Comparisons

one of the most common statistical tools

Robert Pitt, Ph.D., P.E., BCEE

Emeritus Cudworth Professor of Urban Water Systems

Department of Civil, Construction, and Environmental Engineering

University of Alabama, Tuscaloosa, AL USA

1

1

Selection of Statistical Tests Based on Probability Distribution and Other Characteristics

Comparing Paired Observations of Data

Parametric tests (data require normality and equal variance)
- Paired Student's *t*-test (more power than non-parametric tests but only if data requirements are met)

Non-parametric tests
- Sign test (no data distribution requirements, some missing data accommodated)
- Friedman's test (can accommodate a moderate number of "non-detectable" values, but no missing values are allowed
- Wilcoxon signed rank test (more power than sign test, but requires symmetrical data distributions)

2

2

Comparing Two Independent Groups of Data

Parametric tests (data require normality and equal variance)
- Independent Student's *t*-test (more power than non-parametric tests, but only if data distribution requirements are met)

Non-parametric tests
- Mann-Whitney rank sum test (probability distributions of the two data sets must be the same and have the same variances, but do not have to be symmetrical; a moderate number of "non-detectable" values can be accommodated)

3

3

Comparing many groups (use multiple comparison tests, such as the Bonferroni *t*-test, to identify which groups are different from the others if the group test results are significant)

Parametric tests (data require normality and equal variance)
- One-way ANOVA for single factor, but for >2 "locations" (if 2 "locations, use Student's *t*-test)
- Two-way ANOVA for two factors simultaneously at multiple "locations"
- Three-way ANOVA for three factors simultaneously at multiple "locations"
- One factor repeated measures ANOVA (same as paired *t* test, except that there can be multiple treatments on the same group)
- Two factor repeated measures ANOVA (can be multiple treatments on two groups)

4

4

Many Groups (cont.)

Non-parametric tests:

- Kurskal-Wallis ANOVA on ranks (use when samples are from non-normal populations or the samples do not have equal variances).

- Friedman repeated measures ANOVA on ranks (use when paired observations are available in many groups).

5

5

Many Groups (cont.)

Nominal observations of frequencies (used when counts are recorded in contingency tables)

- Chi-square (X^2) test (use if more than two groups or categories, or if the number of observations per cell in a 2X2 table are > 5).

- Fisher Exact test (use when the expected number of observations is <5 in any cell of a 2X2 table).

- McNamar's test (use for a "paired" contingency table, such as when the same individual or site is examined both before and after treatment)

6

6

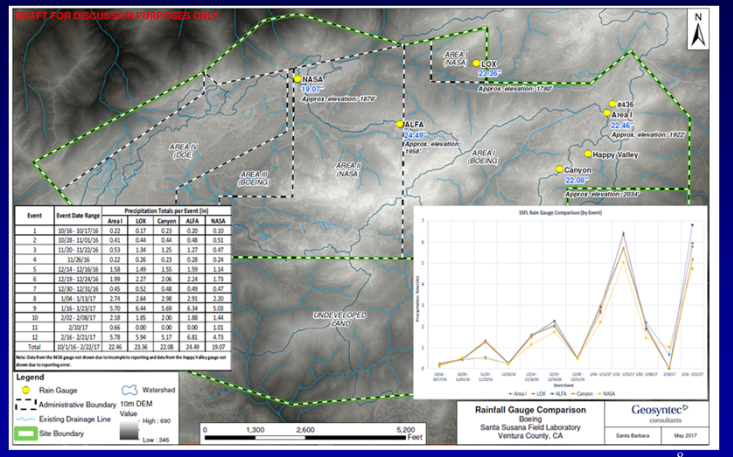
1-way ANOVA

- Is at least one member of a group significantly different from the other members?
- Complement analysis with group box-whisker plot
- This doesn't identify which one(s) is(are) different.
- If a significant member, should be able to recognize from box-whisker plot and with Bonferroni T-test (multiple pair-wise comparisons).

7

7

Variations of Five Rain Gages over Site



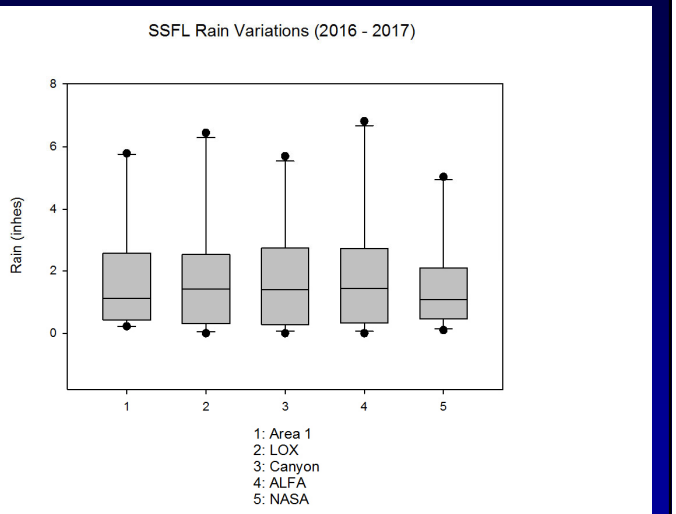
8

8

Rain depths recorded at each rain gage for 12 events:

	event #	area 1	LOX	Canyon	ALFA	NASA
10/16 - 10/17/16	1	0.22	0.17	0.23	0.2	0.1
10/28 - 11/01/16	2	0.41	0.44	0.44	0.48	0.51
11/20 - 11/22/16	3	0.53	1.34	1.25	1.27	0.47
11/26/2016	4	0.22	0.26	0.23	0.28	0.24
12/14 - 12/16/16	5	1.58	1.49	1.55	1.59	1.14
12/19 - 12/24/16	6	1.99	2.27	2.06	2.24	1.73
12/30 - 12/31/16	7	0.45	0.52	0.48	0.49	0.47
01/04 - 01/13/17	8	2.74	2.64	2.98	2.91	2.2
01/16 - 01/23/17	9	5.7	6.44	5.69	6.34	5.03
02/02 -02/08/17	10	2.18	1.85	2	1.88	1.44
2/10/2017	11	0.66	0	0	0	1.01
02/16 -02/22/17	12	5.78	5.94	5.17	6.81	4.73
10/01/16 - 02/22/17	total	22.46	23.36	22.08	24.49	19.07

9



10

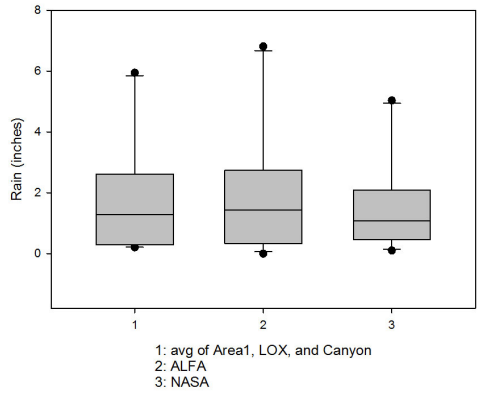
Friedman Repeated Measures Analysis of Variance on Ranks					
Group	N	Missing	Median	25%	75%
Area 1	12	0	1.12	0.42	2.6
LOX	12	0	1.415	0.305	2.547
Canyon	12	0	1.4	0.283	2.75
ALFA	12	0	1.43	0.33	2.742
NASA	12	0	1.075	0.47	2.082
Chi-square= 10.877 with 4 degrees of freedom. (P = 0.028)					
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = 0.028)					
To isolate the group or groups that differ from the others use a multiple comparison procedure.					
All Pairwise Multiple Comparison Procedures (Tukey Test):					
Comparison	Diff of Ranks	q	P		
ALFA vs NASA	24	4.382	0.017		
ALFA vs Area 1	14	2.556	0.369		
ALFA vs Canyon	10.5	1.917	0.656		
ALFA vs LOX	6.5	1.187	0.918		
LOX vs NASA	17.5	3.195	0.158		
LOX vs Area 1	7.5	1.369	0.869		
LOX vs Canyon	4	0.73	0.986		
Canyon vs NASA	13.5	2.465	0.408		
Canyon vs Area 1	3.5	0.639	0.991		
Area 1 vs NASA	10	1.826	0.697		

11

At least one location has statistically significant rainfall difference when examining all 12 events at 5 locations). ALFA and NASA are significantly different (P = 0.017 < 0.05) from each other (but not from other locations).

	Area 1	LOX	Canyon	ALFA	NASA
Area 1	X				
LOX	0.87	X			
Canyon	0.99	0.99	X		
ALFA	0.37	0.92	0.66	X	
NASA	0.7	0.16	0.41	0.017	X

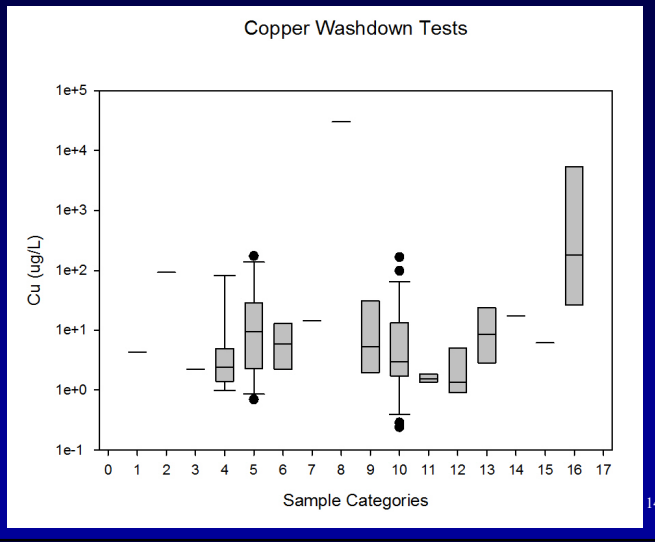
12



	Area 1, LOX, and Canyon combined	ALFA	NASA
Area 1, LOX, and Canyon combined	X		
ALFA	0.33	X	
NASA	0.33	0.012	X

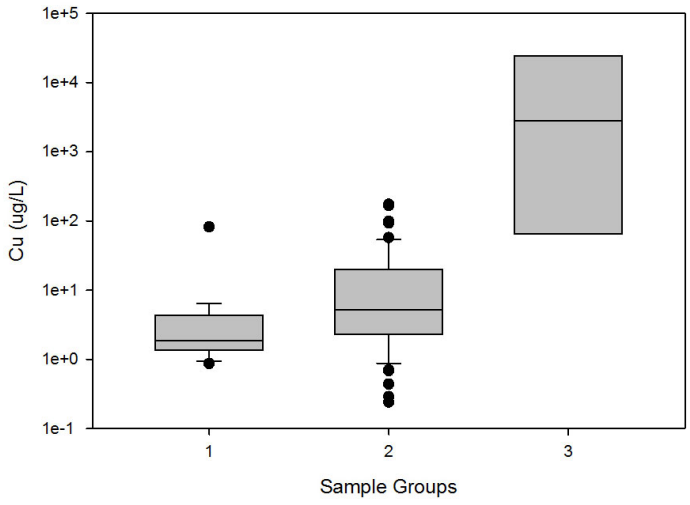
13

Copper Sources



14

14



15

Grouped category:	low	all others	high
Sample Category in Groups:	Al ramp brick wall concrete plaster roof	artificial turf galv bare galv painted galv coated metal bare metal painted rubber wood bare wood painted	barge hull wood treated
number	19	47	4
min	1	0	27
max	81	174	30,334
average	7	21	8,989
median	2	4	2,798
st dev	18	39	14,449
COV	2.7	1.8	1.6

16

16

Kruskal-Wallis One Way Analysis of Variance on Ranks (Cu concentrations)					
Normality Test (Shapiro-Wilk)	Failed	(P < 0.050)			
Group	N	Missing	Median	25%	75%
low	19	0	1.866	1.346	4.301
all others	56	0	5.25	2.293	19.969
high	4	0	2797.907	64.806	24104.41
H = 15.654 with 2 degrees of freedom. (P = <0.001)					
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0.001)					
To isolate the group or groups that differ from the others use a multiple comparison procedure.					
All Pairwise Multiple Comparison Procedures (Dunn's Method) :					
Comparison	Diff of Ranks	Q	P<0.05		
high vs low	47.605	3.771	Yes		
high vs all others	32.518	2.738	Yes		
all others vs low	15.087	2.476	Yes		

17

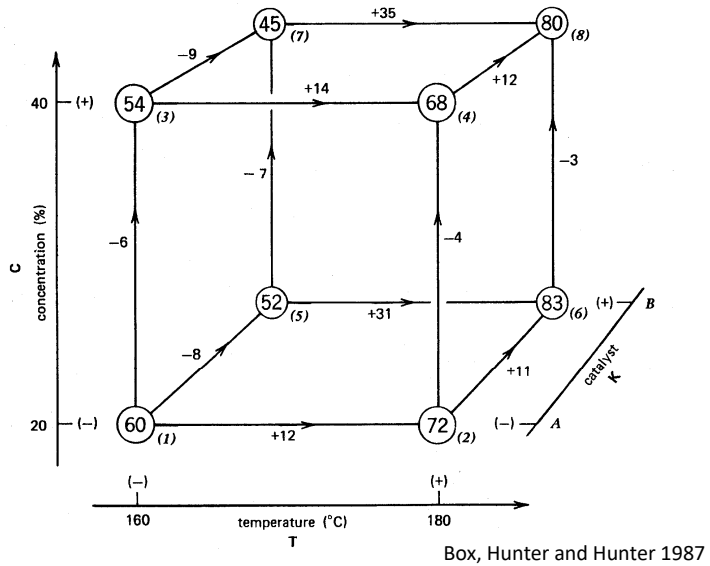
Factorial Analysis

a powerful experimental design and analysis tool

- A basic and powerful tool to identify significant factors and significant interacting factors.
- Use as the first step in sensitivity analysis and model building.
- Far superior to "holding all variables constant except for changing one variable at a time" classical approach (which doesn't consider interactions).
- Should be used in almost all experimental evaluations, especially valuable in controlled laboratory tests, and very useful to organize "environmental" test results.

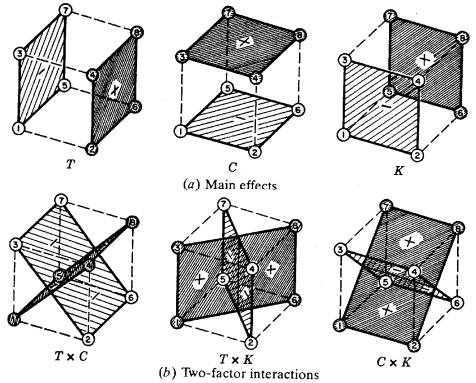
18

18



Box, Hunter and Hunter 1987

19



Box, Hunter and Hunter 1987

20

Ratio of Available Street Dirt Loadings to Total SS Washoff Quantity

the rain intensity and pavement texture were the only significant factors affecting availability of street dirt for washoff

$$I = 0.08 \pm 0.04$$
$$T = -0.08 \pm 0.05$$

$$\hat{Y} = 0.097 + 0.04(I) - 0.04(T)$$

I+T+ (high and rough) : $\hat{Y} = 0.10$
I+T- (high and smooth): $\hat{Y} = 0.18$
I-T+ (low and rough) : $\hat{Y} = 0.02$
I-T- (low and smooth) : $\hat{Y} = 0.10$

21

Sorting Field Data using Full 2^4 Factorial Design for Infiltration Rates

Case	Texture	Uniformity	Organic content	Compaction	Average Fc for test conditions (cm/hr)
1	+	+	+	+	9.1
2	+	+	+	-	20.9
3	+	+	-	+	5.2
4	+	+	-	-	5.8
5	+	-	+	+	110
6	+	-	+	-	282
7	+	-	-	+	1,000
8	+	-	-	-	1,030
9	-	+	+	+	6.7
10	-	+	+	-	46.4
11	-	+	-	+	2.8
12	-	+	-	-	15.8
13	-	-	+	+	7.1
14	-	-	+	-	41.9
15	-	-	-	+	5.5
16	-	-	-	-	8.1

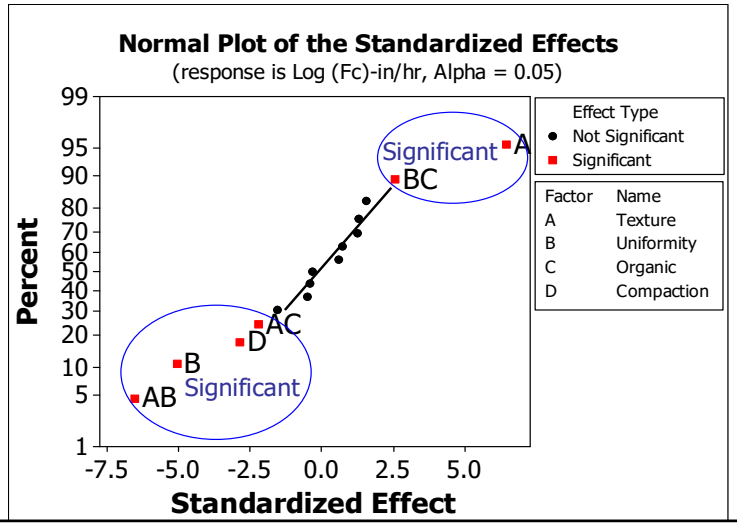
22

Estimated Effects for Fc

Estimated Effects for Fc (cm/hr)	
Texture	290.9
Uniformity	-296.4
Organic	-193.5
Compaction	-37.7
Texture*Uniformity	-298.7
Texture*Organic	-211.0
Texture*Compaction	-15.2
Uniformity*Organic	206.9
Uniformity*Compaction	21.4
Organic*Compaction	-26.9
Texture*Uniformity*Organic	207.1
Texture*Uniformity*Compaction	25.2
Texture*Organic*Compaction	-12.2
Uniformity*Organic*Compaction	17.4

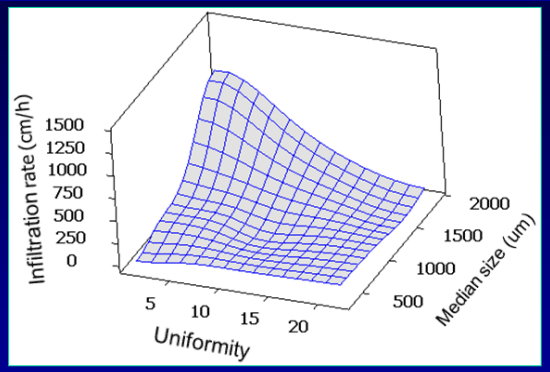
23

Probability Plot of the Calculate Effect Levels to Identify Outliers from Random Effects



24

Resulting Response Surface Model of Significant Effects Affecting Infiltration Rates



Surface plot for uniformity and texture vs. final infiltration rate for low organic content conditions. Higher infiltration rate values were observed for a mixture having low uniformity and higher median size values, as expected.

25

25

Example Pollutant Associations with different Rain Characteristics (rain depth, interevent period and intensity)

The ratio of the calculated effects to the grouped standard errors indicates if the effects are significant. This ratio should be about 3 or greater (for at least a 95% confidence). In this example, the only factor affecting cadmium concentrations is rain, with larger concentrations associated with larger rains.

Constituent	Significant factor (effect/grouped standard error)	High-level interactions (and ratio of effect to grouped standard error)
Cadmium	Rain category (3.06); dry period (2.6, marginal effect)	
Copper	Rain category (4.14); Dry period (3.01)	
Lead	None identified	
Mercury	Dry period (4.11)	Three-way interactions of all factors (5.56)
TCDD	None identified	

26