# An approach to the decomposition of interaction in a factorial experiment with five factors

**Fabiani da Rocha[1], Naine Martins do Vale[1], Leiri Daiana Barili[1], Jefferson Luís Meirelles Coimbra[1\*], Altamir Frederico Guidolin[1] and Juliano Garcia Bertoldo[2]**

[1]Instituto de Melhoramento e Genética Molecular, Centro de Ciências Agroveterinárias, Universidade do Estado de Santa Catarina, Av. Luiz de Camões 2090, 88520-000, Lages, Santa Catarina, Brazil. [2]Programa de Pós-graduação em Recursos Genéticos Vegetais, Centro de Ciências Agrárias, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, Brazil. *Author for correspondence. E-mail: coimbrajefferson@cav.udesc.br

**ABSTRACT.** Factorial experiments are widely employed in agricultural research. In these experiments, inferences related to the interaction between factors are fundamental. However, many researchers are still unable to analyze this type of experiment, and others do not consider the effect of interaction. This study aims to exemplify a scheme for unfolding the degrees of freedom and to demonstrate the relevance of this rearrangement when the interaction is significant. For these purposes, the response variable time of cooking bean grains was measured in minutes using a completely casualized experimental design, with two replications, arranged in a 2 x 2 x 3 x 2 x 2 factorial scheme (fivefold classification). The following factors were evaluated: bean genotypes (2 levels), salt type (2 levels), salt dose (3 levels), hydration time (2 levels) and storage time (2 levels). The results highlight the importance of unfolding the degrees of freedom of the interaction every time it is significant because the possibility of evaluating the interaction between factors leads to conclusions related to the dependence between the factors. These results are more coherent with biological systems, and the example shown provides a solid basis for minimizing errors in factorial experiments.

**Keywords:** dependence between factors, main effect, simple effect.

## Uma abordagem sobre a decomposição da interação em um experimento fatorial com cinco fatores

**RESUMO.** Experimentos fatoriais são amplamente empregados na pesquisa agrícola. Nestas situações, inferências quanto à interação entre os fatores são fundamentais. Porém, muitos pesquisadores apresentam ainda inabilidade quanto à análise deste tipo de experimento e outros simplesmente desconsideram o efeito da interação. Este trabalho teve como objetivos exemplificar um esquema de desdobramento de graus de liberdade e de demonstrar a importância deste rearranjo quando a interação é significativa. Para isto a variável resposta tempo de cocção de grãos de feijão foi mensurada em minutos, utilizando um delineamento experimental inteiramente casualizado, com duas repetições, arranjado num esquema fatorial 2 x 2 x 3 x 2 x 2 (classificação quíntupla). Os fatores avaliados foram: genótipos de feijão (2 níveis), tipo de sal (2 níveis), dose de sal (3 níveis), tempo de hidratação (2 níveis) e tempo de armazenamento (2 níveis). Os resultados obtidos reforçam a importância do desdobramento dos graus de liberdade da interação sempre que esta se mostrar significativa, pois a possibilidade de avaliação da interação entre os fatores permite que sejam feitas conclusões quanto a dependência entre os fatores, e estes resultados são sem dúvida mais coerentes com o sistema biológico; e o exemplo exposto fornece base segura para que os erros envolvendo experimentos fatoriais sejam ao menos minimizados.

**Palavras-chave:** dependência entre os fatores, efeito principal, efeito simples.

## Introduction

In agricultural experiments, when the simultaneous effect of two or more factors on a given characteristic of interest is the focus of a study, factorial experiments are the most common and useful types of experimentation (YASSIN et al., 2002). Here, each subdivision of a factor is considered a level of the factor. The treatments are formed by combinations among the different factor levels (CHEW, 1976a; SILVA, 1999), and the sum of the treatment squares should be partitioned into components corresponding to the main effects of the factors and their interactions.

An assessment of the interaction among factors is fundamental, and when it is significant, it is important to perform the decomposition of the degrees of freedom (NOGUEIRA; CORRENTE, 2000). If the effect of the significant interaction is

ignored, the statistical inferences may later be impaired because when there is interaction, the factors are dependent, and thus, the combinations of treatments may affect the data in various ways. However, many researchers simply disregard the interaction effect, even when it is significant.

Articles reporting mistakes and successes in applying statistical tests show that most errors are related to experiments with more than one factor. The interaction among factors is simply disregarded by some researchers, and tests for the comparison of averages are applied separately (BEZERRA NETO et al., 2002; BERTOLDO et al., 2007, 2008a and b; CARDELLINO; SIEWERDT, 1992; PETERSEN, 1976; SANTOS et al., 1998). This problem has persisted for decades, and it reveals that researchers are unable to analyze factorial experiments. Despite this problem, only one survey on errors has been reported. However, a question arises: what is missing that would allow this process to be carried out concisely? The answer is the following: examples of applications that help researchers perform the correct partitioning of the degrees of freedom every time the interaction is significant.

Therefore, the present work has the following aims: i) to exemplify a scheme for unfolding the degrees of freedom while considering a third-order interaction and ii) to demonstrate the importance of rearranging the degrees of freedom of the interaction comparatively with the isolate analysis of each factor to provide further information to researchers.

## Material and methods

The treatment factors evaluated with their respective levels were the following: i) bean genotypes: Pérola and IPR Uirapuru; ii) salt type: KCl and NaCl; iii) salt dose: 0, 10 and 25 g; iv) hydration time: 8 and 16 hours; and v) storage time: 45 and 90 days after harvest. The measured response variable was the cooking time of bean grains, in min., according to the methodology proposed by Proctor and Watts (1987). The experiment was conducted in a completely randomized experimental design with two replications arranged in a 2 x 2 x 3 x 2 x 2 factorial scheme (fivefold classification).

The statistical procedures for the data analysis were carried out with the SAS software system by means of the PROC GLM procedure. The analyses involved the following:

a) The analysis generally used (though incorrect for a factorial structure), which considers only the main effect of factors. The analyses were carried out considering the performance of the main effect of

the factors, disregarding the interaction effect. A polynomial adjustment was used for the quantitative factor (salt dose) followed by the regression analysis. The F test was always sufficient for the qualitative factors because there were only two levels in each factor. The statistical model was formed by the following expression:

$$\gamma_{ijlmnk} = \mu + a_i + b_j + c_l + d_m + f_n + e_{iljmnk}$$

where:

$\gamma_{ijlmnk}$ refers to the value of the response variable observed (cooking time, in min.); $\mu$ represents a constant referring to all the data; $a_i$, $b_j$, $c_l$, $d_m$ and $f_n$ refer to the differential effects of the treatments $i, j, l, m$ and $n$ (genotype, salt type, salt dose, hydration time and storage time, respectively); and $e_{iljmnk}$ represents the experimental error.

b) The correct analysis for an experiment with a factorial structure. Here, the statistical analysis was carried out considering the main effect and that of the interaction among the factors. In this situation, the simple effect of the factors was dealt with when the interaction was significant (5% probability by the F test), followed by the regression adjustment. Therefore, the equation of the statistical model for the experiment with five fixed-effect factors is:

$$\gamma_{ijlmnk} = \mu + a_i \lfloor b_j \lfloor c_l \lfloor d_m \lfloor f_n + e_{iljmnk}$$

where:

$\gamma_{ijlmnk}$ represents the response variable (cooking time in min.); $a_i$, $b_j$, $c_l$, $d_m$ and $f_n$ refer to the differential effects of the treatments $i, j, l, m$ and $n$ (genotype, salt type, salt dose, hydration time and storage time, respectively); the symbol "$\lfloor$" indicates that the SAS software system will consider the main effect of factors in the analysis of variance of the data as well as all the interactions among them; and $e_{iljmnk}$ refers to the experimental error.

## Results and discussion

With the results achieved by the individual analysis of variance, significant differences can be observed for the genotype, hydration and salt dose factors in relation to the cooking time variable (Table 1). The researcher could only verify that the genotype of the black group IPR Uirapuru presented a significantly shorter cooking time (25 min.) than the genotype of the carioca group (32 min.), and the hydration time of 16 hours produced a shorter cooking time (27 min.) than a hydration of 8 hours (29 min.). In this case, no test for an average comparison was necessary because,

according to Silva (1999), when a factor presents only two levels (i.e., a degree of freedom), the F test is sufficient and is identical to the $t^2$.

**Table 1.** Summary of the analysis of variance (ANOVA) of the main effects of the factors genotype, salt type, salt dose (g), hydration time (min.) and storage time (days) without verifying the effect of the interaction among the factors for cooking time (min.) with the respective sources of variation (SV), degrees of freedom (DF), sum of squares (SS) and mean square (MS).

| SV | DF | SS | MS | Pr > F |
|---|---|---|---|---|
| Genotype | 1 | 1,410.66 | 1,410.66 | 0.0001 |
| Salt type | 1 | 48.16 | 48.16 | 0.1376 |
| Dose | 2 | 644.02 | 322.01 | 0.0001 |
| Hydration | 1 | 170.66 | 170.66 | 0.0059 |
| Storage | 1 | 80.66 | 80.66 | 0.0557 |
| Error | 89 | 1,909.77 | 21.45 | |
| Total | 95 | 4263.95 | | |

The factor salt dose also presented a significant effect (5% probability by the F test) on cooking time. Because this factor is a quantitative factor, the most appropriate approach is to adjust the regression because the inferences of higher interest refer to the form of the response curve. Thus, through the equation $\hat{Y}$(min.) = $31.5 - 0.97x + 0.03x^2$ (significant parameters at 5% by the t test), where Y represents the time of cooking and the determination coefficient is estimated at 0.37, it can be observed that the cooking time presents a quadratic behavior along the salt dose and that 16.16 g of salt can provide the shortest cooking time.

In the analyses shown, the causes of variation due to the effect of the interaction among factors were not analyzed. Therefore, no conclusions related to the dependence among the effects of the factors evaluated can be discussed. Thus, there can be no speculation regarding the behavior of genotypes in relation to the salt dose. These results can be inconsistent because the possibility of the existence of interaction was disregarded. Each time researchers make this mistake, they may lose one of the main sources of variation (i.e., the interaction among the factors). Thus, they are unable to make an assessment of the possible occurrence of differences in the responses of the levels of a factor in each of the levels of the other factor(s) (BERTOLDO et al., 2007).

Researchers are often interested in testing more than one factor simultaneously in an experiment, where the comparison among the levels of a factor with the level of another factor depends on the interaction among the factors, which means that, if the interaction is present, there is an indication that the differences among the levels of a factor depend on the level of the other factor (CHEW, 1976b). However, due to the inability of researchers to interpret the results involving factorial experiments, erroneous inferences about the treatments investigated have been observed (BEZERRA NETO et al., 2002). A comparison of marginal averages without considering possible interactions among the main effects is an example of such an error, which, undoubtedly, is an inappropriate practice (CARDELLINO; SIEWERDT, 1992).

When reviewing scientific publications of the journal Qualis A, Bertoldo et al. (2008a) noted that 72% of the published works in the factorial category were incorrect when tests of average comparison were conducted. Problems were observed when the effect of the interaction was ignored or when the effect was neglected when it was significant. In these situations, tests of average comparison, such as the Tukey test, were applied separately for each factor. The academic training of researchers, based only on the use of models of situations, without questioning the adequacy of the statistical methods employed or the results achieved, may explain this type of error, which has been observed in dozens of scientific works. Examples in the literature of the study of three or more factors are also lacking. Moreover, in most examples, the factors present only two or three levels. Steel et al. (1997) presented one of the few works that report the detailed unfolding of each stage of an interaction involving three factors (second order).

However, for experiments arranged in a factorial scheme, the derivation of inferences referring to the effects of the factors depends directly on the presence or absence of interaction (SILVA, 1999). If no interaction is detected by means of the analysis of variance, complementary procedures of the main effects of factors are carried out, and the effect of the interaction is disregarded (PERECIN; CARGNELUTTI FILHO, 2008). However, in the presence of significant interaction, an evaluation of the results requires a comparison of the levels of a factor inside the fixed levels of another factor (NOGUEIRA; CORRENTE, 2000). In other words, the inferences about one of the factors depend directly on the level of the other factor. The degrees of freedom of a treatment will be partitioned in components attributed to the main effects of the factors and to the effects of the interaction among factors.

Figure 1 presents the commands used for the statistical analysis in the SAS software system (PROC GLM). Here, the unfolding of the degrees of freedom of a third order interaction were considered along with the polynomial adjustment (the coefficients of the contrasts are different from

one because they were estimated according to the amplitude among the salt doses, with the use of the commands also presented in this figure) and the estimates of the parameters of the equation. As previously highlighted, each stage of the analysis depends on the result of the previous analysis.

Therefore, the result of the ANOVA for the cooking time (achieved by the first program, "i) interaction ANOVA" (Figure 1)) considering the main effect and the interactions among factors revealed the following interactions (Table 2): *i)* genotype *vs.* salt type (3); *ii)* genotype *vs.* dose (5); *iii)* genotype *vs.* storage (17); *iv)* salt type *vs.* dose *vs.* storage (22); and *v)* salt type *vs.* dose *vs.* hydration *vs.* storage (30). All factors were involved in the interactions. Therefore, it is not advisable make any considerations for the main effects of factors. The correct procedure is to make a comparison among the levels of a factor separately for each combination of the levels of the other factors.

```
Data;
Input Gen$ Tsalt$ Ds Hyd St Rep Coc;
Cards;
Perola      NaCl    0      8      1      1      36
...
IPRUirapuru      Kcl    25    16      2      2      19
;
/* i) ANOVA interaction*/
proc glm;
class Gen Tsalt Ds Hyd St Rep;
model Coc = Gen|Tsalt|Ds|Hyd|St;
lsmeans Gen|Tsalt|Ds|Hyd|St;  run;
/* ii) ANOVA simple effect  - Tsalt*Ds*Hyd*St*/
proc glm;
class Gen Tsalt Ds Hyd St Rep;
model Coc = Gen|Tsalt|Ds|Hyd|St;
lsmeans Tsalt*Ds*Hyd*St / slice=Tsalt*Hyd*St;  run;
/* iii) CONTRAST*/
proc glm;
class Gen Tsalt Ds Hyd St Rep Coc;
model Coc = Ds (Tsalt*Hyd*St);
lsmeans Ds(Tsalt*Hyd*St);
contrast 'c1_L_KCl*Hyd8*St1' Ds(Tsalt*Hyd*St) -0.655  -0.094 0.749;
contrast 'c2_Q_KCl*Hyd8*St1' Ds(Tsalt*Hyd*St)  0.487 -0.811 0.324;
contrast 'c3_L_KCl*Hyd8*St2' Ds(Tsalt*Hyd*St) 0 0 0 -0.655  -0.094 0.749;
contrast 'c4_Q_KCl*Hyd8*St2' Ds(Tsalt*Hyd*St) 0 0 0 0.487 -0.811 0.324;
contrast 'c5_L_KCl*Hyd16*St1' Ds(Tsalt*Hyd*St) 0 0 0 0 0 0 -0.655  -0.094 0.749;
contrast 'c6_Q_KCl*Hyd16*St1' Ds(Tsalt*Hyd*St) 0 0 0 0 0 0 0.487 -0.811 0.324;
contrast 'c7_L_KCl*Hyd16*St2' Ds(Tsalt*Hyd*St) 0 0 0 0 0 0 0 0 0 -0.655  -0.094 0.749;
contrast 'c8_Q_KCl*Hyd16*St2' Ds(Tsalt*Hyd*St) 0 0 0 0 0 0 0 0 0 0 0.487 -0.811 0.324;
contrast 'c9_L_NaCl*Hyd8*St1' Ds(Tsalt*Hyd*St) 0 0 0 0 0 0 0 0 0 0 0 0 -0.655  -0.094 0.749;
contrast 'c10_Q_NaCl*Hyd8*St1' Ds(Tsalt*Hyd*St) 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.487 -0.811 0.324;
contrast 'c11_L_NaCl*Hyd8*St2' Ds(Tsalt*Hyd*St) 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.655  -0.094 0.749;
contrast 'c12_Q_NaCl*Hyd8*St2' Ds(Tsalt*Hyd*St)0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.487 -0.811 0.324;
contrast 'c13_L_NaCl*Hyd16*St1' Ds(Tsalt*Hyd*St) 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.655  -0.094 0.749;
contrast 'c14_Q_NaCl*Hyd16*St1' Ds(Tsalt*Hyd*St) 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.487 -0.811 0.324;
contrast 'c15_L_NaCl*Hyd16*St2' Ds(Tsalt*Hyd*St)0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.655  -0.094 0.749;
contrast 'c16_Q_NaCl*Hyd16*St2' Ds(Tsalt*Hyd*St)0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.487 -0.811 0.324;
run;
/* iv) ADJUSTING THE REGRESSION*/
proc sort; by Tsalt Hyd St;
proc glm; by Tsalt Hyd St;
model Coc= Ds Ds*Ds; run;
proc sort; by Tsalt Hyd St;
proc glm; by Tsalt Hyd St;
model Coc= Ds; run;
```

```
/*Estimate the coefficients of the contrasts*/
proc iml;
t={0 10 25};
C=orpol(t);
print C;
quit;
```

**Figure 1.** Statistical programming of the GLM procedure of the SAS software system for the analysis of the interaction among factors, the simple effect, the assessment of the degree of the polynomials and the adjustment of the regression equation.

**Table 2.** Summary of the analysis of variance considering the main effects of the genotype factors, salt type, doses of salt (g), hydration time (min.) and storage time (days), and the effect of the interaction among the factors on the cooking time (min.), with the respective sources of variation (SV), degree of freedom (DF), sum of squares (SS) and the mean square (MS).

| SV | DF | SS | MS | Pr > F |
|---|---|---|---|---|
| (1) Genotype (Gen) | 1 | 1410.67 | 1410.67 | 0.0001 |
| (2) Salt type (Tsalt) | 1 | 48.17 | 48.17 | 0.0523 |
| (3) Gen*Tsalt | 1 | 155.04 | 155.04 | 0.0008 |
| (4) Dose (Ds) | 2 | 644.02 | 322.01 | 0.0001 |
| (5) Gen*Ds | 2 | 285.65 | 142.82 | 0.0001 |
| (6) Tsalt*Ds | 2 | 74.52 | 37.26 | 0.0560 |
| (7) Gen*Tsalt*Ds | 2 | 21.90 | 10.95 | 0.4134 |
| (8) Hydration (Hyd) | 1 | 170.67 | 170.67 | 0.0005 |
| (9) Gen*Hyd | 1 | 22.04 | 22.04 | 0.1846 |
| (10) Tsalt*Hyd | 1 | 30,38 | 30.38 | 0.1207 |
| (11) Gen*Tsalt*Hyd | 1 | 13,50 | 13.50 | 0.2974 |
| (12) Ds*Hyd | 2 | 34,40 | 17.20 | 0.2532 |
| (13) Gen*Ds*Hyd | 2 | 15.02 | 7.51 | 0.5436 |
| (14) Tsalt*Ds*Hyd | 2 | 30.56 | 15.28 | 0.2940 |
| (15) Gen*Tsalt*Ds*Hyd | 2 | 39.44 | 19.72 | 0.2084 |
| (16) Storage (St) | 1 | 80.67 | 80.67 | 0.0132 |
| (17) Gen*St | 1 | 92.04 | 92.04 | 0.0084 |
| (18) Tsalt*St | 1 | 9.38 | 9.38 | 0.3844 |
| (19) Gen*Tsalt*St | 1 | 13.50 | 13.50 | 0.2974 |
| (20) Ds*St | 2 | 65.40 | 32.70 | 0.0783 |
| (21) Gen*Ds*St | 2 | 24.02 | 12.01 | 0.3801 |
| (22) Tsalt*Ds*St | 2 | 172.31 | 86.16 | 0.0020 |
| (23) Gen*Tsalt*Ds*St | 2 | 16.19 | 8.09 | 0.5188 |
| (24) Hyd*St | 1 | 3.38 | 3.38 | 0.6008 |
| (25) Gen*Hyd*St | 1 | 10.67 | 10.67 | 0.3538 |
| (26) Tsalt*Hyd*St | 1 | 0.17 | 0.17 | 0.9073 |
| (27) Gen*Tsalt*Hyd*St | 1 | 9.38 | 9.38 | 0.3844 |
| (28) Ds*Hyd*St | 2 | 51.19 | 25.59 | 0.1331 |
| (29) Gen*Ds*Hyd*St | 2 | 4.15 | 2.07 | 0.8439 |
| (30) Tsalt*Ds*Hyd*St | 2 | 117.27 | 58.64 | 0.0124 |
| (31) Gen*Tsalt*Ds*Hyd*St | 2 | 14.31 | 7.16 | 0.5593 |
| Error | 48 | 584.00 | 12.17 | |
| Total | 95 | 4263.96 | | |

The third-order interaction (fourth degree) among the factors salt type *vs.* dose *vs.* hydration *vs.* storage (30) was significant, and therefore, the other interactions of the second or first order were inserted in it as well as the main effects of these factors. The other three interactions (genotype *vs.* salt type (3), genotype *vs.* dose (5) and genotype *vs.* storage (17)) also needed to be unfolded because the factor genotype was not included in the interaction of higher degree.

Therefore, it is necessary to diversify all the combinations of the levels of the factors to clarify the effect of each factor and the possible manners in which the effect of a factor can be modified by the variation of other factors (RIBEIRO et al., 2007); this means dealing with the simple effect of a factor (Table 3), which is nothing more than a decomposition of the degrees of freedom and the sum of the squares of the interactions.

For example, the interaction Tsalt*Ds*Hyd*St produces a sum of squares (SS) of 1,532.41 with a respective 23 degrees of freedom (DF) that can be unfolded. These values refer to the addition between the different SS and DF of the main factors and of

all the interactions encompassed by this interaction. In other words, it refers to the total SS and DF among the sources of variation presented in Table 2, numbered as 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 and 30.

**Table 3.** Analysis of variation for the variable cooking time in min. with the decomposition of the variation attributable to the salt dose (fixed dose) for each combination among the levels of the other factors: salt type (KCl and NaCl), hydration time (45 and 90 days) and storage time (45 and 90 days).

| SV | DF | SS | MS | Pr > F |
|---|---|---|---|---|
| Tsalt\|Hyd_St | (7) | (342.77) | (48.96) | |
| Dose\|KCl_Hyd_8_St_45 | 2 | 240.66 | 120.33 | 0.0003 |
| Dose\|KCl_Hyd_8_St_90 | 2 | 109.50 | 54.75 | 0.0162 |
| Dose\|KCl_Hyd_16_St_45 | 2 | 234.00 | 117.00 | 0.0003 |
| Dose\|KCl_Hyd_16_St_90 | 2 | 94.50 | 47.25 | 0.0273 |
| Dose\|NaCl_Hyd_8_St_45 | 2 | 18.66 | 9.33 | 0.4700 |
| Dose\|NaCl_Hyd_8_St_90 | 2 | 298.66 | 149.33 | 0.0001 |
| Dose\|NaCl_Hyd_16_St_45 | 2 | 175.16 | 87.58 | 0.0018 |
| Dose\|NaCl_Hyd_16_St_90 | 2 | 18.50 | 9.25 | 0.4731 |
| (Dose*Tsalt*Hyd*St) | (23) | (1,532.41) | (66.62) | |
| Error | 48 | 584.00 | 12.66 | |

Because the salt dose was the most important factor, and the other factors were supplementary, it was considered fixed, while the other factors varied. The sum of squares of the factors that were varied account for 342.77 of the sum of squares with 7 degrees of freedom, coming from the sum of the sources of variation in Table 2, numbered as 2, 8, 16, 18, 24 and 26. Therefore, the effect of the dose on the time of cooking was evaluated in each one of the combinations among salt type, hydration time and storage time through the commands of Figure 1 - item "ii". It is apparent that the doses of salt affect the cooking time significantly ($p < 0.05$) for most combinations tested, except when the grains were stored for 45 days and then hydrated in an NaCl solution for 8 hours or when the grains were stored for 90 days and hydrated with the same dose of salt for a period of 16 hours (Table 3).

However, two degrees of freedom still remained for each combination of factors, which needed to be unfolded for the analyses of the conclusion. To achieve the results of this phase, the commands explained in the item "iii" of Figure 1 were used. Because the salt dose was a quantitative factor, the most appropriate decision was the regression adjustment. The degree of the polynomial was determined by means of contrasts, which were previously estimated because the doses were not equidistantly separated. Thus, Table 4 presents the summary of the analysis of variance for the tests of significance of the linear and quadratic components in which the DF and SS of the simple effect of each factor were then unfolded for each component. The first dose combination inside the salt type KCl, with

a hydration time of 8 hours and a storage time of 45 days (Dose|KCl_Hyd_8_St_45), had an SS of 240.66 and 2 DF, which were decomposed in one DF for the linear component and one for the quadratic in Table 4. Similarly, 100.31 of SS was partitioned for the linear component, and 140.37 was partitioned for the quadratic component.

**Table 4.** Summary of the analysis of variance for the tests of significance of the linear and quadratic components of the variation attributable to the doses for each salt type (KCl and NaCl), hydration time (8 – Hyd_8 and 16 hours – Hyd_16) and storage time (45 – St_45 and 90 days – St_90) for the cooking time (min.).

| F.V. | DF | Component | SS | MS | Pr > F |
|---|---|---|---|---|---|
| KCl_Hyd_8_St_45 | 1 | Linear | 100.31 | 100.31 | 0.0070 |
| | 1 | Quadratic | 140.37 | 140.37 | 0.0017 |
| KCl_Hyd_8_St_90 | 1 | Linear | 5.08 | 5.08 | 0.5296 |
| | 1 | Quadratic | 104.42 | 104.42 | 0.0061 |
| KCl_Hyd_16_St_45 | 1 | Linear | 6.36 | 6.36 | 0.4820 |
| | 1 | Quadratic | 227.65 | 227.65 | 0.0001 |
| KCl_Hyd_16_St_90 | 1 | Linear | 29.98 | 29.98 | 0.1304 |
| | 1 | Quadratic | 64.53 | 64.53 | 0.0286 |
| NaCl_Hyd_8_St_45 | 1 | Linear | 18.56 | 18.56 | 0.2319 |
| | 1 | Quadratic | 0.10 | 0.10 | 0.9283 |
| NaCl_Hyd_8_St_90 | 1 | Linear | 271.64 | 271.64 | 0.0000 |
| | 1 | Quadratic | 27.04 | 27.04 | 0.1504 |
| NaCl_Hyd_16_St_45 | 1 | Linear | 13.05 | 13.05 | 0.3151 |
| | 1 | Quadratic | 162.11 | 162.11 | 0.0008 |
| NaCl_Hyd_16_St_90 | 1 | Linear | 10.43 | 10.43 | 0.3685 |
| | 1 | Quadratic | 8.07 | 8.07 | 0.4286 |
| Error | 48 | - | 584.00 | 12.66 | |

However, this phase requires that researchers carry out part of the work "manually" because the GLM procedure of the SAS will not estimate the significance of the contrasts correctly. In other words, the problem is in the inadequacy of the F test. Thus, it is necessary to make the estimation manually by applying the correct denominator (NOGUEIRA; CORRENTE, 2000) because, when it is defined in the command *model* that Coc= Ds(Tsalt★Hyd★St) to perform the estimate of the contrast for the linear effect of the doses of salt on the time of cooking (considering KCl as salt, a hydration time of 8 hours and a storage time of 45 days - contrast 'c1_L_KCl★Hyd8★St1' Ds(Tsalt★Hyd★St) -0.655 -0.094 0.749, e.g., Figure 1), the correct sum of squares of the effect of the related treatments will be achieved. However, because the effects of other factors are eliminated from the model, the estimate of the MS of the error will be incorrect. Then, researchers must utilize the sum of squares and the mean square estimated by the program, but it must estimate the significance of the contrasts performed separately using the correct estimate of the experimental error achieved in the analysis of variance in Table 2 (12.16). If researchers try to correct the value estimated for the experimental error by adding factors to the model,

they will eventually achieve an incorrect estimate of the effects of the treatments. Thus, the best solution is to estimate the significance using other tools, such as a spreadsheet.

Therefore, it is possible to verify that the results for cooking time indicate that the significant variations ascribed to the doses of salt presented a quadratic behavior as follows: *i)* when the grains of beans were stored for 45 days and hydrated in a KCl solution for 8 hours*; ii)* when the grains of beans were stored for 90 days and hydrated in a KCl solution for 8 hours; *iii)* when the grains of beans were stored for 45 days and hydrated in a KCl solution for 16 hours; *iv)* when the grains of beans were stored for 90 days and hydrated in a KCl solution for 16 hours; and *v)* when the grains of beans were stored for 45 days and hydrated in an NaCl solution for 16 hours. However, grains stored for 90 days and then submitted to a hydration time of 8 hours with NaCl presented a descending linear behavior for the cooking time variable.

When variation in the linear or quadratic effect was verified, an estimate of the equation parameters was carried out using the commands presented in item "iv" of Figure 1. When these parameters were achieved, it was possible to prepare Figure 2, which shows the behavior of the cooking time over the doses of salt for each of the respective combinations of the other factors. Thus, when the grains were stored for a period of 90 days and then hydrated for 8 hours in a KCl solution before the cooking process, the behavior of the cooking time over the doses of salt was quadratic, and the dose of 12.14 g provided the shortest time of cooking. However, when the grains were stored for the same period but then hydrated in an NaCl solution for eight hours, the cooking time presented a linear behavior along the salt doses. Thus, for up to 25 g of salt, there was a decrease in the time of cooking.

It is demonstrated that the results achieved with a unifactorial analysis, when there is a factorial arrangement of treatments, may be incomplete or even wrong, which means that it is not true that the cooking time has a quadratic behavior along the doses of salt, as reported in the unifactorial analysis. However, it can be said that there is a dependence between the effects of the factors, salt type (Tsalt) *vs.* salt dose (Ds) *vs.* hydration time (Hyd) *vs.* storage time (St). The behavior of the cooking time with the doses of salt depends on the combination of the levels of the other factors. With a simple change of the salt type from KCl to NaCl, with the maintenance of the levels of the other factors, the behavior of the cooking time with the salt dose varied from quadratic to linear, respectively.
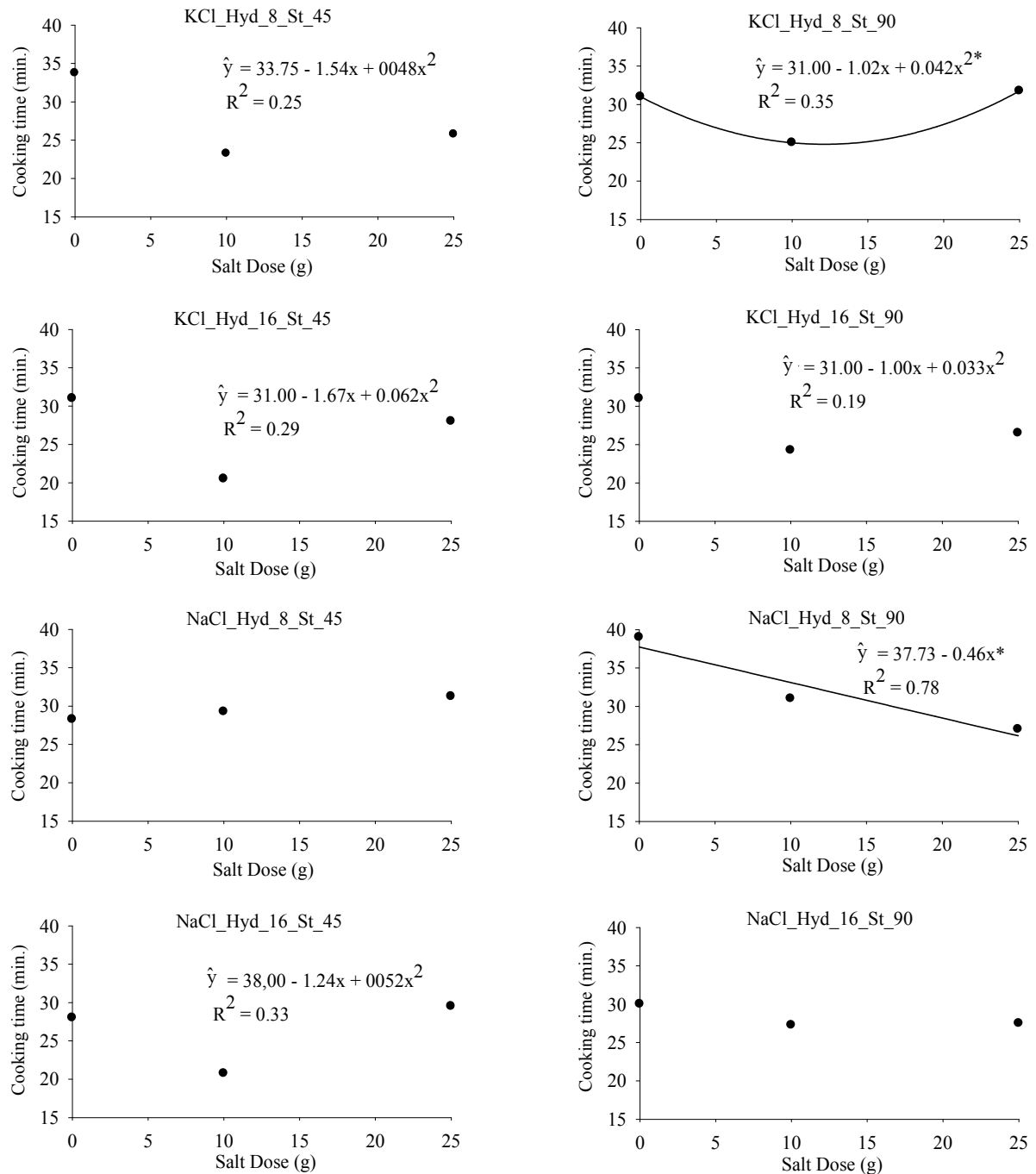
**Figure 2**. Effect of the doses of salt on the cooking time (min.) of the grains of beans combined with two types of salt (KCl and NaCl), two hydration periods (8 (Hyd_8) and 16 hours (Hyd_16)) and two storage times ((45 (St_45) and 90 (St_90) days)).
*Significant parameter by the t test at 5% significance.

In these situations, the effect of the factors acting jointly is different from the sum of the effect of each one acting separately. Therefore, the effect is not simply additive. Continuously, the response to a factor depends on the level of one or more levels of the other factors related to an experiment. Factor dependence determines the interaction, which cannot be found in unifactorial experiments because, in such experiments, the other 'factors' are present in a single level (PERECIN; CARGNELUTTI FILHO, 2008).

In general, it is possible to consider that the interaction investigation can provide results more comprehensive and coherent for experimenters. This fact becomes explicit when considering the analyses in Tables 1 (unifactorial) and 2 (factorial) simultaneously. For example, when analyzed separately, the factor salt type did not present

significance by the F test; it proved to be an important source of variation for the study on interactions, participating significantly in the most significant interactions (Gen★Tsalt, Tsalt★Ds★St and Tsalt★Ds★Hyd★St). A similar situation occurred for the factor storage time. Therefore, if researchers neglect the combination of several factors of treatment in their research work, they may miss relevant information that contributed to the composition of the response variable, or they may reach incomplete or wrong conclusions.

Experiments involving a set of factors are usually more efficient than simple experiments because they allow for broader responses and the evaluation of the interaction among factors, providing an estimate of the variance of the experimental error with more accuracy and increasing the power of the statistical tests (WECHSLER, 1998). In addition, because all the plots are employed in the calculation of the main effects and interactions, the number of replications is high (STEEL et al., 1997). However, the use of factorial experiments presents some disadvantages, such as a sharp increase in the number of treatments, the necessity of confounding when there are many treatments and the difficulties in interpretation due to the increased number of levels and factors (PIMENTEL-GOMES; GARCIA, 2002).

It is also important to point out that the efficiency of an experiment depends mainly on the planning of the experiment, which is a decisive stage for the success of the work. Also, in the case of factorial experiments, as already highlighted, the effect of the treatment is formed by the combination of the levels of the factors. In other words, each experimental unit will receive, in each raffle, one level of one factor. The PROC FACTEX of the SAS (SAS, 2004) is a procedure that can assist with experimental planning, especially for factorial experiments. This procedure providing support of experimental designs without limiting the number of factors and designs one intends to construct.

## Conclusion

When the effect of the interaction among factors is significant, the unfolding of the levels of a factor in relation to the level of another factor allows researchers to make more inferences about the data achieved;

The availability of explanatory examples to researchers may lead to the reduction of the errors found in statistical analyses involving factorial experiments;

An understanding of the factors involved with the response variable allows more complete and appropriate conclusions; and

The experimental planning, a priori, minimizes the problems related to the errors involving factorial experiments.

## References

BERTOLDO, J. G.; COIMBRA, J. L. M.; GUIDOLIN, A. F.; MANTOVANI, A.; VALE, N. M. Problemas relacionados com o uso de testes de comparação de médias em artigos científicos. **Biotemas**, v. 21, n. 2, p. 145-153, 2008a.

BERTOLDO, J. G.; COIMBRA, J. L. M.; GUIDOLIN, A. F.; MIQUELOTO, A.; TOALDO, D. Uso ou abuso em testes de comparações de média: conhecimento científico ou empírico? **Ciência Rural**, v. 38, n. 4, p. 1145-1148, 2008b.

BERTOLDO, J. G.; ROCHA, F.; COIMBRA, J. L. M.; ZITTERELL, D.; GRAH, V. F. Teste de comparação de médias: dificuldades e acertos em trabalhos científicos. **Revista Brasileira de Agrociência**, v. 13, n. 4, p. 441-447, 2007.

BEZERRA NETO, F.; NUNES, G. H. S.; NEGREIROS, M. Z. Avaliação de procedimentos de comparações múltiplas em trabalhos publicados na revista Horticultura Brasileira de 1983 a 2000. **Horticultura Brasileira**, v. 20, n. 1, p. 5-9, 2002.

CARDELLINO R. A.; SIEWERDT, F. Utilização adequada e inadequada dos testes de comparação de médias. **Revista Brasileira de Zootecnia**, v. 21, n. 6, p. 985-995, 1992.

CHEW, V. Comparing treatment means: a compendium. **HortScience**, v. 11, n. 4, p. 348-335, 1976a.

CHEW, V. Uses and abuses of Duncan's multiple range test. **Proceedings of the Florida State Horticultural Society**, v. 89, p. 251-253, 1976b.

NOGUEIRA, M. C. S.; CORRENTE, J. E. Decomposição da interação tripla significativa utilizando o comando CONTRAST do PROC GLM do SAS aplicado a modelo de classificação tripla para dados balanceados. **Bragantia**, v. 59, n. 1, p. 109-115, 2000.

PERECIN, D.; CARGNELUTTI FILHO, A. Efeitos por comparações e por experimento em interações de experimentos fatoriais. **Ciência e Agrotecnologia**, v. 32, n. 1, p. 68-72, 2008.

PETERSEN, G. R. Use and misuse of multiple comparison procedures. **Agronomy Journal**, v. 69, n. 2, p. 205-208, 1976.

PIMENTEL-GOMES, F.; GARCIA, H. G. **Estatística aplicada a experimentos agronômicos e florestais**: exposição com exemplos e orientações pra uso de aplicativos. Piracicaba: Fealq, 2002.

PROCTOR, J. R.; WATTS, B. M. Development of a modified Mattson Bean Cooker procedure based on sensory panel cook ability evaluation. **Canadian**

**Institute of Food Science and Technology Journal**, v. 20, n. 1, p. 9-14, 1987.

RIBEIRO, L. G. M.; ALVES, L. H.; SILVA, M. B. Melhoria da qualidade na produção de peças de grande porte em aço fundido utilizando planejamento de experimentos. **Tecnologia em Metalurgia e Materiais**, v. 4, n. 1, p. 36-41, 2007.

SANTOS, J. W.; MOREIRA, J. A. N.; BELTRÃO, N. E. M. Avaliação do emprego dos testes de comparação de médias na revista Pesquisa Agropecuária Brasileira (PAB) de 180 a 1994. **Pesquisa Agropecuária Brasileira**, v. 33, n. 3, p. 225-230, 1998.

SAS-Statistical Analisys System. **User's guide**. Versão 9.1. Cary: Statistical Analisys System Institute, 2004.

SILVA, J. G. C. A consideração da estrutura de unidades em inferências derivadas do experimento. **Pesquisa Agropecuária Brasileira**, v. 34, n. 6, p. 911-925, 1999.

STEEL, R. G. D.; TORRIE, J. H.; DICKEY, D. **Principles and procedures of statistics**. A biometrical approach. 3rd ed. New York: Mcgraw Hill, 1997.

WECHSLER, F. S. Fatoriais fixos desbalanceados: uma análise mal compreendida. **Pesquisa Agropecuária Brasileira**, v. 33, n. 3, p. 231-26, 1998.

YASSIN, N.; MORAIS, A. R.; MUNIZ, J. A. Análise de variância em um experimento fatorial de dois fatores com tratamentos adicionais. **Ciência e Agrotecnologia**, v. 26, n. 7, p. 1541-1547, 2002.