

# A Gaussian process regression model to predict energy contents of corn for poultry

Abbas Abdullah Baiz,<sup>\*</sup> Hamed Ahmadi,<sup>\*,†,1</sup> Farid Shariatmadari,<sup>\*</sup> and  
Mohammad Amir Karimi Torshizi<sup>\*</sup>

<sup>\*</sup>Department of Poultry Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran; and <sup>†</sup>Bioscience and Agriculture Modeling Research Unit, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran

**ABSTRACT** The present study proposes a Gaussian process regression (**GPR**) approach to develop a model to predict true metabolizable energy corrected for nitrogen (**TMEn**) content of corn samples (as model output) for poultry given levels of feed chemical compositions of crude protein, ether extract, crude fiber, and ash (as model inputs). A 30 corn samples obtained from 5 origins [Brazil (n = 9), China (n = 5), Iran (n = 7), and Ukraine (n = 9)] were assayed to determine chemical composition and TMEn content using chemical analyses and bioassay technique. In addition to GPR model, data were also analyzed by multiple linear regression (**MLR**) model. Results revealed that corn samples of different origins

differ in their gross energy and chemical composition of crude protein, crude fiber, and ash, but no differences were observed for their ether extract and TMEn contents. Based on model evaluation criteria of  $R^2$  and root mean square error (**RMSE**), the GPR model showed satisfactory performance ( $R^2 = 0.92$  and  $RMSE = 33.68$  kcal/kg DM) in predicting TMEn and produced relatively better prediction values than those produce by MLR ( $R^2 = 0.23$  and  $RMSE = 104.85$  kcal/kg DM). The GPR model may be capable of improving our aptitude and capacity to precisely predict energy contents of feed ingredients to formulate optimal diets for poultry.

**Key words:** corn metabolizable energy, Gaussian process regression, prediction model

2020 Poultry Science 99:5838–5843

<https://doi.org/10.1016/j.psj.2020.07.044>

## INTRODUCTION

Nowadays, several high performance corn varieties with wide range of grain characteristics and composition have been developed through genetic improvements (Collins et al., 2001). In common practice in North America and West Asia, corn is utilized as the key energy source for poultry, comprising 40 to 60% of the poultry diet (Klopfenstein et al., 2013). An important criteria for describing quality of a feed ingredient is its true metabolizable energy corrected for nitrogen (**TMEn**) value (Latshaw and Freeland, 2008; Dudley-Cash, 2009). The biological procedure used to determine the TMEn value of grain sample is time consuming and relatively expensive. An accurate and precise procedure for estimating TMEn value given feed chemical composition would be helpful; therefore, the nutritionists are

usually interested in applying an easy to use prediction model that may predict the TMEn values. Predictive models are widely accepted as informational tools to support rapid and cost effective assessment of feed. Using the adult rooster as a test birds, many researchers have shown that the energy content of corn was related with its chemical composition (Sibbald, 1976; Ertl and Dale, 1997; Latshaw and Freeland, 2008). Variations in protein content, fat, fiber, ash, and starch are observed in analysis of corn samples (Ertl and Dale, 1997; Collins et al., 2001). Traditionally, multiple linear regression (**MLR**) method are applied as a basic model for feedstuff evaluations. It has also been reported that the method of artificial neural network (**ANN**) and support vector machine models are useful to estimate TMEn contents of feedstuffs when enough data are available (Ahmadi et al., 2008; Ahmadi and Rodehutsord, 2017).

In recent years, Gaussian process (**GP**) based method such as GP regression (**GPR**) model has attracted much scientific attention in the area of computational data modeling (Swain et al., 2016; Schulz et al., 2018). Both the ANN and GPR methods have been reported as efficient tools in developing forecasting models and estimating predictions. The GPR model is used to solve

© 2020 Published by Elsevier Inc. on behalf of Poultry Science Association Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Received February 22, 2020.

Accepted July 17, 2020.

<sup>1</sup>Corresponding author: [hamed.ahmadi@modares.ac.ir](mailto:hamed.ahmadi@modares.ac.ir)

regression problems with GP and requires fewer data samples for modeling than those data needed for ANN model fitting (Richardson et al., 2017; Schulz et al., 2018; Huttunen et al., 2019; Li et al., 2019). The GPR processing also uses as a flexible model that may describe nonlinear relationship between inputs and the response variables as well as uncertainties in the data (Roberts et al., 2013; Swain et al., 2016; Tonner et al., 2017). Owing to methodological advantages, GPR model has been commonly used in many fields such as energy, electric power, meteorology, and medicine (Roberts et al., 2013; Swain et al., 2016; Richardson et al., 2017; Tonner et al., 2017). Because the GP models showed both suitable practical performance and appropriate analytical properties motivates the current work. To the best of our knowledge, GPR application in animal nutrition and feed evaluation fields was not reported yet.

The objective of this study was to 1) estimate the chemical composition and TMEn content of corn samples of different origins and test their variation and 2) to assess the ability of GPR model to predict TMEn contents of corn samples (as model output) given their chemical compositions (as model inputs). The performance of GPR model was compared with that of common regression model.

## MATERIALS AND METHODS

### Animal Ethics Statement

The research proposal and used procedures including working with birds were reviewed and approved by “Tarbiat Modares University Committee on Animal and Avian Care and Use” before starting the bio-assay experiments (Code B1398/3007). The agreement was in accordance with the general Guidelines for Experimental Animals developed by the Ministry of Science, Research, and Technology (Tehran, Iran). All efforts were made to minimize the suffering of the experimental birds.

### Bioassays and Data Collections

Thirty samples of corn grains submitted from the Brazil (n = 9), China (n = 5), Iran (n = 7), and Ukraine (n = 9) were assayed for their proximate composition based on standard methods for moisture (AOAC, 2000; method 930.15), CP (method 990.03), EE (method 960.39), CF (method 978.10), and ash (method 942.05). All corn samples were analyzed for gross energy (GE) with an automatic adiabatic oxygen bomb calorimeter (Parr Instrument Co., Moline, IL).

A bioassay according to precision-fed method for cockerel (Sibbald, 1976; Dudley-Cash, 2009) was used for determining the TMEn of the 30 corn samples. In total, 130 healthy Hy-Line roosters (BW = 2.3 ± 0.11 kg, 30 wk of age) were obtained from a local commercial farm. All birds were housed in individual wire cages (45 × 60 × 35 cm) for an acclimation and preconditioning period for 3 wk. A basal diet was fed during

acclimation phase (Table 1). Birds were maintained on a 16-h light schedule and allowed ad libitum access to water. The trays placed under each cage were used for collection of total excreta. Following a period of 24 h without feed, 25 g of the different ground corn samples were fed by intubation to 120 birds (4 per each corn sample). Remaining 10 roosters were fasted to estimate endogenous losses. Total excreta voided over the following 48-h period were collected, dried, and ground for subsequent analyses to estimate their GE contents.

A four-replication assessment was considered for both chemical and bioassay estimations for each corn sample, so a total of 120 chemical composition and TMEn determinations were obtained (30 samples × 4 replications = 120 data lines). All values of chemical and TMEn contents of corn samples were standardized based on respective dry matter (DM) and presented as DM basis.

### Statistical Analysis

Data were analyzed statistically by ANOVA using the general linear model option developed in Minitab software, version 17.0 (Minitab, 2014). Differences were considered significant at  $P < 0.05$ , and significant differences between means were separated by Tukey’s test.

### Building Prediction Model Using MLR

A MLR model was defined as the following general equation,

$$\hat{y}_i = \beta_0 + \sum_{i=1}^n \beta_i X_i + e_i, \quad i=1, 2, \dots, n \quad [1]$$

where:  $\hat{Y}_i$  is the TMEn in the  $i^{th}$  sample,  $X_i$  is the value related to input compositions (CP, EE, CF, and ash

**Table 1.** Basal diet used during a 3-wk acclimation phase for experimental roosters before starting the TMEn assay.

Ingredient	%
Corn	60.80
Soybean meal	33.70
Soybean oil	1.70
Dicalcium phosphate	2.16
Limestone	0.79
Sodium chloride	0.40
Vitamin-trace mineral premix <sup>1</sup>	0.30
HCl Lysine	0.08
DL Methionine	0.07
Calculated analysis	
AME, kcal/kg	2,900.00
CP, %	19.00
Ca, %	0.90
Available P, %	0.45
Lysine, %	1.1

Abbreviation: TMEn, true metabolizable energy corrected for nitrogen.

<sup>1</sup>Provided per kilogram of diet: Co, 0.3 mg; Cu, 5 mg; Fe, 25 mg; I, 1 mg; Mn, 125 mg; Zn, 60 mg; choline chloride, 638 mg; *trans*-retinol, 3.33 mg; cholecalciferol, 60 µg; dl- $\alpha$ -tocopheryl acetate, 60 mg; menadione, 4 mg; thiamine, 3.0 mg; riboflavin, 12 mg; niacin, 35 mg; calcium pantothenate, 12.8 mg; pyridoxine, 10 mg; cyanocobalamin, 0.017 mg; folic acid 5.2 mg; biotin, 0.2 mg; antioxidant, 100 mg; molybdenum, 0.5 mg; selenium, 200 µg.

concentration) in the  $i^{\text{th}}$  sample (assumed to be a known constant measured without error),  $\beta_0$  is the overall intercept,  $\beta_i$  is the linear coefficient for input variables, and  $e_i$  is the residual error assumed to be normal ( $N \sim (0, \sigma^2)$ ). Probability was considered significant when  $P < 0.05$ , and tendencies were mentioned if  $P \leq 0.10$ . Data were fitted to the equation 1 using the REG procedure of SAS software (Freud and Littell, 2000).

## Building Prediction Model Using GPR

The GPR model was developed to build the prediction model using the “fitgpr” function of Matlab version R2019b (Mathworks, 2019). A detailed description of GPR method terminology, development, and application is outside the scope of this article. Suggested references for comprehensive descriptions of the method are Rasmussen and Williams (2006), Schulz et al. (2018), and chapter 6 of MathWorks (2019).

Briefly, a GP is a nonparametric method given a stochastic process considering  $f(x)$  ( $x \in R^d$ ), where  $f(x_1), f(x_2), \dots, f(x_n)$  is a multivariate Gaussian random variable for all combinations of input variables ( $x_1, x_2, \dots, x_n$ ). In this way, the GPR model can be defined by introducing a mean function of the form  $\mu(z) = E(f(x))$  and a covariance function of the form  $k(x, x') = \text{cov}(f(x), f(x'))$ . Considering our case in which the inputs  $x$  are a vector of CP, EE, CF, and ash contents (all inputs values as %) of corn samples, whereas  $y$  is the response variable (TMEn content of corn samples). Hence, the response variable can be modeled as

$$y(x) = h(x)^T \beta + f(x) + \varepsilon \quad [2]$$

where  $h(x)$  is a vector of (deterministic) basis functions,  $\beta$  is a vector of basis function coefficients,  $f(x)$  is a GP with zero mean and covariance function  $k(x, x')$ , and  $\varepsilon$  is an Gaussian noise. The first term of equation 2 denotes mean behavior of the GP model. The GP term builds a nonlinear relationship between input and the response variable as well as elated uncertainties in the data.

Training data comprises of input-output pairs such as  $\{(x_i, y_i); i = 1, 2, \dots, N\}$ . Supposing  $y_i$ 's are output of the considered model (i.e.  $y_i = y(x_i)$ ), and  $X = (x_1, x_2, \dots, x_p)$  are inputs for which the predictions are calculated. Accordingly,  $Y = (y_1, y_2, \dots, y_N)$  and  $Y' = [y(x_1), y(x_2), \dots, y(x_p)]$  are both Gaussian. The conditional distribution of  $Y'$  based on  $Y$  can be defined as

$$p(Y'|Y) = N\left(\mu_{Y'} + \sum_{Y'}^{-1} \sum_Y^{-1} (Y - \mu_Y), \sum_{Y'} + \sum_{Y'}^{-1} \sum_Y^{-1} \sum_{Y'}^{-1}\right) \quad [3]$$

where  $\mu_Y$  and  $\Sigma_Y$  indicate the mean and covariance of  $Y$ , respectively, and  $\Sigma_{Y'Y}$  is the cross-covariance of  $Y$  and  $Y'$ . The means and covariances can be calculated by plugging

in the model (Rasmussen and Williams, 2006; Huttunen et al., 2019), as

$$\mu_{Y'|Y} = h(z')^T \beta + k(X, X') (k(X, X) + \sigma_\varepsilon^2 I)^{-1} (Y - h(X))^T \beta \quad [4]$$

$$\Sigma_{Y'|Y} = k(X', X') - k(X', X) (k(X, X) + \sigma_\varepsilon^2 I)^{-1} k(X, X') \quad [5]$$

where  $h(X')$  and  $k(X', X)$  are simplified symbolizations for the vector and matrix with the components  $h(x_i)$  and  $k(x_i, x_j)$ , respectively. Using the conditional mean function (equation 4), a prediction value for  $Y'$  can be computed along with a confidence approximation given by the conditional covariance (equation 5).

Choosing the functional form of the covariance (kernel) function ( $k(x, x')$ ) is mostly based on assumptions about the main function to be modeled. In this study, the most widely used covariance function namely “squared exponential function” was used. This kernel function can be written as (Roberts et al., 2013):

$$k(x, x') = \sigma^2 e^{-\frac{(x-x')^2}{\lambda^2}} \quad [6]$$

where  $\sigma^2$  is the variance, and  $\lambda$  is the length scales for each input (hyper-parameters).

To apply Gaussian processes in regression fitting, the hyperparameters of selected covariance function have to be optimized with respect to the experimental data (Rasmussen and Williams 2006). In this way, the Matlab's *fitrgp* function estimates hyperparameters of  $\theta$  ( $\beta, \sigma_\varepsilon^2, \sigma^2, \lambda$ 's) by minimizing the negative log-likelihood,

$$\mathcal{L}(\theta) = -\log \left( p(y|X, \theta) = 0.5 y^T \sum_{\theta}^{-1} y + 0.5 \log \det \sum_{\theta} + \frac{n}{2} \log 2\pi \right) \quad [7]$$

where  $\Sigma_{\theta} = k(X, X; \theta) + \sigma_\varepsilon^2 I$ . The default value of parameters in *fitrgp* (Mathworks, 2019) are chosen for optimizing process of hyperparameter in equation 7.

A k-fold cross-validation ( $k = 5$ ) method as described in chapter 24 of Mathworks (2019) was used on the data set to make validation sets. Validation set is applied to monitor performance of the trained GPR model. The fitting performance (in term of root mean square error [RMSE]) on the validation set was considered as an indicator for the generalization ability of model, whereas the model selection was also carried out using this criteria. The “GaussianProcess\_Corn” program coded by the Matlab, and the source code can be downloaded from [https://github.com/hahmadima/GaussianProcess\\_Corn](https://github.com/hahmadima/GaussianProcess_Corn).

## Models Performance Evaluation

Goodness-of-fit criteria were calculated to measure the model fits using following statistics: ratio of variation accounted by the model ( $R^2$ ) which was calculated as  $1 - \text{MSE}/S_y^2$ , where MSE is the mean square error (residual standard deviation) and  $S_y^2$  is the whole variance of the y variable, and the RMSE values which is the root of MSE (Ahmadi, 2017). The ranges are  $0 \leq R^2 \leq 1$  and  $0 \leq \text{RMSE} \leq +\infty$ . Obtained values closer to 1 for  $R^2$  and closer to 0 for RMSE indicate superior model performance. Model adequacy was also examined using plots of residuals (observed minus predicted) against predicted values of y to test for linear prediction bias (St-Pierre, 2001). To compare the experimental and GPR model predicted TMEn values, an additional one-way ANOVA was used to find if they are significantly different.

## RESULTS

The data for chemical composition and TMEn content of corn samples from different origins are summarized in Table 2. There were significant ( $P < 0.05$ ) differences between different origins for CP, CF, and ash contents of corn sample. Corn samples from China had higher ( $P < 0.01$ ) CP contents. The CF contents were lowest ( $P < 0.01$ ) in corn samples from Brazil. The ash contents were lowest ( $P < 0.05$ ) in Ukrainian corn, whereas ash contents in the Iran and China corn samples were intermediate. The GE of corn samples submitted from Ukraine were lowest ( $P < 0.01$ ). There was no difference ( $P > 0.05$ ) between the EE and TMEn content of samples from different origins (Table 2).

Disregarding the origins of corn samples, average and descriptive statistics values (whole data;  $n = 120$ ) of chemical composition (model inputs) and TMEn (model output) obtained from bioassay and chemical analyses are shown in Table 3. The first step in building a prediction model was to estimate the coefficients of equation 1 by least squares method using the matrix of experimental data and to obtain evidence about MLR model fitness in the form of an analysis of variance. The calculated MLR model on the corn data set was obtained as follow:

True metabolizable energy corrected for nitrogen of corn sample (kcal/kg of DM) =  $3906 (\pm 259) + 56.4 (\pm 13.6)$

CP +  $30.2 (\pm 19.1)$  EE -  $122.6 (\pm 39.9)$  CF -  $178.2 (\pm 49.1)$  Ash  $R^2 = 0.23$  RMSE = 104.85 kcal/kg of DM.

The calculated SE values of coefficients are given in parentheses. Values for input variables considered as percent of DM basis. All coefficient estimates except EE were highly different from zero ( $P < 0.05$ ), whereas the EE tend to be significant ( $P = 0.1$ ). The calculated  $R^2$  value indicates that only 23% of the variability in the responses could be explained by the developed regression model, whereas the mean error of the prediction is 104.8 kcal/kg (found by RMSE value) when it is faced with unforeseen data.

The prediction values calculated by the GPR model are illustrated graphically in Figure 1. A visual comparison could be made between actual and predicted values of corn TMEn obtained from both process of training and cross-validation of proposed GPR model. The goodness-of-fit statistical values ( $R^2$  and RMSE) derived from the GPR model to predict the TMEn are also shown in Figure 1. The plot of model predicted (in validation data set) against their respective residual values illustrated in Figure 2. Calculated linear regression on the residual plots (Figure 2) showed no evidence of any linear prediction bias for the models, whereas the calculated  $P$ -value was much larger than 0.05 ( $P = 0.31$ ), suggesting a good agreement between the actual and predicted values of TMEn for the data set which is fed to the developed GPR model. One-way ANOVA results (Table 4) showed no significant difference between experimentally assayed and GPR model predicted TMEn.

## DISCUSSION

The chemical composition were within the range reported in the literature (Ertl, and Dale 1997; Collins et al., 2001; Rodehutsord et al., 2016), but considerable variation was observed between corn from different origins for chemical components of CP, CF, and ash. Differences in chemical contents may be related to differences in agronomic factors and geographical location. Rodehutsord et al. (2016) reported that time of harvest, precipitation, temperature, soil conditions, fertilization, and other agronomic parameters, as well as harvesting and storage conditions, can affect chemical characteristics of the cereal grains, including their energy content, CP, CF, and ash.

**Table 2.** Chemical composition and TMEn content of corn samples from different origins.<sup>1</sup>

Item	Brazil (n = 36)	China (n = 20)	Iran (n = 28)	Ukraine (n = 36)	Pooled SEM (n = 120)	<i>P</i> -value
CP, % of DM	9.19 <sup>b</sup>	9.72 <sup>a</sup>	8.96 <sup>b</sup>	8.94 <sup>b</sup>	0.07	<0.01
EE, % of DM	3.26	3.15	3.40	3.27	0.05	0.43
CF, % of DM	3.78 <sup>b</sup>	3.85 <sup>a,b</sup>	4.04 <sup>a</sup>	4.04 <sup>a</sup>	0.03	<0.01
Ash, % of DM	1.55 <sup>a</sup>	1.41 <sup>a,b</sup>	1.41 <sup>a,b</sup>	1.39 <sup>b</sup>	0.02	0.02
GE, kcal/kg of DM	4,380.80 <sup>a</sup>	4,403.31 <sup>a</sup>	4,396.20 <sup>a</sup>	4,336.91 <sup>b</sup>	6.51	0.01
TMEn, kcal/kg of DM	3,780.50	3,815.20	3,780.20	3,764.50	10.70	0.50

<sup>a-b</sup>Within a row, means without a common letter are significantly different ( $P < 0.05$ ).

<sup>1</sup>DM = dry matter; GE = gross energy; CP = crude protein; EE = ether extract; CF = crude fibre; and TMEn = true metabolizable energy corrected for nitrogen. Thirty samples of corn grains submitted from the Brazil ( $n = 9$ ), China ( $n = 5$ ), Iran ( $n = 7$ ), and Ukraine ( $n = 9$ ). All criteria were assayed with 4 replications for each corn sample ( $n = 120$ ).



**Table 3.** Information and descriptive statistics for the data used in prediction modeling process.<sup>1</sup>

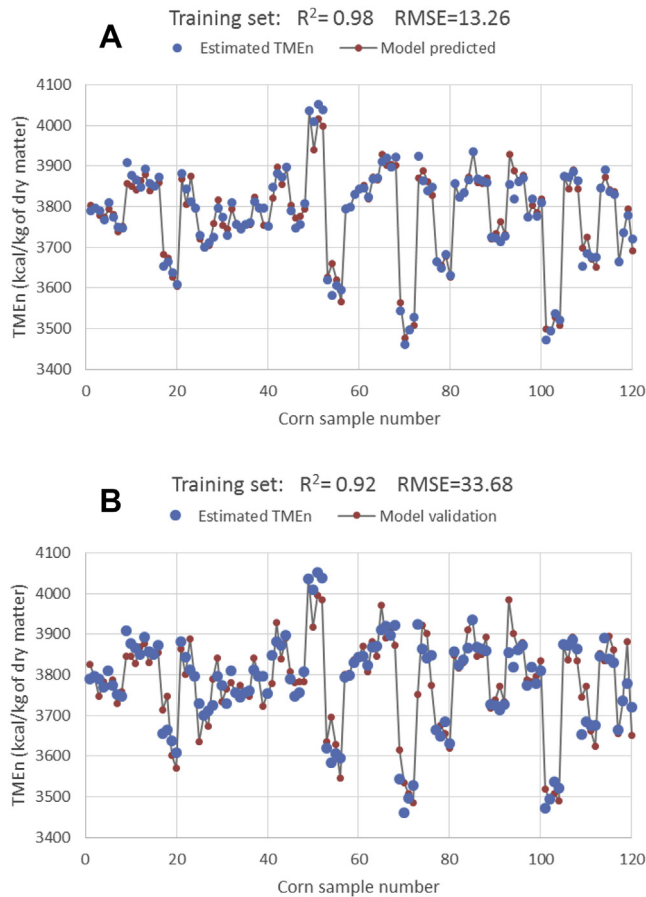
Item	Whole data sets of corn samples (n = 120)			
	Minimum	Maximum	Mean	SD
GE, kcal/kg of DM	4,160.17	4,499.94	4,374.91	71.26
Model inputs				
CP, % of DM	7.50	11.38	9.15	0.712
EE, % of DM	2.46	4.50	3.28	0.518
CF, % of DM	3.29	4.59	3.93	0.297
Ash, % of DM	1.13	2.09	1.45	0.243
Model output				
TMEn, kcal/kg of DM	3,461.60	4,051.60	3,781.40	117.60

<sup>1</sup>DM = dry matter; GE = gross energy; CP = crude protein; EE = ether extract; CF = crude fibre; and TMEn = true metabolizable energy corrected for nitrogen. Thirty samples of corn grains submitted from the Brazil (n = 9), China (n = 5), Iran (n = 7), and Ukraine (n = 9). All criteria were assayed with 4 replications for each corn sample (n = 120).

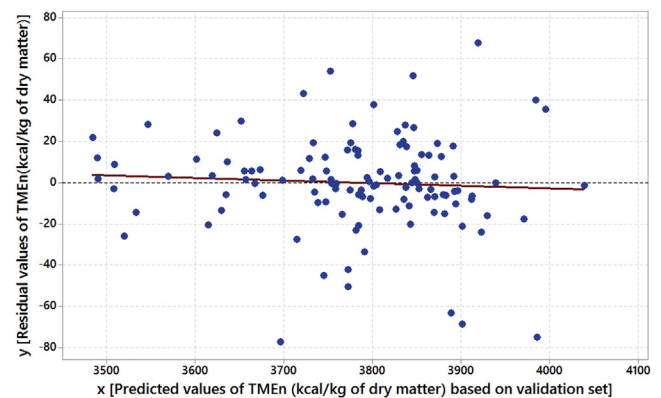
In many situation, the final objective of GPR modeling is to fit models that give more precise and accurate prediction value(s) of the output variable(s). The comparison of actual and predicted output values may define the performance of the prediction model based on the investigated inputs. The proposed GPR

model could suitably predict the TMEn in the validation data set that were not used during the training processes ( $R^2 = 0.92$  and  $RMSE = 33.68$  kcal/kg DM). Moreover, the model had relatively balanced goodness-of-fit statistical values for the 2 subsets of training and cross validation (Figure 1), proposing that the established model had a well generalizability when it faced a totally unseen data set.

To assess the performance of prediction models, the  $R^2$  value is used as a routine measure to judge about the “accuracy” of a model given its predictions, whereas the RMSE is commonly used to indicate the “precision” of a model based on residual (error) examination. Consequently, it is preferred to consider a combination of criteria to determine and/or compare overall performance of prediction process. In modeling TMEn based on concentrations of feed chemical component, the goodness-of-fit in terms of  $R^2$  and RMSE corresponding to MLR, and GPR models showed a higher accuracy and precision of prediction for GPR model as compared with MLR model. The bias value denotes the magnitude of the model over/under estimation with respect to the average of observed values (St-Pierre, 2001). The computed bias were not statistically significant as they could be seen in Figure 2 and corresponding calculated regression line on predicted values vs. residuals in the validation data sets. The superior performance of GPR over MLR model is primarily because of the fact that the conventional MLR requires the arrangement of a linear function to be regressed; thus, the degree of freedom and flexibility of regression equation may be limited. It is believed that the soft computing methods of data modeling would require a very high amount of data samples to build an efficient model. However, GPR models may work well when the data are statistically well distributed in the input domain (Schulz et al., 2018; Li et al., 2019). The main advantage of GPR compared with classical regression tools are 1) the GPR model does not require a prior specification of proper fitting function and 2) the GPR model have a



**Figure 1.** Comparison of experimental and predicted values of true metabolizable energy corrected for nitrogen (TMEn) of corn samples for poultry from the Gaussian process regression model. A) Training set and B) cross validation set (k = 5 for k-fold cross validation process; n = 120 for both sets). Abbreviation: RMSE, root mean square error.



**Figure 2.** Plot of residuals against predicted values (validation set; n = 120) of true metabolizable energy corrected for nitrogen (TMEn) of corn samples for poultry from the Gaussian process regression (GPR) model. The line represents the regression of residuals on GPR model predicted TMEn [ $y = 48.1 - 0.0127 \times \text{predicted TMEn}$ ;  $R^2 = 0.03$ ;  $P = 0.31$ ].

**Table 4.** Comparison between experimental assayed and Gaussian process regression model predicted values of corn TMEn (kcal/kg).

Item	n	Average	SD	95% confidence interval
Experimental TMEn	120	3,781.4	117.6	(3,761.4–3,801.4)
Predicted TMEn	120	3,782.9	104.5	(3,762.9–3,802.9)
Pooled SE	7.17			
P-value	0.92			

Abbreviation: TMEn, true metabolizable energy corrected for nitrogen.

general approximation capability and it can approximate nearly all types of non-linear functions (Roberts et al., 2013).

In conclusion, the present study revealed that corn samples of different origins differ in their chemical composition and GE, but no differences were observed for their TMEn contents. The proposed GPR approach was successfully applied to predict TMEn of corn sample for poultry given levels of their chemical compositions of CP, EE, CF, and ash. When compared with the data from the actual experiment, developed GPR model produce relatively better prediction values in estimating TMEn in corn sample than those produce by conventional regression. The GPR model may be capable of improving our aptitude and capacity to precisely predict energy contents of feed ingredients to formulate optimal diets for poultry.

## ACKNOWLEDGMENTS

**Conflict of Interest Statement:** The authors wish to confirm that there are no known conflicts of interest associated with this publication entitled “A Gaussian process regression for predicting energy contents of corn for poultry”, and there has been no significant financial support for this work that could have influenced its outcome.

## REFERENCES

- Ahmadi, H. 2017. A mathematical function for the description of nutrient-response curve. *PLoS One* 12:e0187292.
- Ahmadi, H., A. Golian, M. Mottaghitlab, and N. Nariman-Zadeh. 2008. Prediction model for true metabolizable energy of Feather meal and poultry Offal meal using Group method of data Handling-Type neural network. *Poult. Sci.* 87:1909–1912.
- Ahmadi, H., and M. Rodehutsord. 2017. Application of artificial neural network and support vector machines in predicting metabolizable energy in Compound feeds for Pigs. *Front. Nutr.* 4:27.
- AOAC. 2000. *Official Methods of Analysis*. 17th ed. AOAC Int., Gaithersburg, MD.
- Collins, N. E., E.T. Moran, Jr, and H. L. Stilborn. 2001. Influence of Yellow dent corn Hybrids having different kernel characteristics yet Similar nutrient composition on Broiler Production. *J. Appl. Poult. Res.* 10:228–235.
- Dudley-Cash, W. A. 2009. A landmark contribution to poultry science—a bioassay for true metabolizable energy in feedingstuffs. *Poult. Sci.* 88:832–834.
- Ertl, D., and N. Dale. 1997. The metabolizable energy of Waxy vs. Normal corn for poultry. *J. Appl. Poult. Res.* 6:432–435.
- Freud, R. J., and R. C. Littell. 2000. *SAS System for Regression*. 3rd ed. SAS Institute, Cary, NC.
- Huttunen, J. M. J., L. Kärkkäinen, and H. Lindholm. 2019. Pulse transit time estimation of aortic pulse wave velocity and blood pressure using machine learning and simulated training data. *Plos Comput. Biol.* 15:e1007259.
- Klopfenstein, T. J., G. E. Erickson, and L. L. Berger. 2013. Maize is a critically important source of food, feed, energy and forage in the USA. *Field Crops Res.* 153:5–11.
- Latshaw, J. D., and K. Freeland. 2008. Metabolizable energy values determined with intact and cecectomized roosters. *Poult. Sci.* 87:101–103.
- Li, L. L., X. B. Zhang, M. L. Tseng, and Y. T. Zhou. 2019. Optimal scale Gaussian process regression model in Insulated Gate Bipolar Transistor remaining life prediction. *Appl. Soft Comput.* 78:261–273.
- Mathworks. 2019. *Statistics and Machine Learning Toolbox User's Guide R2019b*. The MathWorks Inc. Natick, MA.
- Minitab. 2014. *MINITAB Release 17: Statistical Software for Windows*. Minitab Inc., State College, PA.
- Rasmussen, C. E., and C. K. I. Williams. 2006. *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, MA.
- Richardson, R. R., M. A. Osborne, and D. A. Howey. 2017. Gaussian process regression for forecasting battery state of health. *J. Power Sourc.* 357:209–219.
- Roberts, S., M. Osborne, M. Ebden, S. Reece, N. Gibson, and S. Aigrain. 2013. Gaussian processes for time-series modelling. *Philosophical Transactions. Ser. A, Math. Physical, Engineering Sciences* 371:20110550.
- Rodehutsord, M., C. Rückert, H. P. Maurer, H. Schenkel, W. Schipprack, K. E. Bach Knudsen, M. Schollenberger, M. Laux, M. Eklund, and W. Siegert. 2016. Variation in chemical composition and physical characteristics of cereal grains from different genotypes. *Arch. Anim. Nutr.* 70:87–107.
- Schulz, E., M. Speekenbrink, and A. Krause. 2018. A tutorial on Gaussian process regression: modelling, exploring, and exploiting functions. *J. Math. Psychol.* 85:1–16.
- Sibbald, I. R. 1976. A bioassay for true metabolizable energy in feedingstuffs. *Poult. Sci.* 55:303–308.
- St-Pierre, N. R. 2001. Invited Review: Integrating Quantitative Findings from multiple Studies using Mixed model Methodology1. *J. Dairy Sci.* 84:741–755.
- Swain, P. S., K. Stevenson, A. Leary, L. F. Montano-Gutierrez, I. B. N. Clark, J. Vogel, and T. Pilizota. 2016. Inferring time derivatives including cell growth rates using Gaussian processes. *Nat. Commun.* 7:13766.
- Tonner, P. D., C. L. Darnell, B. E. Engelhardt, and A. K. Schmid. 2017. Detecting differential growth of microbial populations with Gaussian process regression. *Genome Res.* 27:320–333.