6 Designs with Split Plots

- Many factorial experimental designs are incorrectly analyzed because the assumption of complete randomization is not true. Many factorial experiments have one or more restrictions on randomization. The major problem is the lack of recognition of these restrictions on randomization by the experimenter.
- Two common factorial designs that contain multiple restrictions on randomization are the **split-plot design** and the **split-split-plot design**.

6.1 Split-Plot Designs

- In a split-plot design, the experimenter is interested in studying the effects of two fixed factors (including the two-factor interaction).
- Let A and B be the two factors of interest with a levels for factor A and b levels for factor B.
- The most common split-plot design has the following structure:
  1. Randomize the a levels of factor A.
     - Assume $a = 3$ and the levels of A are A_1, A_2, and A_3. The 3 levels are randomized, yielding A_2, A_3, A_1.
  2. Randomly assign the levels of factor B to the experimental units within each randomized level of factor A.
     - Assume $B = 4$ and the levels of B are B_1, B_2, B_3, and B_4. The 4 levels are randomized within each of the 3 randomized levels of A.

Randomization of A	Randomization of B			
A_2	B_2	B_4	B_3	B_1
A_3	B_4	B_1	B_2	B_3
A_1	B_2	B_3	B_4	B_1

  3. Collect observations using the randomization structure from Steps 1 and 2.
     - Using the randomization table shown in Step 2, data would be collected in the following treatment order:

Observations 1 to 4	(A_2, B_2),	(A_2, B_4),	(A_2, B_3),	(A_2, B_1),
Observations 5 to 8	(A_3, B_4),	(A_3, B_1),	(A_3, B_2),	(A_3, B_3),
Observations 9 to 12	(A_1, B_2),	(A_1, B_3),	(A_1, B_4),	(A_1, B_1)

  4. Replicate Steps 1, 2, and 3 a total of n times. That is, for each replicate use new randomizations of the a levels of A in Step 1 and the b levels in Step 2.
     - Suppose $n = 2$. The randomizations for the second replicate produced the following treatment order for data collection:

Observations 13 to 16	(A_1, B_1),	(A_1, B_4),	(A_1, B_2),	(A_1, B_3),
Observations 17 to 20	(A_3, B_4),	(A_3, B_2),	(A_3, B_1),	(A_3, B_3),
Observations 21 to 24	(A_2, B_3),	(A_2, B_2),	(A_2, B_4),	(A_2, B_1)

- The ANOVA table has the following structure. The double horizontal lines indicate where the two restrictions on randomization occur. R represents replicates.

ANOVA for a Two-Factor Split-Plot Design

Source of Variation	Sum of Squares	d.f.	Mean Square	F Ratio
R	SS_R	$n-1$	$MS_R = \frac{SS_R}{n-1}$	—
A	SS_A	$a-1$	$MS_A = \frac{SS_A}{a-1}$	$F_A = \frac{MS_A}{MS_{AR}}$
AR	SS_{AR}	$(a-1)(n-1)$	$MS_{AR} = \frac{SS_{AR}}{(a-1)(n-1)}$	—
B	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$F_B = \frac{MS_B}{MS_{BR}}$
BR	SS_{BR}	$(b-1)(n-1)$	$MS_{BR} = \frac{SS_{BR}}{(b-1)(n-1)}$	—
AB	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_{AB} = \frac{MS_{AB}}{MS_{ABR}}$
ABR	SS_{ABR}	$(a-1)(b-1)(n-1)$	$MS_{ABR} = \frac{SS_{ABR}}{(a-1)(b-1)(n-1)}$	—
Error	—	0	—	—
Total	SS_{total}	$abn-1$	——	——

- The sums of squares are calculated exactly the same way as a three-factor factorial design. What is different is that no tests exist for the effects involving Replicate (R) because of the restrictions on randomization.
- The hypotheses for A, B, and AB are the same as for the two-factor factorial design with fixed effects. Because Replicates (R) is a random factor, all interactions involving R (i.e., AR, BR, and ABR) are also random. That is why these effects appear in the denominators of the F-test for A, B, and AB.
- The replicates R are called **whole plots**, factor A is called the **whole plot treatment**, and AR is called the **whole-plot error**.
- Factor B is called the **split-plot treatment** and ABR is called the **split-plot error**.
- The model equation is

$$y_{ijk} = \qquad (41)$$

where ρ_k is the k^{th} replicate effect, α_i is the i^{th} whole plot (factor A) effect, and β_j is the j^{th} split-plot (factor B) effect.

6.1.1 Split-Plot Design Example

- A researcher wanted to study the effects of oven temperature (T) in $^\circ F$ and baking time (B) in minutes on the lifetime (y) of an electrical component.
- The $a = 4$ levels of oven temperature T are $580^\circ, 600^\circ, 620^\circ, 640^\circ$.
- The $b = 3$ levels of baking time B are 5, 10, 15 minutes.
- $n = 3$ replicates were taken. For each replicate, the 4 oven temperatures are randomized and then the baking times are randomized within each temperature level.
- The data are summarized in the following table:

Replicate R	Baking time B	Oven temperature T: 580	600	620	640
I	5	217	158	229	223
	10	233	138	186	227
	15	175	152	155	156
II	5	188	126	160	201
	10	201	130	170	181
	15	195	147	161	172
III	5	162	122	167	182
	10	170	185	181	201
	15	213	180	182	199

$\longleftarrow$ randomize T within R

$\uparrow$ Randomize B within T

- The ANOVA results indicate no statistically significant test results (p-values of .7346, .2415, .8850 for T, B and BT).
- Note the large variability in *Reps R* means: 185.5, 162.6, and 189.3 for Reps I, II, and III. There is also large variability in the means for the interactions that include *Rep*: *R*T, R*B, R*T*B*.
- The mean squares for the interactions involving *Reps* ($MS_{RT}, MS_{RB}, MS_{RTB}$) are large and they appear in the denominators of the F-tests. This is why no statistically significant results occurred.
- Because there are 0 degrees of freedom for the MSE ($r^2 = 1$), we have a **saturated model**. All residuals will equal 0. Therefore, we cannot make diagnostic residual plots (such as a normal probability plot and a residuals vs predicted values plot) for this model.
- To make diagnostic residual plots, I recommend removing the three-factor interaction term ($R * B * T$), run the ANOVA again, and make residual plots for this reduced model.
- There are no problems with the normal probability plot and the residuals vs predicted values plot for this example.
- You use the **TEST** statement in SAS to perform the tests for the fixed effects. The TEST statement has the form

  TEST H=*fixed effect* E=*error term* / HTYPE=3 ETYPE=3.

SPLIT-PLOT DESIGN

The GLM Procedure

Variable: lifetime

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	35	29330.97222	838.02778	.	.
Error	0	0.00000	.		
Corrected Total	35	29330.97222			

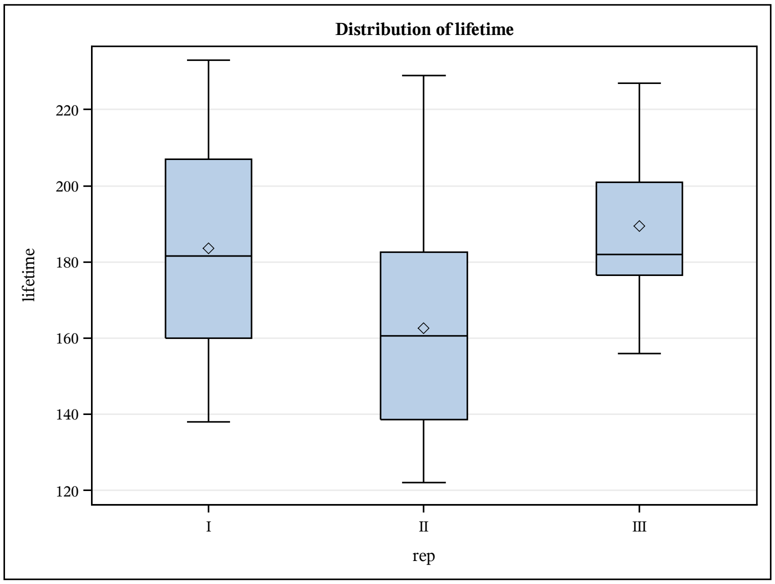
R-Square	Coeff Var	Root MSE	lifetime Mean
1.000000	.	.	178.4722

Source	DF	Type III SS	Mean Square	F Value	Pr > F
rep	2	4748.38889	2374.19444	.	.
baketime	2	566.22222	283.11111	.	.
rep*baketime	4	547.11111	136.77778	.	.
oventemp	3	2059.86111	686.62037	.	.
rep*oventemp	6	9422.72222	1570.45370	.	.
baketime*oventemp	6	1866.22222	311.03704	.	.
rep*baketim*oventemp	12	10120.44444	843.37037	.	.

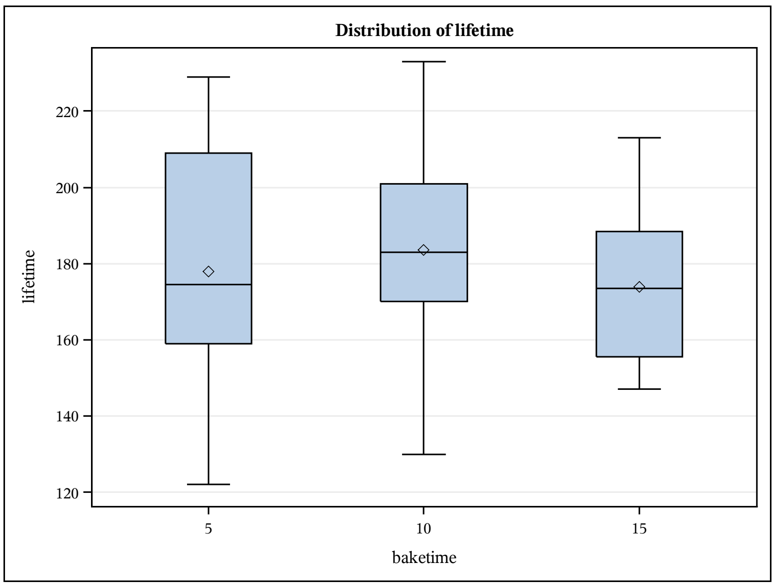
Tests of Hypotheses Using the Type III MS for rep*oventemp as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
oventemp	3	2059.861111	686.620370	0.44	0.7346

Tests of Hypotheses Using the Type III MS for rep*baketime as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
baketime	2	566.2222222	283.1111111	2.07	0.2415

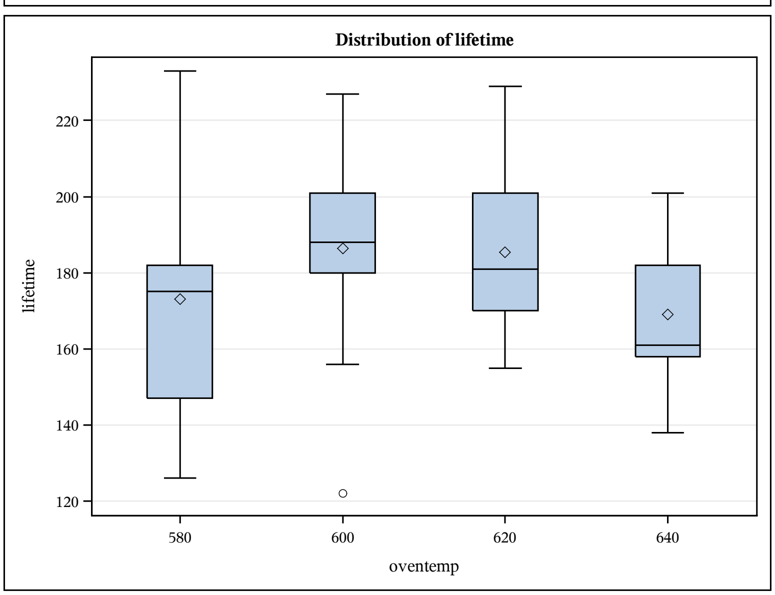
Tests of Hypotheses Using the Type III MS for rep*baketim*oventemp as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
baketime*oventemp	6	1866.222222	311.037037	0.37	0.8850



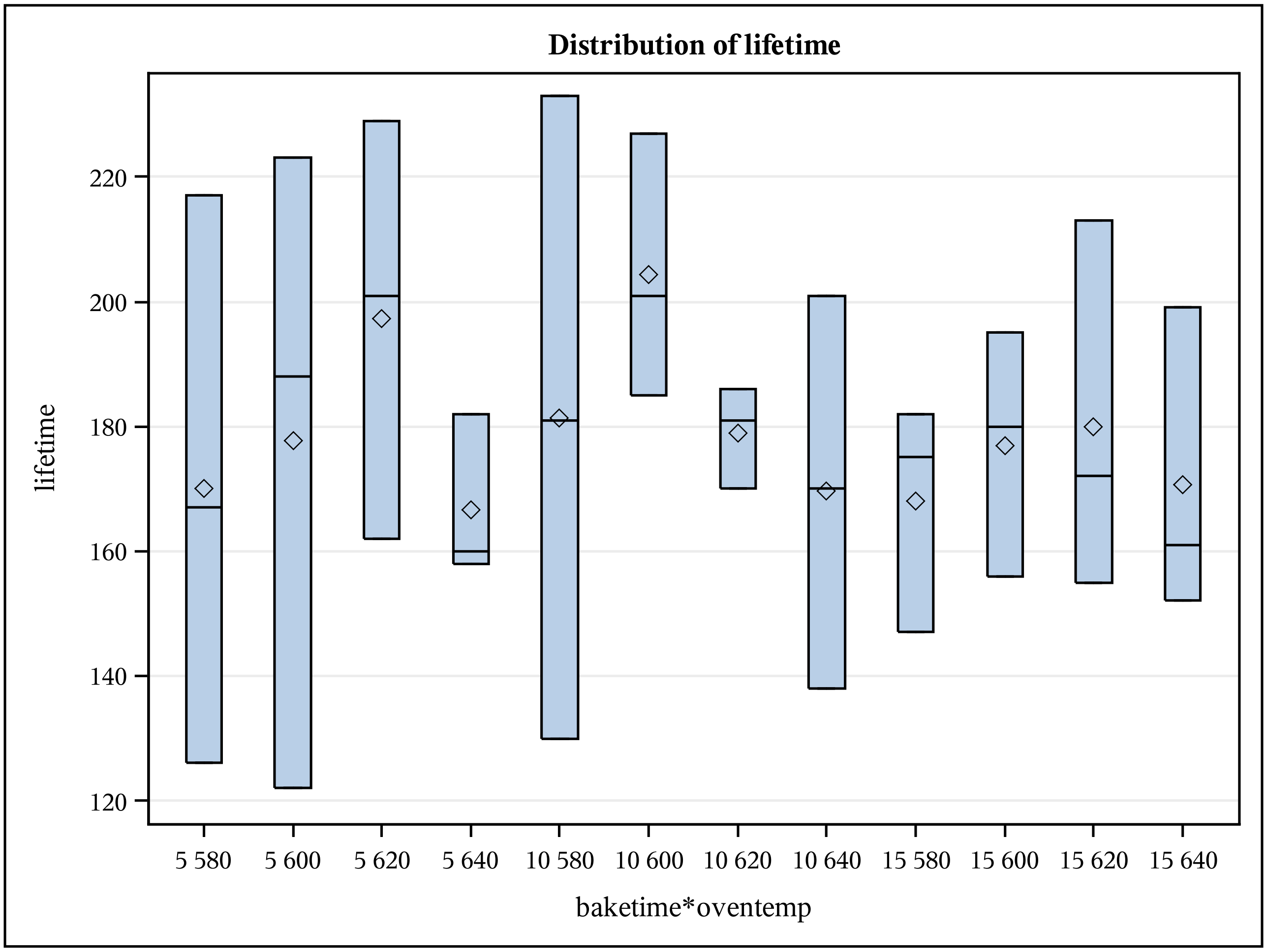
Level of rep	N	lifetime Mean	Std Dev
I	12	183.500000	29.1251219
II	12	162.583333	30.4166770
III	12	189.333333	21.4786716



Level of baketime	N	lifetime Mean	Std Dev
5	12	177.916667	35.3333691
10	12	183.583333	30.5598023
15	12	173.916667	20.7997305



Level of oventemp	N	lifetime Mean	Std Dev
580	9	173.111111	36.1505340
600	9	186.333333	32.3882695
620	9	185.444444	24.5311593
640	9	169.000000	21.2426458



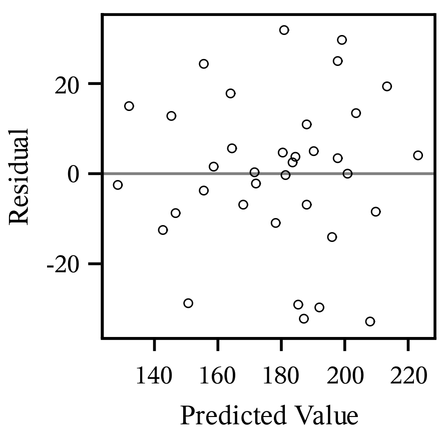
Level of baketime	Level of oventemp	N	lifetime Mean	lifetime Std Dev
5	**580**	3	170.000000	45.5741155
5	**600**	3	177.666667	51.2867754
5	**620**	3	197.333333	33.6501610
5	**640**	3	166.666667	13.3166562
10	**580**	3	181.333333	51.5008091
10	**600**	3	204.333333	21.1974841
10	**620**	3	179.000000	8.1853528
10	**640**	3	169.666667	31.5013227
15	**580**	3	168.000000	18.5202592
15	**600**	3	177.000000	19.6723156
15	**620**	3	180.000000	29.8161030
15	**640**	3	170.666667	24.9466097

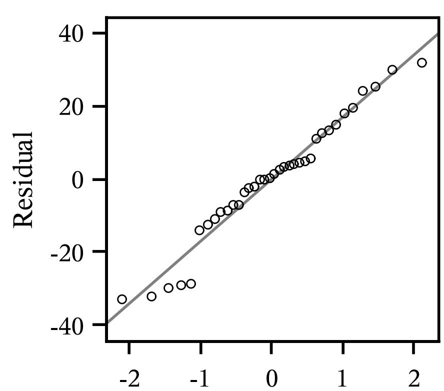
Variable: lifetime

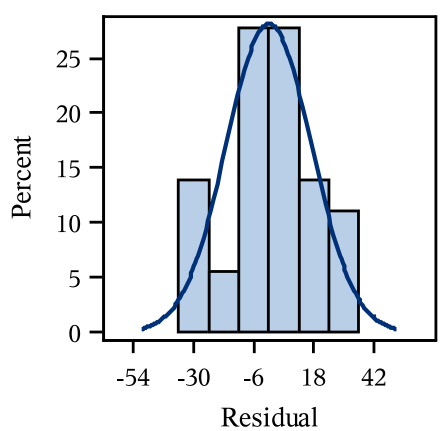
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	23	19210.52778	835.24034	0.99	0.5288
Error	12	10120.44444	843.37037		
Corrected Total	35	29330.97222			

R-Square	Coeff Var	Root MSE	lifetime Mean
0.654957	16.27191	29.04084	178.4722

Source	DF	Type III SS	Mean Square	F Value	Pr > F
rep	2	4748.388889	2374.194444	2.82	0.0994
baketime	2	566.222222	283.111111	0.34	0.7213
rep*baketime	4	547.111111	136.777778	0.16	0.9535
oventemp	3	2059.861111	686.620370	0.81	0.5104
rep*oventemp	6	9422.722222	1570.453704	1.86	0.1692
baketime*oventemp	6	1866.222222	311.037037	0.37	0.8850







6.1.2 SAS Code for Split-Plot Design Example

```
DM 'LOG;CLEAR;OUT;CLEAR;';
ODS GRAPHICS ON;
ODS PRINTER PDF file='C:\COURSES\ST541\SPLIT.PDF';
OPTIONS NODATE NONUMBER;

************************************;
*** EXAMPLE: A SPLIT-PLOT DESIGN ***;
************************************;

DATA in;
  DO rep='I  ', 'II ', 'III';
  DO oventemp=580 TO 640 BY 20;
  DO baketime=5 TO 15 BY 5;
     INPUT lifetime @@;  OUTPUT;
  END; END; END;
  CARDS;
 217 233 175 188 201 195 162 170 213
 158 138 152 126 130 147 122 185 180
 229 186 155 160 170 161 167 181 182
 223 227 156 201 181 172 182 201 199
 ;

PROC GLM DATA=in;
     CLASS rep baketime oventemp;
     MODEL lifetime = rep|baketime|oventemp / SS3;
     TEST H=oventemp E=rep*oventemp / HTYPE=3 ETYPE=3;
     TEST H=baketime E=rep*baketime / HTYPE=3 ETYPE=3;
     TEST H=oventemp*baketime E=rep*oventemp*baketime
          / HTYPE=3 ETYPE=3;
     MEANS rep baketime|oventemp;
TITLE 'SPLIT-PLOT DESIGN';

*******************************************************;
*** DO NOT MAKE PLOTS UNLESS YOU POOL MODEL TERMS   ***;
*** OR YOU REMOVE THE THREE-FACTOR INTERACTION TERM ***;
*******************************************************;
PROC GLM DATA=in PLOTS=(ALL);
     CLASS rep baketime oventemp;
     MODEL lifetime = rep|baketime|oventemp@2 / SS3;
RUN;
```

6.2 Split-Split-Plot Designs

- In a split-split-plot design, the experimenter is interested in studying the effects of three fixed factors (including the interactions).
- Let A B, and C be the three factors of interest with a levels for factor A, b levels for factor B. and c levels for factor C.
- The most common split-split-plot design has the following structure:
  1. Randomize the a levels of factor A.
  2. For each randomized level of factor A, randomize the levels of factor B.
  3. Randomly assign the levels of factor C to the experimental units within each randomized level of factor B.
  4. Collect observations using the randomization structure from Steps 1, 2, and 3.
  5. Replicate Steps 1, 2, 3, and 4 a total of n times. That is, for each replicate use new randomizations of the a levels of A in Step 1, the b levels in Step 2, and the c levels of C in Step 3.
- The replicates R are called **whole plots**, factor A is called the **whole plot treatment**, and RA is called the **whole-plot error**.
- Factor B is called the **split-plot treatment** and ABR is called the **split-plot error**.
- Factor C is called the **split-split-plot treatment** and $ABCR$ is called the **split-split-plot error**.
- The model is

$$\begin{aligned} y_{ijk} &= \mu + \rho_l + \alpha_i + (\rho\alpha)_{il} + \beta_j + (\rho\beta)_{jl} + (\alpha\beta)_{ij} + (\rho\alpha\beta)_{ijl} \\ &\quad + \rho\gamma_{kl} + \alpha\gamma_{ik} + (\rho\alpha\gamma)_{ikl} + \beta\gamma_{jk} + (\rho\beta\gamma)_{jkl} + (\alpha\beta\gamma)_{ijk} + (\rho\alpha\beta\gamma)_{ijkl} + \epsilon_{ijkl} \end{aligned}$$

  where ρ_k is the l^{th} replicate effect, α_i is the i^{th} whole plot (factor A) effect, β_j is the j^{th} split-plot (factor B) effect, and γ_k is the k^{th} split-split-plot (factor C) effect.
- The sums of squares are calculated exactly the same way as a four-factor factorial design. What is different is that no tests exist for the effects involving Replicate (R) because of the restrictions on randomization.
- The hypotheses for A, B, C, AB, AC, BC, and ABC are the same as for the three-factor factorial design with fixed effects. Because Replicates (R) is a random factor, all interactions involving R (AR, BR, CR, ABR ACR, BCR, $ABCR$) are also random. That is why these effects appear in the denominators of the F-test for A, B, C, AB, AC, BC, and ABC.
- The ANOVA table has the following structure. The double horizontal lines indicate where the three restrictions on randomization occur. R represents replicates.

Source of Variation	Sum of Squares	d.f.	Mean Square	F Ratio
R	SS_R	$n-1$	$MS_R = \frac{SS_R}{df(R)}$	—
A	SS_A	$a-1$	$MS_A = \frac{SS_A}{df(A)}$	$F_A = \frac{MS_A}{MS_{AR}}$
AR	SS_{AR}	$(a-1)(n-1)$	$MS_{AR} = \frac{SS_{AR}}{df(AR)}$	—
B	SS_B	$b-1$	$MS_B = \frac{SS_B}{df(B)}$	$F_B = \frac{MS_B}{MS_{BR}}$
BR	SS_{BR}	$(b-1)(n-1)$	$MS_{BR} = \frac{SS_{BR}}{df(BR)}$	—
AB	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{df(AB)}$	$F_{AB} = \frac{MS_{AB}}{MS_{ABR}}$
ABR	SS_{ABR}	$(a-1)(b-1)(n-1)$	$MS_{ABR} = \frac{SS_{ABR}}{df(ABR)}$	—
C	SS_C	$c-1$	$MS_C = \frac{SS_C}{df(C)}$	$F_C = \frac{MS_C}{MS_{CR}}$
CR	SS_{CR}	$(c-1)(n-1)$	$MS_{CR} = \frac{SS_{CR}}{df(CR)}$	—
AC	SS_{AC}	$(a-1)(c-1)$	$MS_{AC} = \frac{SS_{AC}}{df(AC)}$	$F_{AC} = \frac{MS_{AC}}{MS_{ACR}}$
ACR	SS_{ACR}	$(a-1)(c-1)(n-1)$	$MS_{ACR} = \frac{SS_{ACR}}{df(ACR)}$	—
BC	SS_{BC}	$(b-1)(c-1)$	$MS_{BC} = \frac{SS_{BC}}{df(BC)}$	$F_{BC} = \frac{MS_{BC}}{MS_{BCR}}$
BCR	SS_{BCR}	$(b-1)(c-1)(n-1)$	$MS_{BCR} = \frac{SS_{BCR}}{df(BCR)}$	—
ABC	SS_{ABC}	$(a-1)(b-1)(c-1)$	$MS_{ABC} = \frac{SS_{ABC}}{df(ABC)}$	$F_{ABC} = \frac{MS_{ABC}}{MS_{ABCR}}$
$ABCR$	SS_{ABCR}	$(a-1)(b-1)(c-1)(n-1)$	$MS_{ABCR} = \frac{SS_{ABCR}}{df(ABCR)}$	—
Error	—	0	—	—
Total	SS_{total}	$abcn-1$	—	—

Table 16.2: Percent of wetland biomass that is nonweed, by table (T), nitrogen (N), weed (W), and clipping (C).

		W 1		W 2		W 3	
T	N	C 1	C 2	C 1	C 2	C 1	C 2
1	1	87.2	88.8	70.4	75.7	75.9	80.6
	2	80.5	83.8	59.2	61.5	59.5	62.5
	3	76.8	80.8	47.8	49.5	48.4	52.9
	4	77.7	81.5	35.7	37.3	38.3	42.4
2	1	78.2	80.5	65.1	68.3	65.3	66.6
	2	79.8	85.2	57.6	61.4	58.5	61.6
	3	82.4	83.1	50.5	54.0	51.6	54.7
	4	75.5	78.7	39.0	43.9	41.9	45.1

Weed biomass in wetlands

An experiment studies the effect of nitrogen and weeds on plant growth in wetlands. We investigate four levels of nitrogen, three weed treatments (no additional weeds, addition of weed species 1, addition of weed species 2), and two herbivory treatments (clipping and no clipping). We have eight trays; each tray holds three artificial wetlands consisting of rectangular wire baskets containing wetland soil. The trays are full of water, so the artificial wetlands stay wet. All of the artificial wetlands receive a standard set of seeds to start growth.

Four of the trays are placed on a table near the door of the greenhouse, and the other four trays are placed on a table in the center of the greenhouse. On each table, we randomly assign one of the trays to each of the four nitrogen treatments. Within each tray, we randomly assign the wetlands to the three weed treatments. Each wetland is split in half. One half is chosen at random and will be clipped after 4 weeks, with the clippings removed; the other half is not clipped. After 8 weeks, we measure the fraction of biomass in each wetland that is nonweed as our response. Responses are given in Table 16.2.

- There are statistically significant N, W, NW, and C effects (p-value $\leq .05$). All other tests are not significant (p-value $> .05$)
- You use the **TEST** statement in SAS to perform the tests for the fixed effects. The TEST statement has the form

TEST H=*fixed effect* E=*error term* / HTYPE=3 ETYPE=3.

A SPLIT-SPLIT PLOT DESIGN FROM OEHLERT

The GLM Procedure

Variable: biomass

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	47	11615.56667	247.13972	.	.
Error	0	0.00000	.		
Corrected Total	47	11615.56667			

R-Square	Coeff Var	Root MSE	biomass Mean
1.000000	.	.	64.23333

Source	DF	Type III SS	Mean Square	F Value	Pr > F
table	1	14.300833	14.300833	.	.
N	3	3197.055000	1065.685000	.	.
table*N	3	278.950833	92.983611	.	.
C	1	125.453333	125.453333	.	.
table*C	1	0.100833	0.100833	.	.
N*C	3	0.735000	0.245000	.	.
table*N*C	3	3.930833	1.310278	.	.
W	2	7001.255417	3500.627708	.	.
table*W	2	12.325417	6.162708	.	.
N*W	6	929.516250	154.919375	.	.
table*N*W	6	38.092917	6.348819	.	.
C*W	2	0.245417	0.122708	.	.
table*C*W	2	3.200417	1.600208	.	.
N*C*W	6	4.816250	0.802708	.	.
table*N*C*W	6	5.587917	0.931319	.	.

Tests of Hypotheses Using the Type III MS for table*N as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
N	3	3197.055000	1065.685000	11.46	0.0377

Tests of Hypotheses Using the Type III MS for table*W as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
W	2	7001.255417	3500.627708	568.03	0.0018

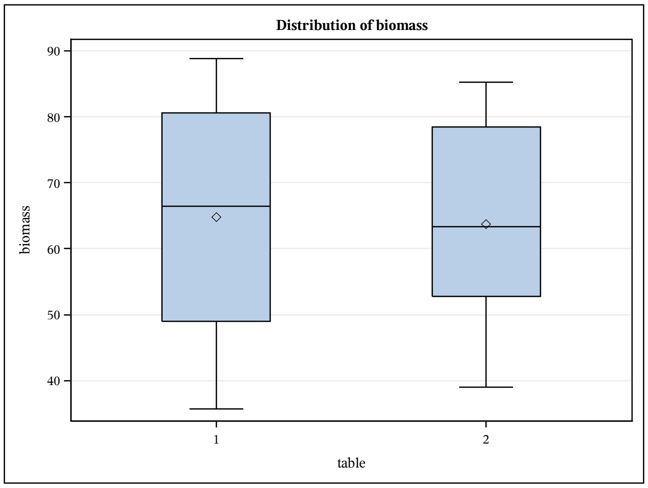
Tests of Hypotheses Using the Type III MS for table*N*W as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
N*W	6	929.5162500	154.9193750	24.40	0.0006

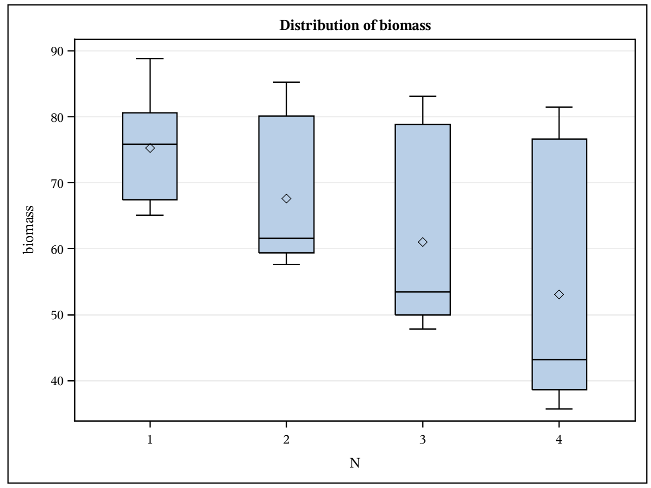
Tests of Hypotheses Using the Type III MS for table*C as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
C	1	125.4533333	125.4533333	1244.17	0.0180

Tests of Hypotheses Using the Type III MS for table*N*C as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
N*C	3	0.73500000	0.24500000	0.19	0.8990

Tests of Hypotheses Using the Type III MS for table*C*W as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
C*W	2	0.24541667	0.12270833	0.08	0.9288

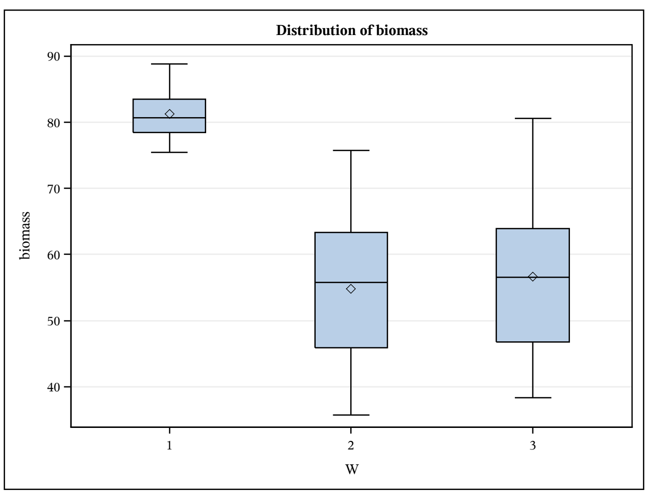
Tests of Hypotheses Using the Type III MS for table*N*C*W as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
N*C*W	6	4.81625000	0.80270833	0.86	0.5693

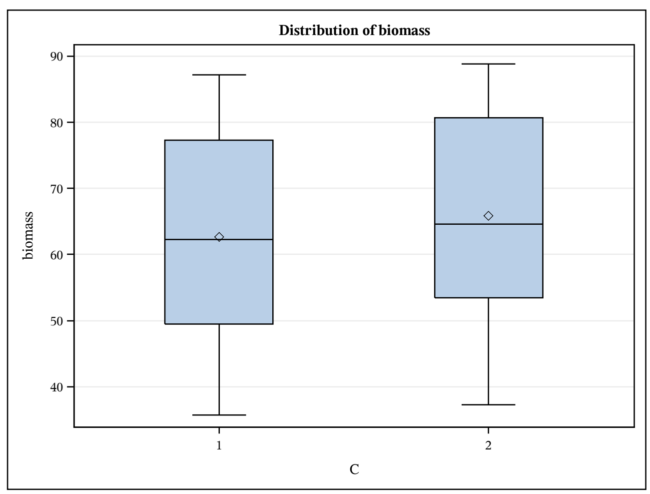




Level of table	N	biomass Mean	biomass Std Dev
1	24	64.7791667	17.2827126
2	24	63.6875000	14.3426185

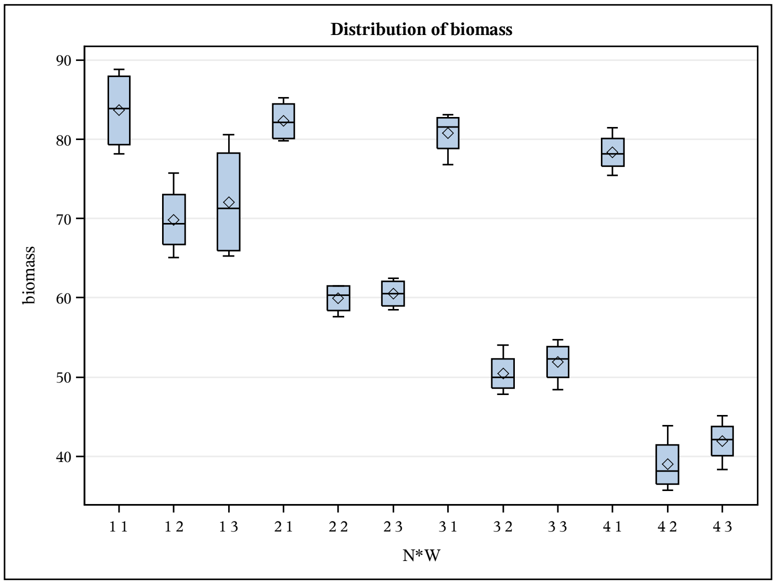
Level of N	N	biomass Mean	biomass Std Dev
1	12	75.2166667	8.2051905
2	12	67.5916667	11.0539551
3	12	61.0416667	14.7900125
4	12	53.0833333	18.8960233



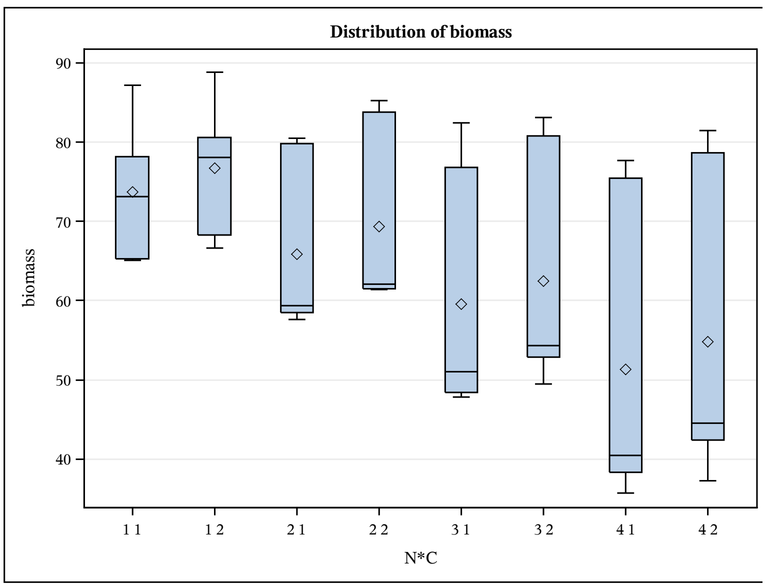


Level of W	N	biomass Mean	biomass Std Dev
1	16	81.2812500	3.6878573
2	16	54.8062500	12.1652219
3	16	56.6125000	12.0841977

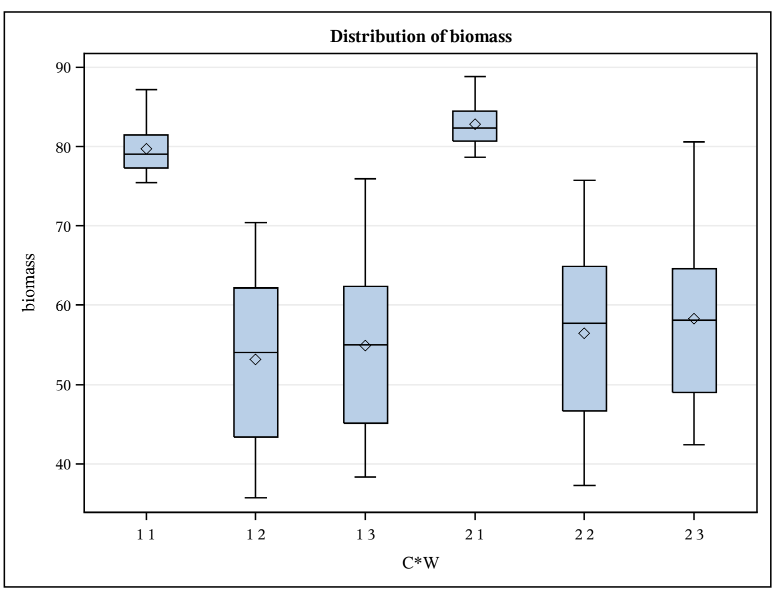
Level of C	N	biomass Mean	biomass Std Dev
1	24	62.6166667	15.8179223
2	24	65.8500000	15.7912469



Level of N	Level of W	N	biomass Mean	Std Dev
1	1	4	83.6750000	5.12339406
1	2	4	69.8750000	4.45299525
1	3	4	72.1000000	7.37518361
2	1	4	82.3250000	2.59149249
2	2	4	59.9250000	1.87860764
2	3	4	60.5250000	1.84458667
3	1	4	80.7750000	2.81942666
3	2	4	50.4500000	2.61597655
3	3	4	51.9000000	2.65706605
4	1	4	78.3500000	2.48931048
4	2	4	38.9750000	3.54906091
4	3	4	41.9250000	2.79568119



Level of N	Level of C	N	biomass Mean	Std Dev
1	1	6	73.6833333	8.5176092
1	2	6	76.7500000	8.3619974
2	1	6	65.8500000	11.0982431
2	2	6	69.3333333	11.7629361
3	1	6	59.5833333	15.6663227
3	2	6	62.5000000	15.1888117
4	1	6	51.3500000	19.6702567
4	2	6	54.8166667	19.7838739



Level of C	Level of W	N	biomass Mean	Std Dev
1	1	8	79.7625000	3.7121182
1	2	8	53.1625000	12.1722093
1	3	8	54.9250000	12.4267856
2	1	8	82.8000000	3.1767009
2	2	8	56.4500000	12.7588625
2	3	8	58.3000000	12.3281328

6.2.2 Split-Split-Plot Design Example with Pooled Errors

- Some statisticians will pool effects involving the interactions involving replicates that have small mean squares.
- In the previous example, the replicates are *tables*. The mean squares for the *table*C*, *table*N*C*, *table*C*W*, and *table*N*C*W* are relatively small and may be pooled together.
- To **pool** effects, you take the sum of the sums of squares and the sum of the degrees of freedom for these effects.

```
Source        DF   Type III SS   Mean Square

table          1     14.300833     14.300833
N              3   3197.055000   1065.685000
table*N        3    278.950833     92.983611

W              2   7001.255417   3500.627708
table*W        2     12.325417      6.162708
N*W            6    929.516250    154.919375
table*N*W      6     38.092917      6.348819

C              1    125.453333    125.453333
table*C        1      0.100833      0.100833 <-- Small MS with table*C
N*C            3      0.735000      0.245000
table*N*C      3      3.930833      1.310278 <-- Small MS with table*N*C
C*W            2      0.245417      0.122708
table*C*W      2      3.200417      1.600208 <-- Small MS with table*C*W
N*C*W          6      4.816250      0.802708
table*N*C*W    6      5.587917      0.931319 <-- Small MS with table*N*C*W
```

- The pooled sums of squares becomes the new SS_E.

$$SS_E = .100833 + 3.930833 + 3.200417 + 5.587917 = 12.82$$

  with $df_E = 1 + 3 + 2 + 6 = 12$. The new $MS_E = 12.82/12 = 1.0683$.

- The new MS_E is used in the denominator of the F-statistic for those effects that had *table*C, table*N*C, table*C*W*, and *table*N*C*W* as the denominator of the F-statistics before pooling.
- MS_E is used in the denominator of the F-statistics for testing significance of effects for C, NC, and NCW:

$$F_C = \frac{MS_C}{MS_E} \qquad F_{NC} = \frac{MS_{NC}}{MS_E} \qquad F_{NCW} = \frac{MS_{NCW}}{MS_E}$$

- All other F-statistics do not change. In this example, the F-statistics

$$F_N = \frac{MS_N}{MS_{table*N}} \qquad F_W = \frac{MS_W}{MS_{table*W}} \qquad F_{NW} = \frac{MS_{NCW}}{MS_{table*N*W}}$$

  are the same as the original analysis.

A SPLIT-SPLIT PLOT DESIGN FROM OEHLERT
POOLING TOGETHER SEVERAL ERROR TERMS

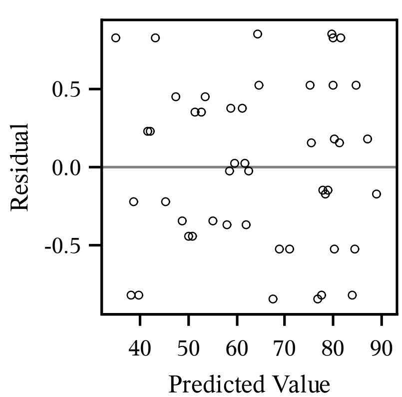
The GLM Procedure

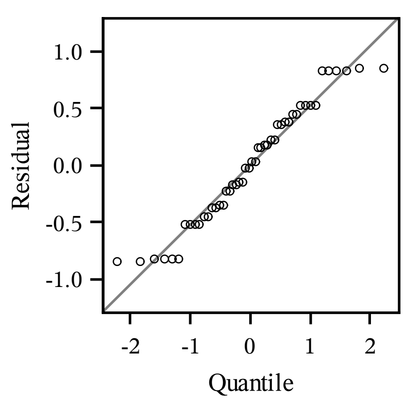
Variable: biomass

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	35	11602.74667	331.50705	310.30	<.0001
Error	12	12.82000	1.06833		
Corrected Total	47	11615.56667			

R-Square	Coeff Var	Root MSE	biomass Mean
0.998896	1.609137	1.033602	64.23333

Source	DF	Type III SS	Mean Square	F Value	Pr > F
N	3	3197.055000	1065.685000	997.52	<.0001
C	1	125.453333	125.453333	117.43	<.0001
N*C	3	0.735000	0.245000	0.23	0.8742
W	2	7001.255417	3500.627708	3276.72	<.0001
N*W	6	929.516250	154.919375	145.01	<.0001
C*W	2	0.245417	0.122708	0.11	0.8925
N*C*W	6	4.816250	0.802708	0.75	0.6203
table	1	14.300833	14.300833	13.39	0.0033
table*N	3	278.950833	92.983611	87.04	<.0001
table*W	2	12.325417	6.162708	5.77	0.0176
table*N*W	6	38.092917	6.348819	5.94	0.0044





Tests of Hypotheses Using the Type III MS for table*N as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
N	3	3197.055000	1065.685000	11.46	0.0377

Tests of Hypotheses Using the Type III MS for table*W as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
W	2	7001.255417	3500.627708	568.03	0.0018

Tests of Hypotheses Using the Type III MS for table*N*W as an Error Term					
Source	**DF**	**Type III SS**	**Mean Square**	**F Value**	**Pr > F**
N*W	6	929.5162500	154.9193750	24.40	0.0006

```
               EXAMPLE: A SPLIT-SPLIT PLOT DESIGN FROM OEHLERT
                    POOLING TOGETHER SEVERAL ERROR TERMS

                              The GLM Procedure
Dependent Variable: biomass

                                  Sum of
 Source           DF      Squares  Mean Square F Value Pr > F

 Model            35  11602.74667    331.50705  310.30 <.0001
 Error            12     12.82000      1.06833      <-- Pooled Error
 Corrected Total  47  11615.56667                  (table*C  + table*C*W
                                                   +table*N*C+ table*N*C*W)

     R-Square     Coeff Var      Root MSE    biomass Mean
     0.998896      1.609137      1.033602        64.23333


 Source     DF   Type III SS   Mean Square  F Value  Pr > F

 N           3   3197.055000   1065.685000
 C           1    125.453333    125.453333   117.43  <.0001  |
 N*C         3      0.735000      0.245000     0.23  0.8742  |
 W           2   7001.255417   3500.627708                   | Fixed effects
 N*W         6    929.516250    154.919375                   |
 C*W         2      0.245417      0.122708     0.11  0.8925  |
 N*C*W       6      4.816250      0.802708     0.75  0.6203  |

 table       1     14.300833     14.300833                   |
 table*N     3    278.950833     92.983611                   | Random effects
 table*W     2     12.325417      6.162708                   |
 table*N*W   6     38.092917      6.348819                   |
```

6.2.3 SAS Code for Split-Split-Plot Design Example

```
DM 'LOG; CLEAR; OUT; CLEAR;';
ODS GRAPHICS ON;
ODS PRINTER PDF file='C:\COURSES\ST541\SSPLOT.PDF';
OPTIONS NODATE NONUMBER;

*****************************************************;
*** EXAMPLE:  SPLIT-SPLIT PLOT DESIGN, Oehlert P.432 ***;
*****************************************************;
DATA in;
  DO table = 1 to 2 ;
  DO N = 1 to 4 ;
  DO W = 1 to 3 ;
  DO C = 1 to 2 ;
     INPUT biomass @@;  OUTPUT;
  END; END; END; END;
  CARDS;
87.2 88.8 70.4 75.7 75.9 80.6
80.5 83.8 59.2 61.5 59.5 62.5
76.8 80.8 47.8 49.5 48.4 52.9
77.7 81.5 35.7 37.3 38.3 42.4
78.2 80.5 65.1 68.3 65.3 66.6
79.8 85.2 57.6 61.4 58.5 61.6
82.4 83.1 50.5 54.0 51.6 54.7
75.5 78.7 39.0 43.9 41.9 45.1
 ;
PROC GLM DATA=in ;
     CLASS table N C W;
     MODEL biomass = table|N|C|W / SS3;
     TEST H=N     E=table*N     / HTYPE=3 ETYPE=3;
     TEST H=W     E=table*W     / HTYPE=3 ETYPE=3;
     TEST H=N*W   E=table*N*W   / HTYPE=3 ETYPE=3;
     TEST H=C     E=table*C     / HTYPE=3 ETYPE=3;
     TEST H=N*C   E=table*N*C   / HTYPE=3 ETYPE=3;
     TEST H=W*C   E=table*W*C   / HTYPE=3 ETYPE=3;
     TEST H=N*W*C E=table*N*W*C / HTYPE=3 ETYPE=3;
     MEANS table N|W|C@2;
TITLE 'A SPLIT-SPLIT PLOT DESIGN FROM OEHLERT';

****************************************************;
*** DO NOT MAKE PLOTS UNLESS YOU POOL MODEL TERMS   ***;
*** OR YOU REMOVE THE THREE-FACTOR INTERACTION TERM ***;
****************************************************;
PROC GLM DATA=in PLOTS=(ALL);
     CLASS table N C W ;
     MODEL biomass = N|C|W table table*N table*W table*N*W  / SS3;
     TEST H=N     E=table*N     / HTYPE=3 ETYPE=3;
     TEST H=W     E=table*W     / HTYPE=3 ETYPE=3;
     TEST H=N*W   E=table*N*W   / HTYPE=3 ETYPE=3;
TITLE2 'POOLING TOGETHER SEVERAL ERROR TERMS';
RUN;
```