**Example .    Heavy Metals in Sewage Sludge**

Sludge is the dried product remaining from processed sewage that contains nutrients beneficial to plant growth. It can be used for fertilizer on agricultural crops provided it does not contain toxic levels of certain elements such as heavy metals. Typically the levels of metals in sludge are assayed by growing plants in media containing different doses of the sludge.

**Research Hypothesis:** A soil scientist hypothesized the concentration of certain heavy metals in sludge would differ among the metropolitan areas from which the sludge was obtained. The variation could be a result of any number of causes including different industrial bases surrounding the areas. If this was true then recommendations for applications on crops would have to be preceded by knowledge about the source of the material. An assay was planned to determine whether there was significant variation in heavy metal concentrations among diverse metropolitan areas.

**Treatment Design:** The investigator obtained sewage sludge from treatment plants located in three different metropolitan areas. Barley plants were grown in a sand medium to which the sludge was added as a fertilizer. The sludge was added to the sand at three different rates: 0.5, 1.0, and 1.5 metric tons/acre. The factorial arrangement for the treatment design consisted of one qualitative factor, "city," with three levels and one quantitative factor, "rate," with three levels.

**Experiment Design:** Each of the nine treatment combinations were assigned to four replicate containers in a completely randomized design. The containers were arranged completely at random in a growth chamber. At a certain stage of growth the zinc content in parts per million was determined for the barley plants grown in each of the containers. The data are shown in Table 6.7

**Table 6.7**  Zinc content (ppm) of barley plants grown in media containing sludge at three rates from three metropolitan areas

| City and Rate (MT/a) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| A | | | B | | | C | | |
| 0.5 | 1.0 | 1.5 | 0.5 | 1.0 | 1.0 | 0.5 | 1.0 | 1.5 |
| 26.4 | 25.2 | 26.0 | 30.1 | 47.7 | 73.8 | 19.4 | 23.2 | 18.9 |
| 23.5 | 39.2 | 44.6 | 31.0 | 39.1 | 71.1 | 19.3 | 21.3 | 19.8 |
| 25.4 | 25.5 | 35.5 | 30.8 | 55.3 | 68.4 | 18.7 | 23.2 | 19.6 |
| 22.9 | 31.9 | 38.6 | 32.8 | 50.7 | 77.1 | 19.0 | 19.9 | 21.9 |

Source: J. Budzynski, Department of Soil and Water Science, University of Arizona.

## SAS code for Levene's Test and Weighted Least Squares

```
DM 'LOG; CLEAR; OUT; CLEAR;';

ODS GRAPHICS ON;
ODS PRINTER PDF file='C:\COURSES\ST541\WGTLS2.PDF';
OPTIONS NODATE NONUMBER;

DATA in;
DO treatmnt = 'A_0.5', 'A_1.0', 'A_1.5',
              'B_0.5', 'B_1.0', 'B_1.5',
              'C_0.5', 'C_1.0', 'C_1.5';
DO rep = 1 to 4;
   INPUT zincppm @@; OUTPUT;
END; END;
cards;
26.4 23.5 25.4 22.9   25.2 39.2 25.5 31.9   26.0 44.6 35.5 38.6
30.1 31.0 30.8 32.8   47.7 39.1 55.3 50.7   73.8 71.1 68.4 77.1
19.4 19.3 18.7 19.0   23.2 21.3 23.2 19.9   18.9 19.8 19.6 21.9
;
PROC GLM DATA=in PLOTS=(DIAGNOSTICS);
  CLASS treatmnt;
  MODEL zincppm = treatmnt / SS3;
  MEANS treatmnt / HOVTEST=LEVENE;
  LSMEANS treatmnt / ADJ=BON LINES;
TITLE 'LEVENE''S TEST EXAMPLE FOR TWO-FACTOR FACTORIAL';
RUN;

DATA in;
DO trtplant = 'A', 'B', 'C';
DO rate = .5, 1.0, 1.5;
DO rep = 1 to 4;
   INPUT zincppm @@; OUTPUT;
END; END; END;
cards;
26.4 23.5 25.4 22.9   25.2 39.2 25.5 31.9   26.0 44.6 35.5 38.6
30.1 31.0 30.8 32.8   47.7 39.1 55.3 50.7   73.8 71.1 68.4 77.1
19.4 19.3 18.7 19.0   23.2 21.3 23.2 19.9   18.9 19.8 19.6 21.9
;
PROC SORT DATA=IN; BY trtplant rate;
PROC MEANS DATA=in noprint; BY trtplant rate;
    VAR zincppm;
    OUTPUT OUT=wset VAR=var_ppm;

DATA wset; SET wset;
    wgt = 1/var_ppm;
DROP _FREQ_ _TYPE_;

PROC PRINT DATA=wset;
TITLE 'SAMPLE VARIANCES AND WEIGHTS FOR EACH TREATMENT';

DATA in; MERGE in wset; BY trtplant rate;

PROC GLM DATA=in PLOTS=(DIAGNOSTICS);
    WEIGHT wgt;
    CLASS trtplant rate ;
    MODEL zincppm = trtplant|rate / SS3;
    LSMEANS trtplant*rate / ADJUST=BON LINES;
TITLE 'WEIGHTED LEAST SQUARES EXAMPLE FOR TWO-FACTOR FACTORIAL';
RUN;
```

# LEVENE'S TEST EXAMPLE FOR TWO-FACTOR FACTORIAL

## The GLM Procedure

**Variable: zincppm**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| **Model** | 8 | 9475.515000 | 1184.439375 | 61.75 | <.0001 |
| **Error** | 27 | 517.865000 | 19.180185 | | |
| **Corrected Total** | 35 | 9993.380000 | | | |

| R-Square | Coeff Var | Root MSE | zincppm Mean |
|---|---|---|---|
| 0.948179 | 12.95716 | 4.379519 | 33.80000 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| **treatmnt** | 8 | 9475.515000 | 1184.439375 | 61.75 | <.0001 |

| Levene's Test for Homogeneity of zincppm Variance ANOVA of Squared Deviations from Group Means | | | | | |
|---|---|---|---|---|---|
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **treatmnt** | 8 | 10353.2 | 1294.2 | 2.26 | 0.0539 |
| **Error** | 27 | 15445.9 | 572.1 | | |



Distribution of zincppm

183
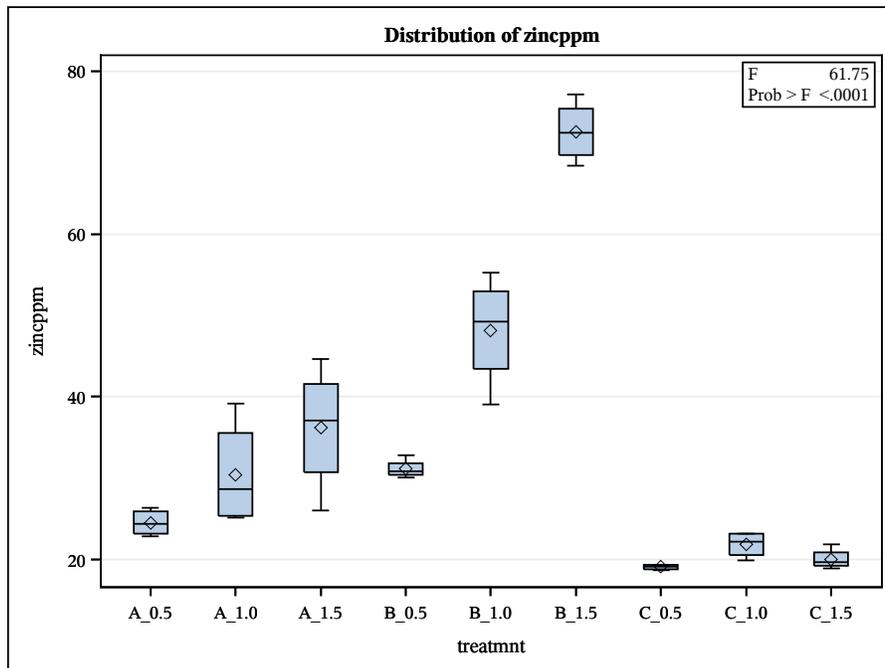
## Fit Diagnostics for zincppm
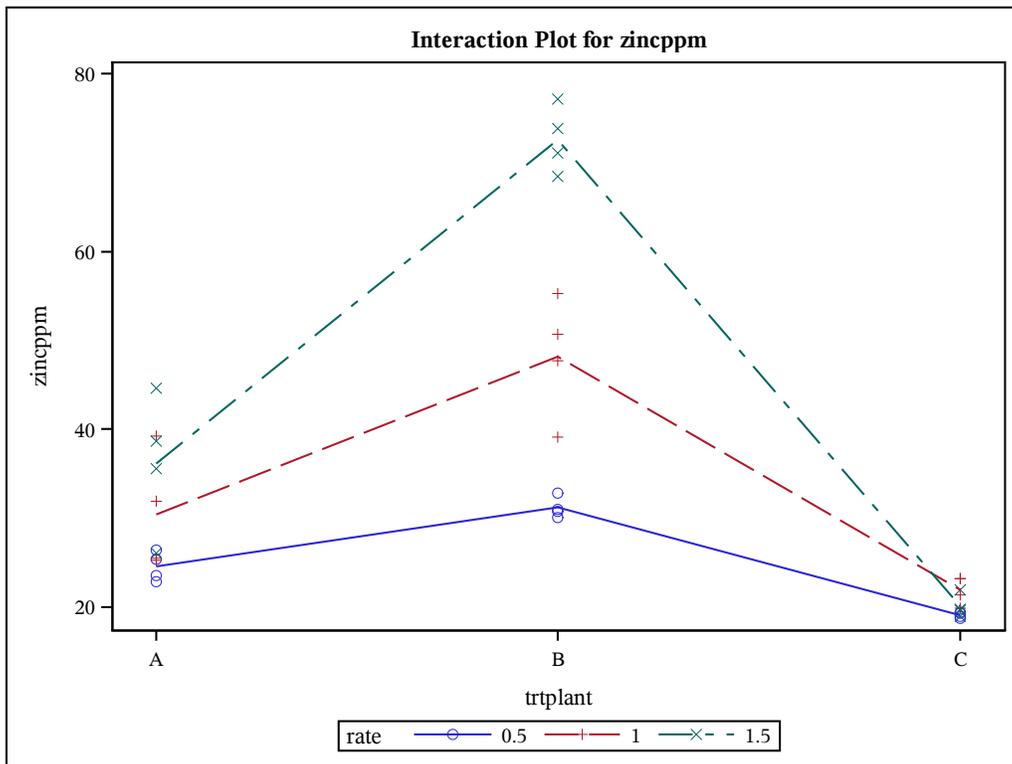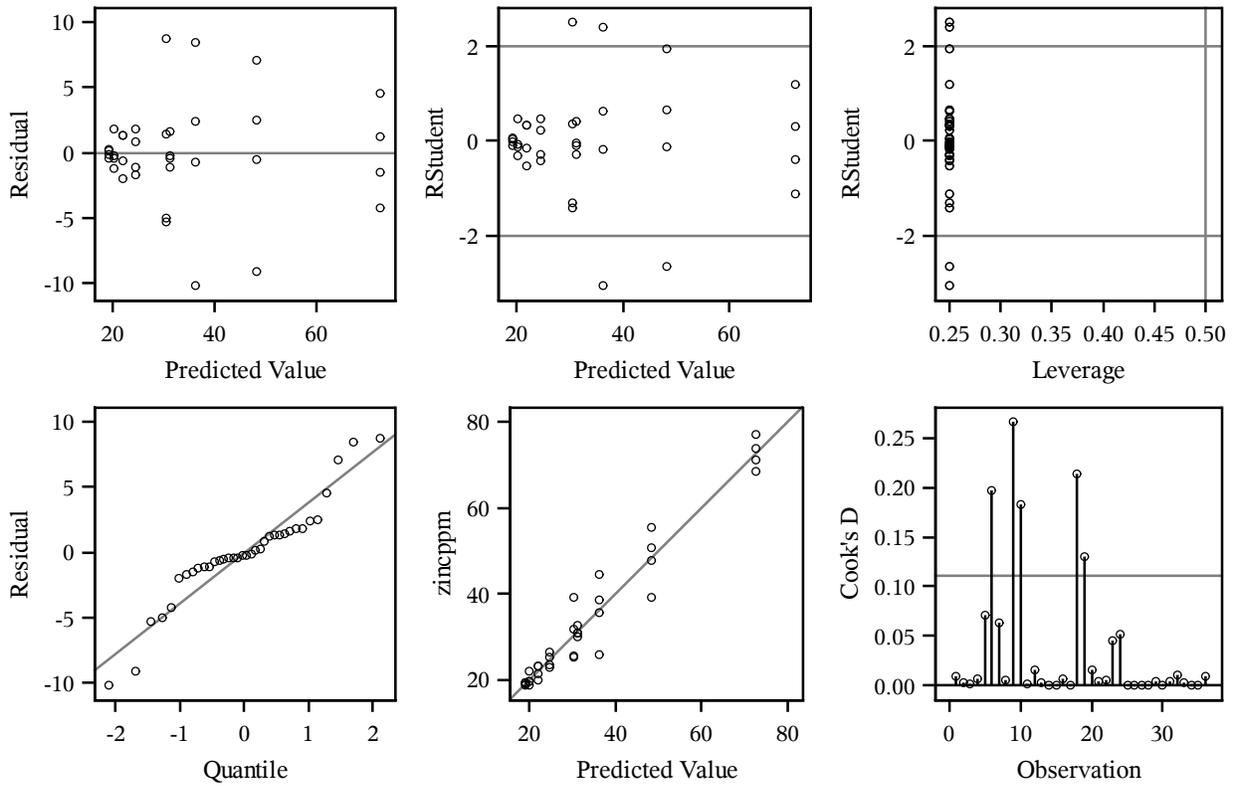


We will now look at a weighted least squares analysis. Based on the residual vs predicted values plot, it is evident that there is a homogeneity of variance problem across the 9 factorial treatment combinations. Later, we will address the question: 'Will a transformation fix the problem?" by finding a transformation using the empirical method.

# SAMPLE VARIANCES AND WEIGHTS FOR EACH TREATMENT

| Obs | trtplant | rate | var_ppm | wgt |
|-----|----------|------|---------|---------|
| 1 | A | 0.5 | 2.6567 | 0.3764 |
| 2 | A | 1.0 | 43.5767 | 0.0229 |
| 3 | A | 1.5 | 60.2825 | 0.0166 |
| 4 | B | 0.5 | 1.3225 | 0.7561 |
| 5 | B | 1.0 | 46.5733 | 0.0215 |
| 6 | B | 1.5 | 13.8600 | 0.0722 |
| 7 | C | 0.5 | 0.1000 | 10.0000 |
| 8 | C | 1.0 | 2.5800 | 0.3876 |
| 9 | C | 1.5 | 1.6700 | 0.5988 |

# WEIGHTED LEAST SQUARES EXAMPLE FOR TWO-FACTOR FACTORIAL

## The GLM Procedure

**Dependent Variable: zincppm**

**Weight: wgt**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|--------|----|----------------|-------------|---------|--------|
| Model | 8 | 1325.792420 | 165.724052 | 165.72 | <.0001 |
| Error | 27 | 27.000000 | 1.000000 | | |
| Corrected Total | 35 | 1352.792420 | | | |

| R-Square | Coeff Var | Root MSE | zincppm Mean |
|----------|-----------|----------|--------------|
| 0.980041 | 4.864267 | 1.00000 | 20.55808 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|--------|----|-------------|-------------|---------|--------|
| trtplant | 2 | 521.3323131 | 260.6661565 | 260.67 | <.0001 |
| rate | 2 | 167.8410304 | 83.9205152 | 83.92 | <.0001 |
| trtplant*rate | 4 | 398.7605331 | 99.6901333 | 99.69 | <.0001 |

| treatmnt | zincppm LSMEAN | LSMEAN Number |
|---|---|---|
| **A_0.5** | 24.5500000 | 1 |
| **A_1.0** | 30.4500000 | 2 |
| **A_1.5** | 36.1750000 | 3 |
| **B_0.5** | 31.1750000 | 4 |
| **B_1.0** | 48.2000000 | 5 |
| **B_1.5** | 72.6000000 | 6 |
| **C_0.5** | 19.1000000 | 7 |
| **C_1.0** | 21.9000000 | 8 |
| **C_1.5** | 20.0500000 | 9 |

Let's compare the results of Bonferroni's MCP from the unweighted and weighted least squares analyses.

| | | | **Least Squares Means for effect treatmnt** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | **Pr > \|t\| for H0: LSMean(i)=LSMean(j)** | | | | | | |
| | | | **Dependent Variable: zincppm** | | | | | | |
| **i/j** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
| **1** | | 1.0000 | 0.0304 | 1.0000 | <.0001 | <.0001 | 1.0000 | 1.0000 | 1.0000 |
| **2** | 1.0000 | | 1.0000 | 1.0000 | 0.0002 | <.0001 | 0.0384 | 0.3685 | 0.0845 |
| **3** | 0.0304 | 1.0000 | | 1.0000 | 0.0217 | <.0001 | 0.0003 | 0.0031 | 0.0006 |
| **4** | 1.0000 | 1.0000 | 1.0000 | | 0.0003 | <.0001 | 0.0208 | 0.2094 | 0.0463 |
| **5** | <.0001 | 0.0002 | 0.0217 | 0.0003 | | <.0001 | <.0001 | <.0001 | <.0001 |
| **6** | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | | <.0001 | <.0001 | <.0001 |
| **7** | 1.0000 | 0.0384 | 0.0003 | 0.0208 | <.0001 | <.0001 | | 1.0000 | 1.0000 |
| **8** | 1.0000 | 0.3685 | 0.0031 | 0.2094 | <.0001 | <.0001 | 1.0000 | | 1.0000 |
| **9** | 1.0000 | 0.0845 | 0.0006 | 0.0463 | <.0001 | <.0001 | 1.0000 | 1.0000 | |

| | | | **Least Squares Means for effect trtplant*rate** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | **Pr > \|t\| for H0: LSMean(i)=LSMean(j)** | | | | | | |
| | | | **Dependent Variable: zincppm** | | | | | | |
| **i/j** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
| **1** | | 1.0000 | 0.2450 | <.0001 | <.0001 | <.0001 | <.0001 | 1.0000 | 0.0067 |
| **2** | 1.0000 | | 1.0000 | 1.0000 | 0.0317 | <.0001 | 0.0695 | 0.6509 | 0.1648 |
| **3** | 0.2450 | 1.0000 | | 1.0000 | 0.9985 | <.0001 | 0.0056 | 0.0453 | 0.0123 |
| **4** | <.0001 | 1.0000 | 1.0000 | | 0.0014 | <.0001 | <.0001 | <.0001 | <.0001 |
| **5** | <.0001 | 0.0317 | 0.9985 | 0.0014 | | <.0001 | <.0001 | <.0001 | <.0001 |
| **6** | <.0001 | <.0001 | <.0001 | <.0001 | <.0001 | | <.0001 | <.0001 | <.0001 |
| **7** | <.0001 | 0.0695 | 0.0056 | <.0001 | <.0001 | <.0001 | | 0.0721 | 1.0000 |
| **8** | 1.0000 | 0.6509 | 0.0453 | <.0001 | <.0001 | <.0001 | 0.0721 | | 1.0000 |
| **9** | 0.0067 | 0.1648 | 0.0123 | <.0001 | <.0001 | <.0001 | 1.0000 | 1.0000 | |

| Bonferroni Comparison Lines for Least Squares Means of treatmnt | | | | | |
|---|---|---|---|---|---|
| **LS-means with the same letter are not significantly different.** | | | | | |
| | | | zincppm LSMEAN | treatmnt | LSMEAN Number |
| | A | | 72.600 | B_1.5 | 6 |
| | | | | | |
| | B | | 48.200 | B_1.0 | 5 |
| | | | | | |
| | C | | 36.175 | A_1.5 | 3 |
| | C | | | | |
| D | C | | 31.175 | B_0.5 | 4 |
| D | C | | | | |
| D | C | E | 30.450 | A_1.0 | 2 |
| D | | E | | | |
| D | F | E | 24.550 | A_0.5 | 1 |
| D | F | E | | | |
| D | F | E | 21.900 | C_1.0 | 8 |
| | F | E | | | |
| | F | E | 20.050 | C_1.5 | 9 |
| | F | | | | |
| | F | | 19.100 | C_0.5 | 7 |

| Bonferroni Comparison Lines for Least Squares Means of trtplant*rate | | | | | |
|---|---|---|---|---|---|
| **LS-means with the same letter are not significantly different.** | | | | | |
| | | zincppm LSMEAN | trtplant | rate | LSMEAN Number |
| | A | 72.600 | B | 1.5 | 6 |
| | | | | | |
| | B | 48.200 | B | 1 | 5 |
| | B | | | | |
| C | B | 36.175 | A | 1.5 | 3 |
| C | | | | | |
| C | | 31.175 | B | 0.5 | 4 |
| C | | | | | |
| C | D | 30.450 | A | 1 | 2 |
| C | D | | | | |
| C | D | 24.550 | A | 0.5 | 1 |
| | D | | | | |
| | D | 21.900 | C | 1 | 8 |
| | D | | | | |
| | D | 20.050 | C | 1.5 | 9 |
| | D | | | | |
| | D | 19.100 | C | 0.5 | 7 |

## SAS code for a transformed response based on the empirical method

```
DM 'LOG; CLEAR; OUT; CLEAR;';

ODS GRAPHICS ON;
ODS PRINTER PDF file='C:\COURSES\ST541\TRANS2.PDF';
OPTIONS NODATE NONUMBER;

DATA in;
DO trtplant = 'A', 'B', 'C';
DO rate = .5, 1.0, 1.5;
DO rep = 1 to 4;
   INPUT zincppm @@; OUTPUT;
END; END; END;
cards;
26.4 23.5 25.4 22.9   25.2 39.2 25.5 31.9   26.0 44.6 35.5 38.6
30.1 31.0 30.8 32.8   47.7 39.1 55.3 50.7   73.8 71.1 68.4 77.1
19.4 19.3 18.7 19.0   23.2 21.3 23.2 19.9   18.9 19.8 19.6 21.9
;
************************************************************;
*** Find the transformation using the empirical method ***;
************************************************************;

PROC SORT DATA=IN; BY trtplant rate;
PROC MEANS DATA=in noprint; BY trtplant rate;
    VAR zincppm;
    OUTPUT OUT=yset MEAN=mean STD=std;
DATA yset; SET yset;
    logstd =LOG(std);   logmean=LOG(mean);
PROC PRINT DATA=yset;
    VAR mean std logstd logmean;
TITLE 'EMPIRICAL SELECTION OF ALPHA';

PROC GLM DATA=yset;
    MODEL logstd=logmean / SS3 solution;
TITLE 'ANOVA TO FIND EMPIRICAL SELECTION OF ALPHA';

************************************;
*** ANOVA AFTER A TRANSFORMATION ***;
************************************;

DATA in; SET in;;  transppm = zincppm**(-.6);

PROC GLM DATA=in PLOTS=(DIAGNOSTICS);
    CLASS trtplant rate;
    MODEL transppm = trtplant|rate / SS3;
    MEANS trtplant rate;
    LSMEANS trtplant*rate / ADJUST=BON LINES;
TITLE 'TWO-FACTOR FACTORIAL WITH TRANSFORMATION';

RUN;
```

SAS output for transformed response based on empirical method

## EMPIRICAL SELECTION OF ALPHA

| Obs | mean | std | logstd | logmean |
|---|---|---|---|---|
| 1 | 24.550 | 1.62993 | 0.48854 | 3.20071 |
| 2 | 30.450 | 6.60126 | 1.88726 | 3.41609 |
| 3 | 36.175 | 7.76418 | 2.04952 | 3.58837 |
| 4 | 31.175 | 1.15000 | 0.13976 | 3.43962 |
| 5 | 48.200 | 6.82447 | 1.92051 | 3.87536 |
| 6 | 72.600 | 3.72290 | 1.31450 | 4.28496 |
| 7 | 19.100 | 0.31623 | -1.15129 | 2.94969 |
| 8 | 21.900 | 1.60624 | 0.47389 | 3.08649 |
| 9 | 20.050 | 1.29228 | 0.25641 | 2.99823 |

## ANOVA TO FIND EMPIRICAL SELECTION OF ALPHA

### The GLM Procedure

**Variable: logstd**

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 1 | 3.95994394 | 3.95994394 | 5.50 | 0.0515 |
| Error | 7 | 5.04249542 | 0.72035649 | | |
| Corrected Total | 8 | 9.00243936 | | | |

| R-Square | Coeff Var | Root MSE | logstd Mean |
|---|---|---|---|
| 0.439875 | 103.5171 | 0.848738 | 0.819901 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| logmean | 1 | 3.95994394 | 3.95994394 | 5.50 | 0.0515 |

| Parameter | Estimate | Standard Error | t Value | Pr > |t| |
|---|---|---|---|---|
| Intercept | -4.670813534 | 2.35887187 | -1.98 | 0.0882 |
| logmean | 1.602374102 | 0.68342858 | 2.34 | 0.0515 |

$\widehat{\alpha} \approx 1.6 \longrightarrow \widehat{\lambda} \approx (1 - 1.6) = -0.6.$

### The GLM Procedure

*Variable: transppm*

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| **Model** | 8 | 0.03150350 | 0.00393794 | 46.99 | <.0001 |
| **Error** | 27 | 0.00226292 | 0.00008381 | | |
| **Corrected Total** | 35 | 0.03376642 | | | |

| R-Square | Coeff Var | Root MSE | transppm Mean |
|---|---|---|---|
| 0.932983 | 6.915990 | 0.009155 | 0.132373 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| **trtplant** | 2 | 0.02438919 | 0.01219460 | 145.50 | <.0001 |
| **rate** | 2 | 0.00488593 | 0.00244297 | 29.15 | <.0001 |
| **trtplant*rate** | 4 | 0.00222838 | 0.00055709 | 6.65 | 0.0007 |



Fit Diagnostics for transppm

Interaction Plot for transppm

rate   —○— 0.5   —+— 1   —✕— 1.5
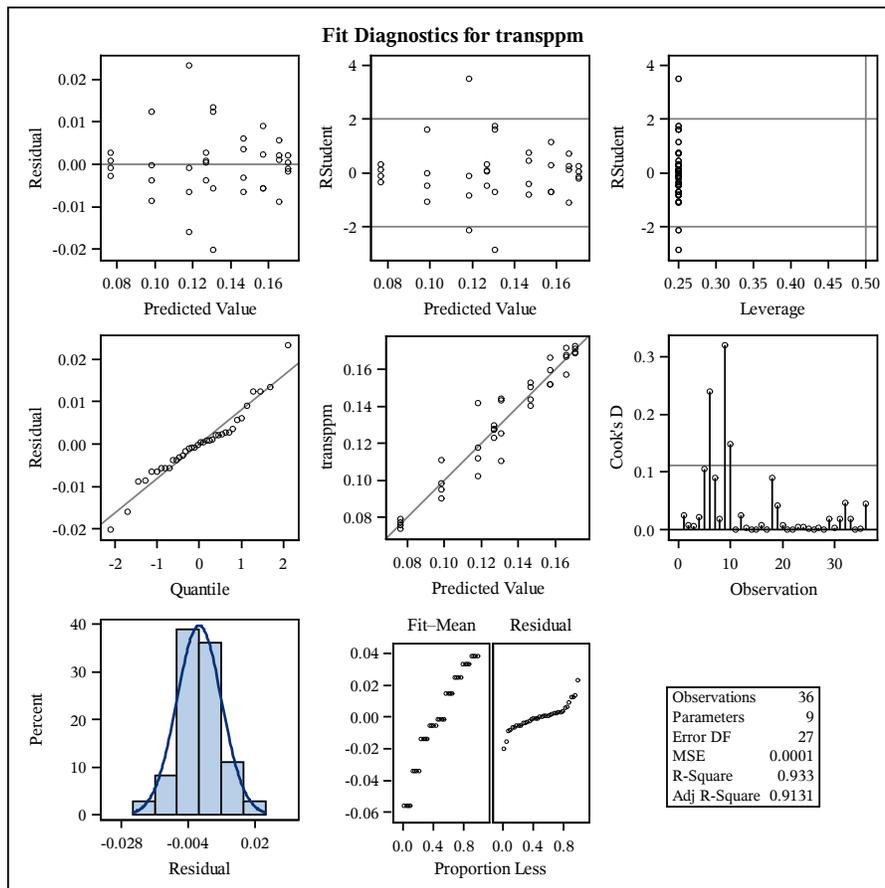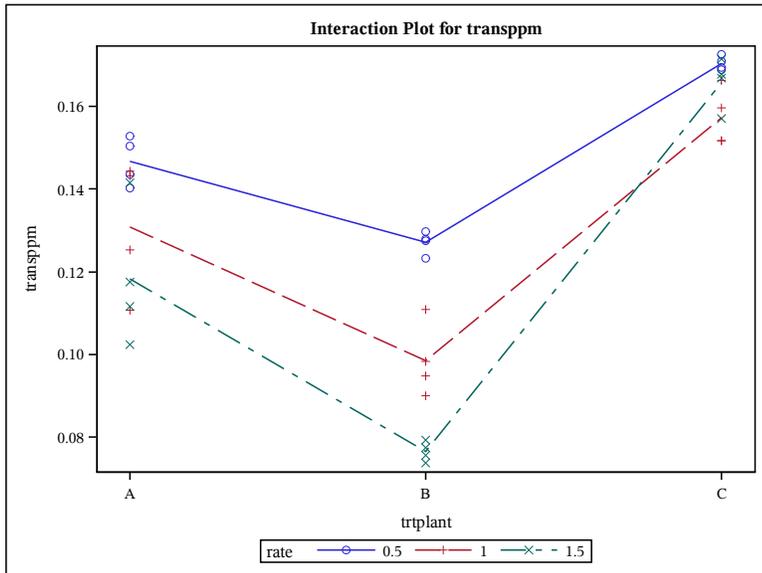
*Least Squares Means*
*Adjustment for Multiple Comparisons: Bonferroni*

| trtplant | rate | transppm LSMEAN | LSMEAN Number |
|---|---|---|---|
| A | 0.5 | 0.14677635 | 1 |
| A | 1 | 0.13085324 | 2 |
| A | 1.5 | 0.11828981 | 3 |
| B | 0.5 | 0.12703539 | 4 |
| B | 1 | 0.09852037 | 5 |
| B | 1.5 | 0.07653291 | 6 |
| C | 0.5 | 0.17038174 | 7 |
| C | 1 | 0.15725129 | 8 |
| C | 1.5 | 0.16571326 | 9 |

| Bonferroni Comparison Lines for Least Squares Means of trtplant*rate | | | | | | |
|---|---|---|---|---|---|---|
| **LS-means with the same letter are not significantly different.** | | | | | | |
| | | transppm LSMEAN | trtplant | rate | | LSMEAN Number |
| | A | 0.170381736 | C | 0.5 | | 7 |
| | A | | | | | |
| B | A | 0.165713255 | C | 1.5 | | 9 |
| B | A | | | | | |
| B | A | 0.157251293 | C | 1 | | 8 |
| B | | | | | | |
| B | C | 0.146776346 | A | 0.5 | | 1 |
| | C | | | | | |
| D | C | 0.130853238 | A | 1 | | 2 |
| D | C | | | | | |
| D | C | 0.127035386 | B | 0.5 | | 4 |

| | | | | | | |
|---|---|---|---|---|---|---|
| D | | | | | | |
| D | E | 0.118289808 | A | 1.5 | | 3 |
| | E | | | | | |
| F | E | 0.098520369 | B | 1 | | 5 |
| F | | | | | | |
| F | | 0.076532906 | B | 1.5 | | 6 |