Research Paper , 1–??
2018

Effectiveness of Different Non-Additivity Tests

G. Lee*
University of Guelph, ON N1G 2W1, CAN
**email:* glee03@uoguelph.com

Summary: The goal of this study was to find the tests that are the most effective when used on three different data sets for detecting non-additivity when only one observation exists because there is no replication. The tests that were used for this study was the Tukey's Test, Mandel's Test, Tukey's Modified Test, Tussell Test, LBI Test and the Johnson and Graybill Test. After being tested on the three different data sets it can be concluded that the Mandel Test was the least effective out of all the test and the others were effective in their own right however the Johnson and Graybill, LBI and Tussell Test were the most effective in detecting non-additivity.

Key words: Non-Additivity; Single Observation Interaction; Methodologies

1. Introduction

The difference between a non-additive model and an additive model is that, with a non-additive model there is an interaction term that is included between two factors while the additive models does not include an interaction term (Alin and Kurt, 2006). This interaction can be tested in a variety of different ways depending on the environment of the model. The most common way of testing for this interaction is using a two-way ANOVA, this is the most common test when looking for non-additivity (Alin and Kurt, 2006). The issue with this test although it is effective is that there must be more than one observation per cell. This requirement is mandatory for the two-way ANOVA to be effective when testing for interaction. This issue of a data set having only one observation per cell can be overlooked by various other types of methodologies that were implemented to test for non-additivity within this specific environment. These different methodologies are unique in their own ways and yield different effectiveness when compared with each other.

The first test to detect a interaction between two variables when only one source of replication exists was the Tukey's one-degree-of-freedom test. This is one of the most notable test that is observed when understanding interaction with one observation per cell, but Tukey's model does not propose any particular structure of interaction (Alin and Kurt, 2006). Other methodologies can either go more in-depth when looking for a higher order interaction or be more general in terms of finding a interaction, this will be further discussed within this study. In particular there will be a focus around the following tests: LBI, Johnson and Greybill, Mandell, Tukey, Modified Tukey, and the Tussell test. These tests will be performed on different Data sets to come to a appropriate conclusion on the effectiveness of each test and the intricacies behind them. The tests will be carried out within the programming language , R, and the package known as, "additivityTests", will be used to carry out each test mentioned.

2. Identification of Multiple Tests for Non-Additivity

It is imperative that we closely look at the basis and underlying reasoning behind each test before we can conclude the effectiveness of each test. The tests that will be examined within this study are the LBI, Johnson and Greybill, Mandell, Tukey, Modified Tukey, and the Tussell test.

2.1 *Tukey's One Degree of Freedom*

The first test that will be examined is the Tukey's One Degree of Freedom test. This test was developed by John W. Tukey, he had implemented a way to examine and look for non-additivity in the form of analyzing rows and columns of the data in which they all sum to 0(Tukey, 1949). The two main situations that lead to a reaction when using this test from swelling is one or more observations are unusually discrepant or the effect of rows and columns are not additive based on the analysis Tukey (1949). This test was pioneered very early compared to the other tests and some tests have even tried to improve on Tukey's methodology of detecting non-additivity. Tukey's Test can be modelled the following way where the null hypothesis is Ho: Lambda = 0 (Alin and Kurt, 2006):

$$Y_{ij} = \mu + \alpha_i + \beta_j + \lambda\alpha_i\beta_j + \varepsilon_{ij} \qquad \begin{cases} i = 1, \ldots, a \\ j = 1, \ldots, b \end{cases}$$

$$\sum_i \alpha_i = 0, \quad \sum_j \beta_j = 0$$

The model assumes normality of eij as well as one of the a or b should be greater than 2 for the Tukey's one-degree-of-freedom test to hold true, this is because Tukey's test realizes on the need for a matrix that has to have a structure that is greater than 1x1. (Alin and Kurt, 2006). Figure 1 also describes the formulas in regards to the analysis of variance, the test statistic F is what determines the rejection of the hypothesis under Ho: lambda = 0. By using this model

The analysis of variance table for Tukey's one degree of freedom test for non-additivity

Source of variation	Sum of squares	Degrees of freedom	Mean square (MS)	F
Factor A	$SS_A = b\sum_i(\bar{Y}_{i.} - \bar{Y}_{..})^2$	$a - 1$	$\frac{SS_A}{a-1}$	$\frac{MS_A}{MS_E}$
Factor B	$SS_B = b\sum_j(\bar{Y}_{.j} - \bar{Y}_{..})^2$	$b - 1$	$\frac{SS_B}{b-1}$	$\frac{MS_B}{MS_E}$
Non-additivity	$SS_{AB} = \frac{\left(\sum_i\sum_j Y_{ij}(\bar{Y}_{i.} - \bar{Y}_{..})(\bar{Y}_{.j} - \bar{Y}_{..})\right)^2}{\sum_i(\bar{Y}_{i.} - \bar{Y}_{..})^2\sum_j(\bar{Y}_{.j} - \bar{Y}_{..})^2}$	1	SS_{AB}	$\frac{MS_{AB}}{MS_E}$
Error	$SS_E = SS_T - SS_{AB} - SS_A - SS_B$	$(a-1)(b-1)-1$	$\frac{SS_E}{(a-1)(b-1)-1}$	
Total	$SS_T = \sum_i\sum_j(Y_{ij} - \bar{Y}_{..})^2$	$ab - 1$		

Figure 1. Analysis of Variance for the Tukey's One Degree of Freedom Test (Alin and Kurt, 2006)

developed by Tukey, it had led to a fundamental basis of testing for non-additivity through the analysis of not relying on a structure because it tests for 0 on the coefficient lambda within the model.

2.2 *Mandel's Test*

The Mandel Test was created by John Mandel for the purpose of detecting non-additivity when only one replication exists, it was first purposed in 1961 through his paper called, "Non-Additivity in Two-Way analysis of Variance". Mandel proposed a model that is far more general when compared to Tukey's One degree of freedom because of the way it tests for interaction in regards to columns and rows (Mandel, 1961). This generalization of Tukey's test had resulted in either testing for the row interaction or the column interaction and not both, which the following parameters represent(Alin and Kurt, 2006):

$$\tau_{ij} = \theta_i\beta_j \longrightarrow \text{ interaction depending on column effect}$$
$$\tau_{ij} = \Re_j\alpha_i \longrightarrow \text{ interaction depending on row effect}$$

The Mandel test only focuses on one of the above interactions, the full model can be shown as the following when testing for column effect (Alin and Kurt, 2006):

$$Y_{ij} = \mu + \alpha_i + \beta_j + \theta_i\beta_j + \varepsilon_{ij} \quad \begin{cases} i = 1, \dots, a \\ j = 1, \dots, b \end{cases}$$

$$\sum_i \alpha_i = 0, \quad \sum_j \beta_j = 0, \quad \sum_i \theta_i = 0$$

Furthermore Mandel's suggestion was that for any row or column it could be quantifiable as (Alin and Kurt, 2006):

$$E(Y_{ij} - \bar{Y}_{.j}) = \alpha_i + \theta_i\beta_j$$

This suggestion ultimately changes the model for testing interaction on columns effects to (Alin and Kurt, 2006):

$$Y_{ij} = \mu_i + \phi_i\beta_j + \varepsilon_{ij} \quad \begin{cases} i = 1, \dots, a \\ j = 1, \dots, b \end{cases}$$

This change in the model leads to a hypothesis test on(Alin and Kurt, 2006)

$$H_0\colon \phi_i = 1$$

This change and generalization of Tukey's test looks specifically at quantifying the row or column effects rather than both at the same time was Mandel's modification towards Tukey's original test, the analysis of variance for Mandel's test can also be seen in Figure 2. This is why it was considered a more general model when comparing to the original.

2.3 *Simecek and Simeckova's Modification on Tukey's Test*

This test was developed by Petr Simecek and Marie Simeckova, it was proposed within the paper, "Modification of Tukey's additivity test" in 2012, where it would be more suitable in testing for more general interaction schemes (Simecek and Simeckova, 2012). This test will be referred to as the modified Tukey test within the experiment. They had proposed a separate model that would stem from Tukey's original model, but had tweaked the model ever so slightly to improve on different environments. The purposed "sub-model" that the two had created was the following (Simecek and Simeckova, 2012):

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}.$$

In this sub-model that was proposed the row effects, alpha(i) and the column effects beta(j) are both calculated the same way as Tukey's original model, but the sub-model is composed of parameters in a linear fashion (Simecek and Simeckova, 2012). The proposed test takes this sub-model and tests against the original Tukey model by a likelihood ratio test (Simecek and Simeckova, 2012).

The analysis of variance table for Mandel's test for non-additivity

Source of variation	Sum of squares	Degrees of freedom	Mean square (MS)	F
Mean	$ab\bar{Y}_{..}^2$	1		
Factor A	$SS_A = b\sum_i (\bar{Y}_{i.} - \bar{Y}_{..})^2$	$a-1$	$\frac{SS_A}{a-1}$	$\frac{MS_A}{MS_{Res}}$
Factor B	$SS_B = a\sum_j (\bar{Y}_{.j} - \bar{Y}_{..})^2$	$b-1$	$\frac{SS_B}{b-1}$	$\frac{MS_B}{MS_{Res}}$
Slopes	$SS_{slopes} = \sum_i (\hat{\phi}_i - 1)^2 \sum_j (\bar{Y}_{.j} - \bar{Y}_{..})^2$ where $\hat{\phi}_i = \frac{\sum_j Y_{ij}(\bar{Y}_{.j} - \bar{Y}_{..})}{\sum_j (\bar{Y}_{.j} - \bar{Y}_{..})^2}$	$a-1$	$\frac{SS_{slopes}}{a-1}$	$\frac{MS_{slopes}}{MS_{Res}}$
Residual	$SS_{Res} = \sum_i \sum_j \left[(Y_{ij} - \bar{Y}_{i.}) - \hat{\phi}_i(\bar{Y}_{.j} - \bar{Y}_{..})\right]^2$	$(a-1)(b-2)$	$\frac{SS_{Res}}{(a-1)(b-2)}$	
Total	$SS_T = \sum_i \sum_j Y_{ij}^2$	ab		

Figure 2. Analysis of Variance for Mandel's Test (Alin and Kurt, 2006)

2.4 *Johnson and Graybill Test*

The Johnson and Graybill test was created by Dallas E. Johnson and Franklin A. Graybill in 1972 where they proposed a new model that could test for interaction when no replication exists. Within their paper they discuss Tukey's model, but point out that Tukey's model did not have a particular type of interaction in terms of its model (Johnson and Graybill, 1972). Johnson and Graybill had proposed that the interaction may not always be a function of the treatment and block effects like Tukey had proposed and offered a alternative (Johnson and Graybill, 1972). The alternative model they had suggested was (Johnson and Graybill, 1972):

$$\left.\begin{array}{l} y_{ij} = \mu + \tau_i + \beta_j + \lambda\alpha_i\gamma_j + \epsilon_{ij} \\ \epsilon_{ij} \text{ distributed } NID(0, \sigma^2) \\ \Sigma\,\tau_i = \Sigma\,\beta_j = \Sigma\,\alpha_i = \Sigma\,\gamma_j = 0; \\ \Sigma\,\alpha_i^2 = \Sigma\,\gamma_j^2 = 1. \end{array}\right\} \begin{array}{l} i = 1, 2, \cdots, t \\ j = 1, 2, \cdots, b \end{array}$$

Using this model, a hypothesis test on Lambda = 0 which states there is no interaction would be the null hypothesis, if the test were to show there is a interaction there is a procedure in which a suitable estimate of the error variance should be obtained and the pattern of interaction must be determined (Johnson and Graybill, 1972).

2.5 *The Tussell Test*

The Tussell test was created by Fernando Tussell and purposed in his paper, "Testing for interaction in two-way ANOVA tables with no replication" in 1990. Tussell had proposed a new test that had involved restricting only a few rows or columns from a rectangular unreplicated two way ANOVA table and a manner reminiscent of a standard profile analysis on the parallelism of complete rows and columns would be tested for (Tussell, 1990). The model that Tussell had proposed was(Tussell, 1990):

$$y_{ij} = \beta + \beta_{i.} + \beta_{.j} + \epsilon_{ij}$$

Where i represents the rows and j represents the columns within the sample observations. The null hypothesis is the same as the other tests in that no interaction exists. Tussell's test works by trying to annihilate the B, Bi, and Bj of yij by performing successive linear operations on the rows or columns of Y (Tussell, 1990). Through linear operations, the test will end up with centered and independent vectors with a diagonal covariance matrix if there is no interaction present (Tussell, 1990).

2.6 *The LBI Test*

The locally best invariant test was proposed by Robert J. Boik, on testing for non-additivity on a environment which only one observation per cell exists (Boik, 1993). The main difference the locally best invariant test (LBI Test) differs from the rest of the tests are that the test does not call for a particular structural alternative and there does not need to be hypothesis which results in the test to be less subject to model specifications (Boik, 1993).

3. Non-Additivity Experiments

To test the effectiveness of each different test that have been discussed, different data sets were subject to these tests that exhibit a sign of interaction. The goal of this experiment was to find which tests would reject the additivity hypothesis on multiple data sets. For this experiment we used three different data sets, each data set will be further explained in terms of factors and the response variables that were examined. The data was analyzed within , R, and the package, "additivityTests" was used to test each method on the different data sets.

Table 1
Soil Experiment on the effectiveness of blast furnace slags

		Yield of Corn in Bushels Per Acre	
Treatment	Sandy Loam (I)	Sandy Clay Loam (II)	Loamy Sand (III)
No Treatment	11.1	32.6	63.3
Course Slag	15.3	40.8	65.0
Medium Slag	22.7	52.1	58.8
Agricultural Slag	23.8	52.8	61.4
Agricultural Limestone	25.6	63.1	41.1
Agricultural Slag + Minor Elements	31.2	59.6	78.1
Agricultural Limestone + Minor Elements	25.8	55.3	60.2

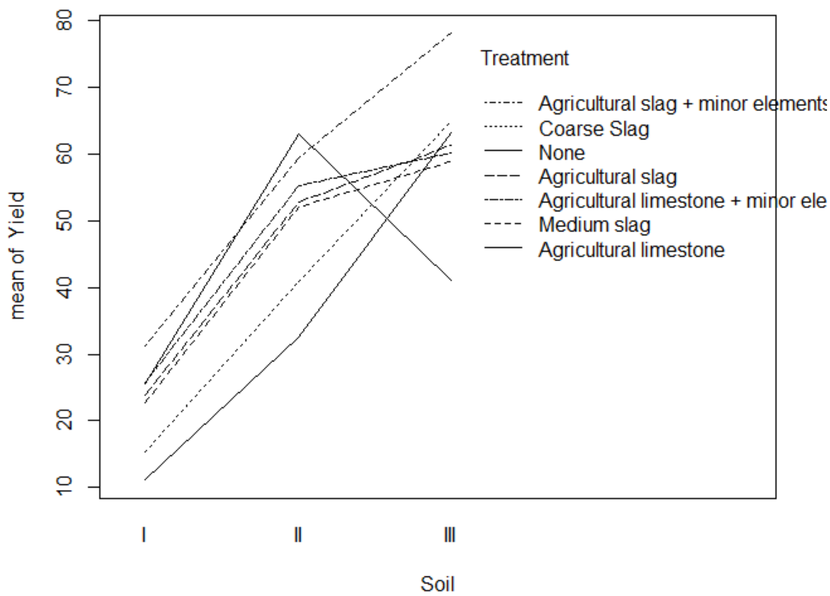

Figure 3. Interaction Plot of Treatment and Soil with the Yield as the Response Variable.

3.1 *The Soil Experiment*

The Soil experiment involved measuring agricultural liming materials in particular, blast furnace slags and determining their effectiveness. The factors associated with this data set were the three soil types which are sand loam (I), sand clay loam (II) and loamy sand (III) and the treatments applied on the soil. The 7 levels of the treatment factor consisted of:

- No Treatment
- Coarse Slag
- Medium Slag
- Agricultural Slag
- Agricultural Limestone
- Agricultural Slag + Minor Elements
- Agricultural Limestone + Minor Elements

At 4000 lbs per acre these treatments were applied to each different type of soil. The response that was to be measured for this experiment was the corn yield in bushels per acre. Table 1 exhibits the data used for the experiment.

This data set was considered to be in the middle in terms of size when looking at the structure of the matrix. A interaction is suggested within this data set between the two factors of treatment and soil type because of the results from a interaction plot of the soil type and the treatments which is exhibited by Figure 3.

Table 2
Results of Each Test for the Soil Experiment at the 5% Level

Test	Results
Tukey	No Rejection
Mandel	No Rejection
Modified Tukey	No Rejection
Johnson and Graybill	Additivity Rejected
Tusell	Additivity Rejected
LBI	Additivity Rejected

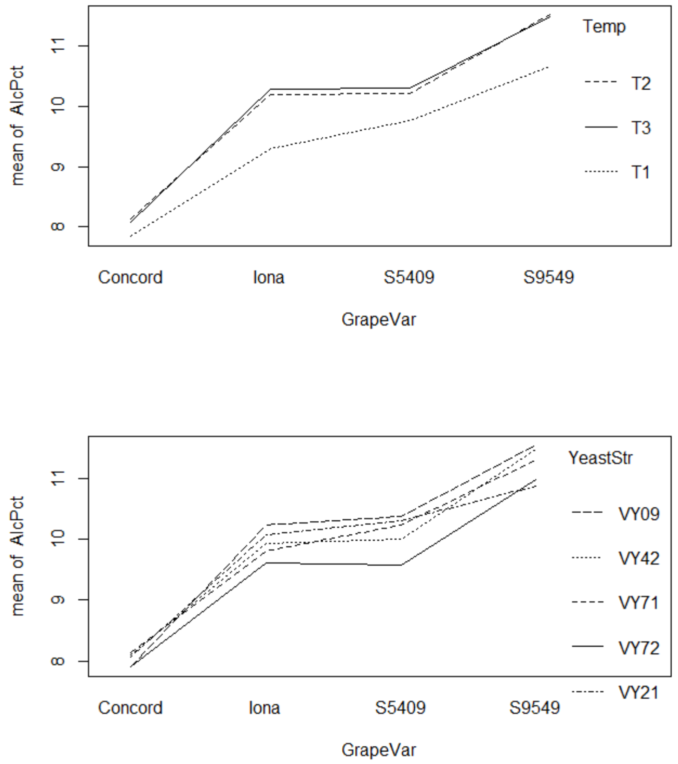

Figure 4. Interaction Plots for the Grape Variety where the Response is Alcohol Percentage.

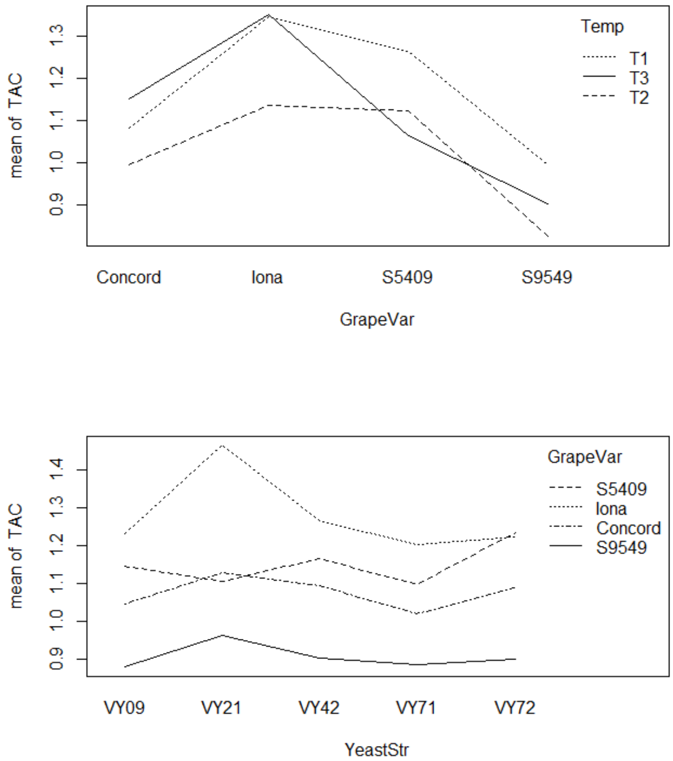

Figure 5. Interaction Plots for the Grape Variety where the Response is Total Alcohol Content.

The results from the tests shown in Table 2, can conclude that for the soil experiment data only the Johnson and Graybill, Tusell and LBI test were proven to be effective in detecting a interaction effect that exists between the two factors hence detecting non-additivity.

3.2 *The Wine Experiment*

The second set of data that was used was a experiment based around wine in which the main response variables that was examined was the alcohol percentage and the total alcohol content of the wine, it was developed by Agnes C. Kormendy (Kormendy, 1956). Unlike the Soil Experiment, there are two response variables therefor multiple stacks were created to test the effectiveness of each test, this was considered to be the largest data set out of the three. Not only did this data set involve two response variables, but there were also three factors involved with the experiment as well. The three factors and their respective levels consisted of:

- Grape Variety (4 Levels): Concord, Iona, S9549, S5409
- Yeast Strain (5 Levels): VY09, VY21, VY42, VY71, VY72
- Temperature of Fermentation (3 Levels): 80F, 50F, 80F for 24 hours then 50F thereafter

The values regarding this data set can be seen from Table 7, this data set was split into two groups where each group would represent one of the two response variables which are represented in Table 7 as AlcPct and TAC.

3.2.1 *Alcohol Percentage of Wine.*

Grape Variety and Alcohol Percentage. This is the first response variable that was examined in the Wine data set, it is represented as AlcPct within Table 7. A interaction plot was created around the variable Grape Variety to find if any interaction exists with the other two factors which is represented by Figure 4. There is a clear interaction with Grape Variety and Yeast Strain and a slight interaction the occurs with Grape Variety and Temperature. The stack for Grape Variety / Yeast Strain was a matrix consisting of 15 rows and 4 columns while the stack for Grape Variety

Table 3

Results for interaction around the Grape Variety for AlcPct at the 5% Level

Test	YeastStr	Temp
Tukey	Additivity Rejected	No Rejection
Mandel	No Rejection	No Rejection
Modified Tukey	Additivity Rejected	No Rejection
Johnson and Graybill	No Rejection	Additivity Rejected
Tusell	No Rejection	Additivity Rejected
LBI	No Rejection	Additivity Rejected

Table 4

Results for interaction around the Grape Variety for TAC at the 5% Level

Test	YeastStr	Temp
Tukey	No Rejection	No Rejection
Mandel	No Rejection	No Rejection
Modified Tukey	No Rejection	No Rejection
Johnson and Graybill	Additivity Rejected	No Rejection
Tusell	Additivity Rejected	No Rejection
LBI	Additivity Rejected	No Rejection

/ Temperature was a matrix consisting of 20 rows and 3 columns. The results for both stacks are shown in Table 3 where some tests were able to detect non-additivity while others could not.

3.2.2 *Total Alcohol Content of Wine.*

Grape Variety and Total Alcohol Content. This is the second response variable that was examined in the Wine data set, it is represented as TAC Within Table 7. Just like the first response variable, a interaction plot was created around the factor Grape Variety to find if any interaction exists with the other two factors. This interaction plot is represented by Figure 5, it is clear that there is a interaction that exists between the Grape Variety and the other two respective factors. The matrix formation when testing was exactly the same as the previous stacks for the first response variable, it was a 15x4 for Grape Variety / Yeast Strain and 20x3 for Grape Variety / Temperature. The results from Table 4 show that all Grape Variety and Temperature factors could not find any non-additivity while some tests were able to find non-additivity in the pairing of Grape Variety and Yeast Strain.

3.3 *The Poison Experiment*

The Poison Experiment was obtained from the OzDASL library, it involved finding the survival time on cows when they were exposed to different poisons and then treated with a cure. The original experiment had used blocks of cows in which four cows were applied to each block, for this experiment only one block was examined to see the effectiveness of the tests. This is considered the smallest data frame out of the three, which results in having the smallest structure in matrix form. The data for the Poison Experiment can be shown in Table 5, in which the Poison is a factor that contains 3 levels and Treatments is a factor that contains 4 levels. When this data is put into a stack, the matrix has 4 rows and 3 columns. A interaction plot was created to examine if any interaction does occur between the two factors. Figure 6 shows the interaction plot, we can conclude that there may be a sign of interaction around the third Poison, but it is not quite clear. Table 6 shows the results of each test for this data set, it appears that the Tukey, Mandel and Modified Tukey had picked up non-additivity while the other could not detect any non-additivity.

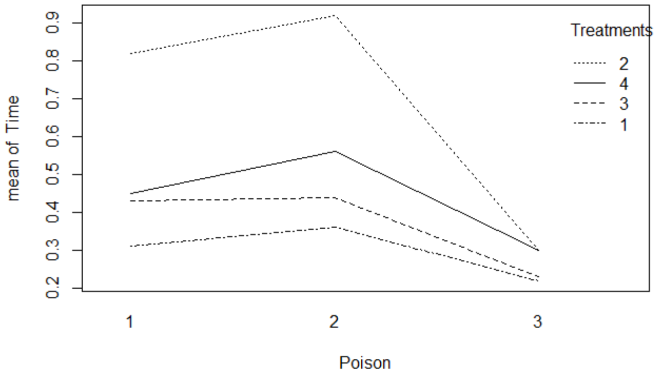

Figure 6. Interaction Plot of Poison and Treatments with the Survival Time as the Response Variable.

Table 5

Survival Time of Cows subject to different Poisons and Treatments

Poison	Treatment	Time
1	1	0.31
1	2	0.82
1	3	0.43
1	4	0.45
2	1	0.36
2	2	0.92
2	3	0.44
2	4	0.56
3	1	0.22
3	2	0.30
3	3	0.23
3	4	0.30

Table 6

Results of Each Test for the Poison Experiment at the 5% Level

Test	Results
Tukey	Additivity Rejected
Mandel	Additivity Rejected
Modified Tukey	Additivity Rejected
Johnson and Graybill	No Rejection
Tusell	No Rejection
LBI	No Rejection

Table 7

Alcohol Percentage and Total Alcohol Percentage of Wine

YeastStr	Temp	GrapeVar	AlcPct	TAC
VY09	T1	Concord	7.6	1.05
VY09	T1	Iona	10.1	1.19
VY09	T1	S9549	11.4	0.98
VY09	T1	S5409	10.3	1.17
VY09	T2	Concord	8.1	0.98
VY09	T2	Iona	10.2	1.14
VY09	T2	S9549	11.6	0.72
VY09	T2	S5409	10.4	1.1
VY09	T3	Concord	8	1.11
VY09	T3	Iona	10.4	1.36
VY09	T3	S9549	11.6	0.94
VY09	T3	S5409	10.4	1.17
VY21	T1	Concord	7.9	1.06
VY21	T1	Iona	9.8	1.83
VY21	T1	S9549	10	1
VY21	T1	S5409	9.9	1.19
VY21	T2	Concord	8.1	1.04
VY21	T2	Iona	10.1	1.14
VY21	T2	S9549	11.2	0.96
VY21	T2	S5409	10.6	1.12
VY21	T3	Concord	8.2	1.29
VY21	T3	Iona	10.3	1.42
VY21	T3	S9549	11.4	0.93
VY21	T3	S5409	10.4	1.01
VY42	T1	Concord	7.9	1.15
VY42	T1	Iona	9.6	1.22
VY42	T1	S9549	11.3	1.03
VY42	T1	S5409	9.7	1.35
VY42	T2	Concord	8.2	0.98
VY42	T2	Iona	10.3	1.14
VY42	T2	S9549	11.6	0.78
VY42	T2	S5409	10.1	1.13
VY42	T3	Concord	8.2	1.15
VY42	T3	Iona	9.9	1.44
VY42	T3	S9549	11.5	0.9
VY42	T3	S5409	10.2	1.02

Table 8

Alcohol Percentage and Total Alcohol Percentage of Wine - Continued

YeastStr	Temp	GrapeVar	AlcPct	TAC
VY71	T1	Concord	8	1.06
VY71	T1	Iona	8.9	1.26
VY71	T1	S9549	10.6	1.03
VY71	T1	S5409	10.2	1.14
VY71	T2	Concord	8.3	0.97
VY71	T2	Iona	10.1	1.15
VY71	T2	S9549	11.7	0.79
VY71	T2	S5409	10.3	1.12
VY71	T3	Concord	8.1	1.03
VY71	T3	Iona	10.4	1.2
VY71	T3	S9549	11.6	0.84
VY71	T3	S5409	10.2	1.03
VY72	T1	Concord	7.8	1.09
VY72	T1	Iona	8.1	1.23
VY72	T1	S9549	10	0.93
VY72	T1	S5409	8.7	1.47
VY72	T2	Concord	8	1
VY72	T2	Iona	10.3	1.11
VY72	T2	S9549	11.5	0.87
VY72	T2	S5409	9.7	1.14
VY72	T3	Concord	7.9	1.18
VY72	T3	Iona	10.4	1.33
VY72	T3	S9549	11.4	0.9
VY72	T3	S5409	10.3	1.09

4. Discussion

The results of the experiment had proven the effectiveness of each test when used against three different data sets that differ structurally when put into their respective stacks. The Poison Experiment was the smallest data set and the Soil Experiment was considered to be larger than the Poison Data set , but was smaller then the Wine Data set. Based on the results of each test, it is clear that the least effective test that have been discussed was the Mandel Test. The Mandel Test only detected non-additivity in one instance which was the Poison data set, this may have occurred because the Mandel test had taken a approach to generalize the method which may have proven to be less effective (Mandel, 1961). It was not surprising that the LBI Test and the Johnson and Graybill Test had similar results as it had already been proven that they both were on par on their effectiveness for detecting non-additivity(Boik, 1993). The LBI Test, Tussell Test and Johnson and Graybill Test all shared the same results for the experiments conducted meaning when one of the three detected non-additivity the other two would find detection as well. These three tests had the greatest success compared to the other tests, but failed to detect additivity in the Poison experiment which was considered to have the smallest structure out of the three. The Tukey and Modified Tukey did however manage to detect non-additivity in the Poison experiment where no other test seemed to be successful. The Tukey and modified Tukey also tested detection within the largest set of data which was the Wine experiment. We can conclude from this study that the Mandel test was the least effective, the Tukey and Modified Tukey were effective in specific situations and the tests that were the most effective across the different data sets were the LBI, Tussell and Johnson and Graybill Tests.

Acknowledgements

The author thanks Professor Umphrey for guidance throughout this study and insight into the topic at hand.

References

Alin, A. and Kurt, S. (2006). Testing non-additvity (interaction) in two-way anova tables with no replication. *Statistical Methods in Medical Research* **15,** 63–85.

Boik, R. J. (1993). Testing additvity in two-way classifications with no replications:the locally best invariant test. *The Journal of Applied Statistics* **20,** 41–55.

Johnson, D. E. and Graybill, F. A. (1972). An analysis of a two-way model with interaction and no replication. *The Journal of American Statistical Association* **67,** 862–868.

Kormendy, A. C. (1956). Low temperature fermentation of wines. *Ontario Agriculutral College, Guelph* **1,** 48.

Mandel, J. (1961). Non-additivity in two-way analysis of variance. *The Journal of American Statistical Association* **56,** 878–888.

Simecek, P. and Simeckova, M. (2012). Modification of tukey's additivity test. *Journal Statistical Planning and Inference* **143,** 197–201.

Tukey, J. W. (1949). One degree of freedom for non-additivity. *Biometrics* **5,** 232–242.

Tussell, F. (1990). Testing for interaction in two-way anova tables with no replication. *Computational Statistics and Data Analysis* **10,** 29–45.