Using the Mixed Procedure to Analyze a Fractional Factorial Split-Plot Experiment

J.M. Lindenauer, M.S., C.Q.E.

Abstract

Introduction: The use of freon in airplane wing panel manufacturing is widespread. It is thought that cooling a drill bit with a blast of freon leads to a smoother hole surface. Since the price of freon doubles each year, and it's related to ozone layer depletion, there is a desire to reduce its use. The airplane manufacturer was interested in whether or not freon could be replaced by a lubricant. The goal of this experiment was to test whether the lubricant was at least the equivalent to freon for hole smoothness.

Purpose: This paper will describe the statistical modeling and analysis of this fractional factorial split-plot experiment using the SAS® Mixed procedure. Two methods of modeling the data will be examined. The first method is very similar to what many “automatic” statistical packages may do.

Method: Proc Mixed was used to analyze the experiment as a completely randomized and an incomplete block (split-plot) design structure. Practical techniques to analyze the experiment were demonstrated.

Summary: Although Proc Mixed was used to analyze the experiment, correctly specifying the statistical model is still of the utmost importance in getting good answers. This paper demonstrated that model misspecification, even with the latest software, can lead to problems.

Key Words: design of experiments, mixed models, split-plot, anova.

Introduction

The manufacturing research and development organization of an airplane manufacturer was interested in determining whether or not a lubricant could replace freon in the manufacturing of wing panels. The key quality characteristic used in determining whether or not freon could be replaced by the lubricant is hole smoothness. It has been shown in accelerated life testing that the smoother the surface of a drilled hole, the longer the fatigue life. It is thought that cooling a drill bit with a blast of freon leads to a smoother hole surface.

Hole smoothness is measured by a device called a profilometer. The instrument is similar to the needle and arm of a stereo turntable. The needle rides along the surface of the hole and measures smoothness (or roughness). This device has inherent variability that was previously discovered by measurement system analysis. It was decided that the average of six hole smoothness measurements would be used as the data point for each experimental run.

The specimens used were 18 inch by 6 inch aluminum coupons. These coupons were either upper wing panel (UWP) or lower wing panel (LWP) aluminum material. A wing panel is made by fastening a skin to a stringer. Each coupon was made by fastening a skin and stringer thickness piece together for both the UWP and LWP material. The Rockwell hardness of the coupons was tested to insure uniformity within each panel type. The coupons had holes drilled through them. To best emulate the actual manufacturing process, a Gemcor Driv-matic was used to drill the holes. This was one of the Gemcor's used in actual production. The drill was either sprayed with a blast of freon or lubricant. No fasteners were put into the drilled holes. The measure of interest is the hole smoothness of the skin portion of the coupon.

The engineers were interested in making this process as robust as possible. During the production process, a new drill will wear out. The machine operator will notice excessive "chatter" or "exit burrs" and stop the process to change drill bits. But before this occurs the quality of the drill is probably bad. It was decided to look at "good" quality and "bad" quality drills. An additional factor the engineers are interested in is the effect of the drill's diameter. A 5/16 inch and 3/8 inch diameter were examined.

Experimental Design

There were four, two-level factors in the experiment. The factors and their levels are:

MATLTYPE (material type) - Upper wing panel or lower wing panel
SPRAYTYP (type of chemical sprayed on the drill) - Freon or lubricant
DRILLQUA (drill quality) - Good or bad (based on a pre-tested index)
DRILLDIA (drill diameter) - 3/8 inch or 5/16 inch

The engineers believe that there was little chance of any three or four factor interactions. There was a slight possibility of two factor interactions. It was decided that a resolution IV design, 2^{4-1} , was acceptable. Since there may be some coupon to coupon variation and the measurement device was variable, the experiment was replicated three times (REP=3). The treatment structure was a fractional factorial arrangement with 8x3=24 runs. The design generator was,

I = MATLTYPE*SPRAYTYP*DRILLQUA*DRILLDIA

The particular aliasing structure for this design generator was,

DRILLDIA = MATLTYPE*SPRAYTYP*DRILLQUA
MATLTYPE*SPRAYTYP = DRILLQUA*DRILLDIA
MATLTYPE*DRILLQUA = SPRAYTYP*DRILLDIA
MATLTYPE*DRILLDIA = SPRAYTYP*DRILLQUA

Table 1: Aliasing structure of the 2^{4-1} experiment.

The large size experimental unit in this study is a coupon. The fact that a coupon must either be made out of UWP or LWP aluminum material restricts the randomization of the experiment. The coupon becomes the block or whole plot. The design structure is an incomplete block. An incomplete block design occurs when the number of treatment combinations is greater than the number of experimental units in a block. In this case we have four factors but can only apply three of the factors per experimental unit.

A coupon of either LWP or UWP received the treatment combinations of the other three factors four times per large size experimental unit. This was done three times for a total of six coupons (three LWP and three UWP).

Another potential restriction to the randomization is that the machine can only have the freon dispenser or lubricant dispenser attached to the drilling apparatus. There was a tube that is attached and removed, for each dispenser, after a treatment combination’s six holes are made. Since the six holes are averaged, this was not considered a restriction. Further analysis of the six holes as repeated measures may prove fruitful at a later date. There is also a smaller size experimental unit: the drilled hole. The hole is the entity to which the drill quality and drill diameter treatments are applied within each material type and spray type. This design structure will be analyzed as split-plot.

The desired response for this experiment is the hole smoothness index (SMOOTH) as measured by a profilometer. The lower the hole smoothness index the better.

The experiment was run as follows: A coupon of either UWP or LWP material was fixed to the Gemcor with a clamp. A drill of good or bad quality and large or small diameter was put in. The freon dispenser or lubricant dispenser was attached to the drilling apparatus. Six consecutive holes were drilled at random on the coupon. The coupon was then removed and cleaned off. Table 2 shows the randomized sequence of treatment combinations for the first set of runs.

Run #	MatlTyp	SprayTyp	DrillQua	DrillDiam
1	UWP	LUBE	BAD	SMALL
2	LWP	FREON	BAD	SMALL
3	LWP	FREON	GOOD	LARGE
4	UWP	LUBE	GOOD	LARGE
5	UWP	FREON	GOOD	SMALL
6	LWP	LUBE	BAD	LARGE
7	UWP	FREON	BAD	LARGE
8	LWP	LUBE	GOOD	SMALL

Table 2: Order of treatment combinations for the first set of runs of the designed experiment.

Figure 1 shows an example of a LWP coupon's drilling pattern. The coupon was replaced and a different treatment combination was set up. When all the treatment combinations were completed, the experiment was replicated two more times. The hole smoothness was measured by an engineer after the completion of the experiment.

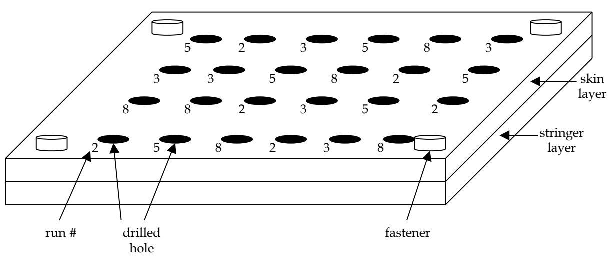

Figure 1: Example of the pattern of treatment combinations for a lower wing panel material coupon.

Analysis

Since the experiment was run with restrictions on randomization due to manufacturing and engineering constraints, the estimability of effects is limited. The 2^{4-1} design no longer has the main effects clear of any aliasing. Clearly this will have an effect on how the statistical model is developed and on the expected mean squares (EMS) of the analysis of variance. Table 3 shows the ANOVA table. Notice that whole plot factor MATLTYPE has it's own error term (labeled whole plot error or R*M in Table 3). This is very important for the correctness of this split-plot design and the associated F test. A quick examination of the expected mean squares for MATLTYPE and R*M (REP*MATLTYPE) shows that the F test must be MS_{MATL} / MS_{R*M} .

Source	**df**	**Expected Mean Squares**
REP	3-1 = 2	$8\sigma^2_R + 4\sigma^2_{RM} + \sigma^2_\varepsilon$
MATLTYPE	2-1 = 1	$12\Phi_M + 4\sigma^2_{RM} + \sigma^2_\varepsilon$
R*M (whole plot error)	(3-1)(2-1) = 2	$4\sigma^2_{RM} + \sigma^2_\varepsilon$
SPRAYTYP	(2-1) = 1	$12\Phi_S + \sigma^2_\varepsilon$
M*S	(2-1)(2-1) = 1	$6\Phi_{MS} + \sigma^2_\varepsilon$
DRILLQUA	(2-1) = 1	$12\Phi_{DQ} + \sigma^2_\varepsilon$
DRILLDIA	(2-1) = 1	$12\Phi_{DD} + \sigma^2_\varepsilon$
S* DQ	(2-1)(2-1) = 1	$6\Phi_{SDQ} + \sigma^2_\varepsilon$
S* DD	(2-1)(2-1) = 1	$6\Phi_{SDD} + \sigma^2_\varepsilon$
R*S+R*DQ +R*DD +R*M*S +R*S*DQ +R*S*DD (sub-plot error)	(3-1)(2-1) +(3-1)(2-1) +(3-1)(2-1) +(3-1)(2-1) +(3-1)(2-1)(2-1) +(3-1)(2-1)(2-1) = 12	σ^2_ε

Table 3: ANOVA table for this fractional factorial split-plot experiment with the associated EMS.

The problem that arose in the original analysis of this experiment was that an engineer decided to use an statistical package that was "easier" to use. The "automatic" analysis assumed that the experiment was run as a completely randomized design. The results were not to implement the use of the lubricant because of the UWP and LWP differences, even though SPRAYTYP showed that lubricant performed better overall. They believed that the material differences warranted further investigation.

The assumption of a completely randomized design was not the case here, as is the case with most industrial experiments. Clearly there is a restriction on randomization as outlined in the previous section. Fortunately the statistical analysis was redone correctly using SAS.

The statistical model of the completely randomized design for this experiment is displayed below.

$$Y_{ijkmn} = \mu + M_j + S_k + MS_{jk} + DQ_m + DD_n + SDQ_{km} + SDD_{kn} + \varepsilon_{i(jkmn)}$$

The following SAS program using the Mixed procedure best represents what an "automatic" analysis does for the above statistical model.

```
***PROGRAM #1;
proc mixed data=f.freondoe;
  class REP MATLTYPE SPRAYTYP DRILLQUA
        DRILLDIA;
  model SMOOTH = REP MATLTYPE SPRAYTYP
                 DRILLQUA DRILLDIA
                 SPRAYTYP*MATLTYPE
                 SPRAYTYP*DRILLQUA
                 SPRAYTYP*DRILLDIA;
run;
```

The output of this SAS program (PROGRAM #1) follows in Table 4.

```
Edited SAS Output from first model without
using the split-plot
```

Source	NDF	DDF	Type III F	Pr > F
REP	1	15	0.08	0.7774
MATLTYPE	**1**	**15**	**8.31**	**0.0114**
SPRAYTYP	1	15	7.35	0.0161
MATLTYP*SPRAYTYP	1	15	2.82	0.1141
DRILLQUA	1	15	1.26	0.2795
DRILLDIA	1	15	8.16	0.0120
SPRAYTYP*DRILLQU	1	15	1.48	0.2433
SPRAYTYP*DRILLDI	1	15	19.00	0.0006

Table 4: Results of the above SAS program (Program #1). Notice that the factor MATLTYPE is statistically significant at the α=0.05 level, but the MATLTYPE*SPRAYTYP two-factor interaction is not.

An interesting plot was the REP*MATLTYPE displayed in Figure2. Notice that the lower wing panel's second replication (REP=2) is much higher for the LSMEAN of SMOOTH. This difference could be related to the measurement error or to coupon to coupon variation. It is experimental error, yet without this interaction used as an error term, it looked like there was a significant MATLTYPE main effect. The engineers said that this difference is not out of the ordinary for coupons of the same material type and Rockwell Hardness. Therefore the data associated with these runs was not removed.

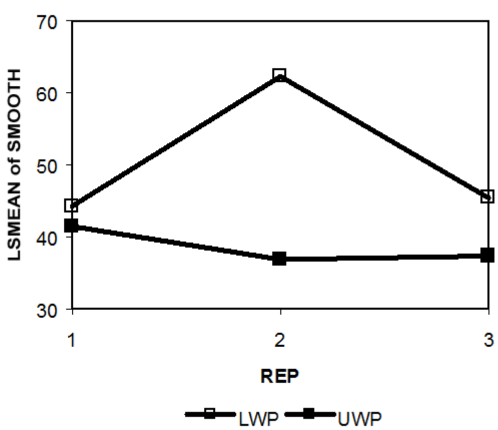

Figure 2: Plot of the REP*MATLTYPE interaction term.

There is a way to use the REP*MATLTYPE interaction as a random effect error term for the MATLTYPE main effect: The incomplete block design structure or split-plot.

The statistical model for the incomplete block (split-plot) for this experiment is displayed below.

$$Y_{ijkmn} = \mu + R_i + M_j + \delta_{i(j)} + S_k + MS_{jk} + DQ_m + DD_n$$
$$+ SDQ_{km} + SDD_{kn} + \varepsilon_{i(jkmn)}$$

The term $\delta_{i(j)}$ has been added to the model. This term is the whole plot error term ($\varepsilon_{i(jkmn)}$ is the sub-plot error term). The following SAS program takes into account the restriction on randomization and the EMS in Table 3, along with the above statistical model, to perform the correct analysis on this experiment's data. The lsmeans statement tests the marginal means of the treatments, including the interactions.

```
***PROGRAM #2;
proc mixed data=f.freondoe;
   class REP MATLTYPE SPRAYTYP DRILLQUA
         DRILLDIA;
   model smooth = REP MATLTYPE SPRAYTYP
                  DRILLQUA DRILLDIA
                  SPRAYTYP*MATLTYPE
                  SPRAYTYP*DRILLQUA
                  SPRAYTYP*DRILLDIA;
   random REP*MATLTYPE;
   lsmeans REP MATLTYPE MATLTYPE*SPRAYTYP
           / pdiff;
run;
```

The whole plot error term is REP*MATLTYPE and is specified in the random statement. What this will do is to divide the MATLTYPE by the appropriate error term (REP*MATLTYPE) as defined in the EMS of Table 3. The output of this SAS program (PROGRAM #2) follows in Table 5.

```
Edited SAS Output using the split-plot
```

Source	NDF	DDF	Type III F	Pr > F
REP	2	2	0.54	0.6474
MATLTYPE	**1**	**2**	**3.14**	**0.2184**
SPRAYTYP	1	12	12.77	0.0038
MATLTYP*SPRAYTYP	1	12	4.89	0.0471
DRILLQUA	1	12	2.19	0.1649
DRILLDIA	1	12	14.17	0.0027
SPRAYTYP*DRILLQU	1	12	2.56	0.1354
SPRAYTYP*DRILLDI	1	12	33.01	0.0001

Table 5: Results of the above SAS program (Program #2). Notice that the factor MATLTYPE is not significant, but the MATLTYPE*SPRAYTYP two-factor interaction is.

The first difference in the statistical analysis is the factor MATLTYPE is significant for the first result and not significant in the second result. Secondly, the two factor interaction MATLTYPE*SPRAYTYP is not significant in the first result but is significant in the second result. It should be noted here that the MATLTYPE*SPRAYTYP is confounded with the DRILLQUA*DRILLDIA interaction (see aliasing structure in Table 1). The engineers believe that the DRILLQUA*DRILLDIA interaction is improbable.

The lsmeans statement in PROGRAM #2 is also very important. First, most "automatic" analysis programs don't calculate the marginal means correctly or don't calculate them at all. Let's examine the most interesting result of the experiment: The MATLTYPE main effect and MATLTYPE*SPRAYTYP two-factor interaction effect. Table 6 shows that the UWP with the lubricant applied to the drill bit is driving the interaction. This interaction results in significantly smoother holes. In fact, it accounts for the large difference in the UWP and LWP levels of the MATLTYPE factor ([47.78+29.36]/2 = 38.57). Table A-1 in the Appendix shows the significance of the t-tests of each of these comparisons.

```
Edited SAS Output for LSMEANS of Hole
smoothness
```

Effect	MATLTYPE	SPRAYTYP	LSMEANS
MATLTYPE	lwp		50.67
MATLTYPE	uwp		38.57
MATLTYPE*SPRAYTYP	lwp	freon	52.83
MATLTYPE*SPRAYTYP	lwp	lube	48.50
MATLTYPE*SPRAYTYP	uwp	freon	47.78
MATLTYPE*SPRAYTYP	uwp	lube	29.36

Table 6: The lsmeans values for the marginal means of MATLTYPE and the MATLTYPE*SPRAYTYP two-factor interaction.

Figure 3 displays the plot of the MATLTYPE*SPRAYTYP interaction. The plot shows lubricant does an equal to or better job for both the UWP and LWP materials.

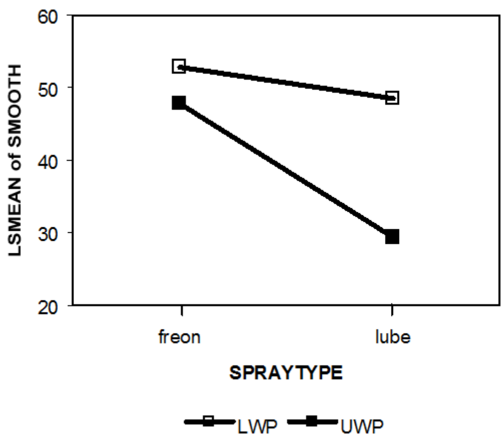

Figure 3: Plot of the MATLTYPE*SPRAYTYP interaction.

From the results above it can be seen that the lubricant can replace freon in the production of these stack sizes of wing panels. In fact, the lubricant was put into the production process for these stack sizes on the Gemcors. The stack size is the thickness of the skin plus stringer. This stack size was considered small to medium. Further analysis was to be performed on the large stack sizes.

Summary

Although Proc Mixed was used to analyze the experiment, correctly specifying the statistical model is still of the utmost importance in getting good answers. This paper demonstrated that a misspecified model, even with the latest software, can lead to problems. Most statistical packages that claim to be easy to use may cause problematic answers simply because the analysts is led to quick answers. Caution is advised.

Clearly, in this case, the inappropriate model would have had negative consequences. This could have led to erroneous decisions regarding the wing panel material and the spray type. The split-plot analysis technique is applicable in many instances of industrial experimentation where restrictions on randomization occurs. Some say that 95% of all experiments outside the lab have randomization restrictions, or blocking, and should be modeled as a split-plot.

Appendix

Edited Output for LSMEANS of Hole smoothness

Effect	MATL	SPRAY	MATL	SPRAY	Diff	Pr>|t|
MATL	lwp		uwp		12.10	0.2184
M*S	lwp	freon	lwp	lube	4.33	0.3548
M*S	lwp	freon	uwp	freon	5.06	0.5149
M*S	lwp	freon	uwp	lube	23.47	0.0089
M*S	lwp	lube	uwp	freon	0.72	0.9252
M*S	lwp	lube	uwp	lube	19.14	0.0259
M*S	uwp	freon	uwp	lube	18.42	0.0015

Table A-1: The lsmeans differences of the marginal means, and associated significance tests, are shown for MATLTYPE and the MATLTYPE*SPRAYTYP two-factor interaction.

References

Box, George E.P, Hunter, William G. and Hunter, J. Stuart, *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*, John Wiley and Sons, Inc., NY, 1978.

Box, George E. P. and Jones, Stephen P., “Split-Plot Designs for Robust Product Experimentation”, *Journal of Applied Statistics, Volume 19*, 1992.

Cochran, William C. and Cox, Gertrude M., *Experimental Designs, Second Edition*, John Wiley and Sons, Inc., NY, 1957.

Hicks, Charles R., *Fundamental Concepts in the Design of Experiments, Fourth Edition*, Oxford University Press, Inc., NY, 1993.

Littell, Ramon C., "Analysis of Unbalanced Mixed Model Data", *Proceedings of the Twenty-First Annual SAS Users Group International Conference*, Cary, NC: SAS Institute Inc., 1996. 1688pp.

Littell, Ramon C., Milliken, George A., Stroup, Walter W., and Wolfinger, Russell D., *SAS® System for Mixed Models*, Cary, NC: SAS Institute Inc., 1996. 633pp.

Lucas, James M. and Hazel, Malcolm C., “Running Experiments with Multiple Error Terms: How an Experiment Is Run Is Important”, *ASQC’s 51st Annual Quality Congress Proceedings*, ASQC, Milwaukee, WI, 1997.

Lucas, James M. and Ju, Huey L., “Split Plotting and Randomization in Industrial Experiments”,1992 - *ASQC Quality Congress Transaction- Nashville*, pg. 374-382, ASQC, Milwaukee, WI, 1992.

Milliken, George A. and Johnson, Dallas E., *Analysis of Messy Data, Volume 1: Designed Experiments*, Van Nostrand Reinhold, NY, 1984.

Montgomery, Douglas C., *Design and Analysis of Experiments, Fourth Edition*, John Wiley and Sons, Inc., NY, 1997.

SAS Institute Inc., *SAS/STAT® Software: Changes and Enhancements through Release 6.12,* Cary, NC: SAS Institute Inc., 1997. 1167pp.

Acknowledgments

The author would like to acknowledge Dick Kleinknecht, Ph.D., Senior Statistician, Weyerhaeuser Company and Robert Mee, Ph.D., Professor, University of Tennessee, for reviewing and editing the paper.

Contact Information

Jon M. Lindenauer, Jr.
Weyerhaeuser
Mailstop WTC 1B20
P.O. Box 2999
Tacoma, WA 98477-2999
e-mail: lindenj@wdni.com