Journal of the American Pomological Society 76(1): 27-35 2022

A Note on the Analysis and Interpretation of Designed Experiments with Factorial Treatment Structure

Richard P. Marini[1]

Additional index words: analysis of variance, interaction, marginal means, multiple comparisons, simple effects

Abstract

Agricultural researchers often use factorial treatment structures, where treatments consist of combinations of two or more levels of two or more factors. Factorial experiments are more efficient than performing experiments involving one factor at a time. They also allow researchers to study the effect of each factor on the response variable, as well as the effects of interactions between factors on the response variable. When interactions are significant, proper interpretation of results is often complicated. Over the years, several post-analysis of variance (ANOVA) techniques have been used to interpret results. A partial data set for a 2 x 2 x 4 factorial arrangement of treatments in a randomized complete block design was used to demonstrate and compare three commonly used post-ANOVA methods when the three-way interaction is significant. In the presence of interaction, there may be situations where marginal means (main effects means) can be compared but slicing the data set without separating the data usually provides the most information and allows correct interpretation of the results. The advantage of slicing is that all the data are used for the analysis and the effect of one factor can be evaluated while holding the other factors fixed.

Horticulturists often perform experiments with a factorial treatment design, where there are two or more independent factors (independent variables), and all levels of each factor are combined with all levels of every other factor. Sometimes, for various reasons, the factorial treatment design may be modified. Since the number of treatment combinations increases rapidly as the number of factors increases, a fractional may be used where only a subset of all possible treatment combinations are used (Ribeiro et al., 2019). Augmented factorials are factorial experiments that also include one or more additional treatments and were discussed by Marini (2003) and Piepho et al. (2006). Factorial experiments are more efficient (Fisher, 1926) and more powerful (Chin and Lee, 2008; Czitrom, 1999) than studying one factor at a time while other factors are kept fixed. This paper will be limited to a discussion of only complete factorials. The primary reason for considering factorial experiments is to determine if the effect of one level of a factor on the response variable depends on the level of another factor. However, when interactions are significant post-analysis of variance (ANOVA) procedures may be complicated. For this reason, some authors choose to ignore significant interactions and discuss only main effects. Even worse, some researchers may not even recognize the factorial nature of the experiment and simply perform a one-way ANOVA on all treatment combinatons. If the *P*-value for treatment is significant, a multiple comparison is used for all possible pair-wise comparisons of the treatment means. However, the reason that we perform

[1] Department of Plant Science, The Pennsylvania State University, University Park, PA 16802 Email: rpm12@psu.edu. This work was supported by the USDA National Institute of Food and Agriculture Federal Appropriations under Project PEN04625 and accession number 1013400.

factorial experiments is because we want to know if two or more factors interact. Ignoring interactions, or comparing simple effect means (cell means) may lead to misinterpretation of the results. Although still complicated, advances in statistical software packages over the past several decades have facilitated the interpretation of interactions.

Types of independent variables. Factorial experiments may involve only quantitative factors, such as temperature, several levels of fertilizer, or several concentrations of a growth regulator, and are best analyzed with multiple regression (Chew, 1977; Hinkelmann and Kempthorne, 1994). Experiments involving only qualitative factors, such as cultivars, rootstocks, and types of potting media, are analyzed with ANOVA and will be discussed in more detail in this paper. Experiments involving both qualitative and quantitative factors can be analyzed with analysis of covariance (ANCOVA) as described by Marini and Ward (2012).

Testing hypotheses with ANOVA. Interpretation of experiments with more than three factors is often very complex and such experiments should be avoided if possible. To demonstrate analysis of a 2 x 4 x 2 factorial experiment, a partial data set from a raspberry high tunnel experiment will be used; where two cultivars ('Josephine' and 'Polka') were grown in high tunnels covered with four plastics (GIN, KLP, TUFF, and UVB) for two years (2016 and 2017). The experimental design was a randomized complete block (RCBD), with three blocks. The RCBD is common in agricultural research, but factorial treatments can be used with other experimental designs. The data were modified to produce a three-way interaction. With this factorial treatment structure, we can test the main effects of the three factors, three two-way interactions to test the additive relationship between each pair of factors, and the three-way interaction to test the additive effects of all three factors.

The following is another way of describing the seven hypotheses that can be tested with a three-way ANOVA.

1. Yields for the two cultivars are equal ('Josephine' = 'Polka'),
2. Yields for the four plastics are equal (GIN = KLP = TUFF = UVB),
3. Yields for the two years are equal (2016 = 2017),
4. The effects of cultivar and plastic are additive (the effect of a cultivar does not depend on the plastic; this is the cultivar x plastic interaction),
5. The effects of cultivar and year are additive (the effect of a cultivar does not depend on the year; this is the cultivar x year interaction),
6. The effects of plastic and year are additive (the effect of a plastic does not depend on the year; this is the plastic x year interaction),
7. The effects of cultivar, plastic and year are additive (the effect of a cultivar does not depend on the combination of plastic and year; the effect of a plastic does not depend on the combination of cultivar and year; the effect of a year does not depend on the combination of plastic and cultivar; this is the cultivar x plastic x year interaction).

Identifying interactions. Three different approaches that are sometimes used in agricultural research to interpret interactions will be compared. The following analyses can be performed with most of the better statistical software packages that can appropriately analyze mixed models, but for this paper SAS's GLIMMIX procedure will be used (SAS Inst. Inc. 2013). Before any statistical analyses are performed it is a good idea to "get to know your data" with some scatter plots to observe patterns in the data and to identify unusual observations that are sometimes called "outliers". Descriptive statistics (N, means, maximum, minimum, variances, etc.) can help identify errors in coding, unusual observations, and possible violation of the assumption of homogenous treatment variances. Visualizing data is often easier with plots of treatment means (Fig. 1). PROC GPLOT

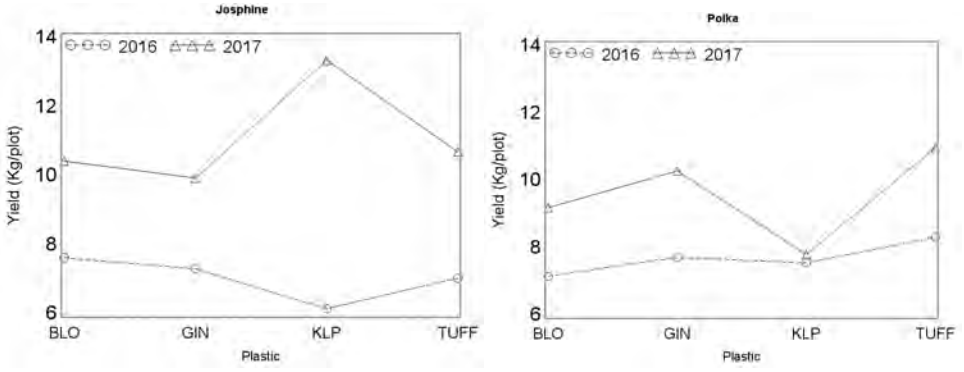

Figure 1. Plots of means for 2 two raspberry cultivars grown in high tunnels under four types of plastic for two years. The symbols represent the mean for each combination of cultivar ('Josephine' on the left and 'Polka' on the right), plastic and year, also called simple means or cell means.

was used to produce the graphs in Figure 1. Yields were generally higher in 2017 than in 2016 and yields tended to be slightly higher for 'Josephine' than for 'Polka' in 2017, but not in 2016, indicating a possible cultivar by year interaction. For 'Josephine' yields were lowest for KLP in 2016, but highest in 2017 and the opposite was true for 'Polka' under KLP in 2017. Plastic affected yield in 2017 to a greater extent than in 2016 regardless of cultivar. Taken together there is visual evidence of a three-way interaction because the lines for the two cultivars are not parallel for cultivars within years or years within each cultivar.

Following graphical examination of the data, we can use formal statistical tests to verify our preliminary interpretation. In this case we can perform a three-way ANOVA with block as a random effect. Depending on the results from the ANOVA, different post-ANOVA analyses may be used to help interpret the results.

1. If no interactions are significant, but one or more main effects are significant, then all pairs of marginal means (main effect means) for a factor can be compared or means can be compared against a control or some other treatment.
2. If the highest-order interaction (in this case the 3-way interaction) is not significant, but one or more lower-order interaction(s) (in this case the 2-way interactions) are significant, simple effects or cell means can be compared because marginal means may not properly represent the treatment differences at the various levels of a factor. Cell means can be compared only as simple effects, where all but one factor is fixed at a certain level.
3. If the highest-order interaction is significant, then the marginal means and simple effects associated with lower-order interactions become less interesting, even if they are significant. The three-way interaction can be dissected by comparing simple effects, to test the effect of one factor on the response variable while holding the other two factors constant.

In the horticultural literature, several post-ANOVA tests have been used when interactions were significant. Some approaches are more appropriate than others and some provide more information than others, but the interpretation of results often depends on the approach used. Below, some of these approaches will be demonstrated to dissect the significant three-way interaction. Many researchers hope that higher-level interactions are not significant because results are difficult to interpret. Most researchers consider

P-values < 0.05 to be significant. However, the American Statistical Association recently stated that "well-reasoned statistical arguments contain much more than the value of a single number and whether that number exceeds an arbitrary threshold", such as $P < 0.05$ (Wasserstein and Lazar, 2016). Since we perform factorial experiments specifically because we want to know if there is interaction, I consider a *P*-value of 0.08 as adequate evidence to reject the null hypothesis that the factors are additive because I want to explore the interaction. Hinkelmann and Kempthorne (2005) suggested that interactions with *P*-values as high as 0.1 may be investigated by looking at simple effects.

One-way ANOVA on all treatment combinations. Table 1 shows the SAS code to perform the three-way ANOVA with PROC GLIMMIX, along with some of the output. The "Covariance Parameter Estimates" table shows the estimates of the variance component parameters. The estimate of the block variance component is 0.01124 and the estimate of the error variance component is 2.5545 (labelled "Residual"). Options can be added to the model statement to request confidence intervals. The second table is the ANOVA showing the sources of variation or "Effect", the degrees of freedom for the numerator and the denominator, the F-value, and the probability of a greater F. The ANOVA table shows that the main effect of year is significant ($P < 0.0001$), but there is insufficient evidence to reject the null hypothesis that means for plastics or cultivars are equal. The only two-way interaction that is significant is year x cultivar ($P = 0.0242$). In cases

Table 1. SAS code and output for a three-way analysis of variance in a randomized complete block design using the GLIMMIX procedure.

```
Proc GLIMMIX;
  Class block year cultivar plastic;
  Model yield = year plastic cultivar plastic*year plastic*cultivar
      cultivar*year cultivar**plastic*year;
  Random block;
  Run;
```

Parameter Estimates		
Cov Parm	**Estimate**	**Standard Error**
block	0.01124	0.1758
Residual	2.5545	0.6596

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
year	1	30	41.62	<.0001
plastic	3	30	0.38	0.7656
cultivar	1	30	0.90	0.3495
year*cultivar	1	30	5.64	0.0242
cultivar*plastic	3	30	1.96	0.1407
year*plastic	3	30	0.41	0.7484
year*cultivar*plastic	3	30	2.99	0.0464

Table 2. SAS code plus output when a three-way factorial in a randomized complete block design is analyzed as a one-way ANOVA with 16 treatments using the GLIMMIX procedure.

```
Title 'One-way ANOVA with 16 treatments (trt)';
  Proc GLIMMIX;
  Class block trt;
  Model yield = trt;
  Random block;
  LSmeans trt / adjust = tukey lines;
Run;
```

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
trt	15	30	4.36	0.0003

	Cultivar	Plastic	Yield
2016	Josephine	GIN	7.3 b [z]
		KLP	6.1 b
		UVB	7.6 b
		TUFF	7.0 b
	Polka	GIN	7.7 b
		KLP	7.5 b
		UVB	7.1 b
		TUFF	8.3 b
2017	Josephine	GIN	9.9 ab
		KLP	13.3 a
		UVB	10.4 ab
		TUFF	10.7 ab
	Polka	GIN	10.3 ab
		KLP	7.8 b
		UVB	9.2 ab
		TUFF	11.0 ab

[z] LSmeans followed by common letters do not differ at the 5%, by Tukey-Kramer test.

such as this, where the highest-level interaction is significant ($P = 0.0464$), the main effects and the lower-level interactions are usually of less interest than understanding the three-way interaction.

Agricultural researchers sometimes ignore significant interactions and only consider marginal means. Although such an approach is usually discouraged, Hinkelmann and Kempthorne (2005) suggest that tests for main effects are useful and meaningful when codirectional interaction is present, but not when antidirectional interaction is present. Codirectional interaction occurs when the change in the response is in the same direction and antidirectional interaction occurs when the change in response is in the opposite direction.

In the horticultural literature we can find several ways to dissect interactions, and each has advantages and disadvantages. One way is to create a new variable by recoding each combination of treatments and perform a one-way ANOVA on the 16 treatment combinations, followed by performing a multiple comparison, such as the Tukey-Kramer test on the 16 means (Table 2). The table for the Covariance Parameter Estimates is not

shown because the estimates are the same as for the three-way ANOVA. As expected, the treatment is significant ($P = 0.0003$). The SAS code in Table 2 generates three additional tables that are not shown because they are large. The first table shows the treatment (trt) Least Squares Means which shows for each treatment the Estimate or Least Squares mean, the standard error of the estimate, the denominator degrees of freedom, the *t*-value to test the hypothesis that the LSmean is equal to zero and its *P*-value. In most cases, we do not care if the LSmeans are equal to zero. The second table in the output is the "Differences of trt Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer", where there is a row for each of the 120 pair-wise comparisons. Results shown in the third table labelled "Tukey-Kramer Grouping for trt Least Squares Means (Alpha=0.05)", requested with the LINES option in the LSMEANS statement, are summarized in Table 2. In the output the treatment combinations were arranged in descending order for yield, so the data were rearranged in Table 2 to facilitate interpretation. Based on the Tukey-Kramer test, 'Josephine' under KLP in 2017 had higher yields than all other treatment combinations except five. This approach allows us to determine which treatment combinations are equal and it is relatively easy to perform and simply requires creating new treatment names for each treatment combination. However, many of the comparisons are not of interest. For example, we probably do not care if yield for 'Josephine' under GIN in 2017 is different than 'Polka' under KLP in 2016. Valid simple effects for this experiment could be 'Josephine' under GIN in 2016 vs. 'Josephine' under GIN in 2017, or 'Josephine' under GIN in 2017 vs. 'Josephine' under KLP in 2017. Another problem with this approach is that pairwise, multiple comparisons are appropriate only for unstructured treatments (Chew, 1976; Gates, 1991; Lowry, 1992; Yossa and Verdegem, 2015), and factorial experiments have structured treatments. Generally, performing multiple comparisons of individual factorial treatments is discouraged, but some statisticians indicated they might condone such an approach if the main effects were not significant (D.B. Duncan) or if their F-ratios were less than two (J.W. Tukey) (Chew, 1977). Older textbooks often suggest using contrasts to make preplanned comparisons of interest (Lentner and Bishop, 1993).

Dividing the data set. A common post-ANOVA technique to interpret interaction involves physically slicing the data to analyze simple main effects by breaking the data set into separate parts. In the case of our three-way factorial, we could perform two two-way ANOVAs on each year, two two-way ANOVAs for each cultivar and four two-way ANOVAs for each plastic. If the two-way interaction is not significant, then main effect means can be compared with a multiple comparison. If the interaction is significant, then one-way ANOVAs can be performed on each level of each variable, followed by a multiple comparison. For factors with only two levels, such as year and cultivar, multiple comparisons are not needed because the ANOVA tests the hypotheses that the two levels are equal. Alternatively, we could perform four one-way ANOVAs to compare plastics within each combination of cultivar and year. For brevity, Table 3 shows results for four ANOVAs to compare plastics within each combination of year and cultivar. Interpretation of results when simple main effects are analyzed is different than when comparing all 16 treatment combinations in Table 2. Notice that the *P*-values for 'Josephine' in 2016 and 'Polka' in 2017 are not significant, indicating that yield was not affected by the four plastics. For 'Polka' in 2016, TUFF had higher yield than UVB. For 'Josephine' in 2017, KLP had higher yields than GIN. According to Schabenberger et al. (2000) there are two disadvantages of physical slicing: 1.) By physically slicing by year and cultivar, each analysis contributes only ¼ of the total information. The error degrees of freedom are reduced, and the individual analyses have

Table 3. SAS code for physically slicing the data set, using the SORT procedure, to perform four one-way ANOVAs in a RCBD. The ANOVAs for each combination of year and cultivar tests the null hypothesis that yield is not affected by plastic. Output for the four ANOVAs is presented in one table.

```
Proc sort; by year cultivar; run;
Proc GLIMMIX; by year cultivar;
   Class block plastic;
   Model yield = plastic;
   Random block;
   LSmeans plastic / adjust = tukey lines;
Run;
```

	2016		2017	
Plastic	Josephine	Polka	Josephine	Polka
GIN	7.3 a [z]	7.7 ab	9.9 b	10.3 a
KLP	6.1 a	7.5 ab	13.3 a	7.8 a
UVB	7.6 a	7.1 b	10.4 ab	9.2 a
TUFF	7.0 a	8.3 a	10.7 ab	11.0 a
P-value from one-way ANOVA				
	0.3280	0.01566	0.0267	0.5828

[z] Values within columns followed by common letters do not differ at the 5% level by Tukey-Kramer test.

less power than the combined analysis where interaction was detected; and 2.) Physically slicing is more work and requires performing at least four ANOVAs which need to be combined for a meaningful joint interpretation.

Slicing without separating the data. A third post-ANOVA approach involves slicing without separating the data (Schabenberger et al., 2000) to make meaningful comparisons of cell means. The significant year x cultivar x plastic interaction can be sliced in the following ways: by year, by cultivar, by plastic, by year x cultivar, by year x plastic, and by cultivar x plastic. The last three ways are most meaningful, because they compare the levels of one factor while holding the other two factors fixed. The SAS code for the three ways of slicing the LSMEANS, along with the cell means for each combination of year and cultivar are shown in Table 4 and cell means for each combination of year and plastic are shown in Table 5. Although the cell means in Tables 4 and 5 are identical to those in Table 3, the letters for the comparisons are different. The SLICEDIFF option in the LSMEANS statement eliminates many of the uninteresting comparisons shown in Table 2. The first LSMEANS statement compares the four plastics within each combination of year and cultivar. The cell means are the same as in Table 3 but results from the multiple comparison is different. For both cultivars in 2016, the plastic covers did not affect yield (Table 4). However, in 2017 for 'Josephine', GIN had lower yields than KLP or UVB. For 'Polka' in 2017, KLP had lower yields than TUFF. The second LSMEANS statement compares the two cultivars within each combination of year and plastic. In 2016 the two cultivars had similar yields for each type of plastic, but in 2017 'Josephine' had higher yields than 'Polka' under KLP. The 16 cell means are shown again in Table 5, along with *P*-values for the comparison of years within each combination of cultivar and plastic. 'Josephine' had significantly higher yields in 2017 for all four plastics. For 'Polka' years were different for only TUFF.

Summary. It is obvious that the interpretation of the results and conclusions often vary depending on the post-ANOVA procedure that is used in the presence of a significant

Table 4. SAS code for slicing the data set without separating the data using the slicediff option in the LSmeans statement. Output from the ANOVA and the LSmeans are rearranged in a way that would be acceptable in a publication.

```
Title 'ANOVA with Slicediff';
Proc GLIMMIX;
  Class block year cultivar plastic;
  Model yield = year plastic cultivar plastic*year cultivar*year
    cultivar*plastic cultivar*plastic*year;
  Random block;
  LSmeans cultivar*plastic*year / slicediff = cultivar*year.
  LSmeans cultivar*plastic*year / slicediff = plastic*year.
  LSmeans cultivar*plastic*year / slicediff = cultivar*plastic.
Run;
```

	2016		2017	
Plastic	Josephine	Polka	Josephine	Polka
GIN	7.3 a [z]	7.7 a	9.9 b	10.3 ab
KLP	6.1 a	7.5 a	13.3 a * [y]	7.8 b
UVB	7.6 a	7.1 a	10.4 a	9.2 ab
TUFF	7.0 a	8.3 a	10.7 ab	11.0 a
P - value from ANOVA				
Year	< 0.0001			
Plastic	0.7656			
Cultivar	0.3495			
Year*Plastic	0.7484			
Year*Cultivar	0.0242			
Year*Cultivar*Plastic	0.0457			

[z] LSmeans within each column (combination of cultivar and year) followed by common letters do not differ at the 5% level by SLICEDIFF, requested with the first LSmeans statement.

[y] Asterisk between cultivars for each combination of year and plastic indicates the two cultivars differ within a year at the 5% level, requested with the second LSmeans statement.

three-way interaction. Analyzing the data as a one-way ANOVA and comparing all 16 cell means indicates that the eight combinations of cultivar and plastic in 2016 were not different, and in 2017 'Josephine' under KLP had higher yields than 'Polka' under KLP (Table 2), but the only significant pair-wise comparison is not very meaningful. When data were physically sliced to perform one-way ANOVAs for each combination of year and cultivar, only two pair-wise comparisons were significant. In 2016 'Polka' under TUFF had higher yields than under UVB, and in 2017 'Josephine' under KLP had higher yields than under GIN (Table 3). Dissecting the three-way interaction with the three LSmeans statements limits the pairwise comparisons to mostly meaningful comparisons. Plastic covers affected yield for both cultivars in 2017 but not in 2016. 'Josephine' had higher yields than 'Polka' under KLP in 2017, but not in 2016. For 'Josephine,' yields were higher in 2017 than in 2016 for all plastics, but for 'Polka' yields for the two years were not different under any of the plastics. The reason for different results is because the standard error of the difference was 0.9246 and 1.3050 for the one-way and the three-way ANOVAs, respectively and both had 30 degrees of freedom in the denominator, but numerator degrees of freedom varied. For the four ANOVAs the standard error of the

Table 5. The effect of year on mean yield for combinations of raspberry cultivar and plastic. Output requested with the third LSmeans statement in Table 4.

	Josephine				Polka			
Year	GIN	KLP	UVB	TUFF	GIN	KLP	UVB	TUFF
2016	7.3	6.1	7.6	7.0	7.7	7.5	7.1	8.3
2017	9.9	13.3	10.4	10.7	10.3	7.8	9.2	11.0
P-value	0.0532[z]	<0.0001	0.0397	0.0087	0.0587	0.8493	0.1283	0.0497

[z] P-values for testing the hypothesis that years are equal within each combination of cultivar and plastic, requested with the third LSmeans statement in Table 4.

difference was 0.7546, 0.2681, 0.8544, and 0.8544 for 'Josephine' in 2016, 'Josephine' in 2017, 'Polka' in 2016 and 'Polka' in 2017 with only 8 degrees of freedom in the denominator. When using post-ANOVA procedures to dissect the significant interactions, authors are encouraged to consider methods that will make the pair-wise comparisons that are most meaningful. This is generally accomplished by comparing the levels of one factor while holding the other factors fixed by slicing the highest-order interaction(s) that is significant.

Literature Cited

Chew, V. 1976. Comparing treatment means: a compendium. HortScience. 11:348-356.

Chew, V. 1977. Comparisons among treatment means in an analysis of variance. USDA-ARS Bul. ARS/H/6, 64 p. https://books.google.com/books?hl=en&lr=&id=6mbefr0NbfwC&oi=fnd&pg=PA1&dq=Chew,+V.+1977.+Comparisons+among+treatment+means+in+an+analysis+of+variance.+USDA-ARS+Bul.+ARS/H/6,+64+p.&ots=dLFw1m4Eqt&sig=ap7LHbasvk8BAfbIdzUBcNRQO9M#v=onepage&q&f=false

Chin, R. and B.Y. Lee. 2008. Principles and practices of clinical trial medicine. Academic Press, London, UK. https://doi.org/10.1016/B978-0-12-373695-6.X0001-4.

Czitrom, V. 1999. One-factor-at-a time versus designed experiments. The Amer. Statistician. 53:126-131. https://www.tandfonline.com/doi/abs/10.1080/00031305.1999.10474445.

Fisher, R. 1926. The arrangement of field experiments. London, Eng. Ministry of Agric. and Fisheries 33:503-513. https://digital.library.adelaide.edu.au/dspace/bitstream/2440/15191/1/48.pdf.

Gates, C.E. 1991. A user's guide to misanalyzing planned experiments. HortScience 26:1262-1265.

Hinkelmann, K. and O. Kempthorne. 1994. Design and analysis of experiments, Vol. 1: Introduction to experimental design. John Wiley and Sons, Inc. New York, N.Y.

Hinkelmann, K. and O. Kempthorne. 2005. Design and analysis of experiments, Vol. 2: Advanced experimental design. John Wiley and Sons, Inc. New York, N.Y.

Lentner, M. and T. Bishop. 1993. Experimental design and analysis. 2nd ed. Valley Book Company. Blacksburg, VA.

Lowry, S.R. 1992. Use and misuse of multiple comparisons in animal experiments. J. Animal Sci. 70:1971-1977. https://doi-org.ezaccess.libraries.psu.edu/10.2527/1992.7061971x.

Marini, R.P. 2003. Approaches to analyzing experiments with factorial arrangements of treatments plus other treatments. HortScience 38:117-120.

Marini, R.P. and D. Ward. 2012. Using analysis of covariance with unequal slopes to increase efficiency and information obtained from designed experiments. J. Amer. Pomol. Soc. 66:91-100.

Piepho, H.P., E.R. Williams, and M. Fleck. 2006. A note on the analysis of designed experiments with complex treatment structure. HortScience 41:44-452.

Ribeiro, P.C.M., M.P. Campos, L.A.S. Pio, and J.S.S. Bueno Filho. 2019. Fractional factorials in a case study nutrition experiment with banana trees. Rev. Bras. Biom. 37:335-349. doi: 10.28951/rbb.v37i3.4.

Schabenberger, O., T.G. Gregoire, J.P Weyerhaeuser, Jr., and F. Kong. 2000. Collections of simple effects and their relationship to main effects and interactions in factorials. The Amer. Statistician 54:210-214.

SAS Inst. Inc. 2013. SAS/STAT 13.1 User's Guide. The GLIMMIX procedure. SAS Inst. Inc., Cary, NC.https://support.sas.com/documentation/onlinedoc/stat/131/glimmix.pdf

Wasserstein, R.L. and N.A. Lazar. 2016. The ASA statement on p-vlaues: context, process, and purpose. The Amer. Statistician 70:129-133. https://doi.org/10.1080/00031305.2016.1154108.

Yossa, R. and M. Verdegem. 2015. Misuse of multiple comparison tests and underuse of contrast procedures in aquaculture publications. Aquaculture 437:344-350. https://doi.org/10.1016/j.aquaculture.2014.12.023.