Neural Comput & Applic (2018) 30:2499–2508
https://doi.org/10.1007/s00521-017-2842-x

ORIGINAL ARTICLE

Mathematical models for response to amino acids: estimating the response of broiler chickens to branched-chain amino acids using support vector regression and neural network models

A. Gitoee[1] · A. Faridi[2] · J. France[3]

Received: 5 March 2015 / Accepted: 5 January 2017 / Published online: 11 January 2017

Abstract Support vector regression (SVR) and neural network (NN) models were used to predict average daily gain (ADG), feed efficiency (FE), and feed intake (FI) of broiler chickens during the starter period. Input variables for construction of the models were levels of dietary protein and branched-chain amino acids (BCAA; valine, isoleucine, and leucine). Starting with 241 lines, the SVR and NN models were trained using 120 data lines and the remainder ($n = 121$) was used as the testing set. The SVR models were developed using different kernel functions including: linear, polynomial (second and third order), radial basis function (RBF), and sigmoidal. In order to evaluate the SVR models, their performance was compared to that of multilayer perceptron (MLP)- and RBF-type NN models. Results indicated that MLP-type NN models were the most accurate in predicting the investigated output variables (R^2 for ADG in training and testing = 0.81 and 0.81; FE = 0.87 and 0.87; FI = 0.68 and 0.62). Among the different SVR kernels, best performance was achieved with the RBF (R^2 for ADG in training and testing = 0.76 and 0.76; FE = 0.85 and 0.87; FI = 0.46 and 0.48) and polynomial (third order) function (R^2 for ADG in training and testing = 0.77 and 0.77; FE = 0.85 and 0.87; FI = 0.46 and 0.39), whose prediction ability was better than that of the RBF-type NN (R^2 for ADG in training and testing = 0.75 and 0.75; FE = 0.82 and 0.82; FI = 0.41 and 0.39) models. Sigmoidal SVR models provided the poorest prediction. The work demonstrates that, through proper selection of kernel functions and corresponding parameters, SVR models can be considered as an alternative to NN models in predicting the response of broiler chickens to protein and BCAA. This type of model should also be applicable in poultry and other areas of animal nutrition.

Keywords Support vector regression · Neural networks · Broiler response · Branched-chain amino acids

1 Introduction

Leucine, isoleucine, and valine have similar structures and are commonly referred to as the branched-chain amino acids (BCAA). Antagonism among these amino acids has been established in avian species such as chickens and turkeys [1]. The interactions among the BCAA have shown that dietary leucine content influences both valine and isoleucine requirements of broilers. Dietary protein is often the most expensive component of a broiler diet. Reducing dietary protein in an effort to decrease production costs can be accomplished via supplementation with crystalline amino acids such as lysine, methionine, and threonine. In low protein diets, adequate dietary valine is critical for supporting optimal growth, feed conversion, and carcass traits [2].

Mathematical modeling stands as a powerful tool to explore information further and to orient future research programs in animal nutrition. Among different mathematical approaches, neural networks (NN) have become popular for predicting and forecasting in a broad area of sciences [3]. Similar to the physiological nervous system, a

✉ J. France
jfrance@uoguelph.ca

1 Department of Animal Science, Islamic Azad University, Sanandaj Branch, Sanandaj, Kurdistan, Iran

2 Gharbdaneh Broiler Chain Production, Sanandaj, Kurdistan, Iran

3 Department of Animal Biosciences, Centre for Nutrition Modelling, University of Guelph, Guelph, ON N1G 2W1, Canada

NN model can 'learn' and therefore be trained to find solutions, recognize patterns, classify data, and forecast future events [4].

A multilayer perceptron (MLP) network, sometimes called a back-propagation (BP) network, is probably the most popular NN for nonlinear mapping and has been referred to as a 'universal approximator' [5]. It consists of an input layer, one (or more) hidden layer and an output layer. The network needs to be trained using a training algorithm (e.g., back propagation, cascade correlation, conjugate gradient). In radial basis function (RBF) networks, as the name implies, a radially symmetric basis function is used as an activation function for the hidden nodes. Training of the network parameters (weights) between the hidden and output layers occurs in a supervised fashion based on target outputs [6]. Despite frequent reports of successful applications of NN for modeling purposes, concerns still exist about their construction and exploitation. The main concern is about network structure. What the optimal NN structure is and how it can be determined are still issues requiring further investigation.

In recent years, the support vector machine (SVM) has been introduced as a new technique for solving a variety of learning, classification, and prediction problems [7]. Support vector regression (SVR), the regression version of SVM, has been developed recently to estimate regression relationships. As with SVM, SVR is capable of solving nonlinear problems using kernel functions [3]. SVR is a nonparametric statistical learning technique in which no assumption is made about the underlying data distribution [7, 8]. Due to its capability to generalize, SVR has attracted the attention of many researchers and has been applied successfully in a range of disciplines [9].

This study set out to predict the response of broiler chickens [average daily gain (ADG), feed efficiency (FE), and feed intake (FI)] to different levels of dietary protein and BCAA. For this purpose, SVR response models with different kernel functions were developed and their predictive performance compared with that of MLP and RBF types of NN model. The reader is referred to two recent textbooks for review of the pros and cons of the two approaches [10, 11].

2 Materials and methods

This study was carried out in four stages: (i) compilation of data required for training and testing; (ii) training of the different SVR and NN models; (iii) testing the models generated with a data set not used in training; and (iv) assessing their performance and comparing the predictive ability of SVR and NN models.

2.1 Data source

Seven data sets were extracted from the literature, providing 241 data lines (pens of birds, average number of birds per pen = 22). The input variables were dietary protein, valine, isoleucine, and leucine all expressed as g/kg of feed, and the output variables were ADG (g/bird per day), FE (g gain to g feed intake), and FI (g/bird per day). The data were selected based on the following criteria: (i) peer-reviewed published papers were used; (ii) the data were collected from studies in which the BCAA concentrations in the diets were used as treatments; (iii) the dietary levels of protein and investigated amino acids were clearly defined; (iv) the ADG, FE, and FI were reported or could be calculated from the published data; and (v) all experimental data were collected from studies conducted during the first 21 days of age of broiler chickens. A brief description of the seven data sources used in this study is given in Table 1. The complete data set was randomly divided into two groups of training and testing (50% each). The same data were used to train the SVR and NN models, and the models developed were tested with identical data. Therefore, differences in the predictive performance of models were regardless of data. The ranges of data used to develop the NN and SVR models are summarized in Table 2, and the correlation matrix for the variables used in the study is shown in Table 3.

2.2 Neural network modeling

The collected experimental data were used to train and test two NN (MLP and RBF) for predicting responses of broiler

Table 1 Data sets used in this study

Data set	Strain	ME (kcal/kg)	Sex	Age	No. of birds	No. of data lines
Burnham et al. [35]	ND	3100	M, F	7–21	1760	162
Waldroup et al. [36]	Cobb	3200	M	1–21	468	14
Farran et al. [37]	R × AA	3200	M	7–21	288	8
Barbour and Latshaw [38]	ND	3200	M, F	1–21	288	12
Farran and Thomas [39]	R × AA	3200	M	1–21	960	30
Tavernari et al. [40]	Cobb	3000	M	8–21	1400	7
Farran and Thomas [41]	R × AA	3200	M	1–21	192	8

ND not defined, *R × AA* Ross × Arbor Acres, *M* male, *F* female

Table 2 Ranges of the data used to develop the SVR and NN models

Variable	Min	Max	Average	Standard deviation
Input				
Protein (g/kg of diet)	102.75	274	190.4	48.54
Isoleucine (g/kg of diet)	3.6	11.6	7.07	1.79
Leucine (g/kg of diet)	9.6	50.3	22.87	10.71
Valine (g/kg of diet)	1	17.2	10.87	3.50
Output				
ADG (g/bird/day)	7.3	36.1	22.8	5.85
FE (g gain to g intake)	0.19	0.77	0.52	0.12
FI (g/bird/day)	26.63	58.6	43.53	6.41

ADG average daily gain, *FE* feed efficiency, *FI* feed intake

Table 3 Correlations among the variables used in this study

	CP	Ile	Leu	Val	ADG	FE	FI
CP	1	0.88	0.62	0.85	0.81	0.74	0.23
Ile	0.88	1	0.43	0.65	0.73	0.78	0.02
Leu	0.62	0.43	1	0.79	0.35	0.12	0.36
Val	0.85	0.65	0.79	1	0.64	0.39	0.48
ADG	0.81	0.73	0.35	0.64	1	0.81	0.39
FE	0.74	0.78	0.12	0.39	0.81	1	−0.17
FI	0.23	0.02	0.36	0.48	0.39	−0.17	1

CP crude protein, *Ile* isoleucine, *Leu* leucine, *Val* valine, *ADG* average daily gain, *FE* feed efficiency, *FI* feed intake

chickens to dietary levels of protein and BCAA. The 241 data lines assembled were shuffled, and 120 were used for the learning process and 121 for testing. The data set was imported into the Statistica Neural Networks software version 8 [12]. The data lines were the same as those used to develop the SVR models. ADG, FE, and FI were considered as dependent variables and dietary protein, valine, isoleucine, and leucine as independent. A common problem in NN training is over-fitting [13]. A network with a large number of weights compared to number of training cases available may achieve a low training error simply through modeling a function that fits the training data well. To circumvent this issue, the optimum architecture (i.e., the number of hidden neurons in the network and training algorithm) was determined using the algorithms integrated within in the 'intelligent problem solver' module of Statistica software [12]. A schematic diagram of the NN models developed in this study is shown in Fig. 1a.

2.3 Support vector regression

Here a brief description of SVR is given, following [14]. Like most linear regression procedures, the SVR algorithm, developed by [8], is based on estimating a linear regression function:

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + b \;\left(\mathbf{w},\, \mathbf{x} \in R^{d}(d\text{−dimensional input space})\right) \tag{1}$$

where $\mathbf{w}$ and b are the slope and offset of the regression line, respectively. The regression function (Eq. 1) is calculated by minimizing:

$$\frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + \frac{1}{n}\sum_{i=1}^{n} c(f(\mathbf{x_i}), y_i) \tag{2}$$

where $\frac{1}{2}\|\mathbf{w}\|^2$ is the term characterizing model complexity (i.e., smoothness of $f(\mathbf{x})$, where $\|\ldots\|$ denotes vector length), and $c(f(\mathbf{x}_i), y_i))$ is the loss function which determines how the distance between $f(\mathbf{x_i})$ and the target value y_i should be penalized. In this primal formulation, several different loss functions are available, but in this paper, we adopt the commonly used ε-insensitive loss function introduced by [7]. This loss function is given by:

$$c(f(\mathbf{x_i}), y_i) = \begin{cases} 0 & \text{if } |y_i - f(\mathbf{x_i})| \le \varepsilon \\ |y_i - f(\mathbf{x_i})| - \varepsilon, & \text{otherwise} \end{cases} \tag{3}$$

Equation 3 defines a tube of radius ε around the hypothetical regression function such that if a data point lies within this tube the loss function equals zero, while if a data point lies outside the tube, the loss is equal to the distance between the data point and the radius ε of the tube.

In this particular case, minimizing Eq. 2 is equivalent to solving the goal programming problem [14, 15]:

$$\begin{aligned} &\text{minimize } \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ &\text{subject to } \begin{cases} y_i - \mathbf{w}^{\mathrm{T}}\mathbf{x_i} - b \le \varepsilon + \xi_i \\ \mathbf{w}^{\mathrm{T}}\mathbf{x_i} + b - y_i \le \varepsilon + \xi_i^* \\ \varepsilon, \xi_i, \xi_i^* \ge 0 \end{cases} \end{aligned} \tag{4}$$

where ξ_i and ξ_i^* are slack variables and the constant $C > 0$ determines the trade-off between model complexity and tolerance of deviations larger than ε. The goal programming problem (Eq. 4) can be expressed in its dual form

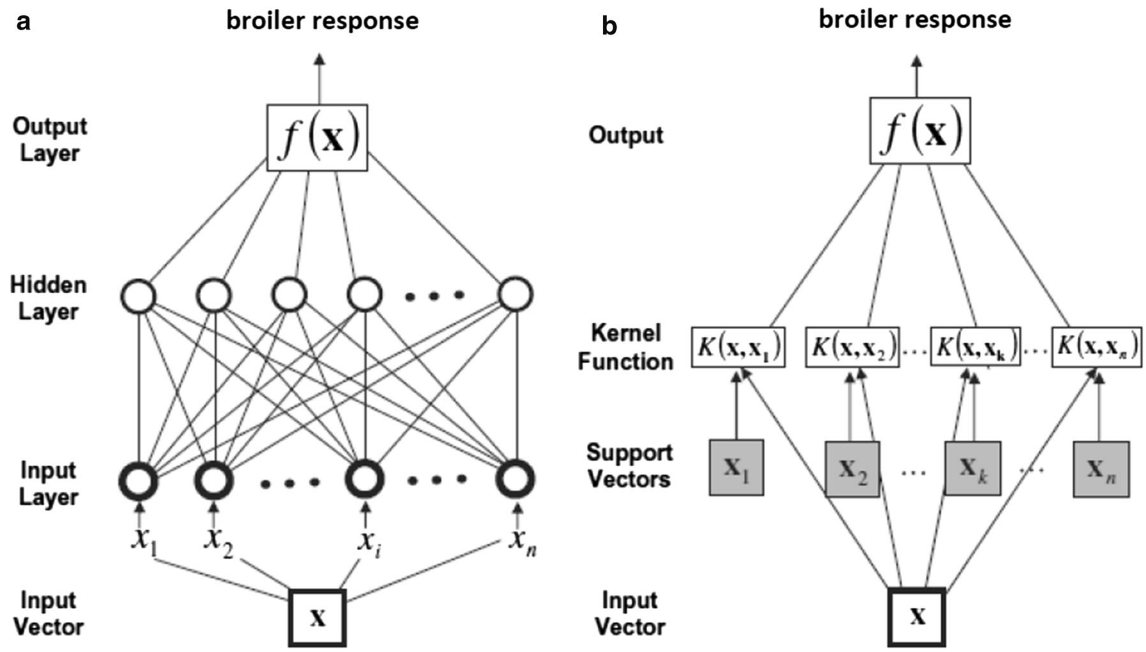

Fig. 1 Schematic diagram of the models used in this study: **a** NN and **b** SVR

using Lagrange multipliers [16]. In this paper, we use the strategy outlined by [7] leading to the solution:

$$f(\mathbf{x}) = \sum_{i=1}^{n} (\lambda_i - \lambda_i^*) K(\mathbf{x_i}, \mathbf{x}) + b \tag{5}$$

where λ_i and λ_i^* $(0 \leq \lambda_i, \lambda_i^* \geq C)$ are the Lagrange multipliers and $K(\mathbf{x_i}, \mathbf{x})$ represents the kernel function [17]. In the context of Eq. 5, data points with nonzero λ_i and λ_i^* values are support vectors. A suitable kernel function makes it possible to map a nonlinear input space to a high-dimensional feature space where linear regression can be performed [7]. Several kernel functions have been proposed in literature. In this paper, different kernel functions were used, viz. linear, RBF, polynomial (second and third order), and sigmoidal. The kernel parameters must be selected appropriately by the user, as generalization performance of the SVR model depends heavily on correct setting of these parameters. For a more detailed description of kernel functions and parameters, the reader is referred to [18]. A schematic diagram of SVR models used in this study is shown in Fig. 1b.

2.3.1 Support vector regression training

SVR was implemented in the Statistica software [12]. The SVR parameters were obtained following a tenfold cross-validation experiment. This algorithm is a way to improve on the holdout method which is the simplest variant of k-fold cross-validation [19]. The data set is divided into k subsets, and the holdout method is repeated k times. Each time, one of these subsets is used as the test set while the remaining subsets are assembled to form a training set. Then the average error across all k trials is computed. The advantage of this method is that it matters less how the data are divided. Every data point gets to be in a test set exactly once and gets to be in a training set $k - 1$ times. The variance of the estimate is reduced as k is increased. The disadvantage of this approach is that the training algorithm has to be rerun from scratch k times, which means it requires k times as much computation to make an evaluation. A variant of this method is to randomly divide the data into a test and training set k different times [20].

2.4 Performance evaluation

Goodness of fit of the NN and SVR models was based on coefficient of determination (R^2), mean square error (MSE), mean absolute error (MAE), and bias [21]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (x_{pi} - x_{ei})^2}{\sum_{i=1}^{n} (x_{pi} - \bar{x})^2}$$

$$\mathrm{MSE} = \frac{1}{n} \sum_{i=1}^{n} (x_{pi} - x_{ei})^2$$

$$\mathrm{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left| x_{pi} - x_{ei} \right|$$

$$\mathrm{Bias} = \frac{1}{n} \sum_{i=1}^{n} (x_{pi} - x_{ei})$$

where x_{pi} is the predicted output for observation i, x_{ei} is the experimental output for observation i, $\bar{x}$ is the average value of the experimental output, $|\ldots|$ denotes modulus, and n is the total number of observations.

3 Results and discussion

The ability of NN models to predict the response of broiler chickens to different dietary nutrients and to predict the energy values of concentrate feedstuffs have been demonstrated previously [22–25]. However, little is known about the suitability or otherwise of SVR models in animal (especially poultry) nutrition. An attempt in this area was a study by [23] of the ability of SVR models to estimate carcass characteristics of broilers influenced by dietary nutrient intakes. In that study, SVR models were developed using only the RBF kernel function and the results showed that support vectors are worth considering as a possible alternative to NN models. As several indices can be used for model evaluation, it is difficult to identify an optimum model that satisfies, at best, every single evaluation method. Therefore, emphasis in the present study was given to the accuracy of models developed in relation to the aforementioned criteria.

3.1 Average daily gain

Kernel parameters for predicting ADG of the broiler chickens in response to protein and BCAA are summarized in Table 4. These values may be considered as initial values when developing future SVR models. The optimal architecture for the MLP-type NN for modeling ADG was found with 4 inputs, 1 output (with linear activation function), and 4 hidden neurons (with the hyperbolic tangent activation function), using a quasi-Newton algorithm [26] for network training. The optimal structure for the RBF-type NN model was found with 4 inputs, 1 output, and 11 hidden neurons. It seems that all the models developed are able to predict ADG of broiler chickens satisfactorily. Scatter plots of predicted versus actual and residual versus predicted values of ADG obtained with the MLP-type NN and linear SVR models are shown in Fig. 2. The MLP-type NN and linear kernel function were considered the best and worst models, respectively, in predicting the broiler chickens' ADG. The regression equations and corresponding low R^2 values for the distribution of residual versus predicted values suggest there is little or no evidence of prediction bias.

Statistics for assessing the performance of each model are shown in Table 5. Based on these criteria, the MLP-type NN models showed a higher coefficient of determination compared to the SVR models. Among SVR models developed with different kernels, best and worst prediction was achieved using the polynomial (third order) and the linear kernel functions, respectively. The performance of SVR models with RBF and polynomial (third order) was better than that of RBF-type NN models (Table 5). Our results showed that the MLP-type NN model is more accurate for predicting ADG than the RBF type. The results revealed good agreement between observed and predicted values of ADG for the training and testing sets. A well-trained model gives balanced values of the evaluation statistics for these two sets, suggesting that over-fitting has not occurred during the training process [24]. In agreement with our results, previous studies have shown the superiority of NN to SVR models in predicting the investigated output variables [27].

Table 4 Kernel function parameters (C, ε, γ) for the SVR models developed to predict average daily gain

Kernel function	C	ε	γ
Linear	11	0.03	–
Polynomial ($d = 2$)	8	0.03	0.25
Polynomial ($d = 3$)	9	0.05	0.3
RBF	8	0.2	0.25
Sigmoidal	10	0.1	0.4

RBF radial basis function

3.2 Feed efficiency

In general, feed costs account for two-thirds of end-table costs in poultry production [28]. In birds, FE can be calculated as gram of gain per gram of feed intake and adequate knowledge and prediction of FE can therefore help provide economic benefits [29].

The kernel parameters to estimate FE are summarized in Table 6. The optimal architecture of the MLP- and RBF-type NN for modeling FE, suggested by intelligent problem solver, was found to be similar to that of the models for ADG. The MLP and RBF models were developed with 4 input variables and had 4 and 11 neurons in the hidden layer, respectively. Scatter plots of predicted versus actual and residual versus predicted values of FE obtained with the MLP-type NN and sigmoidal SVR models are shown in Fig. 3. The regression equations and corresponding low R^2 values for the distribution of residual versus predicted values suggest there is no evidence of prediction bias. As this figure illustrates, the model with the higher R^2 value for the distribution of observed versus predicted FE had the lower R^2 value for the distribution of residual versus predicted values, and vice versa.

The statistics used to assess performance of the models are shown in Table 7, and these indices indicate similar trends in goodness of fit across models. Overall performance appears better for FE than for ADG (Tables 5, 7).

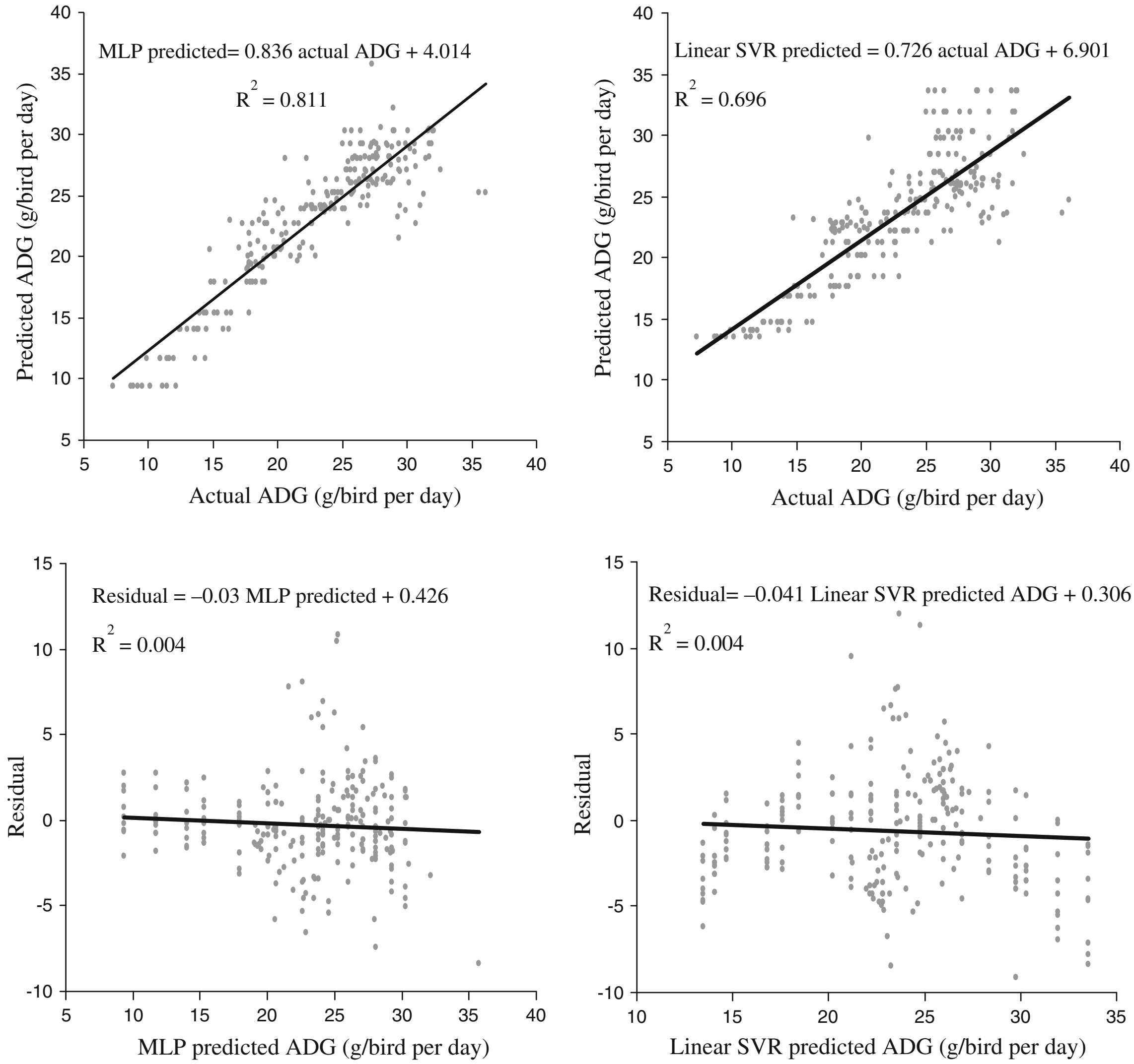

Fig. 2 Scatter plots of predicted versus actual (*top row*) and residual versus predicted (*bottom row*) values of ADG obtained by best (MLP-type NN) and worst (linear SVR) models developed (all 241 data lines represented)

Table 5 Accuracy of the SVR and NN models developed to predict average daily gain

Model	Training set				Testing set			
	R^2	MSE	MAE	Bias	R^2	MSE	MAE	Bias
Linear SVR	0.71	10.6	2.6	−0.34	0.69	11.7	2.62	−0.98
Polynomial SVR ($d = 2$)	0.75	9.1	2.21	0.36	0.73	9.53	2.26	−0.27
Polynomial SVR ($d = 3$)	0.77	8.45	2.15	0.29	0.77	8.42	2.17	−0.34
RBF SVR	0.76	8.9	2.3	0.08	0.76	8.5	2.19	−0.45
Sigmoidal SVR	0.71	11.14	2.65	0.22	0.73	9.9	2.45	−0.31
MLP-type NN	0.81	6.7	1.83	0.01	0.81	6.86	1.88	−0.54
RBF-type NN	0.75	8.81	2.29	0	0.75	8.7	2.25	−0.33

MSE mean square error, *MAE* mean absolute error, *SVR* support vector regression, *RBF* radial basis function, *NN* neural network

For the training set, best performance was achieved with the MLP-type NN model ($R^2 = 0.87$) followed by SVR with RBF and third-order polynomial kernel functions ($R^2 = 0.85$), second-order polynomial SVR- and RBF-type NN ($R^2 = 0.82$), linear SVR ($R^2 = 0.76$), and sigmoidal SVR ($R^2 = 0.71$), respectively. For the testing set, highest

Table 6 Kernel function parameters (C, ε, γ) for the SVR models developed to predict feed efficiency

Kernel function	C	ε	γ
Linear	10	0.01	–
Polynomial ($d = 2$)	10	0.02	0.25
Polynomial ($d = 3$)	11	0.04	0.3
RBF	12	0.05	0.25
Sigmoidal	8	0.02	0.3

RBF radial basis function

accuracy was achieved for MLP-type NN and SVR models with RBF and third-order polynomial kernel functions ($R^2 = 0.87$). Accuracy in prediction, as well as the flexibility of constructed models, affirms the strong effect of the selected dietary nutrients (protein, valine, isoleucine, and leucine) on FE in broiler chickens [30, 31]. In other words, the statistical analyses indicated that it is possible to use dietary protein and BCAA to predict the FE of broilers accurately.

3.3 Feed intake

Lehninger [32] classifies leucine as a ketogenic amino acid as it yields the ketone body acetoacetate. Valine is a glycogenic amino acid as its degradation yields succinate, an intermediate in the TCA cycle. Isoleucine is classified as both a ketogenic and glycogenic amino acid as it is degraded in the body to produce acetyl-COA and succinate. Ketone bodies act as an inhibitory signal in both the central and peripheral nervous systems for feed intake of poultry [33]. Moreover, [34] demonstrates that excess leucine severely depresses feed consumption and weight gain of chicks fed a low protein diet. For these reasons, it was

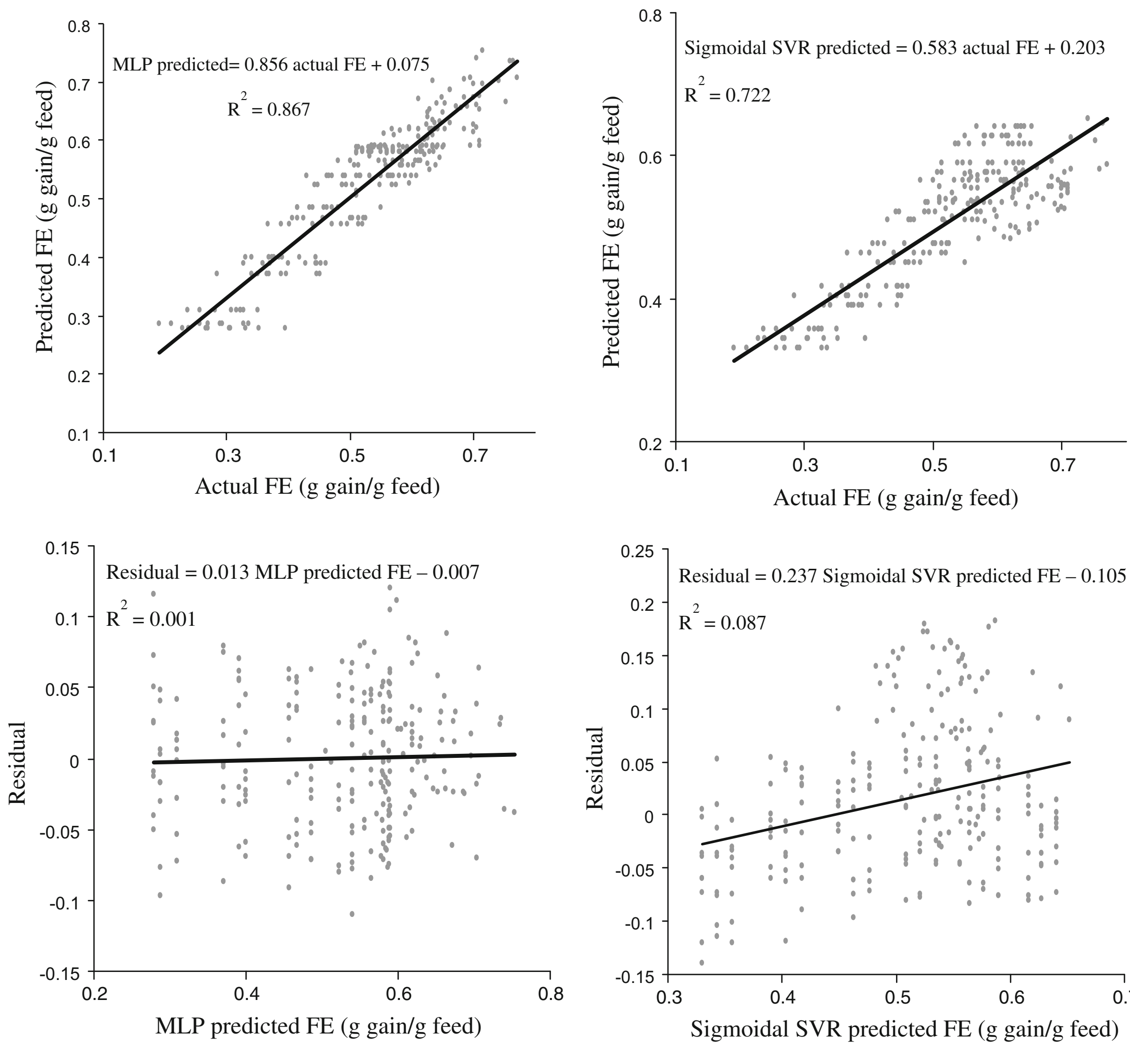

Fig. 3 Scatter plots of predicted versus actual (*top row*) and residual versus predicted (*bottom row*) values of feed efficiency (FE) obtained by best (MLP-type NN) and worst (sigmoidal SVR) models developed (all 241 data lines represented)

Table 7 Accuracy of the SVR and NN models developed to predict feed efficiency

Model	Training set				Testing set			
	R^2	MSE	MAE	Bias	R^2	MSE	MAE	Bias
Linear SVR	0.76	0.004	0.045	0.003	0.82	0.003	0.048	0.015
Polynomial SVR ($d = 2$)	0.82	0.003	0.04	−0.003	0.86	0.002	0.04	0.005
Polynomial SVR ($d = 3$)	0.85	0.002	0.038	−0.002	0.87	0.002	0.039	0.002
RBF SVR	0.85	0.002	0.039	−0.004	0.87	0.002	0.04	0.004
Sigmoidal SVR	0.71	0.005	0.065	0.011	0.74	0.005	0.057	0.02
MLP-type NN	0.87	0.002	0.035	0.00	0.87	0.002	0.041	0.00
RBF-type NN	0.82	0.003	0.042	0.00	0.82	0.003	0.045	0.013

MSE mean square error, *MAE* mean absolute error, *SVR* support vector regression, *RBF* radial basis function, *NN* neural network

Table 8 Kernel function parameters (C, ε, γ) for the SVR models developed to predict feed intake

Kernel function	C	ε	γ
Linear	3	0.4	–
Polynomial ($d = 2$)	10	0.4	0.25
Polynomial ($d = 3$)	9	0.4	0.25
RBF	10	0.05	0.3
Sigmoidal	10	0.2	0.3

RBF radial basis function

assumed that FI of broilers could be predicted in response to different levels of dietary protein and BCAA. However, results showed that overall performance of the models was not as good as for those developed to predict ADG and FE. Our selected input variables (protein and BCAA) only covered a maximum 0.68 of FI variation (R^2 for MLP-type NN). The kernel parameters for predicting FE are summarized in Table 8. The MLP- and RBF-type NN models developed were composed of 5 and 11 hidden neurons, respectively. For the FI models, as for ADG and FE, best performance was achieved with MLP-type NN and SVR with a RBF kernel function. The prediction ability of all SVR models, except for the sigmoidal kernel function, was higher than for the RBF-type NN model. The behavior of the MLP-type NN (as the best fitting) and sigmoidal SVR (as the worst fitting) models is shown in Table 9. Figure 4 shows scatter plots of predicted versus actual and residual versus predicted values of FI obtained with the MLP-type NN and sigmoidal SVR models. With the FI models, as with the FE models, the model with the higher R^2 value for distribution of observed versus predicted FE provided the lower R^2 value for distribution of predicted versus residual values.

4 Conclusions

In this paper, we suggest SVR as a means of estimating the response of broilers to dietary protein and BCAA. The SVR approach was compared to MLP- and RBF-type NN models. Based on the results of this study, we conclude that it is feasible to apply SVR and NN models to predict the performance response of broiler chickens (in terms of ADG, FE, and FI) to dietary protein and BCAA. Models derived using these two approaches (SVR and NN) offer an alternative to those obtained from the standard statistical meta-analysis employed in animal nutrition [42, 43] and

Table 9 Accuracy of the SVR and NN models developed to predict feed intake

Model	Training set				Testing set			
	R^2	MSE	MAE	Bias	R^2	MSE	MAE	Bias
Linear SVR	0.42	25.17	4.27	−0.43	0.33	29.5	4.35	−1.83
Polynomial SVR ($d = 2$)	0.43	24.02	4.35	0.015	0.37	26.3	4.1	−1.3
Polynomial SVR ($d = 3$)	0.46	22.9	4.29	0.19	0.39	24.95	4.05	−0.99
RBF SVR	0.46	22.8	3.9	−0.06	0.48	21.5	3.8	−0.99
Sigmoidal SVR	0.27	32.2	4.8	0.09	0.20	33.3	4.6	−1.15
MLP-type NN	0.68	13.25	2.98	0.02	0.62	15.45	2.45	−0.31
RBF-type NN	0.41	24.73	4.39	0.00	0.39	24.76	4.1	−1.01

MSE mean square error, *MAE* mean absolute error, *SVR* support vector regression, *RBF* radial basis function, *NN* neural network

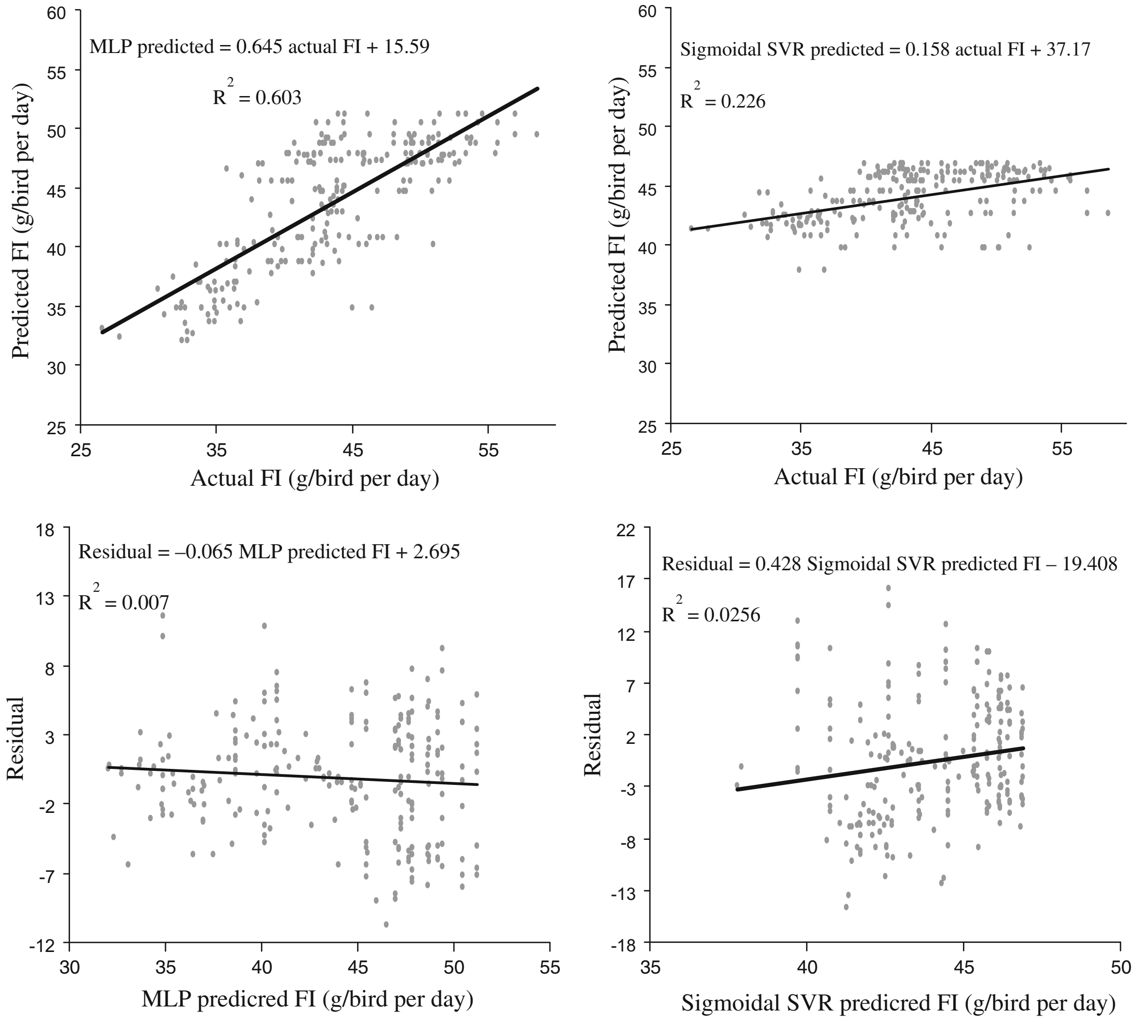

Fig. 4 Scatter plots of predicted versus actual (*top row*) and residual versus predicted (*bottom row*) values of feed intake (FI) obtained by best (MLP-type NN) and worst (sigmoidal SVR) models developed (all 241 data lines represented)

might well provide general models for common usage. Further research, however, is needed to explore these assertions. MLP-type NN models appear to provide a better option than SVR to estimate the output variables investigated. Among the different kernel functions applied here, the RBF and polynomial (third order) functions gave better performance than other SVR kernel functions and RBF-type NN models. However, there remains a paucity of literature information on application of SVR in poultry nutrition and more studies are needed to examine further the use of these models.

Acknowledgements The Canada Research Chairs program (National Science and Engineering Council, Ottawa) is thanked for part funding.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Smith TK, Austic RE (1978) The branched-chain amino acid antagonism in chicks. J Nut 108:1180–1191
2. Corzo A, Dozier WA III, Kidd MT (2008) Valine nutrient recommendations for Ross × Ross 308 broilers. Poult Sci 87:335–338
3. Cheng CS, Chen PW, Huang KK (2011) Estimating the shift size in the process mean with support vector regression and neural networks. Expert Syst Appl 38:10624–10630
4. Bishop CM (1996) Neural networks for pattern recognition. Oxford University Press, Oxford
5. Wu CH (1997) Artificial neural networks for molecular sequence analysis. Comput Chem 21:232–256
6. Haykin S (1998) Neural networks: a comprehensive foundation, 2nd edn. Prentice Hall PTR, Upper Saddle River
7. Vapnik V (1995) The nature of statistical learning theory. Springer, New York
8. Vapnik V (1998) Statistical learning theory. Wiley, New York
9. Anand A, Suganthan PN (2009) Multiclass cancer classification by support vector machines with class-wise optimized genes and probability estimates. J Theor Biol 259:533–540

10. Heaton J (2015) Artificial intelligence for humans, volume 3: deep learning and neural networks. Heaton Research Inc., Chesterfield
11. Deng N, Tian Y, Zhang C (2013) Support vector machines. Chapman & Hall, Boca Raton
12. StatSoft (2009) Statistica data analysis software system, version 7.1. StatSoft Inc., Tulsa
13. Lawrence S, Giles CL, Tsoi AC (1997) Lessons in neural network training: over-fitting may be harder than expected. Artif Intell 97:540–545
14. Ustun B, Melssen WJ, Oudenhuijzen M, Buydens LMC (2005) Determination of optimal support vector regression parameters by genetic algorithms and simplex optimization. Anal Chim Act 544:292–305
15. Thornley JHM, France J (2007) Mathematical models in agriculture, 2nd edn. CAB International, Wallingford
16. Bertsekas DP (1999) Nonlinear programming, 2nd edn. Athena Scientific, Cambridge
17. Belousov AI, Verzakov SA, VanFrese J (2002) A flexible classification approach with optimal generalization performance: support vector machines. Chemom Intell Lab Syst 64:15–25
18. Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, Cambridge
19. Arlot S, Celisse A (2010) A survey of cross-validation procedures for model selection. Stat Surv 4:40–79
20. Polat K, Gunes S (2007) Breast cancer diagnosis using least square support vector machine. Digit Signal Process 17:694–701
21. Nazghelichi T, Aghbashlo M, Kianmehr MH (2011) Optimization of an artificial neural network topology using coupled response surface methodology and genetic algorithm for fluidized bed drying. Comput Electron Agric 75:84–91
22. Faridi A, Golian A, France J (2012) Evaluating the egg production behavior of broiler breeder hens in response to dietary nutrient intake during 31–60 weeks of age through neural network models. Can J Anim Sci 92:473–481
23. Faridi A, Sakomura NK, Golian A, Marcato SM (2012) Predicting body and carcass characteristics of 2 broiler chicken strains using support vector regression and neural network models. Poult Sci 91:3286–3294
24. Faridi A, Golian A, Heravi Mousavi A, France J (2014) Bootstrapped neural network models for analyzing the responses of broiler chicks to dietary protein and branched chain amino acids. Can J Anim Sci 94:79–85
25. Mariano FCMQ, Paixão CA, Lima RR, Alvarenga RR, Rodrigues PB, Nascimento GAJ (2013) Prediction of energy values of feedstuffs for broilers using meta-analysis and neural networks. Animal 7:1440–1445
26. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical recipes: the art of scientific computing, 3rd edn. Cambridge University Press, New York
27. King SL, Bennett KP, List S (2000) Modeling non-catastrophic individual tree mortality using logistic regression, neural networks, and support vector methods. Comput Electron Agric 27:401–406
28. Lemme A, Frackenpohl U, Petri A, Meyer H (2006) Response of male BUT big 6 turkeys to varying amino acid feeding programs. Poult Sci 85:652–660
29. Faridi A, Mottaghitalab M, Rezaee F, France J (2011) Narushin-Takma models as flexible alternatives for describing economic traits in broiler breeder flocks. Poult Sci 90:507–515
30. Park BC, Austic RE (2000) Isoleucine imbalance using selected mixtures of imbalancing amino acids in diets of the broiler chick. Poult Sci 79:1782–1789
31. Corzo A, Kidd MT, Dozier WA III, Vieira SL (2007) Marginality and needs of dietary valine for broilers fed certain all-vegetable diets. J Appl Poult Res 16:546–554
32. Lehninger AL (1981) Oxidative degradation of amino acids. Biochemistry, 2nd edn. Worth Publishers, New York, pp 559–586
33. Sashihara K, Miyamoto M, Ohgushi A, Denbow DM, Furuse M (2001) Influence of ketone body and the inhibition of fatty acid oxidation on the food intake of the chick. Br Poult Sci 42:405–408
34. Fisher H, Griminger P, Leveille GA, Shapiro R (1960) Quantitative aspects of lysine deficiency and amino acid imbalance. J Nutr 71:213–220
35. Burnham D, Emmans GC, Gous RM (1992) Isoleucine requirements of the chicken: the effect of excess leucine and valine on the response to isoleucine. Br Poult Sci 33:71–87
36. Waldroup PW, Kersey JH, Fritts CA (2002) Influence of branched-chain amino acid balance in broiler diets. Int J Poult Sci 1:136–144
37. Farran MT, Brabour EK, Ashkarian VM (2003) Effect of excess leucine in low protein diet on ketosis in 3-week-old male broiler chicks fed different levels of isoleucine and valine. Anim Feed Sci Technol 103:171–176
38. Barbour G, Latshaw JD (1992) Isoleucine requirements of broiler chicks as affected by the concentration of leucine and valine in practical diets. Br Poult Sci 33:561–568
39. Farran MT, Thomas OP (1990) Dietary requirements of leucine, isoleucine, and valine in male broilers during the starter period. Poult Sci 69:757–762
40. Tavernari FC, Lelis GR, Vieira RA, Rostagno HS, Albino LFT, Oliveira Neto AR (2013) Valine needs in starting and growing Cobb (500) broilers. Poult Sci 92:151–157
41. Farran MT, Thomas OP (1992) Valine deficiency: 1. The effect of feeding a valine-deficient diet during the starter period on performance and feather structure of male broiler chicks. Poult Sci 71:1879–1884
42. St-Pierre NR (2001) Integrating quantitative findings from multiple studies using mixed model methodology. J Dairy Sci 84:741–755
43. Faridi A, Gitoee A, France J (2015) A meta-analysis of the effects of nonphytate phosphorus on broiler performance and tibia ash concentration. Poult Sci 94:2753–2762