JOURNAL OF CHEMOMETRICS
J. Chemometrics 2002; **16**: 294–304
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/cem.722

Exploring complex interactions in designed data using GEMANOVA. Color changes in fresh beef during storage

Rasmus Bro[1]* and Marianne Jakobsen[2]

[1]Food Technology, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark
[2]Food Chemistry, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark

Received 15 May 2001; Accepted 20 February 2002

Data from a severely reduced experimental design are investigated in order to obtain detailed information on important factors affecting the changes in quality of meat during storage under different conditions. It is possible to model the response, meat color, using traditional ANOVA (analysis of variance) techniques, but the exploratory and explanatory value of this model is somewhat restricted owing to the number of factors and the fact that several interactions exist. For those reasons, it is not possible to visualize the model in a simple way and therefore not possible to have a clear overview of the total variation in the data. Using a recently suggested alternative to traditional analysis of variance, GEMANOVA (generalized multiplicative ANOVA), it is possible to analyze the data effectively and obtain a more interpretable solution that enables a simple overview of the whole sampling domain. Whereas traditional analysis of variance typically seeks a model with main effects and as few and simple interactions and cross-products as possible, the GEMANOVA model seeks to describe the data primarily by means of higher-order interactions, albeit in a straightforward way. The two approaches are thus complementary. It is shown that the GEMANOVA model is simple to interpret, primarily because the GEMANOVA structure is in agreement with the nature of the data. It is shown that the GEMANOVA model used is mathematically unique, which leads to attractive simplified ways of interpreting the model. The results presented are the first published results where the GEMANOVA model is not simply equivalent to an ordinary PARAFAC model, thus taking full advantage of the additional structural power of GEMANOVA. A new algorithm for fitting the GEMANOVA model is developed and is available from the authors.

KEYWORDS: PARAFAC; interactions; analysis of variance; ANOVA; multilinear; response surface

1. INTRODUCTION

Modified atmosphere packaging (MAP) is widely used to extend the shelf-life of fresh meat. Gas flushing is used to replace the air surrounding the meat with an atmosphere containing elevated levels of oxygen and carbon dioxide. Normally an atmosphere containing 20%–30% CO_2 and 70%–80% O_2 is used for retail packaging of fresh beef. An elevated level of carbon dioxide is used to retard microbial growth and an elevated level of oxygen is used to prolong color stability [1]. Color is the most important factor for consumer preferences when purchasing fresh meat. The attractive red meat color stems from the oxygenated derivative (oxymyoglobin) of the meat pigment. A high oxygen level stabilizes the color, but inevitably the pigment will oxidize to metmyoglobin, resulting in an undesirable brown meat color. A high oxygen level is also expected to enhance other deteriorative processes in meat, such as lipid oxidation [2,3].

Color stability of fresh meat is influenced by a large number of factors, some of a biochemical nature, some due to handling during the slaughter process. Storage and packaging conditions can also influence the color shelf-life of meat [4,5]. This study focuses on describing to what extent different storage and packaging conditions affect the color stability of fresh beef. Different levels of storage time, temperature, oxygen content in the package headspace and extent of light exposure were investigated. Development of mathematical models describing color changes during storage can help identify the important factors and find critical levels for these factors. Further, such models can form the basis for designing an optimal modified atmos-

*Correspondence to: R. Bro, Food Technology, Department of Dairy and Food Science, The Royal Veterinary and Agricultural University, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark.
E-mail: rb@kvl.dk
Contract/grant sponsor: LMC (Centre for Advanced Food Studies) FØTEK program.
Contract/grant sponsor: AQM; Contract/grant number: FR9801214.

phere composition, also affecting other important quality-deteriorative processes in meat.

Large variations in color stability between meat of different origin can influence the empirical mathematical models. Different muscle types show large variability owing to different myoglobin content and different metabolic type of activity. Meat from animals of different age, breed, feeding, etc. will also result in differences in color stability [5].

The data analyzed here come from an experiment using *longissimus dorsi* muscles from three different animals. Supplementary studies have been performed investigating meat from seven different animals and two different muscle types: *longissimus dorsi* (high color stability) and *semimembranosus* (intermediate color stability). Comparison of GEMANOVA models for these new data sets will be treated in a future publication.

1.1. Chemometrics

A large body of tools is available for analyzing experimentally designed data. Usually, analysis-of-variance tools are used [6–10]. In these tools, each continuous factor is treated as one independent variable. For example, an experiment may include a factor *temperature* with values 10, 20 and 30 °C. Qualitative variables are typically split into a number of binary variables; for example, *substrate added* with levels 1 and 0 corresponding to substrate being added or not. If the qualitative variable has more than two levels, it is typically artificially divided into more variables; for example, the three variables *transportation by car* with levels 0 and 1, *transportation by bicycle* with levels 0 and 1 and *transportation by train* with levels 0 and 1. These artificial variables are then related so that if one is at level 1, the remaining ones have to be at level 0.

Thus both quantitative and qualitative variables are treated as pseudo-quantitative variables, which makes straightforward modeling of the correlation between the response(s) and the designed factors possible. Assume that there are I experiments and J variables. Further assume, for simplicity, that there is only one response whose values are held in an I-vector $\mathbf{y}$ with the ith element holding the response for the ith experiment. Let the $I \times J$ matrix $\mathbf{X}$ hold the factors; e.g. the first column in $\mathbf{X}$, $\mathbf{x}_1$, holds the I temperature readings, the second column the *substrate added* variable, the third column the *transportation by car* variable, etc. Then the problem of assessing the effect of each variable on the response can be expressed as a standard multiple linear regression problem

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{t} \tag{1}$$

where $\mathbf{t}$ is an I-vector of residuals holding the part of the response $\mathbf{y}$ that cannot be explained by the experimental design and $\mathbf{b}$ is a J-vector of regression coefficients (b_j) that provides information on the effects. Usually both $\mathbf{X}$ and $\mathbf{y}$ are centered and possibly scaled prior to regression in order to level out differences in offset and magnitude. The relative sizes of the regression coefficients (with their associated uncertainties) provide the relative importance of each variable.

The above analysis-of-variance model is *additive* in each independent variable. This means that the effect of each factor is independent of the variation in other factors. Raising the level of (possibly centered and scaled) variable 1 from level 1 to 5 implies an increase in the response by $5b_1$. In some situations, such a simple additive model of *main effects* is not realistic because the factors do not affect the response independently. A common, natural and well-working remedy to this is to allow *interactions* between the factors. For example, an interaction can be made between *temperature* and *substrate added* by adding a column in $\mathbf{X}$ which is simply the elementwise product of *temperature* and *substrate added*. With this approach, simple interactions such as two-way interactions can be handled efficiently. More complex, higher-order interactions are seldom used because they are difficult to interpret. The present data represent, however, a situation in which it is reasonable to assume that most of the variation in the data is caused by such higher-order interactions. Hence the data pose a challenge to the traditional analysis approach. It will be shown that the GEMANOVA approach recently suggested [11,12] provides a more natural basis for complicated data with higher-order interactions. The GEMANOVA model leads to models from which the underlying characteristics of the problem can be easily understood and visualized.

2. MATERIALS AND METHODS

2.1. Meat samples

Longissimus dorsi muscles of fresh beef were used. The muscles were matured in vacuum bags for 2 weeks at 2 °C. Subsequently the muscles were trimmed free of external fat and cut into 1.5 cm thick steaks. Meat from three different animals was used and left and right parts from each animal were treated independently, giving a total of six muscles. Meat from the different animals was handled in the same way after slaughter, e.g. maturing time, transportation conditions, etc., and all animals passed the normal quality control at the slaughterhouse. The animals may, however, be of different age, breed etc.

2.2. Storage and packaging conditions

Storage time, temperature, time of light exposure and oxygen content in the package headspace (balanced with carbon dioxide) were varied. The meat samples were placed in polystyrene trays and flow-packed using a laminated packaging material with an oxygen transmission rate of 40 cm^3 m^{-2} day^{-1} atm^{-1}. The samples were kept in refrigerated storage and exposed to light (Philips Fluotone TLD 18 W/830 yielding 1000 lux at the packaging surface) for 0%, 50% or 100% of the storage time. The temperature was monitored continuously during storage at several places in the refrigerator using data-loggers (TINY Talk II-Temp Loggers, RS Radio Parts, Copenhagen, Denmark). The meat samples had a volume of 80 ml and the headspace in the packs was 750 ml.

2.3. Instrumental analysis

On days 0, 3, 7, 8 and 10, color was measured on the meat surface, immediately after opening the package, using a Minolta Colorimeter CR-300 (Minolta, Osaka, Japan)

Table I. Levels of factors in designed experiments for which *a*-color response is measured

Variable	Levels
Storage time (days)	0, 3, 7, 8, 10
Temperature (°C)	2, 5, 8
O_2 content in headspace (%)	40, 60, 80
Exposure time to light (%)	0, 50, 100
Muscle no.	1, 2, 3, 4, 5, 6

Table II. Other characteristics of design

Variable	
No. of samples in full design	810
No. of samples in reduced design	324
Missing elements in array (%)	60

measuring the *L, a, b* co-ordinates (CIELAB color system). Red color was expressed as the *a*-value, and only this response value was used in the analysis. A high *a*-value represents a red color of the meat sample. The measurement was repeated on five randomly selected locations for each sample and the average used.

2.4. Experimental errors

The experimental error in the instrumental analysis is an important measure which can be used to evaluate the results from the model. The measurement error (incorporating both instrumental and sampling error) was determined by measuring at different locations on the same muscle five times. This yielded a pooled standard deviation of 2.2 and, as the data used are averaged over the replicates, the standard deviation of this average is approximately $2.2/\sqrt{5} = 1.0$. This is thus the error approximately expected to be the lowest attainable estimated error in the ANOVA model.

2.5. Experimental design (Tables I and II)

The full factorial design constitutes $5 \times 3 \times 3 \times 3 \times 6 = 810$ (storage × temperature × oxygen × light × muscle) combinations all to be measured in five replicates. Owing to the limited number of samples that can be taken from an individual muscle [2], a reduced design was chosen using a modified quadratic D-optimal design with 324 settings (40% of the 810 initial combinations). The experiment was performed six times (during 8 weeks) using meat from three different animals, two samples from each (left- and right-side muscle of each animal). The measurements were randomized over muscles. The actual measurements on different days are not performed on the same physical piece of meat but rather on samples from the same muscle. On day 0 there is no effect of storage and packaging conditions yet, and therefore all variable combinations are assigned the same color *a*-value calculated as the mean of five analyzed replicates on two samples from each of the six muscles.

2.6. Chemometrics

Models were developed using MATLAB™ ver. 5.3 for Windows and the N-way Toolbox [13]. The experiments were designed using the software MODDE™ (Umetrics AB, 1999) and initially analyzed in data analytical software The Unscrambler (Camo AS, 1999). The algorithm for fitting the GEMANOVA model has been described earlier, but the current data turned out to be difficult to handle, primarily owing to the huge amount of missing data. Therefore a new algorithm has been developed which is described in the Appendix. The algorithm is available from http://www.models.kvl.dk.

3. THEORY

3.1. Background

A new model for analyzing data from experimental designs has been proposed by Bro [11,12,14,15]. A similar idea was also presented by van Eeuwijk [16,17] for three-way data specifically. That model, however, is based on an extension of the singular value decomposition and consequently inherits the rotational ambiguity of this. This is not, in general, the case for the model suggested by Bro. This model is called GEMANOVA (generalized multiplicative analysis of variance) and is suited for analysis of data from factorial or fractional factorial designs in which the main variation is caused by higher-order interactions. The theory behind GEMANOVA can be found in the original references, but essential parts are explained in the following.

Consider an experimental design with three different factors varying on *I, J* and *K* levels respectively and one response. Let the response for the first factor at level *i*, the second factor at level *j* and the third factor at level *k* be denoted y_{ijk}. It then follows that the responses obtained from a full factorial design can be held in a three-way array $\underline{Y}$ ($I \times J \times K$) with elements y_{ijk}.

Although traditional analysis-of-variance models can be made to fit data from experimental designs well, the use of several interactions can make the interpretation difficult and thus the usefulness restricted. Conceptually, traditional analysis-of-variance models start from main effects and seek to keep the number of interactions as low as possible and of the lowest possible order. This is motivated by the interpretational complications arising from such interactions. As understanding is often one of the main aspects of analysis of variance, interactions are generally not 'hoped for'. In GEMANOVA the approach is reversed. The GEMANOVA model focuses on the interactions and on making practical models for these.

A three-way interaction for the above data $\underline{Y}$ is usually modeled as d_{ijk} in traditional analysis of variance of qualitative data, meaning that for any combination of the first, second and third factors an individual effect is estimated. This is problematic because the number of effects is then the same as the number of responses. No reduction in complexity is gained and thus no simple interpretation or understanding can be made. However, it is known from multivariate analysis that e.g. an ordinary two-way matrix with typical elements d_{ij} will seldom be full pseudo-rank. Mostly, a well-fitting low-rank decomposition can be made using e.g. principal component analysis. Thus, for a two-way interaction, the structure can often be simplified considerably using a bilinear decomposition.

Fisher and MacKenzie [18] worked on the influence of different manure treatments on different potato varieties. They noted that even though a standard two-factor experiment could be reasonably described by ordinary main effects a_i and b_j plus random noise t_{ij}, i.e.

$$y_{ij} = a_i + b_j + t_{ij} \quad i = 1, \ldots, I, \quad j = 1, \ldots, J \tag{2}$$

the model did not shed much light on the underlying phenomena causing the variations in the data. Instead, they proposed a *multiplicative* model

$$y_{ij} = a_i b_j + t_{ij} \quad i = 1, \ldots, I, \quad j = 1, \ldots, J \tag{3}$$

Even though the model did not provide a significantly better fit, the nature of the model was much more in line with the nature and understanding of the data, and therefore more interpretable and informative.

The usefulness of the model in Equation (3) as compared to a model where the interaction is modeled in the ordinary way as d_{ij} is that the effects of the two factors are *separated* and thus can be interpreted independently, much in the same fashion that main effects are interpreted independently. The difference, though, is that the joined effect is a product rather than a sum of the individual effects. Thus the multiplicative model of interactions is useful because it logically separates the interactions into entities pertaining to the individual factors.

As an intuitive simplified example of the usefulness of a multiplicative model, consider the physical activity of rats as a function of feed and water at critical levels. If the amount of feed is increased to a higher level, the activity of the rats will increase. However, if the level of water is low, the absolute increase in activity will be small compared to the increase observed when the amount of water is high. This is an example of a response which can be meaningfully approximated as a multiplicative effect. The effect of water is dependent on the level of feed and *vice versa*, but the *relative* effect may be independent. Even if it is possible to fit a model of such empirical data using only an independent main effect of water and a main effect of feed, the multiplicative model is preferred because it is more closely related to the nature of the data. In practice, a combination of such multiplicative terms and lower-order effects is usually beneficial, but the example serves to show that traditional analysis of variance conceptually starts from main effects, whereas sometimes it is reasonable to start from the interactions and only add lower-order effects to the extent that such effects are necessary.

Several authors have described different analysis-of-variance methods that include multiplicative terms for *two*-factor experiments [19–22]. In these studies, modifications of the traditional two-factor analysis-of-variance model are shown where main effects and interactions are estimated through the regular equations. Afterwards the interaction term is decomposed by PCA. Several examples are given where this approach gives more descriptive models.

3.2. GEMANOVA—the basics

Realizing that a two-way interaction is often appropriately modeled by a low-rank bilinear decomposition, it is natural to extend this idea also to higher-order interactions. Such ideas have already been proposed in a simplified way where lower-order interactions are ignored [23]. Consider a three-factor experiment and a three-way interaction of the response. The corresponding traditional analysis-of-variance model reads

$$y_{ijk} = d_{ijk} + t_{ijk} \quad i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K \tag{4}$$

where d_{ijk} is the interaction term and t_{ijk} the residual. There may of course also be other effects such as main effects or two-way interactions, but this is ignored here to keep the exposition simple. Note, in relation to e.g. a global offset/grand mean, that such a grand mean cannot be calculated beforehand and subtracted in the GEMANOVA model. In the case where a grand mean is needed, this is estimated and treated similarly to other effects.

It seems useful to consider if the three-way interaction in Equation (4) can be modeled by a low-rank trilinear decomposition. Instead of Equation (4), the data can be modeled, for the simplest situation, by a one-component trilinear model as

$$y_{ijk} = a_i b_j c_k + t_{ijk} \quad i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K \tag{5}$$

This model is actually *much* simpler than the non-decomposed three-way interaction model. Whereas the model d_{ijk} consists of IJK parameters, the one-component model in Equation (5) consists of only $I + J + K$ parameters. If the model fits the data well, a substantial reduction in complexity has been obtained and interpretation is simpler because there are fewer parameters and because relative effects are expressed separately for each factor.

The decomposition of the three-way array into one trilinear three-way interaction term is the starting point for three-factor GEMANOVA. If the data are higher-order, e.g. a four-way array arising from a four-factor factorial design, a four-way quadrilinear model is used ($a_i b_j c_k d_l$). In a study of enzymatic activity of polyphenol oxidase [12] a five-factor experimental design was evaluated. For these particular data it was found that the five-way array of responses of enzymatic activity, y_{ijklm}, could be excellently modeled by a five-way interaction term of the form $a_i b_j c_k d_l e_m$. Thus a simple and interpretable model was obtained. However, it cannot in general be expected that a response array can be modeled by one simple interaction term. Therefore in GEMANOVA the analysis proceeds by adding new terms to the model until a satisfactory fitting model is obtained.

There are many possible ways to add effects to a model. For the model in Equation (5) it can be the case that the addition of a main effect for the first factor is feasible, thus leading to a model

$$y_{ijk} = a_i b_j c_k + d_i + t_{ijk} \quad i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K \tag{6}$$

It is also possible to add an overall offset, a two-way

interaction or further develop the three-way interaction by adding a new trilinear term:

$$y_{ijk} = a_i b_j c_k + d_i e_j f_k + t_{ijk} \quad i = 1,\ldots,I, \quad j = 1,\ldots,J, \quad k = 1,\ldots,K \tag{7}$$

In practice, terms are added in such a way that it leads to the highest increase in fit (possibly corrected for the complexity of the term and possibly evaluated by cross-validation or similar methods). The search for additional effects is then conducted by adding either a trilinear (three-way interaction), bilinear (two-way interaction) or main effect or a global offset to the current model. This is continued until additional terms do not help in improving the model. In Equations (5)–(7) the actual content of the different parameters, e.g. a_i, will change from model to model, but the notation is maintained in all models for simplicity. When a new effect is added to a model, the whole model as such has to be refitted in order to obtain the least squares solution.

3.3. Technical aspects

Finding the useful effects has been presented here in a stepwise approach. This sequential search for optimal models cannot be guaranteed to lead to the globally optimal solution because the effects are not independent. It is of course also possible, although computationally demanding, to search for an appropriate model by trying all possible combinations of effects.

In assessing the significance of different postulated models, it is possible to correct the fitted sum-squared errors by the number of parameters in the model (subtracting one for each independent scale). This provides approximate mean squared errors, but as the validity of the approximation is uncertain, it is generally advised to defer this approach and instead use e.g. cross-validation. That is, the postulated model is fitted *IJK* times, each time leaving out one element. As the elements/experiments are usually independent, the fitted model and the left-out elements are also independent. The left-out element can be predicted from the fitted model parameters, and an independent estimate of the residual is obtained. The residual sum of squares obtained from cross-validated residuals requires no correction for degrees of freedom and can be used to evaluate whether one postulated model performs better than another.

Fractional factorial designs can also be modeled by GEMANOVA models. The response of a factorial design with *R* factors on *I* levels can be described as an *R*-way array with dimension *I* in each mode. If the number of levels differs for the different factors, the appropriate dimensions differ as well. The response of a fractional factorial design with *R* factors on *I* levels can likewise be described as an *R*-way array with dimension *I* in each mode. In this case, however, some of the elements in this array will be missing. Therefore a fractional design can be handled exactly as a factorial design as long as the algorithm used to fit the model can handle such missing elements. The algorithm proposed in the Appendix does so.

In some cases, effects are sought which do not simply correspond to a multilinear component with fixed loadings. For example, it may be useful to have a traditional interaction effect of the type a_{ij} where the effect has two indices rather than one index as in the GEMANOVA models discussed thus far. Such effects can also be estimated but have not been implemented in the GEMANOVA algorithm offered here.

The algorithm used to fit the GEMANOVA model is based on a PARAFAC alternating least squares algorithm. The inclusion of lower-order interactions, however, makes the estimation non-trivial compared to standard PARAFAC. In GEMANOVA it is not possible to estimate different effects in a model independently as it is in ordinary analysis of variance. Many times in ordinary analysis of variance the effects are constrained to be independent by construction. These constraints are inconsequential for the fit of the model. For the GEMANOVA model, however, applying such constraints will either change the model or simply reduce the fit. This is related to the fact that the PARAFAC model is unique up to scaling and permutation [24–26]. Details of the algorithm are given in the Appendix.

4. RESULTS FROM TRADITIONAL ANALYSIS OF VARIANCE

A traditional analysis of variance was performed on the data. The factors are storage time (x_i), temperature (x_j), oxygen (x_k), light (x_l) and muscle (x_m) which is qualitative. The response color is held in y_{ijklm}. The data were scaled and centered and all cross-products and interactions added. After removing insignificant variables, the following ANOVA model was obtained through cross-validated partial least squares regression (two components):

$$y_{ijklm} = \mu + ax_i + bx_j + cx_k + dx_i x_j + ex_i x_l + t_{ijklm} \quad i = 1,\ldots,I, \quad j = 1,\ldots,J, \quad k = 1,\ldots,K, \quad l = 1,\ldots,L, \quad m = 1,\ldots M \tag{8}$$

where

- color at storage level *i*, temperature *j*, light level *k*, oxygen level *l*, muscle *m*: y_{ijklm},
- overall offset: μ,
- storage main effect: *a*,
- temperature main effect: *b*,
- light main effect: *c*,
- interaction between storage and temperature: *d*,
- interaction between storage and light: *e*,
- residuals: t_{ijklm}.

This model is fitted well and provides good cross-validated errors as shown in Figure 1. The correlation as well as the root mean squared error of cross-validation (RMSECV) is provided, where the RMSECV is defined as

$$\mathrm{RMSECV} = \sqrt{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{l=1}^{L}\sum_{m=1}^{M}(y_{ijklm} - \hat{y}_{ijklm})^2}{324}} \tag{9}$$

The summation is only performed over those elements in the array for which the experiments have been performed.

It is evident from the effects shown in Figure 2 that an increase in the level of any of the significant factors leads to a decrease in color. This is in agreement with expectations, as

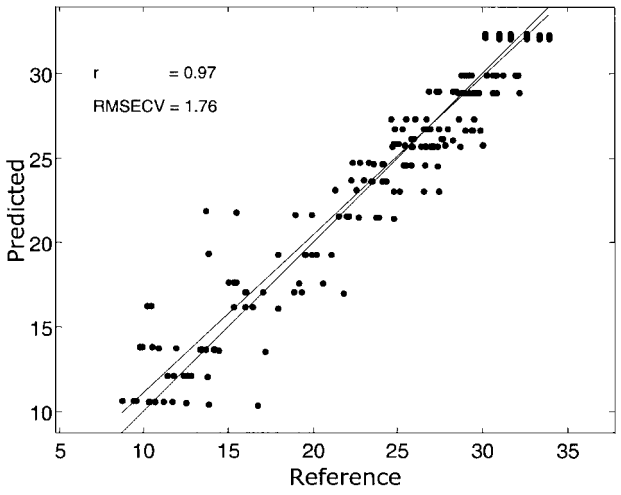

Figure 1. Results of leave-one-out cross-validation using the model in Equation (8) and data averaged over replicates (324 samples).

color is known to deteriorate with increasing time, temperature and light exposure [2,27].

It is also evident that storage time has the largest effect on color in the given experimental domain and that light has a relatively low effect. Oxygen is excluded from the model because it does not have any significant effect. This is somewhat surprising, since general practice is to maintain a high level of oxygen (>70%) because lower levels are assumed to lead to excessive degradation of quality/color. Muscle does not have any significant effect either.

With three factors and two interactions, visualization of the total experimental domain is difficult. Even though all effects have the same sign and response surfaces can be drawn, several of these have to be made for a fixed setting of one factor when plotting the response surfaces for the two remaining factors. In fact, the difficulty of visualizing and communicating the results to a broader, less technical audience initiated this search for a more accessible model for these experimental data.

Before continuing with the GEMANOVA results, the results of some alternative traditional models will be briefly discussed. The GEMANOVA model intrinsically treats all factors as qualitative factors, yielding an effect for each level of a factor. It is reasonable, therefore, to consider the outcome of a traditional ANOVA model where all factors, not only animal, are treated as qualitative. Such a model leads to similar predictions (RMSECV = 1.81) as those described above but using more parameters. The same effects are found except that the interaction between storage and light exposure is not significant, but the model is much more difficult to interpret, especially owing to the interaction between storage and temperature, which requires 15 parameters.

Another obvious model to consider is a model where the logarithmic transform of the response is used instead of the response itself. Such a model is also of similar quality (RMSECV = 1.79) and with similar interpretation as the model discussed in detail above. Thus the logarithmic transform does not provide any gain in interpretability compared to the model of the non-transformed data. That such seemingly different models provide similar quantitative results again stresses that there are many routes to take and that statistical significance of the model and/or its effects is not necessarily the only aspect to consider. The usefulness of the model is also related to how well it can be interpreted and understood by the practitioner using the results.

5. RESULTS FROM GEMANOVA MODELING

5.1. Determining the appropriate model

Even though the original set of responses makes an eight-way array (storage × temperature × oxygen × light × animal × left–right muscle × replicate × color$L/a/b$), the averaged data set used for analysis is a five-way array where the animal and left–right muscle modes have been concatenated and only *a*-color is used. The *a*-color is thus a

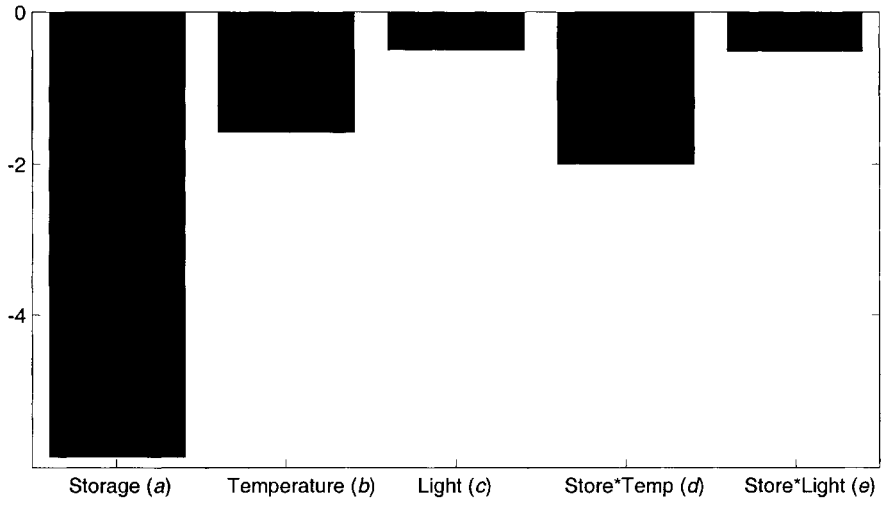

Figure 2. Significant effects in analysis-of-variance model of color. Effects are shown in terms of scaled and centered factors and response.

Table III. Results from leave-one-out cross-validation of different GEMANOVA models using data averaged over the five replicates (*data* $5 \times 3 \times 3 \times 3 \times 6$)

Model	No. of parameters	RMSEC	RMSECV	r^2 CV	Type
$a_i b_j c_k d_l e_m$	20	2.84	2.97	0.89	Multiplicative
$a_i b_j c_k d_l e_m + f_l$	23	2.55	2.73	0.92	Multiplicative + $\text{main}_{\text{oxygen}}$
$a_i b_j c_k d_l e_m + f_k$	23	2.50	2.67	0.92	Multiplicative + $\text{main}_{\text{light}}$
$\boldsymbol{a_i b_j c_k d_l + e_m}$	**20**	**1.60**	**1.75**	**0.97**	**Multiplicative + $\text{main}_{\text{animal}}$**
$a_i b_j c_k d_l e_m + f_j$	23	1.58	1.74	0.97	Multiplicative + $\text{main}_{\text{temp}}$
$a_i b_j c_k d_l e_m + f_i$	25	1.57	1.74	0.97	Multiplicative + $\text{main}_{\text{storage}}$
$a_i b_j c_k d_l e_m + f_m$	26	1.38	1.58	0.97	Multiplicative + $\text{main}_{\text{animal}}$
$a_i b_j c_k d_l e_m + f_i g_j h_k r_l s_m$	40	1.28	1.50	0.98	Multiplicative

function of storage time, temperature, oxygen, light and muscle.

In establishing the GEMANOVA model, it is necessary to determine which effects to include. This is done by first evaluating different alternatives to a one-component five-way PARAFAC model. From this, different alternative effects are added, and in Table III the results from different GEMANOVA models are listed. The average prediction error is given in terms of the RMSECV performed as a leave-one-out cross-validation where each of the 324 elements is left out in turn. The root mean square error of calibration (RMSEC) using fitted values and the correlation are also given in the table.

It is seen that a simple one-component GEMANOVA model yields an unsatisfactory RMSECV of 2.97, while the five last models in the table all have acceptable RMSECV values.

While the two last models have lower errors, the model $a_i b_j c_k d_l + e_m$ is an interesting alternative. The RMSECV error of this model is similar to the traditional ANOVA model and it has few parameters. In this model there is one multiplicative effect of storage, oxygen, temperature and light, but not dependent on muscle, and there is second effect of muscle only. This model is chosen for further investigation because it turns out to provide a convenient and intuitive understanding of the main variation in the data. Two alternative models have lower RMSECV values, but for the two-component model a so-called degeneracy is observed, meaning that the parameters are not well defined [28,29]. In that model the first and second effects of any of the modes (e.g. a_i and f_i in Table III) are almost identical in shape. The other alternative model is similar to the suggested one, but in addition to the independent muscle effect there is a contributing muscle effect in the multiplicative term. Although this refinement leads to a better model in terms of RMSECV, the improvement is not assessed to warrant the added complexity of the model.

Hence it is determined that an adequate and simple GEMANOVA model for these data is given by

$$y_{ijklm} = a_i b_j c_k d_l + e_m + t_{ijklm}$$
$$i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K,$$
$$l = 1, \ldots, L, \quad m = 1, \ldots, M \qquad (10)$$

where the notation is as before. In Figure 3 the predictions obtained from cross-validation are shown. Note that this model has a completely different structure from the ANOVA model. Hence these two models, although of similar predictive quality, have widely different structures. A choice between them is thus mainly based on which one offers the most understandable description of the variation in the data.

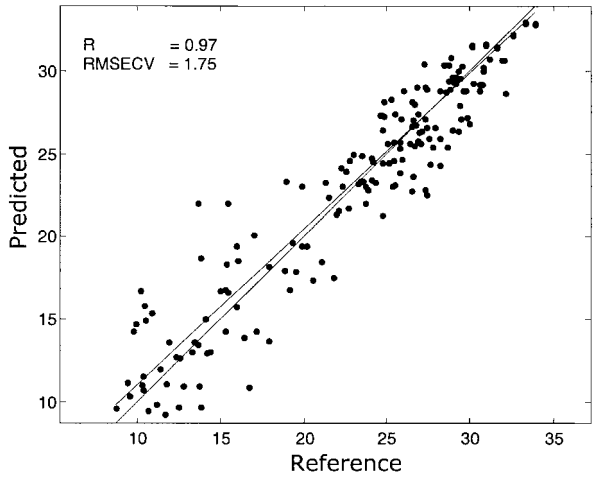

Figure 3. Results of leave-one-out cross-validation using the GEMANOVA model in Equation (10).

5.2. Interpretation of the GEMANOVA model

In order to interpret the model, the parameters are depicted in Figure 4. Here the parameter a_i, $i = 1,\ldots,5$, is the storage effect, b_j, $j = 1,2,3$, is the temperature effect, c_k, $k = 1,2,3$, is the light effect and d_l, $l = 1,2,3$, is the oxygen effect. The separate muscle effects e_m are shown in Figure 5. The interpretation of the model is simple. The multiplicative term in Equation (10), $a_i b_j c_k d_l$, is zero, thus absent, at time zero owing to a_1 being zero. Thus the initial level is given specifically by the separate muscle effect. Every muscle has a starting color level of approximately 32 as shown in Figure 5 and, as can be seen, the estimated muscle effect is almost identical to the initial color measured. For all other settings of the factors the estimated response is simply the starting level plus the product of the four effects read from ordinates in Figure 4. Some important points easily deduced from the plot are as follows.

- The effect of storage is zero at time zero. This is not a constraint imposed in the model, but an empirical finding

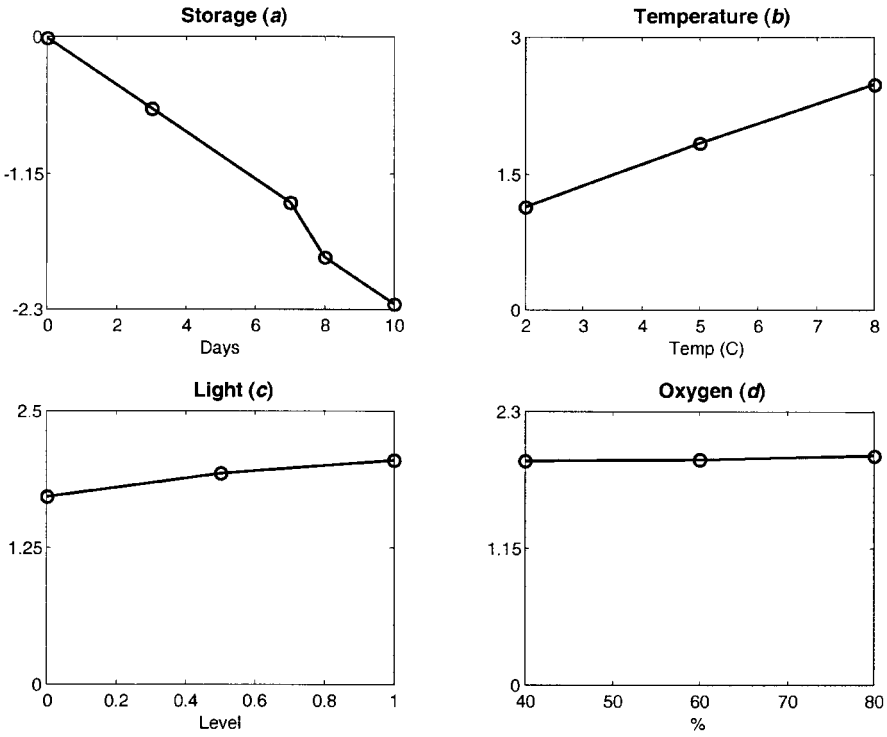

Figure 4. Parameters from multiplicative effect in GEMANOVA model. The estimated response at specified levels of the four factors equals the product of the corresponding effects plus the muscle term, which varies between 31 and 33.

from the estimated parameters. It confirms that the model is sensible and that it provides a physically intuitive model of the data.

- Any change in color is negative, hence a decrease from the starting point, because the product of the four parameters $a_i b_j c_k d_l$ consists of one negative number (storage) and three positive numbers.
- The changes are relative. For example, going from temperature 2 to 8 °C, it is seen that the loading b_j increases from 1.2 to 2.4 Therefore the temperature effect will be that the overall decrease is twice as high at 8 °C as it is at 2 °C, regardless of all other factors.
- The relative effect of temperature is linear.
- The effect of light is small, but there seems to be a small increase (hence decrease in color) with amount of light. The significance of this effect is treated below.
- The effect of oxygen is insignificant, thus supporting the surprising conclusion from the traditional analysis of variance that even a value as low as 40% oxygen does not significantly alter the color development during storage.

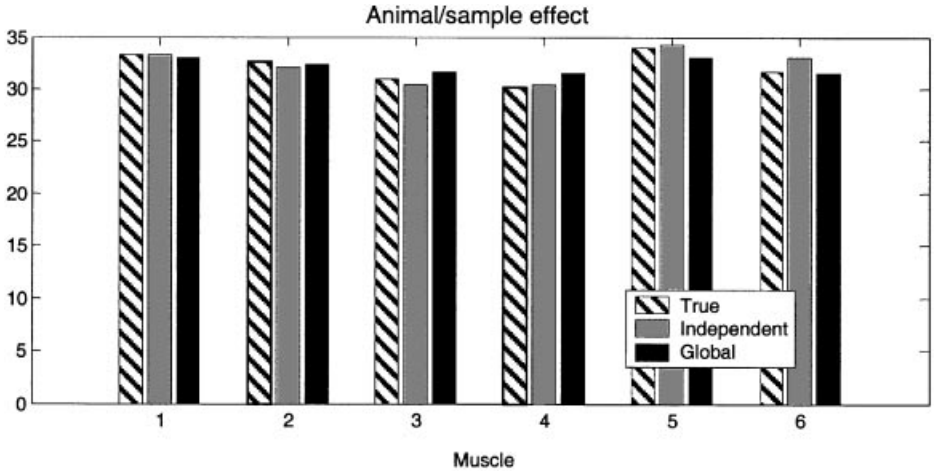

Figure 5. Muscle effects for each individual muscle. The reference color at time zero is given first in each set of bars, then the effects estimated from the six individual models, Equation (11), and finally the effects found in the model of all data, Equation (10). As can be seen, all three coincide nicely.

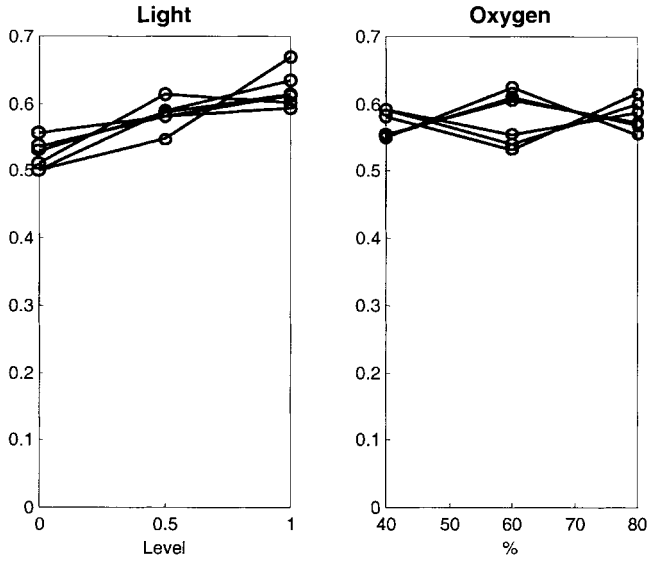

Figure 6. Parameters for light and oxygen effects from six individual GEMANOVA models of data for each muscle.

It is important to verify whether the increasing effect of light is real and whether the absence of the oxygen effect is real. This can be verified e.g. by estimating the model from independent data sets. The data were split into six subgroups by taking as one data set all elements x_{ijklm} for $m = 1$, etc. That is, the data from each muscle are modeled separately. These six data sets are independent and from Equation (10) the model of each with m fixed is seen to follow

$$y_{ijkl} = a_i b_j c_k d_l + e + t_{ijkl}$$
$$i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K, \quad l = 1, \ldots, L \tag{11}$$

where the subscript on e is now removed because in each model there is only one muscle.

The difference now being that the model is estimated for fixed m, a four-way model is fitted with one constant term e. This is then done six times, one for each subset of data. The resulting six estimates of the effects of oxygen and light are shown in Figure 6. From the similarity of these independent models it is observed that there is a slight but significant (linear) increase from the light effect, while there is no oxygen effect.

Looking at the muscle effects (Figure 5), it is seen that the estimated parameters from the six individual models as well as from the model of all data are estimates of the starting color for the specific muscle. Thus the GEMANOVA model is able to capture the small differences in starting color between the different muscles. The reason why the GEMANOVA model is able to uniquely identify this starting point is that the model is in fact equal to a constrained two-component PARAFAC model, which is unique under mild conditions [26,30]. This is explained in detail in Reference [31] and is a remarkable aspect of the GEMANOVA model which is in stark contrast to traditional ANOVA models. In the traditional ANOVA model the parameters are only unique owing to 'arbitrary' mathematically imposed constraints and it is not possible, for example, to *estimate* the offset uniquely from such an ANOVA model. For example, in a traditional ANOVA model a global offset would be identical to the average color of all experiments (hence less than the starting color). On hindsight, one can argue that the initial color is more suitable as a general offset, but this hindsight is not required in the GEMANOVA model. The offset comes out uniquely from the estimated parameters.

5.3. Refined model

The uniqueness properties of GEMANOVA coupled with the simple interpretation of Figure 5 as well as the constant level of the oxygen effect suggest the following GEMANOVA model for the data. The starting level can simply be subtracted from the data, and the effect of oxygen can be removed by fixing it to one in the model estimation. The absolute level of oxygen in Figure 4 of approximately two is of no concern when removing the oxygen effect (effectively by setting it to one). In a multiplicative model this difference in level is automatically retained in other parts of the multiplicative term when oxygen is left out. This leads to a three-way model of the five-way data:

$$y_{ijklm} - \mu_m = a_i b_j c_k + t_{ijklm}$$
$$i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K,$$
$$l = 1, \ldots, L, \quad m = 1, \ldots, M \tag{12}$$

where μ_m is either the actual color level at time zero or the estimated muscle effect. If μ_m is estimated from the data, this leads to a model close to the above model with an RMSECV of 1.73, and if μ_m is set equal to the starting color, the RMSECV is slightly higher. This again supports the observation that oxygen has an insignificant effect.

The refined model shows that, apart from starting level, which is sample-specific, the functional shape of the decrease as a function of the storage time, temperature and light is a simple one-component PARAFAC model $a_i b_j c_k$. This multiplicative part holds for all samples, whereas the starting level is specific. It would not have been possible to deduce these observations directly from the analysis of variance with traditional tools. It is the specific uniqueness of the GEMANOVA model that enables the finding that the muscle effect is simply the estimated color at time zero.

6. CONCLUSION

With the use of the GEMANOVA model, interesting new and simple results have emerged in the analysis of storage of beef. Through the appropriateness and uniqueness of the GEMANOVA model, an intuitive and sensible model is obtained. The refined model of Equation (12) for color changes in fresh beef is easy to interpret. E.g.

- The model emphasizes the importance of keeping a low storage temperature.
- Time of exposure to light has only a minor effect.
- There is no effect of oxygen concentration in the applied interval from 40% to 80%.

Although the last of these is somewhat surprising, as the oxygen concentration normally used is 70%–80%, it is in

agreement with the results of Jakobsen and Bertelsen [2], who found a stable interval for color stability between 55% and 80% oxygen. Reducing the oxygen level opens up the possibility of using more carbon dioxide or nitrogen and thereby decreasing other deteriorative reactions in meat, e.g. lipid oxidation.

Acknowledgements

This work was sponsored by the FØTEK program through LMC (Centre for Advanced Food Studies). The authors are grateful to Pernille N. Jensen for excellent technical assistance as well as to AGA A/S, Danisco Flexible ON and Danish Crown, Hvidovre for providing materials and packaging facilities for the experiments. R. Bro gratefully acknowledges support provided by LMC (Center for Advanced Food Studies) and AQM (FR9801214; Advanced Quality Monitoring in the Food Production Chain) supported by the ministries of research.

APPENDIX. IMPROVED GEMANOVA ALGORITHM

There are many ways to implement an algorithm for fitting the GEMANOVA model. Owing to the complexity of the problems in this paper, where there are 60% missing data, special care is needed in order to avoid local minima solutions. Hence, rather than the simple algorithms used earlier, a new algorithm was developed which is based on gradually approaching the true solution by means of simpler problems that are less susceptible to numerical problems. The algorithm will be explained in general terms first and then the important steps are explained in detail.

The basic rationale in the algorithm is that the GEMANOVA model can be viewed as a constrained PARAFAC model. For example, for a three-way data set, one multiplicative GEMANOVA effect equals one PARAFAC component, $a_i b_j c_k$. If a GEMANOVA effect is sought of the type $a_i c_k$, then this corresponds to an ordinary PARAFAC component $a_i b_j c_k$ where the parameters b_j ($j = 1,\ldots,J$) are all constrained to be one. Thus GEMANOVA can be fitted by a PARAFAC model with certain parameters constrained to be one.

Algorithm GEMANOVA for N-way array of responses

1. Choose the number of effects, R, and the modes whose parameters are to be fixed to one for each effect. These constraints are defined by a binary matrix $\mathbf{G}$ ($N \times R$). An element g_{nr} is one when the rth effect is constant in mode n.
2. Fit an ordinary R-component PARAFAC model four times starting from different random starting points. Pick the best-fitting of these four models and let its parameters be the starting point for the next step.
3. Fit an R-component PARAFAC model where the fixed or constant effects are gradually forced towards one.
4. Fit an R-component PARAFAC model where the constraints of parameters being one are imposed exactly.

As can be seen, the algorithm is relatively complicated, involving several steps of fitting models that are gradually approaching the final model. This has, however, proven to be a fruitful approach for difficult problems.

In step 3 the gradual enforcement of the constraints is implemented in the following way. A weight w defines the degree to which the constraints are imposed. The higher the value, the stronger the constraint is imposed. Six models are fitted ranging from not imposing the constraints at all to almost imposing the constraints exactly. The value of w ranges equidistantly in six steps from 0 to SSY/300, where SSY is the sum of squares of elements in the response array. Within the algorithm the following special precautions are taken compared to ordinary alternating least squares implementations of PARAFAC [14,24]. Before updating the parameters in a given alternating least squares cycle, all component vectors for a given effect (except fixed ones) are scaled such that they have equal norm. Otherwise, the gradual imposing of constraints lead to ill-defined problems.

The parameters in a given mode are estimated iteratively and 'effectwise', corresponding to estimating one of the columns in a PARAFAC component matrix at a time. This is done by subtracting the sum of all other current estimated effects from the response array and using the residual. Thus, for updating the gth effect, the residual e_{ijk} (adding more subscripts for higher orders) is defined as

$$e_{ijk} = y_{ijkl} - \sum_{r \neq g}^{R} a_{ir} b_{jr} c_{kr} \qquad i = 1,\ldots,I, \quad j = 1,\ldots,J, \quad k = 1,\ldots,K \tag{13}$$

The conditional problem is then to find the solution to

$$e_{ijk} = a_{ig} b_{jg} c_{kg} + t_{ijk} \qquad i = 1,\ldots,I, \quad j = 1,\ldots,J, \quad k = 1,\ldots,K \tag{14}$$

with t_{ijk} being variation not explained by the interim model. Estimating the parameters one mode at a time, conditional on the remaining estimates, is a simple least squares problem where each parameter is independent of all others [11,31]. For example, estimating the parameters a_i ($i = 1,\ldots,I$) can be based on the least squares problem

$$\mathbf{E} = \mathbf{q}\mathbf{a}^{\mathrm{T}} + \mathbf{T} \tag{15}$$

where $\mathbf{E}$ is an $JK \times I$ matrix holding the elements e_{ijk} properly arranged, $\mathbf{T}$ is the corresponding matrix of unexplained variation, $\mathbf{q}$ is a JK-vector holding the product $b_{jg}c_{kg}$ as its elements and $\mathbf{a}$ is an I-vector holding the sought parameters a_i. As can be seen, the optimal estimate for each a_i is independent of the remaining elements of $\mathbf{a}$ (see Lemma 1 in Reference [31]).

Updating the parameters one mode at a time provides an improved update for the gth effect. If the parameters are not fixed, an ordinary least squares regression is performed based directly on Equation (15), and if there are missing elements in the corresponding column of $\mathbf{E}$, these can simply be skipped in the equations because of the independence between the elements in $\mathbf{a}$. If, on the other hand, the parameters in e.g. $\mathbf{a}$ are to be fixed to one, a penalty depending on w is added for deviating from this value. An R-vector of ws is appended to $\mathbf{E}$ and an element w is appended to $\mathbf{q}$. This part of the underdetermined system hence requires $wa_i = w$, hence that a_i equals one. Owing to the

variance in **E** and **q**, this will not force **a** to be exactly one, but depending on the size of w the constraint is imposed to a certain degree.

In step 4 the final solution of step 3 is used for an algorithm imposing the constraints exactly. This algorithm is simple. It is an unconstrained PARAFAC-ALS algorithm implemented as indicated above. However, parameters fixed to one are initially set to one and simply not updated during the iterations. This will lead to a convergent solution of the problem.

REFERENCES

1. Taylor AA. Gases in fresh meat packaging. *Meat World* 1972; **5**: 3–6.
2. Jakobsen M and Bertelsen G. Colour stability and lipid oxidation of fresh beef. Development of a response surface model for predicting the effects of temperature, storage time, and modified atmosphere composition. *Meat Sci.* 2000; **54**: 49–57.
3. Zhao Y, Wells JH and McMillin KW. Applications of dynamic modified atmosphere packaging systems for fresh red meats: review. *J. Muscle Foods* 1994; **5**: 199–328.
4. Faustman C and Cassens RG. The biochemical basis for discolouration in fresh meat: a review. *J. Muscle Foods* 1990; **1**: 217–243.
5. Renerre M. Factors involved in the discoloration of beef meat. *Int. J. Food Sci. Technol.* 1990; **25**: 613–630.
6. Hirsh RF. Analysis of variance in analytical chemistry. *Anal. Chem.* 1977; **49**: 691A–700A.
7. Hoaglin DC and Welsch RE. The hat matrix in regression and ANOVA. *Am. Statist.* 1978; **32**: 17–22.
8. Latorre G. Analysis of variance and linear models. In *Mathematics and Statistics in Chemistry*, Kowalski BR (ed.). Reidel: Dordrecht, 1984; 377–391.
9. Martens H, Izguierdo L, Thomassen M and Martens M. Partial least squares regression on design variables as an alternative to analysis of variance. *Anal. Chim. Acta* 1986; **191**: 133–148.
10. Weisberg S. *Applied Linear Regression*. Wiley: New York, 1985.
11. Bro R. *Multi-way analysis in the food industry. Models, algorithms, and applications. PhD Thesis*, University of Amsterdam, 1998 (http://www.mli.kvl.dk/staff/foodtech/brothesis.pdf).
12. Bro R and Heimdal H. Enzymatic browning of vegetables. Calibration and analysis of variance by multiway methods. *Chemometrics Intell. Lab. Syst.* 1996; **34**: 85–102.
13. Andersson CA and Bro R. The N-way toolbox for MATLAB. *Chemometrics Intell. Lab. Syst.* 2000; **52**: 1–4.
14. Bro R. PARAFAC. Tutorial and applications. *Chemometrics Intell. Lab. Syst.* 1997; **38**: 149–171.
15. Heimdal H, Bro R, Larsen LM and Poll L. Prediction of polyphenol oxidase activity in model solutions containing various combinations of chlorogenic acid, (−)-epicatechin, O_2, CO_2, temperature and pH by multiway analysis. *J. Agric. Food Chem.* 1997; **45**: 2399–2406.
16. Van Eeuwijk FA. *Between and beyond additivity and non-additivity; the statistical modelling of genotype by environment interaction in plant breeding. PhD Thesis*, University of Wageningen, 1996.
17. Van Eeuwijk FA and Kroonenberg PM. Multiplicative models for interaction in three-way ANOVA, with applications to plant breeding. *Biometrics* 1998; **54**: 1315–1333.
18. Fisher RA and MacKenzie WA. Studies in crop variation. II. The manurial response of different potato variaties. *J. Agric. Sci.* 1923; **13**: 311–320.
19. Gollob HF. A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika* 1968; **33**: 73–115.
20. Hegemann V and Johnson DE. On analyzing two-way ANOVA data with interaction. *Technometrics* 1976; **18**: 273–281.
21. Mandel J. The partitioning of interaction in analysis of variance. *J. Res. NBS* 1969; **73B**: 309–328.
22. Mandel J. A new analysis of variance model for non-additive data. *Technometrics* 1971; **13**: 1–18.
23. Kettenring JR. A case study in data analysis. *Proc. Symp. Appl. Math.* 1983; **28**: 105–139.
24. Harshman RA. Foundations of the PARAFAC procedure: models and conditions for an 'explanatory' multimodal factor analysis. *UCLA Working Papers Phonet.* 1970; **16**: 1–84.
25. Kruskal JB. Three-way arrays: rank and uniqueness of trilinear decompositions with applications to arithmetic complexity and statistics. *Linear Algebra Appl.* 1977; **18**: 95–138.
26. Sidiropoulos ND and Bro R. On the uniqueness of multilinear decomposition of N-way arrays. *J. Chemometrics* 2000; **14**: 229–239.
27. Andersen HJ, Bertelsen G and Skibsted LH. Colour and colour stability of hot processed frozen minced beef. Results from chemical model experiments tested under storage conditions. *Meat Sci.* 1990; **28**: 87–97.
28. Kruskal JB, Harshman RA and Lundy ME. How 3-MFA data can cause degenerate PARAFAC solutions, among other relationships. In *Multiway Data Analysis*, Coppi R, Bolasco S (eds). Elsevier: Amsterdam, 1989; 115–122.
29. Mitchell BC and Burdick DS. Slowly converging PARAFAC sequences: swamps and two-factor degeneracies. *J. Chemometrics* 1994; **8**: 155–168.
30. Kruskal JB. Rank, decomposition, and uniqueness for 3-way and N-way arrays. In *Multiway Data Analysis*, Coppi R, Bolasco S (eds). Elsevier: Amsterdam, 1989; 8–18.
31. Bro R and Sidiropoulos ND. Least squares algorithms under unimodality and non-negativity constraints. *J. Chemometrics* 1998; **12**: 223–247.