Mathematical models for the innovative interpretation of the associative effect

Franci O., Acciaioli A.

in

Antongiovanni M. (ed.).
Exploitation of Mediterranean roughage and by-products

Zaragoza : CIHEAM
Options Méditerranéennes : Série B. Etudes et Recherches; n. 17

1998
pages 73-78

Mathematical models for the innovative interpretation of the associative effect

O. Franci and A. Acciaioli
Dipartimento di Scienze Zootecniche, Universitá di Firenze, Via delle Cascine 5, 50144 Florence, Italy

SUMMARY - The more recent methods of analysis of additive and associative effects of feeds use the continuous system. If there are two components, wich are, at most, complementary the test of linearity or curvilinearity of regression analysis allows to estimate the presence of their associative effect. For more components, the Response Surface Methodology (RSM) associated to multiple regression analysis is employed. Statistical procedures are discussed for fitting a RSM to experimental data and an outline is made of the mathematical process for finding the stationary point. The specific case of a three-component system, whose sum is the whole, is examined. Examples, even in graphic format, are produced to show the application of continuous analysis to studies on the associative effects between feeds.

Key words: Associative effect, feed, mathematical model, response surface, multiple regression, triangular graph.

RESUME - *"Modèles mathématiques pour une interprétation novatrice de l'effet associatif". Les méthodes les plus récentes d'analyse des effets additifs et associatifs des aliments utilisent les systèmes continus. Si les composants en jeu sont deux au maximum, complémentaires entre eux, le test de linéarité ou de curvilinéarité de la régression permet l'évaluation de la présence d'effet associatif. Quand l'expérimentation intéresse plusieurs composants on utilise la Méthodologie de la Surface de Réponse (RSM) associée à l'analyse de la régression multiple. Les procédures statistiques pour adapter la RSM aux données expérimentales sont discutées et le processus mathématique pour individualiser le point stationnaire est indiqué. On a examiné aussi le cas particulier d'un système à trois composants, dont la somme est l'entier. Enfin, des exemples, même en forme graphique, sont présentés pour montrer l'application de l'analyse continue aux études sur les effets associatifs entre aliments.*

Mots-clés : *Effet associatif, aliment, modèle mathématique, surface de réponse, régression multiple, graphique triangulaire.*

Introduction

The study on additive and interactive effects resulting from the use of different feeds categories, single feeds or simple nutritive principles in diets for animals is carried out by means of the discrete analysis or the continuous one. As far as the first method is concerned, several classes corresponding to the levels of the considered factor are constituted and the contrasts between the relative means are statistically analysed. If two or more factors are involved, a more complex factorial design is performed and the interaction between the factors can be tested. However, in this analysis a seesaw trend of responses, statistically significant, can often emerge which makes difficult and feeble the interpretation of the results, particularly when a definite relationship dose-response is expected. The continuous analysis is recently preferred since, by means of a suitable design of experiment, it makes possible to quantify, in infinitesimal manner, the additive and interactive response of the factor(s) along the whole considered range.

Continuous analyses

Monofactorial design

When an only factor is considered (e.g. percentage of protein, percentage of concentrate in the ration), the use of analysis of simple linear regression appears easy and the relative results are directly intelligible. It happens even if the variation of a factor determines the opposite variation of another, as complement to 100, since the regression analysis regards, obviously, one independent variable only.

The expected additivity of the factor (dose-response relationship) is tested by the analysis of linearity of the regression, whereas the interactive effect (associative effect with the complementary component) is estimated by the significance of the quadratic coefficient.

Example

An example for this approach can be drawn from the work of Berge and Dulphy (1991). The authors studied the effect of interaction between concentrate and roughage on the digestibility of the diet. The results of analysis are graphically reported in Fig. 1 where the parabolic trend of the digestibility is evident. The value observed on experimental animals fed diet of intermediate composition significantly deviated from the expected value, obtained additively as weighed mean of digestibility of the roughage and that of the concentrate.

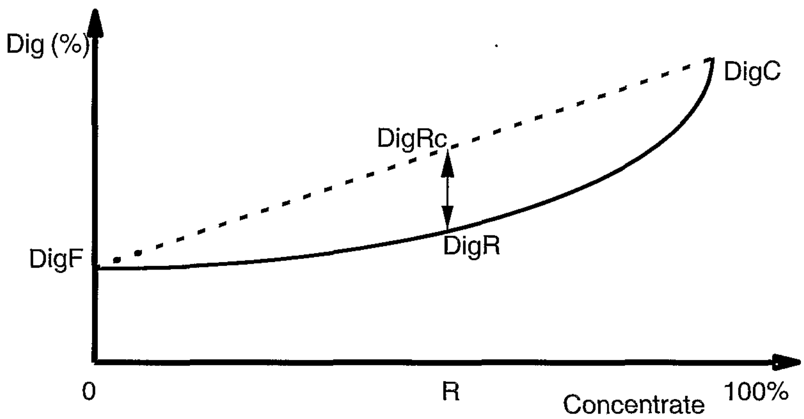

Fig. 1. Deviation from linearity for digestibility of ration R (DigR) in respect to average digestibility of concentrate (DigC) and forage (DigF) (from Berge and Dulphy, 1991).

Polyfactorial design

When the investigation concerns the relationship among several quantitative variables (e.g. two or more ingredients or nutritive principles in the ration), in an effort to optimize a response, the *Response Surface Methodology (RSM)* technique can be applied. This technique is an integration of experimental strategies, mathematical methods and statistical inference and allows for the simultaneous variation of several factors to find the level that will give the most interesting response.

The theory and the technique of RSM have been described in a number of papers or reviews (Box and Wilson, 1951; Box, 1954; Mead and Pike, 1975) and the RSM method has been often applied to animal science researches, particularly in animal nutrition studies (e.g. Yoshida *et al.*, 1969; Roush *et al.*, 1979; Toyomizu *et al.*, 1982, 1993).

The fitting of a response surface to experimental data is an extension of multiple regression. It is assumed that there is a mathematical function that describes a response. The empirical model chosen is usually a second order polynomial and is fitted to the experimental data by regression methods. It is of advantage if the experiment is designed so that the centre of the surface will be as close to the theoretical optimum response as possible. In this case enough information will be provided without an excessive number of treatments. Box and Wilson (1951) developed a central composite rotatable design for fitting second order response surfaces.

Statistical analysis

The statistical analysis is performed by the standard multiple regression procedure to approximate the coefficients; if there are two independent variables, the following model is employed:

$$Y = a + b_1X_1 + b_2X_2 + c_1X_1^2 + c_2X_2^2 + dX_1X_2 \qquad \text{Model 1}$$

where Y is the response; a is the intercept; b_i, c_i and d are the coefficients; and X_i are the variables.

Standard analysis of variance procedure is used to fit the equation to the experimental data. The selection methods (e.g. stepwise) can be used to select the variables of second order and/or the interactive variable which remain in a restricted final model. However also the whole model can be performed and the significances obtained from the analysis of variance indicate if the surface is quadratic or linear. In the latter case, there is neither maximum nor minimum for the experimental area. On the contrary, when an order higher than the second is significant, it is an indication that the experimental design covers a too large region of a factor.

Once the surface is fitted with a second order equation the *stationary point* and the *yield at the stationary point* can be determined. As far as the stationary point is concerned it represents the point on the response surface where the slope of the surface is equal to zero. The vector **x** for the stationary point is $\mathbf{x} = -0.5\mathbf{A}^{-1}\mathbf{b}$ where **A** is the symmetrized matrix with diagonal elements equal to the coefficients of the quadratic terms and off-diagonal elements equal to half the coefficient of the cross-product; **b** is the vector of linear coefficients. The matrix equation for the expected yield at the stationary point is $y = a + 0.5\mathbf{x}'\mathbf{b}$ where a is the intercept (Model 1) and **x**' is the transpose of **x**.

It should be noted that the stationary point is not necessarily the point that maximizes the response. If both coefficients of second order are negative, the stationary point is at a maximum value; if they are positive, there is a minimum; if the coefficients have mixed signs, the point represents a saddle point. More details on mathematical process and derivation of the matrix formulae are found in Myers (1976).

Three-component system

The basic principles of the response surface methodology are applied with some variation in studies on animal nutrition. Particularly, some authors (Toyomizu *et al.*, 1982, 1993; Franci *et al.*, 1997) used the RSM in a three-component system experiment. When the response to three dietary factors such as protein, fat and carbohydrate content, whose sum is 100, is to be embodied on the rectangular co-ordinates, one of the three factors is inevitably eliminated. The composition expressed as a percentage of three dietary factors, whose sum is the whole, can be plotted as a point on Roozemboom's triangular graph (Fig. 2). A given point of the diagram (equilateral triangle) has a constant sum (e.g. 100) of the three perpendicular heights from each side and the sum is equal to the height of the triangle. Another property of the triangular diagram is that the line drawn through a vertex represents all the combinations in which the ratio of two ingredients remains constant whereas the proportion of the third varies from zero to the maximum.

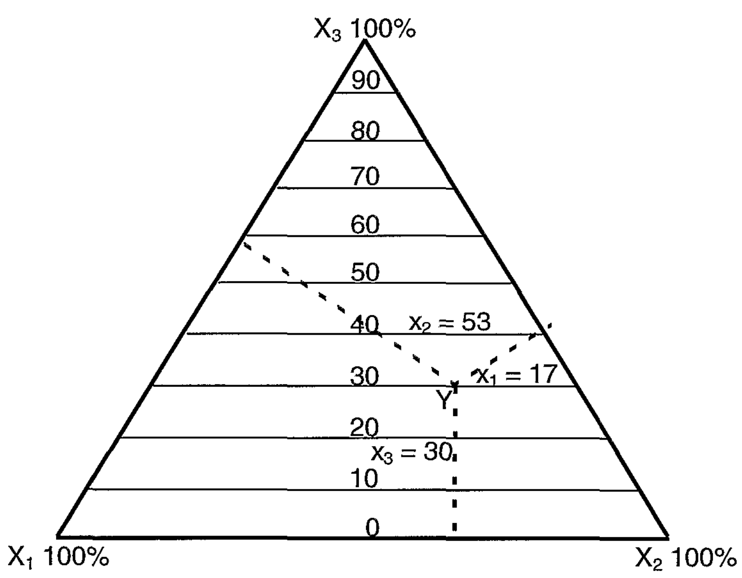

Fig. 2. Triangular co-ordinate graph showing the composition of the diet Y as a mixture of the three ingredients (X_1, X_2 and X_3), whose sum is 100.

In a system with three components, related each other such as $X_1 + X_2 + X_3 = 100$, the independent variables are reduced since specifying each two out of the three factors fixes the value of the third ($X_3 = 100 - X_1 - X_2$).

To study the additive and the interactive effects of three dietary ingredients ($X_1 + X_2 + X_3 = 100$), Franci *et al.* (1997) proposed the following model:

$$Y = a + b_1X_1 + b_2X_2 + c_1(X_1X_2) + c_2(X_1X_3) + c_3(X_2X_3) \qquad \text{Model 2}$$

In this kind of analysis, the intercept (a) represents the response of the diet entirely constituted by the ingredient X_3 which is removed from the model, apart from the interactive components, to allow the independence of the variables; the partial linear coefficients (b_i) measure the effect of substitution of the i^{th} ingredient for X_3 and therefore its additive effect; the coefficients c_i measure the effect of interaction of first order between the ingredients and therefore indicate their associative effect. Standard analysis of variance procedures are used to determine the significance of the coefficients but, apart from the significances, the model is left in the complete form to describe the behaviour of the dependent variable, also in graphic form.

To allow a 3D graphic representation, only two independent variables can be used and the original equation is transformed into a quadratic one by placing $X_3 = 100 - X_1 - X_2$. Therefore the Model 2 is changed to the following form, analogous to the Model 1:

$$Y = a + (b_1+100c_2)X_1 + (b_2+100c_3)X_2 - c_2(X_1)^2 - c_3(X_2)^2 + (c_1-c_2-c_3)X_1X_2$$

Example

Franci *et al.* (1997) studied the effect of lucerne hay (H), wheat straw (S) and maize gluten feed (G) on the performances of lambs. Nine diets, composed with different proportion of H, S and G, were given to 27 animals during a 50 days growth period. A model analogous to the Model 2 above mentioned was performed, where the linear component of H was not considered. Figure 3 shows the pattern of the feed combinations and indicates the behaviour of the third variable H, which is not reported in the two orthogonal axes.

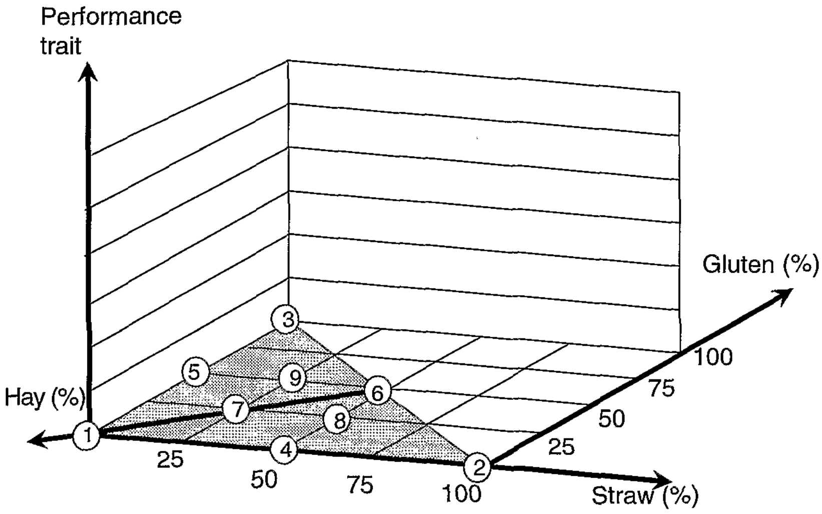

Fig. 3. Surface of feed combinations: diet 1 = HHHH, diet 2 = SSSS, diet 3 = GGGG, diet 4 = HHSS, diet 5 = HHGG, diet 6 = SSGG, diet 7 = HHSG, diet 8 = HSSG, diet 9 = HSGG.

Table 1 reports the parameters of the relationship between the empty body weight (EBW) daily gain and the diet composition. Table 2 reports the estimated means relative to the nine ingredients combinations used in the experiment.

Table 1. Parameters of multiple regression equations for empty body weight daily gain (g) of lambs (Franci *et al.*, 1997). S (straw), G (maize gluten feed) and H (hay) are expressed as percent unit

Intercept	Regression coefficient ± SE					RSD	R^2
	S	G	S*G	S*H	G*H		
229.3	-2.05†	0.81†	0.043†	0.011	-0.013	34.8	0.87
	±0.28	±0.28	±0.009	±0.009	±0.009		

†Significant coefficient ($P < 0.05$)

Table 2. Estimated means (±SE) of empty body weight daily gain (g) for the nine diets given to lambs (Franci *et al.*, 1997)

Diet								
HHHH†	SSSS	GGGG	HHSS	HHGG	SSGG	HHSG	HSSG	HSGG
229	23	309	153	236	274	221	206	262
±20	±20	±20	±18	±18	±18	±10	±10	±10

†Each letter (H, G and S) indicates 25% of the relative ingredient (hay, gluten or straw) in the diet

It is evident that daily gain was negatively influenced by S (minus 2 g per % unit of substitution of S for H) and positively influenced by G. Associative effect was found only for S*G which increased the performance.

The graphic representation, performed by the G3D PROC of SAS (1988), is shown in Fig. 4. The relative quadratic equation, obtained by transformation of the original as above specified, was:

$$EBW = 229.3 - 0.95*S - 0.49*G - 0.011*S^2 + 0.013*G^2 + 0.045*S*G$$

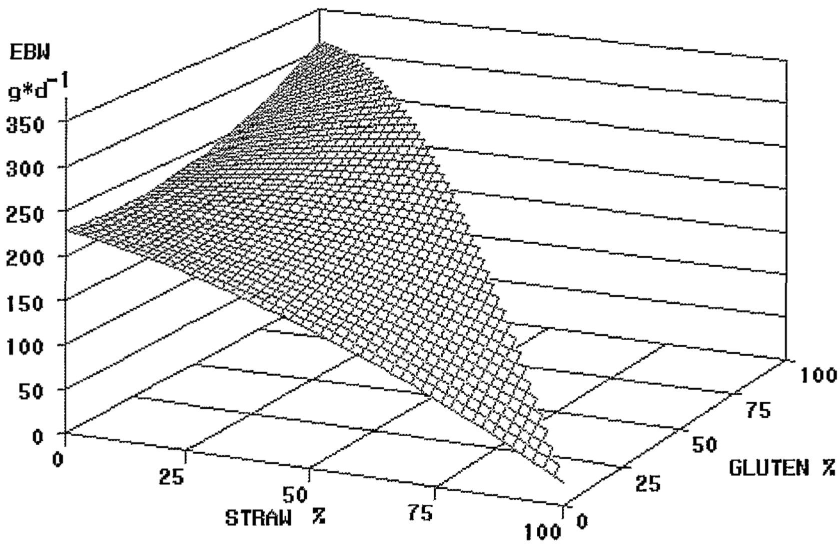

Fig. 4. Response surface of empty body weight (EBW) daily gains (g d^{-1}) in relation to straw and gluten content of the diet.

A saddle point, or minimax, due to the mixed signs of the second degree coefficients, is located at about S = 35% and G = 25% (corresponding to H = 40%). However, the continuous pattern shows that the highest performances can be achieved, along the line of H equal to zero, without resorting to the maximum dose of G, but merely by replacing G with modest amount of S (about 20%).

References

Berge, P. and Dulphy, J.P. (1991). Etude des interactions entre fourrage et aliment concentré chez le mouton. II. Facteurs de variation de la digestibilité. *Ann. Zootech.*, 40: 227-246.

Box, G.E.P. (1954). The exploration and exploitation of response surfaces: Some general considerations and examples. *Biometrics*, 10: 16-60.

Box, G.E.P. and Wilson, K.G. (1951). On the experimental attainment of optimum conditions (with discussion). *J. Roy. Statist. Soc., Ser. B*, 13: 1-45.

Franci, O., Antongiovanni, M., Acciaioli, A., Bruni, R. and Martini, A. (1997). Response Surface Analyses of the associative effects of lucerne hay, wheat straw and maize gluten feed on growing lambs. *Anim. Feed Sci. Technol.* (in press).

Mead, R. and Pike, D.J. (1975). A review of response surface methodology from a biometrics viewpoint. *Biometrics*, 31: 803-851.

Myers, R.H. (1976). *Response surface methodology*. Edwards Bros., Inc., Ann Arbor, MI.

Roush, W.B., Petersen, R.G. and Arscott, G.H. (1979). An application of response surface methodology to research in poultry nutrition. *Poultry Sci.*, 58: 1504-1513.

SAS (1988). *SAS/GRAPH user's guide. Release 6.03*. SAS Institute, Inc., Cary, NC.

Toyomizu, M., Akiba, Y., Horiguchi, M. and Matsumoto, T. (1982). Multiple regression and response surface analyses of the effect of dietary protein, fat and carbohydrate on the body protein and fat gains in growing chicks. *J. Nutr.*, 112: 886-896.

Toyomizu, M., Kimura, S. and Tomita, Y. (1993). Response surface analyses of the effects of dietary protein, fat and carbohydrate on feeding and growth pattern in mice from weaning to maturity. *Anim. Prod.*, 56: 251-259.

Yoshida, M., Hoshii, H. and Morimoto, H. (1969). Age difference in the pattern of growth response to the change in the dietary energy and protein levels. *Jpn. Poultry Sci.*, 6: 7-21.