ASSESSMENT OF HYPOTHESES TEST FOR THE COMPONENTS OF MEAN SQUARE ERROR OF PREDICTION

Edenio DETMANN[1]
Hugo Colombarolli BONFA[1]
Paulo Roberto CECON[2]
Fabyano Fonseca e SILVA[1]

▪ ABSTRACT: The analysis of mean square error of prediction is helpful to compare measured values with values simulated by mathematical models. Such analysis is based on the orthogonal decomposition of this quantity into three components which will indicate the probable constraints of the model concerning bias, unequal variance, and incomplete covariation when contrasted to actual values. However, such analysis has been carried out as a descriptive procedure without an adequate hypotheses test framework. Thus, we aimed to develop single hypothesis test to evaluate the statistical significance of mean square error of prediction components based on likelihood ratio test and χ^2 distribution. This proposal was evaluated by using simulated populations and was applied to a dataset obtained by simulating characteristics of cattle diets using two different models. We concluded that this test might help the modeler to focus on the real significant constraints of his model and to work on doing the necessary modifications on its mathematical structure in order to improve the accuracy and precision of the simulated values.

▪ KEYWORDS: Inference; mathematical models; simulation; validation.

1 Introduction

A mathematical model is an equation or a set of equations which represents the behavior of a system (THORNLEY and FRANCE, 2007). Modelling should be seen as a central and integral part of scientific method. However, model evaluation is not a wholly process. Models can be perceived as hypotheses expressed in mathematics and should therefore be subject to the usual process of hypothesis evaluation (FRANCE and KEBREAB, 2008).

The model adequacy evaluation is an essential step in the modeling process because it indicates the levels of precision and accuracy of the model predictions. This is an important phase either to build up confidence in the current model or to allow selection of alternative models (TEDESCHI, 2006). An ideal testing procedure is one in which the model is used to anticipate behavior of the simulated system under circumstances not previously studied (BALDWIN, 1995) and using a different dataset (GAUCH Jr. et al,

[1]Universidade Federal de Viçosa – UFV, Departamento de Zootecnia, Viçosa, MG, Brasil, e-mail: *detmann@ufv.br; hugo.bonfa@ufv.br; fabyanofonseca@ufv.br*
[2] Universidade Federal de Viçosa – UFV, Departamento de Estatística, Viçosa, MG, Brasil, e-mail: *cecon@ufv.br*

2003). Thus, if the model anticipation is poor, further model refinement is required (BALDWIN, 1995).

As summarized by Tedeschi (2006), several techniques are available to assess model adequacy to ensure impartiality during the decision process of accepting or rejecting the suitability of a given mathematical model. Among those techniques, the analysis of the mean square error of prediction (*MSEP*) and its components have been applied by several modelers in order to depict different aspects of the model failure to reproduce the real world (KOBAYASHI and SALAM, 2000; GAUCH Jr. et al., 2003; TEDESCHI, 2006).

The analysis of MSEP and its components was firstly applied by Theil (1966) and posteriorly refined by other authors (BIBBY and TOUTENBURG, 1977; KOBAYASHI; SALAM, 2000; GAUCH Jr. *et al*., 2003; KOBAYASHI, 2004). Such analysis allows orthogonally decomposing the MSEP into three components, which in turn will indicate the probable constraints of the model concerning bias, unequal variance, and incomplete covariation when contrasted to actual values.

However, in spite of providing good information with regards model accuracy and precision, the components of *MSEP* have been interpreted just as descriptive statistics. Such approach sometimes may produce restricted conclusions because the information about significance of each component is not available. Therefore, the information produced from *MSEP* components could be more helpful when using an adequate hypothesis test besides the mathematical calculations themselves. Under this context, the objective of this work was to develop and present a single hypothesis test to assess the statistical significance of components of the mean square error of prediction.

2 Derivation of the method

The prediction error of a model when contrasted to an actual (or measured) value is calculated as:

$$e_i = x_i - y_i, \qquad (1)$$

where e_i is the prediction error for the *i*-th data, and x_i and y_i are the simulated values and measured values, respectively.

It must be emphasized that calculation of prediction error is performed similarly for both linear and non-linear models, as it is a simple calculation of the distance between predicted and observed values. Therefore, all assumptions to be considered hereafter are applied to the prediction error itself and are independent of the nature of the model applied to predict the real world (linear or non-linear).

According to Kobayashi and Salam (2000), the *n* prediction errors obtained in the evaluation or validation of a model can be summarized by the *MSEP*, whose estimator obtained by the maximum likelihood method is:

$$MSEP = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2 = \frac{1}{n}\sum_{i=1}^{n}e_i^2 \cdot \qquad (2)$$

The $MSEP$ is proportional to the dissimilarity between predicted values and measured values. Nonetheless, it is not able to qualitatively point out which are the main constraints of the simulated values obtained by the model. To overcome this situation, the *MSEP* should be decomposed to improve the understanding about the discrepancies between predicted and measured values. Initially, the *MSEP* (Equation 2) can be partitioned into two components (KOBAYASHI and SALAM, 2000) as follow:

$$MSEP = (\bar{x} - \bar{y})^2 + \frac{1}{n}\sum_{i=1}^{n}[(x_i - \bar{x}) - (y_i - \bar{y})]^2, \quad (3)$$

where $\bar{x}$ and $\bar{y}$ are the means of x_i and y_i (i = 1, 2, ..., *n*), respectively.

The first term of the right side of Equation (3) represents the bias of the simulation from the measurements and can be denoted as square of bias (*SB*), namely:

$$SB = (\bar{x} - \bar{y})^2. \quad (4)$$

The second term of right side of Equation (3) is the difference between the simulation and the measurement with regards to the deviations from the means and is denoted as mean square variation (*MSV*), namely:

$$MSV = \frac{1}{n}\sum_{i=1}^{n}[(x_i - \bar{x}) - (y_i - \bar{y})]^2. \quad (5)$$

Higher *MSV* values indicate that the model failed to predict the variability of the measurement around the mean. Furthermore, it must be emphasized that SB and MSV are orthogonal and can be addressed separately (KOBAYASHI and SALAM, 2000).

The *MSV* may be further partitioned into two different components (KOBAYASH and SALAM, 2000). The final resultant of such partitioning is:

$$MSV = (SD_s - SD_m)^2 + 2 \times SD_s \times SD_m \times (1 - r), \quad (6)$$

where SD_s and SD_m are the standard deviations of simulated and measured values, respectively, and r is the correlation coefficient between the simulated and measured values. It should be noted that all calculations of variances and covariance must be performed using the total number of values (n) as denominator because prediction errors, and not experimental errors, are evaluated. In this case, the sum of all prediction errors are not necessarily equal to zero, which only must be observed when simple deviations around a mean of a sample are evaluated.

The first term of the right side of Equation (6), which is named *SDSD*, indicates the difference in the magnitude of fluctuation between the simulation and measurement, which means that there was an unequal variation of predicted and measured values (BIBBY and TOUTENBURG, 1977; KOBAYASHI and SALAM, 2000). Accordingly, this term specifies that the model failed to reproduce the dispersion of the values around the mean at the same extent of that observed for real values. Namely:

$$SDSD = (SD_s - SD_m)^2 . \tag{7}$$

The second term of the right side of Equation (6), is essentially the lack of positive correlation weighed by the standard deviations, denoted as *LCS*. Higher *LCS* values mean that the model failed to simulate the pattern of the fluctuation across the *n* measurements (KOBAYASHI and SALAM, 2000), or that there is an incomplete covariation between predicted and observed values (BIBBY and TOUTENBURG, 1979). Namely:

$$LCS = 2 \times SD_s \times SD_m \times (1 - r) . \tag{8}$$

From the Equations (3), (4), (6), (7), and (8) it may be perceived that *MSEP* is orthogonally partitioned into three different components as follow:

$$MSPE = SB + SDSD + LCS . \tag{9}$$

Under the assumption that prediction errors (Equation 1) can be studied according to the properties of the normal distribution, the *MSEP* (Equation 2) could be interpreted as the variance of prediction errors:

$$MSEP = \hat{\sigma}_e^2 . \tag{10}$$

From this, it is assumed that *MSEP* can be studied following the properties of χ^2 distribution (SEARLE et al., 1992). Namely:

$$\frac{n \times \hat{\sigma}_e^2}{E(\hat{\sigma}_e^2)} \sim \chi_n^2 . \tag{11}$$

In spite of components *SDSD* and *LCS* are not entirely independent as they have the same constituents (Equations 7 and 8), the orthogonal partition of *MSEP* allows addressing separately for each constituent. Under the assumption that *MSEP* is a variance and that the different constituents can be study separately, it can be assumed that each constituent of *MSEP* can be considered as a variance component and would present the same distributional properties of the total variance. From this, it can be written:

$$\sigma_e^2 = \sigma_{SB}^2 + \sigma_{SDSD}^2 + \sigma_{LCS}^2. \quad (12)$$

From Equation (12) it is assumed that variance σ_e^2 corresponds to a parametric space Ω composed by three components which are presupposed independent to each other.

The general hypotheses to be addressed to this problem encompass the evaluation of significance associated with each component of the parametric space Ω. Generically, the null and alternative hypotheses are written as follow:

$$H_0 : \sigma_j^2 = 0$$
$$H_a : \sigma_j^2 > 0, \quad (13)$$

where σ_j^2 is the *j*-th component of the parametric space Ω to be tested.

By considering the hypotheses aforementioned (Equation 13), it is mandatory to define two different parametric spaces:

$$\Omega = \sigma_e^2 \quad (14)$$

$$\Omega^R = \sigma_e^2 - \sigma_j^2, \quad (15)$$

where Ω corresponds to whole parametric space in which all components are present, and Ω^R is the restricted parametric space in which the component to be tested is not taken into account.

Therefore, the general hypotheses described in Equation (13) are based on the evaluation of the similarity of the restricted and whole parametric spaces. If there is similarity of these parametric spaces, H_0 is considered to be true (Equation 13a) and the evaluated component is supposed to be non-significant.

The statistics of the likelihood ratio test for this problem would be (RAO, 1973; REGAZZI and SILVA, 2004; ARCHONTOULIS and MIGUEZ, 2015):

$$\Lambda = \left(\frac{{}^R\hat{\sigma}_e^2}{\hat{\sigma}_e^2} \right)^{\frac{n}{2}} = \left(\frac{\hat{\sigma}_e^2 - \hat{\sigma}_j^2}{\hat{\sigma}_e^2} \right)^{\frac{n}{2}}, \quad (16)$$

where ${}^R\hat{\sigma}_e^2$ is the estimate of MSEP when the restriction established by the general hypotheses is taken into account (Equations 13 and 15).

For large samples, the probability distribution of [-2 × $\log_e$ (Λ)] is approximately χ^2 with v degrees of freedom, being v the number of components in the parametric space Ω minus the number of components in the parametric space Ω^R. Therefore, according to the hypotheses showed in Equation (13), v = 1 when each component is tested individually. In this way, it can be written:

$$-2 \times \log_e(\Lambda) = -n \times \log_e\left(\frac{\hat{\sigma}_e^2 - \hat{\sigma}_j^2}{\hat{\sigma}_e^2}\right) \sim \chi^2_{(1)}. \qquad (17)$$

3 Evaluation of the statistical efficiency of the method

In order to evaluate the statistical efficiency of the method, two different scenarios for each component of MSPE were simulated. The first scenario was simulated to check the ability of the method to point out a significance for each component when the null hypothesis (Equation 13a) is false. On the other hand, the second scenario aimed to check if the method is able to point out correct decision when the null hypothesis (Equation 13a) is true. For each scenario, populations composed by one million records each (n = 1,000,000) were simulated according to a bivariate normal distribution. The different scenarios were established by varying the mean (μ), variance (σ^2) and correlation (ρ) of the normal bivariate distribution according to the values expressed in Table 1.

Table 1 - Characteristics of the bivariate populations simulated to evaluate the statistical efficiency of the method

	Measured values		Predicted values		
Component	μ	σ^2	μ	σ^2	ρ
			Scenario 1 – H_0 is false		
SB	1.0	1.0	1.5	1.0	1.0
SDSD	1.0	1.0	1.0	2.5	1.0
LCS	1.0	1.0	1.0	1.0	0.9
			Scenario 2 – H_0 is true		
SB	1.0	1.0	1.0	2.5	0.9
SDSD	1.0	1.0	1.5	1.0	0.9
LCS	1.0	1.0	1.5	2.5	1.0

SB, square of bias; *SDSD*, the difference in the magnitude of fluctuation between the simulation and measurement; *LCS*, the lack of positive correlation weighed by the standard deviations; H_0, null hypothesis; μ, population mean; σ^2, variance of the population; ρ, correlation of the population.

All simulations and statistical evaluations were performed using the statistical software R, version 3.2.3. The function "mvrnorm" in the MASS package (version 7. 3-45) was used to simulate the bivariate normal populations.

To evaluate the frequencies of correct or incorrect decisions within each scenario, one hundred random samples of each population were taken, creating one hundred datasets with one hundred records each (n=100). These datasets were individually submitted to the hypotheses test (Equation 17) by considering α = 0.05. The results were computed by each component, and the frequency of acceptance and rejection of the null hypothesis was computed for both scenarios (Table 2). It is clarified that this simulation allowed obtaining populations that agreed to what was planned and that random samples presented, on average, characteristics that were very close to those previously planned to performed the evaluation of the statistical efficiency of the method.

The frequencies of acceptance of null hypothesis when this is true varied from 97 to 100%, averaging 99% (Table 2). This value represents an indicator of confidence coefficient of the test and its complement (1%) was lower than the nominal α value proposed for the simulation. On the other hand, the frequencies of rejection of null hypothesis when this is false varied from 95 to 100%, averaging 97%. This value is an indicator of the power of the test and its complement (3%) represents an approaching to the type II error (KAPS and LAMBERSON, 2004), which could be assumed of low magnitude.

In this sense, from simulations results, it can be concluded the method proposed here seems present adequate levels of confidence and adequate power to point out which of the MSEP components are statistically significant.

Table 2 - Frequencies (%) of the decisions based on hypotheses test carried out using the method on the information obtained from simulated populations in different scenarios

Decision	Scenario: H_0 is true	Scenario: H_0 is false
	SB	
Accept	97	5
Reject	3	95
	SDSD	
Accept	100	4
Reject	0	96
	LCS	
Accept	100	0
Reject	0	100

SB, square of bias; *SDSD*, the difference in the magnitude of fluctuation between the simulation and measurement; *LCS*, the lack of positive correlation weighed by the standard deviations.

4 Example of application of the method

To demonstrate the method here proposed, it was utilized a dataset obtained from the simulation of the content of total digestible nutrients in cattle diets from the chemical composition of the feeds and using two different mathematical models (n = 107; DETMANN *et al*., 2008; Table 3).

First, the homoscedasticity of the *MSEP* obtained with both models (Table 3) can be evaluated by using the F statistics, as follow:

$$\hat{F} = \frac{> MSEP}{< MSEP} = \frac{MSEP_I}{MSEP_{II}} = \frac{6505.8}{3156.2} = 2.0613 \qquad (18)$$

The F statistics obtained in Equation (18) presents 107 degrees of freedom ($d.f. = n$) for both numerator and denominator, which corresponds to an approximate P value of 0.0001. Thus, one can infer that Model II presented lower (or more homogeneous) MSEP compared to Model I.

The evaluation of the SB component for both models (Table 3, Equation 17) gives:

$$\hat{\chi}^2_{SB} = -2 \times \log_e(\Lambda) = -n \times \log_e\left(\frac{MSEP - SB}{MSEP}\right)$$

$$\hat{\chi}^2_{SB(I)} = -107 \times \log_e\left(\frac{6505.8 - 1390.0}{6505.8}\right) \cong 25.71 \quad \textit{(P value < 0.001)} \qquad (19)$$

$$\hat{\chi}^2_{SB(II)} = -107 \times \log_e\left(\frac{3156.2 - 76.4}{3156.2}\right) \cong 2.62 \quad \textit{(P value = 0.105)}$$

Table 3 - Decomposition of the mean square error of prediction obtained for the content of total digestible nutrients (g/kg of dry matter) in cattle diets simulated by two different models (n = 107; Detmann *et al*., 2008)

Item	Observed values	Predicted values	
		Model I	Model II
Mean	645.1	607.8	653.8
Standard deviation	68.8	73.7	49.7
Linear correlation	-	0.4976	0.6026
MSPE	-	6505.8	3156.2
SB	-	1390.0	76.4
SDSD	-	24.1	365.2
LCS	-	5091.7	2714.6

MSEP, mean square error of prediction; *SB*, square of bias; *SDSD*, the difference in the magnitude of fluctuation between the simulation and measurement; *LCS*, the lack of positive correlation weighed by the standard deviations.

The evaluation of the *SDSD* component for both models (Table 3, Equation 17) gives:

$$\hat{\chi}^2_{SDSD} = -2 \times \log_e(\Lambda) = -n \times \log_e\left(\frac{MSEP - SDSD}{MSEP}\right)$$

$$\hat{\chi}^2_{SDSD(I)} = -107 \times \log_e\left(\frac{6505.8 - 24.1}{6505.8}\right) \cong 0.39 \quad \textit{(P value = 0.532)} \qquad (20)$$

$$\hat{\chi}^2_{SDSD(II)} = -107 \times \log_e\left(\frac{3156.2 - 365.2}{3156.2}\right) \cong 13.15 \quad \textit{(P value < 0.001)}$$

The evaluation of the *LCS* component for both models (Table 3, Equation 17) gives:

$$\hat{\chi}^2_{LCS} = -2 \times \log_e(\Lambda) = -n \times \log_e\left(\frac{MSEP - LCS}{MSEP}\right)$$

$$\hat{\chi}^2_{LCS(I)} = -107 \times \log_e\left(\frac{6505.8 - 5091.7}{6505.8}\right) \cong 163.30 \quad \textit{(P value < 0.001} \qquad (21)$$

$$\hat{\chi}^2_{LCS(II)} = -107 \times \log_e\left(\frac{3156.2 - 2714.6}{3156.2}\right) \cong 210.43 \quad \textit{(P value < 0.001)}$$

By assuming α = 0.05, it can be inferred that Model I presents constraints associated with bias (*SB*) and incomplete covariation (*LCS*) (P<0.05), despite of produces simulated data with a range of variation similar to that one observed for actual values. On the other hand, Model II simulates accurately the real values, but it presents constraints associated with simulating equal variance (*SDSD*) and also produces simulated data with incomplete covariation (*LCS*) with real data (P<0.05).

From the results presented in Equations (19), (20), and (21), it was possible to point out that evaluated models present different characteristics with regards the simulation constraints. Such information would be helpful to the modelers guide the efforts to improve the mathematical structure of each model.

5 Discussion

The simulation vs. measurement comparison based on MSEP is straightforward, where the whole *MSEP* indicates the overall deviation of the model output from the measurement, and the *MSEP* components will represent the different aspects of the overall deviation. Thus, for direct comparisons between model output and measurement, the MSEP-based analysis seems better than the commonly practiced correlation-regression analysis (KOBAYASHI and SALAM, 2000) because such analysis tends to focus more on the fitting of the regression than do on the actual limitations of the model itself. However, a single decomposition of the *MSEP* without guidance based on inductive statistics can be difficult for the true identification of the model's constraints because the modeler will not have tools to make decisions about the actual relevance of each component to the overall value of *MSEP*.

Regazzi and Silva (2004) have used the likelihood ratio test based on χ^2 distribution to evaluate the equality of parameters in non-linear models. The test developed by these authors and the hypotheses test proposed here have been based on the same statistics (Equation 16). They worked with a great number of simulated samples and found that occurrence of type I error decreases and becomes closer to the chosen α value as the sample size increases. This seems to give a theoretical support that the likelihood ratio test here proposed can be useful also to evaluate the significance of the *MSEP* components. On the other hand, other statistical approaches to evaluate components of *MSEP* have been proposed (*e.g.,* KOBAYASHI and SALAM, 2000). Nonetheless, those approaches demand an estimate of error variance which would only be obtained when experimental replications are available. It is rather difficult to obtain as most models are either evaluated or validate by using samples of measured values without any experimental structure (*i.e.,* without replicates). Therefore, the approach here presented is simpler with regards pre-requisites for applying, because an estimate of error variance is not necessary (Equation 17).

The model evaluation or validation is a step that assures if the model yields simulation results in quantitative agreement with results obtained in studies of the real system or, in other words, the model accurately simulates reality and does so for the right reasons (BALDWIN, 1995). The identification and acceptance of wrongness of a model is an important step towards the development of more reliable and accurate models. The assessment of the adequacy of mathematical models is only possible through the use of a combination of several statistical analyses and proper investigation regarding the purposes for which the mathematical model was initially conceptualized and developed for. The usefulness of a model should be assessed for its sustainability for a particular purpose (TEDESCHI, 2006). From these statements, it is understood that association of statistical tests with the algebraic decomposition of *MSEP* might help the modeler to focus on the real significant constraint of his model and to work on doing the necessary modifications on the mathematical structure of the model for improving the accuracy and precision of the simulated values.

In spite of the results obtained here, it must be emphasized that the hypothesis test based on χ^2 distribution should be seen as a relative inference tool rather than an absolute diagnostic of the model. This statement means that the results obtained by its application will point out the main actual constraint of the model and, therefore, which are the main refinements to be done on the mathematical structure of the model. More refined evaluations of the model predictions must be performed by using simultaneously several mathematical approaches for that.

Conclusions

The test of hypothesis based on likelihood ratio can be properly used to evaluate the statistical significance of the components of the mean square error of prediction. This test might help the modeler to focus on the real significant constraints of his model and to work on doing the necessary modifications on its mathematical structure in order to improve the accuracy and precision of the simulated values.

Acknowledgements

The authors gratefully acknowledge to the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) and Instituto Nacional de Ciência e Tecnologia de Ciência Animal (INCT - Ciência Animal) for financial support.

- DETMANN, E.; BONFA, H. C.; CECON, P. R.; SILVA, F. F. Proposição de teste de hipóteses para os componentes do quadrado médio do erro de predição. *Rev. Bras. Biom.*, Lavras, v.35, n.4, p.658-669, 2017.

- *RESUMO: A avaliação do quadrado médio do erro de predição constitui importante ferramenta para a comparação de valores reais com valores simulados por modelos matemáticos. Esta avaliação está baseada na decomposição original do quadrado médio do erro de predição em três componentes, os quais indicarão os prováveis entraves do modelo em relação ao viés,* *variâncias heterogêneas e incompleta covariação em relação aos valores reais. Contudo, esta análise tem sido conduzida como um procedimento descritivo, sem a orientação dada por um teste de hipóteses adequado. Desta forma, objetivou-se desenvolver um teste de hipóteses simples para a avaliação dos componentes do quadrado médio do erro de predição baseando-se no teste da razão de verossimilhança e na distribuição de χ^2. Esta aproximação foi avaliada por intermédio de populações simuladas e aplicada sobre um conjunto de dados obtido pela simulação de características da dieta de bovinos utilizando-se dois diferentes modelos matemáticos. Concluiu-se que o teste proposto pode auxiliar profissionais na área de modelagem para a identificação dos entraves ou limitações reais dos modelos desenvolvidos, orientando-os no desenvolvimento de modificações necessárias na estrutura matemática dos modelos de forma a ampliar a exatidão e a precisão dos valore simulados.*

▪ *PALAVRAS-CHAVE: Inferência; modelos matemáticos; simulação; validação.*

References

ARCHONTOULIS, S.V.; MIGUEZ, F. Nonlinear regression models and applications in agricultural research. *Agronomy Journal*, Madison, v.107, p.786-798, 2015.

BALDWIN, R.L. *Modeling ruminant digestion and metabolism.* London: Chapman & Hall, 1995. 587p.

BIBBY, J.; TOUTENBURG, H. *Prediction and improved estimation in linear models.* New York: John Wiley & Sons, 1977. 188p.

DETMANN, E.; VALADADES FILHO, S.C.; PINA, D.S.; HENRIQUES, L.T.; PAULINO, M.F.; MAGALHÃES, K.A.; SILVA, P.A.; CHIZZOTTI, M.L. Prediction of the energy value of cattle diets based on the chemical composition of the feeds under tropical conditions *Animal Feed Science and. Technology*, Amsterdam, v.143, p.127-147, 2008.

FRANCE, J.; KEBREAB, E. Introduction. In: FRANCE, J.; KEBREAB, E. (Eds.) *Mathematical modelling in animal nutrition.* Wallingford: CAB International, 2008. p.1-11.

GAUCH Jr., H.G.; HWANG, J.T.G.; FICK, G.W. Model evaluation by comparison of model-based predictions and measured values. *Agronomy Journal*, Madison, v.95, p.1442-1446, 2003.

KAPS, M.; LAMBERSON, W. *Biostatistics for animal science*. Wallingford: CABI Publishing, 2004. 445p.

KOBAYASHI, K. Comments on another way of partitioning mean square deviation proposed by Gauch et al. (2003). *Agronomy Journal*, Madison, v.96, p.1206-1208, 2004.

KOBAYASHI, K.; SALAM, M.U. Comparing and measured values using mean squared deviation and its components. *Agronomy Journal*, Madison, v.92, p.345-352, 2000.

RAO, C.R. *Linear statistical inference and its application.* 2ed. New York: John Wiley & Sons, 1973. 656p.

REGAZZI, A.J.; SILVA, C.H.O. Test for parameters equality in nonlinear regression models. I. Data in the randomized complete design. *Revista de Matemática e Estatística*, São Paulo, v.22, p.33-45, 2004.

SEARLE, S.R.; CASELLA, G.; McCULLOCH, C.E. *Variance components.* New York: John Wiley & Sons, 1992. 501p.

TEDESCHI, L.O. Assessment of the adequacy of linear models. *Agricultural Systems*, Amsterdam, v.89, p.225-247, 2006.

THEIL, H. 1966. *Applied economic forecasting*. Amsterdam: North Holland Publishing Co., 1966. 474p.

THORNLEY, J.H.M., FRANCE, J. *Mathematical models in agriculture*. *Quantitative methods for the plant, animal and ecological sciences*. 2ed. Wallingford: CAB International, 2007. 906p.

Received in 20.06.2016

Approved after revised in 15.11.2016