Spanish Journal of Agricultural Research
13(2), e06SC02, 5 pages (2015)
eISSN: 2171-9292
http://dx.doi.org/10.5424/sjar/2015132-6573
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA)

SHORT COMMUNICATION

OPEN ACCESS

Prediction of apparent metabolisable energy content of cereal grains and by-products for poultry from its chemical composition

Beatriz Losada[1], Carlos De Blas[2], Paloma García-Rebollar[2], Pilar Cachaldora[1], Jesús Méndez[1] and Miguel Ibáñez[3]

[1] *COREN, Sociedad Cooperativa Galega, 32003 Ourense, Spain*

[2] *Universidad Politécnica de Madrid, Departamento de Producción Animal, 28040 Madrid, Spain*

[3] *Universidad Politécnica de Madrid, Departamento de Estadística y Métodos de Gestión en Agricultura, 28040 Madrid, Spain.*

Abstract

In order to predict the metabolisable energy content of ninety batches of cereal grains and cereal by-products for poultry, regression models derived from different sample aggregations and using chemical components as independent variables were compared. Several statistics have been calculated to estimate the error of prediction. The results indicate that the highest levels of significance and coefficients of determination were obtained for equations derived from the larger data sets. However, the lowest prediction errors were associated to equations calculated for data or groups of data closer to the ingredient studied.

Additional key words: energy evaluation; raw materials; error of prediction.

Abbreviations used: AMEn (apparent metabolisable energy corrected for nitrogen); DDGS (dry distillers grains and solubles); DM (dry matter); LOO (leave-one-out); NIRS (near infrared reflectance spectroscopy); RBPRED (square root of the mean of squares of the deviations between the observed values and the predicted values when using bootstrap); RPRESS (square root of the mean of squares of the deviations between the observed values and the values predicted by the model when using LOO); RSD (residual standard deviation); RSqE (root mean square error).

Citation: Losada, B.; De Blas, C.; García-Rebollar, P.; Cachaldora, P.; Méndez, J.; Ibáñez, M. (2015). Short communication: Prediction of apparent metabolisable energy content of cereal grains and by-products for poultry from its chemical composition. Spanish Journal of Agricultural Research, Volume 13, Issue 2, e06SC02, 5 pages. http://dx.doi.org/10.5424/sjar/2015132-6573.

Received: 22 Jul 2014. **Accepted:** 10 Mar 2015

Funding: Financial support was provided by Xunta de Galicia (Project PGDIT05RAG014E)

Competing interests: The authors have declared that no competing interests exist.

Correspondence should be addressed to Carlos De Blas: c.deblas@upm.es

In spite of its economic importance, work made on energy evaluation of poultry feeds is quite limited and scarcer than that available for other species, as ruminants or swine. Several regression equations are available at present to predict apparent metabolisable energy corrected for nitrogen (AMEn) contents of compound feeds from its chemical composition (*e.g.* Carpenter & Clegg, 1956; Sibbald *et al*., 1980; Fisher, 1982; Carré *et al*., 1984; EEC, 1986). However, the prediction error of these equations can increase greatly when they are extrapolated to estimate the energy value of single ingredients (Dolz & De Blas, 1992; Losada *et al*., 2010), which would justify the use of specific equations for single raw materials. The use of indirect techniques, as near infrared reflectance spectroscopy (NIRS), for this purpose also has limits when using equations derived from a reduced number of ingredients or from complete diets (Valdes & Leeson, 1992, 1994; Garnsworthy *et al*., 2000). Previous work from our group (Losada *et al*., 2009) has studied several methods of prediction (chemical composition, *in vitro* digestibility and NIRS) of the AMEn values of eleven cereal grains and cereal by-products for poultry. The aim of the current study has been to use this data set to compare different combinations of sample aggregation to predict the AMEn values from its chemical composition, and to discuss the best model to minimize the error of prediction for each single ingredient.

Fifty six batches of five starchy grains: wheat (*Triticum aestivum* L., *Triticum turgidum* L.), barley (*Hordeum vulgare* L.), corn (*Zea mays* L.), sorghum (*Sorghum vulgare* L.) and rye (*Secale cereale* L.) and 34 batches of six cereal byproducts (corn gluten feed, rice bran, wheat bran and dry distillers grains and solubles (DDGS) from wheat, corn and sorghum), were

sampled from the COREN SCG poultry feed manufacturing plant (Orense, Spain) throughout a 3-yr period and obtained in a previous work (Losada *et al*., 2009). Animals were handled according to the principles for the care of animals in experimentation published by the Spanish Royal Decree 1201/2005 (BOE, 2005), and experiments were approved by the Ethics Committee of the Research and Development Department of COREN SCG. Determination of AMEn of the experimental diets was made following the European reference method (Bourdillon *et al*., 1990). Energy values for feed ingredients were determined *in vivo* by using the difference method, as described by Losada *et al*. (2009).

Prediction equations of AMEn *in vivo* values of single and groups of ingredients from chemical analytical data were developed by stepwise regression analysis, by using PROC REG of SAS (2008). The stepwise procedure introduced variables in the model only if they contributed to a significant improvement ($p < 0.05$) in the estimation of the dependent variable. Several combinations of ingredients were used to predict AMEn values of a single raw material. For instance, AMEn for corn grain can be estimated from equations calculated from: i) all the samples studied (n = 90), ii) all data from the several cereal grains studied (n = 57), iii) all the samples of the same ingredient (including corn byproducts, n = 22), iv) all the data obtained with corn grain (n = 12), or v) from the average value obtained for this ingredient. This approach has been also made for other ingredients or groups of ingredients included in this study: corn byproducts, wheat grain and wheat byproducts. To compare these different approaches to predict the AMEn values of all data included in each model, two statistics were calculated: i) the coefficient of determination (R^2), *i.e.* the proportion of the total variation explained by a particular regression equation, and ii) the residual standard deviation (RSD), *i.e.* the square root of the mean of the squares of the deviations between the values actually observed and the values predicted by a particular equation. These two statistics are the more frequently used in the literature to select the best prediction model. Alternatively, when regressions were used to predict the energy value of one of the single ingredients included in a data base, another statistic can be computed: iii) the root mean square error (RSqE), *i.e.* the square root of the mean square deviations for this specific ingredient. However, computing the prediction error with the data used to estimate the model overestimates the predictive ability of the model. Therefore, it is suitable to have a validation sample to judge the true performance of the model. When a proper validation sample is not available, some resampling methods such as cross-validation and bootstrap techniques (Efron & Tibshirani, 1993) can be used. For small datasets, the leave-one-out (LOO) cross validation has been proposed. In this procedure, an observation is removed at each step; the model is estimated with the remaining observations and is used to predict the deleted observation. At the end of process another statistic can be calculated: iv) RPRESS, *i.e.* the square root of the mean of squares of the deviations between the observed values and the values predicted by the model when using LOO. In the bootstrap method a sample is obtained from the data set at random with replacement, having the same size as the original data set, so that the sample may contain repeated observations. The sampling procedure is repeated n times and the average prediction value is utilized to compute: v) RBPRED, the square root of the mean of squares of the deviations between the observed values and the predicted values when using bootstrap. In addition, bias of the model prediction was calculated as the mean value of the differences between the observed and values predicted by the model. All the statistical indicated above were calculated using the package boot (Canty & Ripley, 2013) in R 3.0.2 (R Development Core Team, 2013).

The regression equations obtained for the prediction of AMEn values of the different ingredients and groups of ingredients that reached statistical significance ($p < 0.05$) are shown in Table 1. The first variable entering in the models for the whole data set was dietary neutral detergent fibre content, showing a negative relation (r = –0.784; $p < 0.001$) with energy value of the raw materials studied. This result reflects the low digestive efficiency of fibre constituents in poultry; the figure also indicates that this relationship is not homogeneous across the ingredients studied.

As shown in Table 1, the use of large data bases, as those of all data (n = 90), all cereal grains (n = 56), all cereal byproducts (n = 34), all wheat products (n = 27) or all corn products (n = 22) led to highly significant equations ($p < 0.001$) and relatively high coefficients of determination (from 0.604 to 0.933, 0.800 as average). These results are justified because R^2 estimates the proportion of the total standard deviation (SD) explained by the independent variables included in the model, so that it tends to increase in parallel to the heterogeneity of the ingredients considered in the regression analysis (with total SD varying from 264 to 425 kcal/kg dry matter (DM), 364 as average, in this group of equations). If the R^2 and/or the level of significance are used as the main or the only criteria for model selection, this type of general equations would be preferred. However, as also shown in Table 1, these general equations also tend to have high RSD values

Table 1. Stepwise regression analysis for AMEn (kcal/kg DM) using chemical composition traits (% DM) as predictors

Data	n†	Regression equation‡	R^2	RSD	p
All data	90	3840 (±66.0) – 32.1 (±2.64) NDF	0.616	265	<0.001
		3810 (±58.1) – 36.5 (±2.45) NDF + 28.2 (±5.25) EE	0.718	232	<0.001
		3697 (±52.9) – 11.7 (±4.63) NDF + 57.1 (±6.58) EE – 177 (±29.7) Ash	0.791	196	<0.001
Cereal grains	56	1352 (±256) + 30.1 (±3.81) Starch	0.535	180	<0.001
		2426 (±586) + 19.7 (±6.32) Starch – 216 (±107) Ash	0.568	175	<0.001
		2817 (±595) + 12.2 (±7.03) Starch – 239 (±104) Ash + 59.5 (±27.6) EE	0.604	170	<0.001
Cereal by-products	34	2368 (±53.2) + 46.0 (±5.09) EE	0.719	191	<0.001
		2049 (±110) + 47.9 (±4.51) EE + 12.1 (±3.77) CP	0.789	170	<0.001
		2362 (±217) + 52.5 (±5.21) EE + 11.6 (±3.69) CP – 59.2 (±35.7) Ash	0.807	163	<0.001
Wheat products	27	3902 (±92.4) – 36.1 (±3.22) NDF	0.860	177	<0.001
Corn products	22	4092 (±78.8) – 37.8 (±2.83) NDF	0.899	170	<0.001
		3897 (±91.4) – 43.1 (±2.92) NDF + 60.6 (±19.6) EE	0.933	143	<0.001
All DDGS	16	2560 (±144) + 34.6 (±17.2) EE	0.224	152	0.05
Corn grain	12	–245 (±1744) + 52.0 (±23.6) StSu	0.326	127	0.05
		–2453 (±1493) + 70.8 (±16.3) StSu + 177 (±51.1) CP	0.777	79.6	0.005
Corn byproducts	10	2018 (±126) + 92(±18) EE	0.768	120	<0.001
Barley grain	11	5370 (±1027) – 444 (±203) CF	0.347	185	0.05
Wheat byproducts	14	1400 (±280) + 291 (±70) EE	0.594	167	<0.001
Wheat bran	10	1291 (±432) + 323 (±116) EE	0.491	181	0.02
Corn DDGS	6	5601 (±693) – 93.0 (±23.0) CP	0.803	54	0.02
Sorghum DDGS	6	4278 (±422) – 235 (±72.7) Ash	0.723	119	0.03
		7522 (±1471) – 417 (±95.5) Ash – 67.0 (±29.7) CP	0.897	84	0.03

AMEn, apparent energy metabolisable energy corrected for nitrogen; CP, crude protein; CF, crude fibre; DDGS, dry distillers grains with solubles; EE, ether extract; DM, dry matter; NDF, neutral detergent fibre; R^2, coefficient of determination; RSD, residual standard deviation; StSu, starch+sugars. † Number of data used to develop the different models. ‡ Values into parentheses are standard errors.

(from 143 to 196 kcal AMEn/kg DM, 170 as average), which indicates that the amount of residual variation not explained by the model is also relatively high.

In contrast, the use of specific regressions for single feedstuffs with smaller databases (n = 6 to 13), and lower total SD (from 70.6 to 258 kcal AMEn/kg DM, 160 as average, see Table 1) led to less significant equations (from $p > 0.05$ to $p = 0.005$) and lower R^2 (from 0.347 to 0.897, with 0.663 as average of the equations shown in Table 1), but also to a lower RSD (from 54 to 185 kcal AMEn/kg DM, 117 as average) than the general equations of the first group. If the objective of regression analysis is to minimize the deviations between the values actually observed and the values predicted, this type of equations would be preferable. Moreover, general equations are less precise when they are extrapolated to single ingredients whose chemical composition is far away from the average.

This aspect is considered in the statistics presented in Table 2 (BIAS, RSqE, RBPRED and RPRESS) that serve to compare the different equations when applied to predict the energy value of specific raw materials. Data show that use of the equation derived from all data resulted in the highest mean prediction errors, expressed as the percentage of bias on the average AMEn value determined in the study of Losada *et al*. (2009) (+1.84, +4.66, +1.15 and –2.39% in the case of corn grain, corn byproducts, wheat grain and wheat products equations, respectively). Bias of equations tended to decrease when they were derived from data bases closer to the predicted values, being zero when prediction was made from the information derived from that particular ingredient. Otherwise, the error of prediction (RSqE) tended to decrease when using more specific equations (by 35% when comparing the equations obtained from all data or from values obtained with a single raw material), even when error of prediction was calculated after cross validation or bootstrap (22% as average). In the case of wheat grains, where the total SD of AMEn of the samples studied was relatively low (126 kcal/kg DM) and it was scarcely explained by the independent variables included in the model ($R^2 = 0.224$, $p = 0.09$), little benefit was obtained when using any of the derived equations with respect to the estimation of its energy value from the average AMEn.

Figure 1 shows a plot of residual versus predicted values for the ingredients studied in Table 2 from each of the models considered. In this Figure, the residues should be distributed randomly along the horizontal line passing through the zero residual value. The non-random pattern in the residuals indicates the systematic error in the prediction model. The lines drawn with different patterns correspond to the regression equations between residuals and predicted values in each of the prediction models. Therefore, these figures

Table 2. Error of prediction of the estimation of AMEn (kcal/kg DM) of several ingredients for poultry

Ingredients studied	Data included in the model	Model		Prediction error of the single ingredient studied			
		R^2	RSD	BIAS	RSqE	RBPRED	RPRESS
Corn grain	All data	0.791	198	66.1	147	150	151
	Cereal grains	0.604	170	65.9	135	140	141
	Corn products	0.933	143	5.6	132	144	145
	Corn grains	0.777	79.6	0.0	65.0	113	115
	Average	–	148	0.0	142	152	155
Corn byproducts	All data	0.791	196	123	173	184	188
	Corn products	0.933	143	–6.7	132	159	159
	Corn byproducts	0.790	121	0	108	132	134
	Average	–	236	0	224	246	149
Wheat grain	All data	0.791	196	39.1	123	125	125
	Cereal grains	0.604	170	–42.4	128	132	131
	Wheat products	0.860	177	3.11	134	144	145
	Wheat grains	0.240	119	0	109	124	126
	Average	–	131	0	126	136	136
Wheat byproducts	All data	0.791	196	–61.32	245	255	258
	Wheat products	0.860	177	–2.89	198	212	216
	Wheat byproducts	0.591	167	0.0	154	174	179
	Average	–	251	0.0	242	254	261

AMEn, apparent energy metabolisable energy corrected for nitrogen; BIAS, mean value of the differences between the observed and values predicted by the model; R^2, coefficient of determination; RBPRED, square root of the mean of squares of the deviations between the observed values and the values predicted for a specific ingredient when using bootstrap; RPRESS, square root of the mean of squares of the deviations between the observed values and the values predicted by the model for a specific ingredient when using the leave-one-out cross validation; RSD, residual standard deviation of the regression equation; RSqE, square root of the mean of the squares deviations between the observed values and the values predicted by the model for a specific ingredient.

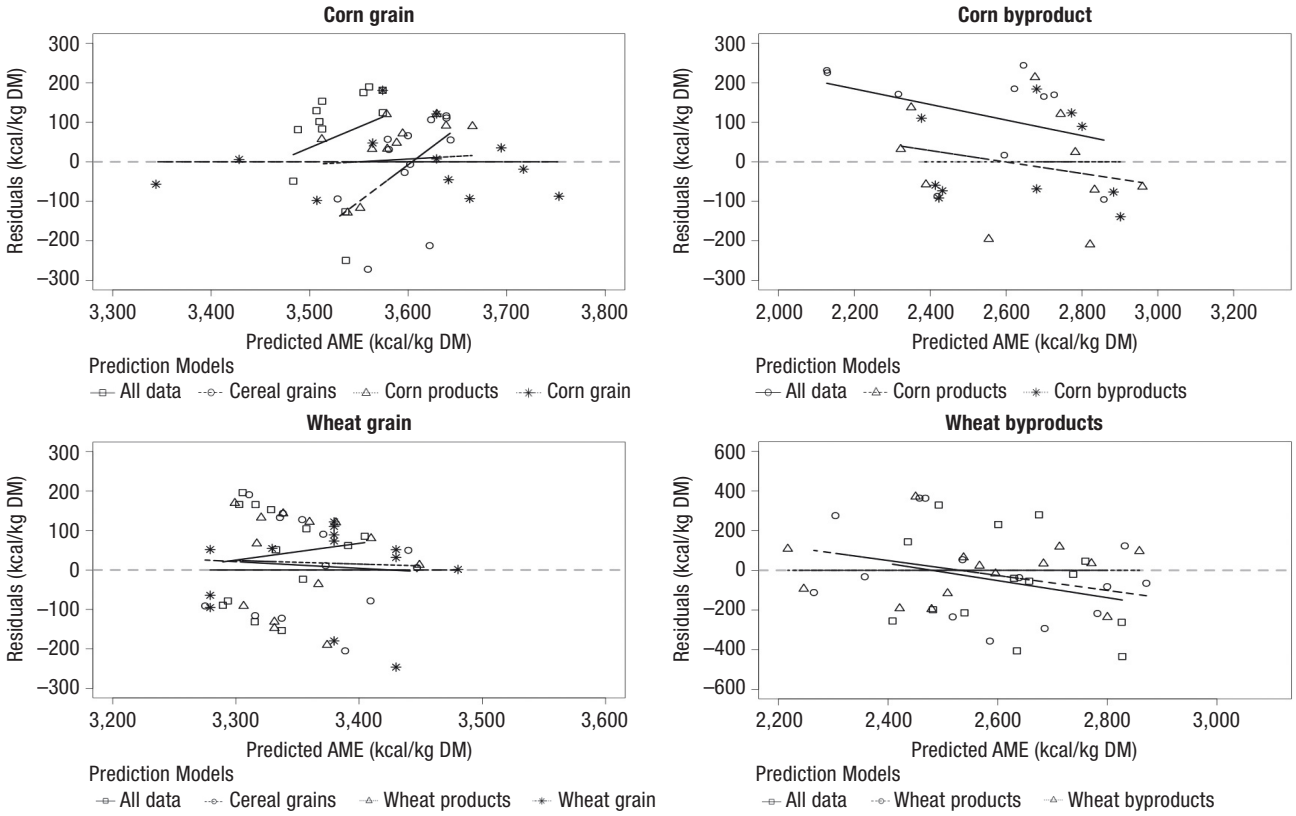

Figure 1. Plot of residual versus predicted values for corn grain, corn by products, wheat grain and wheat by products from each of the prediction models considered.

represent the decomposition of the statistical RSqE of Table 2 in i) the variation of the estimated values by the regression along the zero value (the sum of the squared bias), and ii) the variance of the residuals along the fitted line. The figures serve to visualize that the use of general equations derived from large data sets leads to higher errors of prediction than the more specific equations derived from data obtained from the same ingredient or from groups of similar ingredients.

The current data indicate that the use of the coefficient of determination or the level of statistical significance to select the best model to predict the AMEn of single ingredients for poultry may lead to erroneous conclusions. Minimizing the prediction error should be used as the main criterion to select the best model.

References

BOE, 2005. Real Decreto 1201/2005, de 10 de octubre, sobre protección de los animales utilizados para experimentación y otros fines científicos. Boletín Oficial del Estado 252: 34637-34391.

Bourdillon A, Carré B, Conan L, Duperray J, Huyghebaert G, Leclercq B, Lessire M, McNab J, Wiseman J, 1990. European reference method for the in vivo determination of metabolisable energy with adult cockerels: reproducibility, effect of food intake and comparison with individual laboratory methods. Br Poult Sci 31: 557-565. http://dx.doi.org/10.1080/00071669008417287

Canty A, Ripley B, 2013. Boot: Bootstrap R (S-Plus) Functions. R package version 1.3-9.

Carpenter KJ, Clegg KM, 1956. The metabolisable energy of poultry feeding-stuffs in relation to their chemical composition. J Sci Food Agric 7: 45-51. http://dx.doi.org/10.1002/jsfa.2740070109

Carré B, Prevotel B, Leclercq B, 1984. Cell wall content as a predictor of metabolisable energy value of poultry feedingstuffs. Br Poult Sci 25: 561-572. http://dx.doi.org/10.1080/00071668408454898

Dolz S, De Blas C, 1992. Metabolisable energy of meat and bone meal from Spanish rendering plants as influenced by level of substitution and method of determination. Poult Sci 71: 316-322. http://dx.doi.org/10.3382/ps.0710316

EEC, 1986. Directive de la Commission du 9 avril 1986 fixant la méthode de calcul de la valeur énergétique des aliments composés destinés à la volaille. Journal Officiel des Communautés Européennes 130: 53-54.

Efron B, Tibshirani RJ, 1993. An introduction to the bootstrap. Ed. Chapman & Hall, New York, 436 p.

Fisher C, 1982. Energy values of compound poultry feeds. Occasional Publication No. 2. Institute for Grassland & Animal Production, Poultry Division, Roslin, Midlothian, UK.

Garnsworthy PC, Wiseman J, Fegeros K, 2000. Prediction of chemical, nutritive and agronomic characteristics of wheat by near infrared spectroscopy. J Agric Sci 135: 409-417. http://dx.doi.org/10.1017/S0021859699008382

Losada B, Garcia-Rebollar P, Cachaldora P, Álvarez C, Mendez J, De Blas C, 2009. A comparison of the prediction of apparent metabolisable energy content of starchy grains and cereal by-products for poultry from its chemical components, in vitro analysis or near-infrared reflectance spectroscopy. Span J Agric Res 7: 813-823. http://dx.doi.org/10.5424/sjar/2009074-1096

Losada B, Garcia-Rebollar P, Álvarez C, Cachaldora P, Ibáñez MA, Mendez J, De Blas C, 2010. The prediction of apparent metabolisable energy content of oil seeds and oil seed by-products for poultry from its chemical components, in vitro analysis or near-infrared reflectance spectroscopy. Anim Feed Sci Technol 160: 62-72. http://dx.doi.org/10.1016/j.anifeedsci.2010.06.012

R Development Core Team, 2013. R: A language and environment for statistical computing. Ed. R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.R-project.org.

SAS, 2008. SAS/STAT user's guide, version 9.3. SAS Institute Inc., Cary, NC, USA.

Sibbald IR, Price K, Barrete JP, 1980. True metabolisable energy values for poultry of commercial diets measured by bioassay and predicted from chemical data. Poult Sci 59: 808-811. http://dx.doi.org/10.3382/ps.0590808

Valdes EV, Leeson S, 1992. Near infrared reflectance analysis as a method to measure metabolisable energy in complete poultry feeds. Poult Sci 71: 1179-1187. http://dx.doi.org/10.3382/ps.0711179

Valdes EV, Leeson S, 1994. Measurement of metabolisable energy, gross energy, and moisture in feed grade fats by near infrared reflectance spectroscopy. Poult Sci 73: 163-171. http://dx.doi.org/10.3382/ps.0730163