (2021) proposed an equation for ileal digestibility of energy, where starch, CF, and phytate contents were necessary to achieve a significant prediction and CAID of energy was negatively affected by dietary CF and phytate. In contrast, the present study reported a positive relationship between dietary CF and CAID of energy (Eq. 2 and 16).

The CAID of fat was predicted by dietary content of fat, Ca, P, GE, and $CF^2$ ($R^2 = 0.77$) (Equation 3) and the same predictors were selected when bootstrapping except for $CF^2$ (Equation 17). In bootstrap resampling, $CF^2$ appeared only 11.7 % of the time and was, therefore, not included while developing the model with the selected variables. According to Noblet and Perez (1993), digestibility coefficient of ether extract (EE) depended on dietary EE as well as the square value of EE ($R^2 = 0.70$). A negative correlation between CAID of fat and dietary content of P was observed in both equations. However, dietary Ca content had a positive effect on CAID of fat. This was in contrast to previous studies (Edwards et al., 1960; Griffith et al., 1961; Atteh and Leeson, 1983; Hakansson, 1974; Mutucumarana et al., 2014; Tancharoenrat and Ravindran, 2014) reported a decrease in digestibility of fat with an increase in dietary Ca level. High dietary Ca affects the utilisation of fat through the formation of Ca soaps (Atteh and Leeson, 1983). In the present study, however, Ca in the diets was only from the 10 ingredients and not from the major Ca sources such as limestone and dicalcium phosphate (DCP) which are known for the negative effects on fat and energy digestibility. Moreover, the concentrations of Ca and P were only high in 10 diets with high inclusion of meat and bone meal (55.0–105.2 g/kg DM).

The original equation for CAID of starch (Equation 4) with all predictors was composed of fat, CF, and GE content. However, fat content was not selected as a variable as it appeared only 23.5 % of the time in the 1000 bootstrap resamples when the second stepwise regression was done (Equation 18). The CF and GE contents had a positive and negative relationship, respectively, with CAID of starch. Cerrate et al. (2019) reported that the digestibility coefficient for starch was negatively affected by neutral detergent fibre (NDF) content ($R^2 = 0.10$) in diets without enzyme and positively affected ($R^2 = 0.03$) with enzyme.

Precision of the equation for CAID of Ca was improved when $CF^2$, fat:CF, and starch:CF were considered. In contrast, the content of starch and fat:CF were eliminated from the model when variables from bootstrapping were included even though both had been selected more than 30 % of the time in bootstrapping (Equation 19). The $R^2$ of Equation 19 was low compared to Equation 5. In both equations, CAID of Ca was not influenced by dietary Ca level. This agrees with Mutucumarana et al. (2014) who suggested that Ca digestibility was not influenced by dietary Ca concentration.

Current results show that CAID of P was predicted by Equations 6 and 20. In both equations, a negative, curvilinear effect of CF was observed, and CAID of P decreased 0.02 units with every unit increase in dietary P. Even though N and starch contents were selected more than 30 % of the time in bootstrap resamples, they were not included in Equation 20 as they hadn't met the 0.15 significance level for entry into the model. The CF was an important variable for prediction of CAID of DM. Results of the study by Mtei et al. (2019) demonstrated that CAID of DM, starch, fat, NDF, and GE were influenced by dietary fibre content, and it differed among different bird types (layers, broilers, and pullets).

In the current study, IDC of N increased linearly with dietary N content. In agreement with this, digestible protein in poultry (Tahir et al., 2008; Cerrate et al., 2019) and pigs (Noblet and Perez, 1993; Shi and Noblet, 1993) improved by adding dietary protein. The RMSE value of IDC of N increased when variables selected by bootstrapping were used with the removal of CF and starch:CF and the replacement of starch. Both equations also indicated a negative effect of ash on amount of N lost, where 0.04 g of N was lost per g increment in ash content of the diet.

The starch:CF, CF, and GE had a positive relationship, and $CF^2$ had a negative relationship for the prediction of IDC of energy. Furthermore, results from Noblet and Perez (1993) in pigs agree with the positive relationship of dietary GE content with IDC of energy. The IDC of fat increased by 0.6 units when there was a unit increase in dietary content of fat. In agreement with this, Wiseman and Salvador (1991), Shi and Noblet (1993), and Cerrate et al. (2019) reported that fat digestibility improved linearly by adding dietary fat. According to Noblet and Perez (1993), digestible EE was predicted with either EE content alone ($R^2 = 0.97$) or with EE and NDF content ($R^2 = 0.97$), where EE content had a positive relationship with digestible EE in both equations and NDF content had a negative relationship with digestible EE.

In the current study, the best predictors for IDC of starch were dietary content of starch and GE. Cerrate et al. (2019) stated that digestible content of starch could be estimated using dietary starch content. They reported a 0.94 decrement in digestible starch content for each unit increment in dietary starch ($R^2 = 0.94$). In the current study and that by Cerrate et al. (2019), digestible starch had a linear relationship with dietary starch content.

The best parameter to predict IDC of Ca was the dietary content of Ca. Consequently, dietary Ca content had a negative impact on IDC of P, with a positive effect of ash and a negative curvilinear effect of CF. As described earlier, these equations can be used as a general guide as the concentration of Ca and P were very high in diets containing high amounts of MBM.

In this study, we also found that the inclusion of CF and its interactions as a variable in prediction equations increased the $R^2$ value. Even though the CF represents a small portion of the total fibre present in the ingredients, the impact of CF can be seen in most of the developed equations. There is potential for further improvement of the developed equations by adding NDF, ADF, or non-starch polysaccharides (NSP). However, it is important to consider the cost for the analysis to best select the model with practical application.

The major limitation of this study was the diets used, which represented a wide range of chemical compositions due to varying inclusion level of ingredients. This ensured the development of better prediction equations assisting in finding the relationship between the variables in regression equations. However, compared to the diets used in the current study, commercial diets will be complete and nutritionally balanced. Therefore, in practice, the equations should be used with caution depending on the chemical composition of the feeds to be studied. Moreover, further research is needed to assess the generalizability of our findings to commercial settings.

In conclusion, prediction equations can be developed for CAID and IDC of nutrients in diets using chemical composition by including interactions between dietary chemical components. Measures of goodness of fit ($R^2$, Adj. $R^2$, RMSE, and AIC) of the original stepwise regression models and after the bootstrap exercise for estimated regression coefficients for CAID of N, energy, and DM, as well as IDC of energy, starch, and Ca, were the same with no differences in parameter estimates. This indicated the stability performance for the stepwise regression models among the bootstrap samples. Not much difference was observed between the performance of the original model and the model constructed with the selected variables using bootstrapping. However, this study gives a useful insight into the variable selection approach as stepwise selection is just a single model without any information about its stability and this might be useful to select the best variables when constructing prediction models with greater confidence. Rigorous analysis and external validation with a new data set should require ensuring the use of the equations in practical situations.

## CRediT authorship contribution statement

**S. Thiruchchenthuran**: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing. **N. Lopez-Villalobos**: Methodology, Software, Formal analysis, Writing – review & editing. **F. Zaefarian**: Writing – review & editing, Supervision, Funding acquisition. **M.R. Abdollahi**: Writing – review & editing, Supervision, Funding acquisition. **T.J. Wester**: Writing – review & editing, Supervision. **N.B. Pedersen**: Writing – review & editing, Funding acquisition. **A.C. Storm**: Writing – review & editing, Funding acquisition. **A.J. Cowieson**: Writing – review & editing, Funding acquisition. **P.C.H. Morel**: Conceptualization, Methodology, Formal analysis, Investigation, Writing – review & editing, Supervision, Project administration.

## Funding

The first author was funded by the Accelerating Higher Education Expansion and Development (AHEAD) operation (Grant number: AHEAD/PhD/R3/Agri/380), a world bank funded project, Ministry of Education, Sri Lanka.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

Alvarenga, R., Zangeronimo, M., Pereira, L., Wolp, R., Almeida, E., 2013. Formulation of diets for poultry; the importance of prediction equations to estimate the energy values. Arch. Zootec. 62, 1–11. ⟨http://www.redalyc.org/articulo.oa?id=49558826001⟩.

Alvarenga, R.R., Rodrigues, P.B., Zangeronimo, M.G., Oliveira, E.C., Mariano, F.C., Lima, E.M., Garcia, A.A., Jr, Naves, L.P., Nardelli, N.B., 2015. Validation of prediction equations of energy values of a single ingredient or their combinations in male broilers. Asian-Australas. J. Anim. Sci. 28, 1335–1344. https://doi.org/10.5713/ajas.14.0339.

AOAC. 2005. Official Methods of Analysis. $18^{th}$ ed. Association of Official Analytical Chemists, Arlington, VA.

AOAC. 2016. Official Methods of Analysis. $20^{th}$ ed. Association of Official Analytical Chemists, Arlington, VA.

Atteh, J., Leeson, S., 1983. Effects of dietary fatty acids and calcium levels on performance and mineral metabolism of broiler chickens. Poult. Sci. 62, 2412–2419. https://doi.org/10.3382/ps.0622412.

Austin, P.C., Tu, J.V., 2004b. Bootstrap methods for developing predictive models. Am. Stat. 58, 131–137. https://doi.org/10.1198/0003130043277.
Austin, P.C., Tu, J.V., 2004a. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. J. Clin. Epidemiol. 57, 1138–1146. https://doi.org/10.1016/j.jclinepi.2004.04.003.
Bertolini, R., Finch, S.J., Nehm, R.H., 2022. Quantifying variability in predictions of student performance: Examining the impact of bootstrap resampling in data pipelines. Comput. Educ. Artif. Intell. 3, 100067 https://doi.org/10.1016/j.caeai.2022.100067.
Brunelli, A., 2014. A synopsis of resampling techniques. J. Thorac. Dis. 6, 1879–1882 https://doi.org/10.3978%2Fj.issn.2072-1439.2014.09.09.
Campbell, G.L., Salmon, R.E., Classen, H.L., 1986. Prediction of metabolizable energy of broiler diets from chemical analysis. Poult. Sci. 65, 2126–2134. https://doi.org/10.3382/ps.0652126.
Castilho, R.A., Pozza, P.C., Oliveira, N.T.E.D., Sangali, C.P., Langer, C.N., Nunes, R.V., 2015. Equations to predict the metabolizable energy of meat and bone meal for growing pigs. Cienc. Agrotecnol. 39, 565–573. https://doi.org/10.1590/S1413-70542015000600003.
Cerrate, S., Ekmay, R., England, J., Coon, C., 2019. Predicting nutrient digestibility and energy value for broilers. Poult. Sci. 98, 3994–4007. https://doi.org/10.3382/ps/pez142.
Chowdhury, M.Z., Turin, T.C., 2021. Validating prediction models for use in clinical practice: concept, steps and procedures focusing on hypertension risk prediction. Hypertens 7, 21–29. https://doi.org/10.15713/ins.johtn.0221.
Edwards, H.M., Dunahoo, W.S., Carmon, J.L., Fuller, H.L., 1960. Effect of protein, energy and fat content of the ration on calcium utilization. Poult. Sci. 39, 1389–1394. https://doi.org/10.3382/ps.0391389.
Efron, B., 1979. Bootstrap methods: another Look at the Jackknife. Ann. Stat. 7, 1–26. https://doi.org/10.1007/978-1-4612-4380-9_41.
Griffith, F., Grainger, R., Begin, J.J., 1961. The effect of dietary fat and cellulose on apparent calcium digestibility in growing chickens. Poult. Sci. 40, 1492–1497. https://doi.org/10.3382/ps.0401492.
Hakansson, J., 1974. Factors affecting the digestibility of fats and fatty acids in chicks and hens. Swed. J. Agric. Res. 4, 33–47.
Harrell, F.E., 2015. Regression modeling strategies, second ed. Springer International Publishing. https://doi.org/10.1007/978-3-319-19425-7.
Honda, K., Kamisoyama, H., Isshiki, Y., Hasegawa, S., 2009. Effects of dietary fat levels on nutrient digestibility at different sites of chicken intestines. J. Poult. Sci. 46, 291–295. https://doi.org/10.2141/jpsa.46.291.
Mateos, G., Camara, L., Fondevila, G., Lazaro, R., 2019. Critical review of the procedures used for estimation of the energy content of diets and ingredients in poultry. J. Appl. Poult. Res. 28, 506–525. https://doi.org/10.3382/japr/pfy025.
Mtei, A., Abdollahi, M., Schreurs, N., Girish, C., Ravindran, V., 2019. Dietary inclusion of fibrous ingredients and bird type influence apparent ileal digestibility of nutrients and energy utilization. Poult. Sci. 98, 6702–6712. https://doi.org/10.3382/ps/pez383.
Mutucumarana, R.K., Ravindran, V., Ravindran, G., Cowieson, A.J., 2014. Influence of dietary calcium concentration on the digestion of nutrients along the intestinal tract of broiler chickens. J. Poult. Sci. 51, 392–401. https://doi.org/10.2141/jpsa.0140022.
Noblet, J., Perez, J., 1993. Prediction of digestibility of nutrients and energy values of pig diets from chemical analysis. J. Anim. Sci. 71, 3389–3398. https://doi.org/10.2527/1993.71123389x.
Nunez, E., Steyerberg, E.W., Nunez, J., 2011. Regression modeling strategies. Rev. Esp. Cardiol. (Engl. Ed. ) 64, 501–507. https://doi.org/10.1016/j.rec.2011.01.017.
Oliveira, N.T.E. d, Pozza, P.C., Castilha, L.D., Pasquetti, T.J., Langer, C.N., 2019. Metabolisable energy prediction in energy feedstuffs and evaluation of the stepwise validation procedure using bootstrapping. Rev. Cienc. Agron. 50, 131–139. https://doi.org/10.5935/1806-6690.20190016.
Pedersen, N.B., Zaefarian, F., Storm, A.C., Ravindran, V., Cowieson, A.J., 2021. Mathematical prediction of ileal energy and protein digestibility in broilers using multivariate data analysis. Poult. Sci. 100, 101106 https://doi.org/10.1016/j.psj.2021.101106.
Royston, P., Sauerbrei, W., 2009. Bootstrap assessment of the stability of multivariable models. Stata. J. 9, 547–570. https://doi.org/10.1177/1536867×0900900403.
SAS., 2016. SAS software, version 9.4. In SAS Institute Cary, NC, USA.
Sauerbrei, W., 1999. The use of resampling methods to simplify regression models in medical statistics. Appl. Stat. 48, 313–329. https://doi.org/10.1111/1467-9876.00155.
Sauerbrei, W., Schumacher, M., 1992. A bootstrap resampling procedure for model building: application to the Cox regression model. Stat. Med. 11, 2093–2109. https://doi.org/10.1002/sim.4780111607.
Scalon, J.D., Freire, S.M., Cunha, T.A., 1998. Validation of models for predicting the use of health technologies. Med. Decis. Mak. 18, 311–319. https://doi.org/10.1177/0272989×9801800309.
Sheikhhasan, B.S., Moravej, H., Shivazad, M., Ghaziani, F., Esteve-Garcia, E., Kim, W.K., 2020a. Prediction of the total and standardized ileal digestible amino acid contents from the chemical composition of soybean meals of different origin in broilers. Poult. Sci. 99, 4947–4957. https://doi.org/10.1016/j.psj.2020.06.033.
Sheikhhasan, B.S., Moravej, H., Ghaziani, F., Esteve-Garcia, E., Kim, W.K., 2020b. Relationship between chemical composition and standardized ileal digestible amino acid contents of corn grain in broiler chickens. Poult. Sci. 99, 4496–4504. https://doi.org/10.1016/j.psj.2020.06.013.
Shi, X., Noblet, J., 1993. Contribution of the hindgut to digestion of diets in growing pigs and adult sows: effect of diet composition. Livest. Prod. Sci. 34, 237–252. https://doi.org/10.1016/0301-6226(93)90110-4.
Short, F.J., Gorton, P., Wiseman, J., Boorman, K.N., 1996. Determination of titanium dioxide added as an inert marker in chicken digestibility studies. Anim. Feed Sci. Technol. 59, 215–221. https://doi.org/10.1016/0377-8401(95)00916-7.
Smith, B., Hassen, A., Hinds, M., Rice, D., Jones, D., Sauber, T., Iiams, C., Sevenich, D., Allen, R., Owens, F., 2015. Predicting the digestible energy of corn determined with growing swine from nutrient composition and cross-species measurements. J. Anim. Sci. 93, 1025–1038. https://doi.org/10.2527/jas.2014-7807.
Steyerberg, E.W., Harrell, F.E., 2016. Prediction models need appropriate internal, internal-external, and external validation. J. Clin. Epidemiol. 69, 245–247. https://doi.org/10.1016/j.jclinepi.2015.04.005.
Steyerberg, E.W., Harrell Jr, F.E., Borsboom, G.J., Eijkemans, M., Vergouwe, Y., Habbema, J.D.F., 2001. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J. Clin. Epidemiol. 54, 774–781. https://doi.org/10.1016/S0895-4356(01)00341-9.
Tahir, M., Saleh, F., Ohtsuka, A., Hayashi, K., 2008. An effective combination of carbohydrases that enables reduction of dietary protein in broilers: importance of hemicellulase. Poult. Sci. 87, 713–718. https://doi.org/10.3382/ps.2007-00340.
Tancharoenrat, P., Ravindran, V., 2014. Influence of tallow and calcium concentrations on the performance and energy and nutrient utilization in broiler starters. Poult. Sci. 93, 1453–1462. https://doi.org/10.3382/ps.2013-03817.
Wiseman, J., Salvador, F., 1991. The influence of free fatty acid content and degree of saturation on the apparent metabolizable energy value of fats fed to broilers. Poult. Sci. 70, 573–582. https://doi.org/10.3382/ps.0700573.
Yegani, M., Swift, M.L., Zijlstra, R.T., Korver, D.R., 2013. Prediction of energetic value of wheat and triticale in broiler chicks: a chick bioassay and an in vitro digestibility technique. Anim. Feed Sci. Technol. 183, 40–50. https://doi.org/10.1016/j.anifeedsci.2013.03.010.
Zaefarian, F., Cowieson, A.J., Pontoppidan, K., Abdollahi, M.R., Ravindran, V., 2021. Trends in feed evaluation for poultry with emphasis on in vitro techniques. Anim. Nutr. 7, 268–281. https://doi.org/10.1016/j.aninu.2020.08.006.