Repeated Measures

Split-Plot In Time Analysis

<u>Treatments</u>:

Blocks (B)
Whole plots (W)
Split plots (S)

Split is conceptual, as treatment is applied to entire plot (or experimental unit) which is sampled in time (repeated measure).

<u>Linear additive model</u>:

$$Y_{ijk} = \mu + \underbrace{B_i + \delta_{(i)} + W_j + BW_{ij} + \omega_{(ij)}}_{\text{Space}} + \underbrace{S_k + BS_{ik} + WS_{jk} + BWS_{ijk}}_{\text{Time}}$$

Split-Plot in Time

Grass-Legume Mixture Example

<u>Treatments</u>:

Blocks (B) 4 reps

Treatment - (T) 4 grass-legume mixtures

1) no legume control
2) birdsfoot trefoil
3) alfalfa
4) kura clover

Date - Repeated Measure -(D) 4 harvest dates

1) spring
2) early summer
3) mid summer
4) late summer

Split-Plot in Time

Grass-Legume Example

- Field layout is the same as a RCBD
- Entire plot is harvested at each interval and allowed to regrow before next harvest
- Need to assume that measurements collected from a plot at one harvest are uncorrelated with measurements collected from same plot at other harvests

Block \ Plot	1	2	3	4
1	1	3	2	4
2	3	4	1	2
3	3	2	4	1
4	1	3	4	2

Grass-Legume Example

Expected Mean Squares

	Source	b R i	t F j	d F k	EMS
	B_i	1	t	d	$d\sigma^2_\omega + td\sigma^2_\delta + td\sigma^2_B$
	$\delta_{(i)}$	1	t	d	$d\sigma^2_\omega + td\sigma^2_\delta$
	T_j	b	0	d	$d\sigma^2_\omega + d\sigma^2_{BT} + bd\Phi(T)$
Error a	BT_{ij}	1	0	d	$d\sigma^2_\omega + d\sigma^2_{BT}$
	$\omega_{(ij)}$	1	1	d	$d\sigma^2_\omega$
	D_k	b	t	0	$t\sigma^2_{BD} + bt\Phi(D)$
Pool ↓	BD_{ik}	1	t	0	$t\sigma^2_{BD}$
	TD_{jk}	b	0	0	$\sigma^2_{BTD} + b\Phi(TD)$
Error b	BTD_{ijk}	1	0	0	σ^2_{BTD}

Split-Plot in Time

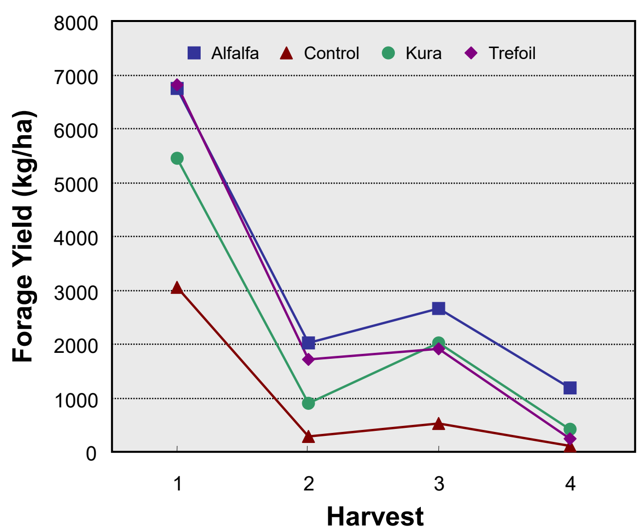
Grass-Legume Mixture Example

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Blk	3	505996.5	168665.5	1.22	
Trt	3	40986209	13662070	253.96	<.0001
Error a	9	484157.4	53795.3	0.39	
Date	3	2.41E+08	80241422	578.19	<.0001
Trt*Date	9	15762104	1751345	12.62	<.0001
Error b	36	4996075	138779.9		

This analysis is OK as long as our assumptions hold. What happens if repeated measurements are correlated?

Split-Plot in Time

Grass-Legume Mixture Example



Split-Plot in Time

Grass-Legume Mixture Example

trt*date Effect Sliced by date for yield

Date	DF	SS	Mean Square	F Value	Pr > F
1	3	36884379	12294793	88.59	<.0001
2	3	7402380	2467460	17.78	<.0001
3	3	9688544	3229515	23.27	<.0001
4	3	2773011	924337	6.66	.00011

$$SED = \sqrt{\frac{2[(4-1)(138779.9)+53795.3]}{4(4)}} = 242.42$$

$$LSD = 2.028(242.42) = 491.63$$

Grass-Legume Mixture Example

Differences of Least Squares Means

Effect	trt	date	_trt	_date	Estimate	Error	DF	t Value	Pr > |t|
trt*date	Alfalfa	1	Control	1	3677.25	242.42	36	15.17	<.0001
trt*date	Alfalfa	1	Kura	1	1288	242.42	36	5.31	<.0001
trt*date	Alfalfa	1	Trefoil	1	-80	242.42	36	-0.33	0.7433
trt*date	Control	1	Kura	1	-2389.25	242.42	36	-9.86	<.0001
trt*date	Control	1	Trefoil	1	-3757.25	242.42	36	-15.5	<.0001
trt*date	Kura	1	Trefoil	1	-1368	242.42	36	-5.64	<.0001
trt*date	Alfalfa	2	Control	2	1723.75	242.42	36	7.11	<.0001
trt*date	Alfalfa	2	Kura	2	1123	242.42	36	4.63	<.0001
trt*date	Alfalfa	2	Trefoil	2	296.25	242.42	36	1.22	0.2296
trt*date	Control	2	Kura	2	-600.75	242.42	36	-2.48	0.018
trt*date	Control	2	Trefoil	2	-1427.5	242.42	36	-5.89	<.0001
trt*date	Kura	2	Trefoil	2	-826.75	242.42	36	-3.41	0.0016
trt*date	Alfalfa	3	Control	3	2135.5	242.42	36	8.81	<.0001
trt*date	Alfalfa	3	Kura	3	641.25	242.42	36	2.65	0.012
trt*date	Alfalfa	3	Trefoil	3	759.5	242.42	36	3.13	0.0034
trt*date	Control	3	Kura	3	-1494.25	242.42	36	-6.16	<.0001
trt*date	Control	3	Trefoil	3	-1376	242.42	36	-5.68	<.0001
trt*date	Kura	3	Trefoil	3	118.25	242.42	36	0.49	0.6287
trt*date	Alfalfa	4	Control	4	1077	242.42	36	4.44	<.0001
trt*date	Alfalfa	4	Kura	4	767.5	242.42	36	3.17	0.0031
trt*date	Alfalfa	4	Trefoil	4	938	242.42	36	3.87	0.0004
trt*date	Control	4	Kura	4	-309.5	242.42	36	-1.28	0.2099
trt*date	Control	4	Trefoil	4	-139	242.42	36	-0.57	0.5699
trt*date	Kura	4	Trefoil	4	170.5	242.42	36	0.7	0.4864

Comparing Two Sample Means
Standard Errors (of the difference)

For uncorrelated measurements:

$$S_{\bar{d}} = \sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

Business as usual

For correlated measurements:

$$S_{\bar{d}} = \sqrt{(s_p^2 - \text{cov}_{12})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

Houston we have a problem

Repeated Measures
Covariance

$$\sigma_{XY}^2 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{n-1}$$

Repeated Measures
Correlation

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2}$$

$$r = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y}) / (n-1)}{\sqrt{\sum (X_i - \overline{X})^2 / (n-1)} \sqrt{\sum (Y_i - \overline{Y})^2 / (n-1)}}$$

Repeated Measures
Standard Error of a Mean Difference
Correlated Measures

$$S_{\bar{d}} = \sqrt{\frac{2(s_p^2 - s_{12})}{n}}$$

$$S_{\bar{d}} = \sqrt{s_p^2 - s_{12}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$S_{\bar{d}} = \sqrt{\left(\frac{s_1^2 - s_{12}}{n_1} + \frac{s_2^2 - s_{12}}{n_2}\right)}$$

Repeated Measures

Variance-Covariance Matrix

$$Var\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} Var[Y_1] & Cov[Y_1,Y_2] & Cov[Y_1,Y_3] \\ Cov[Y_2,Y_1] & Var[Y_2] & Cov[Y_2,Y_3] \\ Cov[Y_3,Y_1] & Cov[Y_3,Y_2] & Var[Y_3] \end{bmatrix}$$

Repeated Measures

Univariate Approaches

Split-Plot in Time Assumptions:

1. Variances are homogenous over time
2. All pairs of observations:
   - not correlated, or
   - equally correlated, or
   - have same variance of a difference (Huynh-Feldt Condition)

Repeated Measures
Univariate Approaches

Valid when:

$$Var\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}$$

or

$$Var\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} \sigma^2+\sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2+\sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2+\sigma_1 \end{bmatrix}$$

Split-Plot in Time
Adjusting df

Source	df	adjusted df
Blk	3	
Trt	3	
Error a	9	
Date	3	1
Trt*Date	9	3
Error b	36	12

Conservative test for repeated factor – divide df for each factor associated with time (subplots) by the df for the main effect of the subplot treatment.

Split-Plot in Time

Grass-Legume Mixture Example

Adjusted Subplot df

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Blk	3	505996.5	168665.5	1.22	
Trt	3	40986209	13662070	253.96	<.0001
Error a	9	484157.4	53795.3	0.39	
Date	1	2.41E+08	2.41E+08	578.19	<.0001
Trt*Date	3	15762104	5254035	12.62	0.0005
Error b	12	4996075	416339.6		

Reducing the subplot df changes the F-distribution to which the calculated F values are compared making these tests more conservative.

Repeated Measures

Univariate Approaches

Other Univariate Approaches:

- Analysis by Time
- Analysis of Differences
- Response Indices

Repeated Measures

Response Index Example – Corn Residue

Treatments:

Blocks (B) 4 reps

Tillage (T) 2

1) Chisel plow
2) No-till

Days - Repeated Measure -(D) 9

7, 14, 21, 29, 43, 57, 71, 83, 90

Variable measured

Percent initial residue biomass remaining

Data from Alex Cleverenga (2020)

Repeated Measures

Response Index Example – Corn Residue

SAS code for split-plot in time analysis

```
proc glm;
class block tillage days;
model prctmass = block tillage
 block*tillage days tillage*days;
test h=tillage e= block*tillage;
means days / lsd linestable;
lsmeans tillage*days;
ods output lsmeans=means;
run;
```

Repeated Measures

Response Index Example – Corn Residue

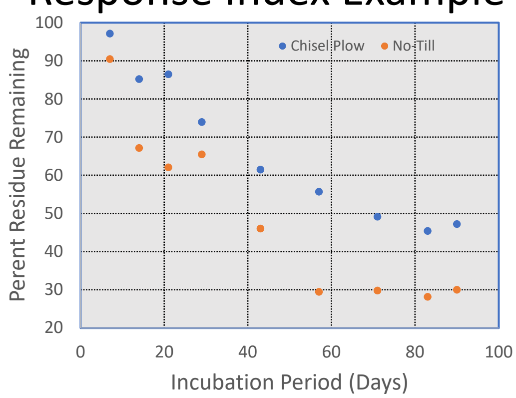
ANOVA for split-plot in time analysis

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Block	2	421.52062	210.76031	4.66	
Tillage	1	4115.41670	4115.41670	14.03	0.0645
Error a	2	421.33734	210.66867	4.65	
Days	8	20494.10819	2561.76352	56.59	<.0001
Tillage*Days	8	482.81903	60.35238	1.33	0.2644
Error b	31	1403.37879	45.27028		

- Tillage is important (large F combined with low error df support using α = 0.10)
- Days of incubation are also important
- There is no interaction between tillage and days ∴ the effect of tillage was consistent over time

Repeated Measures

Response Index Example – Corn Residue



LSD

Group	Mean	N	Days
A	93.798	6	7
B	76.204	6	14
B	74.287	6	21
B	69.723	6	29
C	53.764	6	43
D	42.579	6	57
D	38.61	6	90
D	37.546	5	71
D	36.765	6	83

- Disappearance of residue DM appears to follow exponential decay ∴ first-order kinetics apply
- Can determine the rate (k) of decay under each tillage practice by ln transforming the % remaining and using linear regression
- Only time intervals where degradation has occurred can be used
- LSD indicates time intervals > 57 should be excluded

Repeated Measures

Response Index Example – Corn Residue

SAS code for regression analysis

```
data means; set means;
if days > 57 then delete;
lnprct = log(lsmean);
day = input(days, best5.);
run;

proc mixed;
class tillage;
model lnprct = tillage tillage*day / solution noint;
estimate 'Chisel intercept' tillage 1;
estimate 'Chisel rate' tillage*day 1;
estimate 'No-till intercept' tillage 0 1;
estimate 'No-till rate' tillage*day 0 1;
estimate 'Chisel v No-till intercept' tillage 1 -1;
estimate 'Chisel v No-till rate' tillage*day 1 -1;
run;
```

Using the means output from GLM to create dataset with ln-transformed variable

Using PROC MIXED to calculate and compare regression parameters

Repeated Measures

Response Index Example – Corn Residue

Estimates					
Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Chisel intercept	4.6436	0.07074	8	65.64	<.0001
Chisel rate	-0.01134	0.002129	8	-5.32	0.0007
No-till intercept	4.6053	0.07074	8	65.10	<.0001
No-till rate	-0.01986	0.002129	8	-9.33	<.0001
Chisel v No-till intercept	0.03839	0.1000	8	0.38	0.7112
Chisel v No-till rate	0.008523	0.003011	8	2.83	0.0221

- The relative degradation rate differed between tillage systems and was nearly twice as fast under No-till than Chisel plow
- The units of k are 1/days but can be interpreted as percent per day by multiplying by 100; 1.1%/day vs 2.0%/day
- Using ln-linear regression over the time intervals provides much more insight into the process than evaluating the mean responses

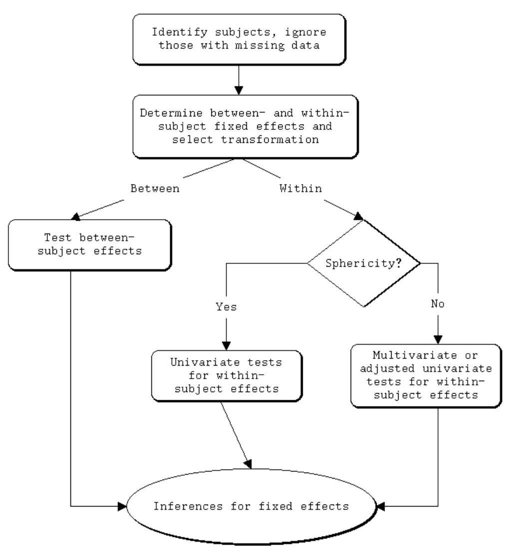
Repeated Measures

Multivariate Approach

$$Var\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}$$

Repeated Measures

GLM Approach



Source: Russ Wolfinger and Ming Chang, SAS Institute Inc., Cary, NC