Split Plots Pros and Cons:
Dealing with a Hard-to-Change Factor

There are many attendees today! *To avoid disrupting the Voice over Internet Protocol (VoIP) system, I will mute all. Please use the Questions feature on GotoWebinar which we will answer during the presentation.*
-- Pat

Presented by Pat Whitcomb, Founder
Stat-Ease, Inc., Minneapolis, MN
pat@statease.com

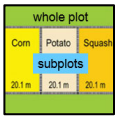
Agenda
Split Plot Pros and Cons

1. Restricting randomization
2. Factorial split plot
   (DNA Amplification)
3. Combined mixture process split plot
   (Reverse Phase HPLC)
4. Summary

Agenda
Split Plot Pros and Cons

1. **Restricting randomization**
2. Factorial split plot
   (DNA Amplification)
3. Combined mixture process split plot
   (Reverse Phase HPLC)
4. Summary

Split-Plot Designs



Problem:

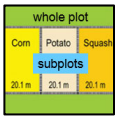
Often in designed experiments some factors, e.g., temperature, are more difficult or expensive to vary than others. In some cases, conducting a completely randomized design isn't practical.

Solution:

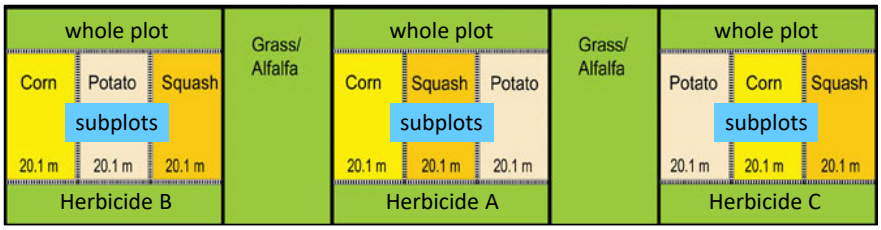
Restrict the randomization so it is practical to conduct the design. If you have to sort a factor to make a DOE easier to run, this restriction in randomization results in a "split-plot" design.

Do you have factors that are difficult to randomize?

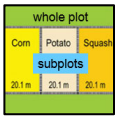
Split-Plot Designs



The "split-plot" design originated in the field of agriculture. Experimenters applied one treatment (*e.g. herbicide*) to a large area of land, called a "**whole plot**" and other treatments (*e.g. crop*) to smaller areas of land within the whole plot called "**subplot**".



Split-Plot Designs



Split-Plot Designs:

- Split plots have two types of factors: "Hard-to-change" (HTC) applied to the whole plots and "Easy-to-change" (ETC) applied to the subplots.
- The randomization of the HTC factor is restricted.
- Split plots naturally arise in many DOE studies.
- Building and analyzing a split-plot design is tricky, unless you have a good DOE package.

You are in luck, split plots are easy using Design-Expert software!

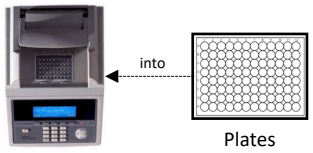
Agenda
Split Plot Pros and Cons

1. Restricting randomization
2. **Factorial split plot**
   (DNA Amplification)
3. Combined mixture process split plot
   (Reverse Phase HPLC)
4. Summary

DNA Amplification
2^{9-1} Factorial run as Split-Plot

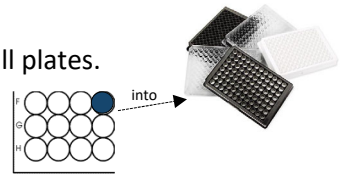
Discussions with microbiologists identify 9 factors to study for the DNA amplification DOE:

- 3 factors are associated with the configuration of the thermocycler.



- 6 factors come with sample preparation on an automated liquid handling system.

The experiment is done with 96-well plates.



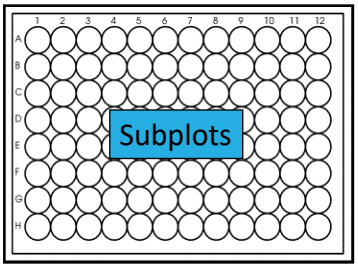
DNA Amplification
2^{9-1} Factorial run as Split-Plot

Key Points:

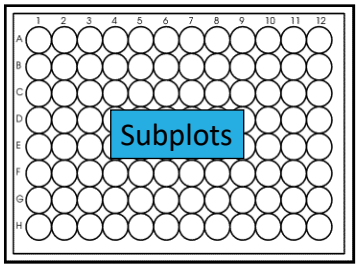
- Each combination of the 3 thermocycler factors (2^3 = 8 combinations) must be run on a different plate; therefore 8 plates are required.
- Use a 2^{9-1} fractional factorial (256 runs) for the 9 factors: 3 thermocycler + 6 sample preparation.

DNA Amplification
2^{9-1} Factorial run as Split-Plot

1. The 3 thermocycler factors, which are HTC, are varied between plates. They are “whole-plot” (or “whole-plate”).
2. The 6 sample preparation factors—easy to change (ETC) vary within the plates. They are the “subplot” (or “subplate”) factors.



Whole Plot



Whole Plot

DNA Amplification
2^{9-1} Factorial run as Split-Plot

3 Thermocycler factors:

Factor (HTC)	−1 level	+1 level	Factor Type
Anneal Temperature	55°C	65°C	Inter-plate
Denature Temperature	90°C	97°C	Inter-plate
Denature Time	10 sec	20 sec	Inter-plate

6 Sample preparation factors:

Factor (ETC)	−1 level	+1 level	Factor Type
Forward [FOR] primer	200 nM	900 nM	Reagent
Reverse [REV] primer	200 nM	900 nM	Reagent
DNA probe	100 nM	400 nM	Reagent
MgCl2	3 nM	9 nM	Reagent
Tween (PCR) buffer	0.005%	0.020%	Reagent
Taq Gold polymerase	0.1 units/µL	0.25 units/µL	Reagent

2^{9-1} Factorial Split-Plot
8 whole plots each with 32 subplots

Anneal temperature	Denature temperature	Denature time	Combinations of the 6 sample preparation factors
55 deg C	90 deg C	10 sec	32 sample preps on a plate
65 deg C	90 deg C	10 sec	32 sample preps on a plate
55 deg C	97 deg C	10 sec	32 sample preps on a plate
65 deg C	97 deg C	10 sec	32 sample preps on a plate
55 deg C	90 deg C	20 sec	32 sample preps on a plate
65 deg C	90 deg C	20 sec	32 sample preps on a plate
55 deg C	97 deg C	20 sec	32 sample preps on a plate
65 deg C	97 deg C	20 sec	32 sample preps on a plate

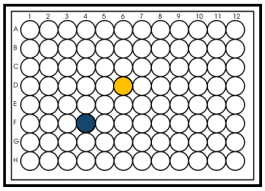
2^{9-1} Factorial Split-Plot
8 whole plots each with 32 subplots

Summary:

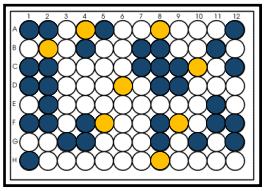
- If you wanted to change the thermocycler factors every run (randomize every cell change), you would need 256 plates, instead of just 8 (8 plates * 32 cells per plate = 256 cells).
- This would amount to having only 1 DOE cell on the plate at a time, a waste of resources.
- Using the split-plot structure minimizes the number of plate changes.

DOE cell
QC cell

Completely Randomized



Split-Plot Design



DNA Amplification
29-1 Factorial run as Split-Plot

96-well plate factors:
D: Forward Primer
E: Reverse Primer
F: DNA Probe
G: $MgCl_2$
H: Tween (PCR) buffer
J: Taq polymerase

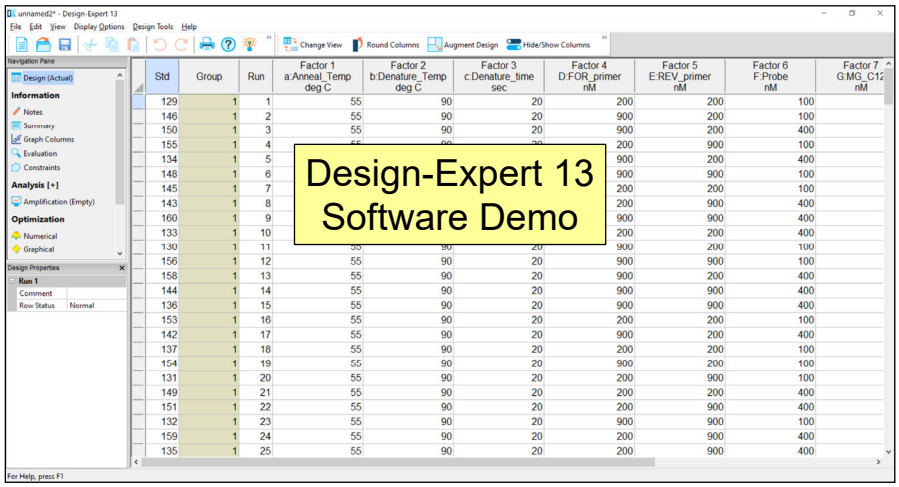
SUB-PLOTS within PLATES

Thermocycler factors:
A: Anneal Temperature
B: Denature Temperature
C: Denature Time

WHOLE PLOT
across PLATES

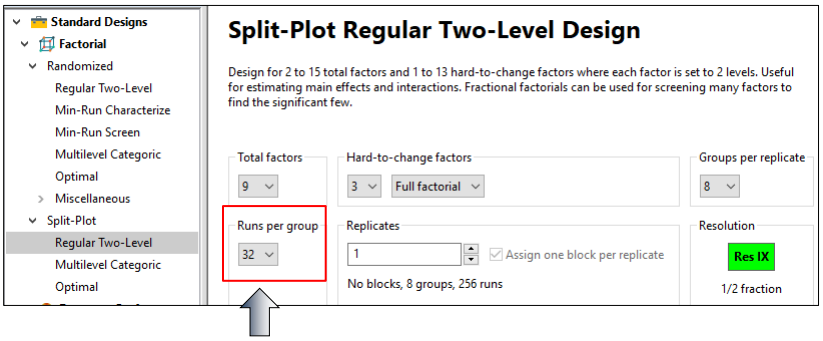
With only 8 whole-plot combinations (plates), we can get our 256 runs with 256/8 = 32 cells per plate. That's 32 times faster!

DNA Amplification
Building the 2^{9-1} Factorial Split-Plot



2^{9-1} Factorial Split-Plot *Instructor-led*
Building the Design *(page 1 of 2)*

1. Open "**PCR starter.dxpx**" and re-build the design using previous info for a split-plot, regular, two-level, fractional factorial design for **9** total factors, **3** hard-to-change factors done with **32** runs (change from 16) per group.



2^{9-1} Factorial Split-Plot
Building the Design *(page 2 of 2)*

2. Check the factor names and levels already entered:

	Name	Units	Change	Type	Low	High
a [Numeric]	Anneal_Temp	deg C	Hard	Numeric	55	65
b [Numeric]	Denature_Temp	deg C	Hard	Numeric	90	97
c [Numeric]	Denature_time	sec	Hard	Numeric	10	20
D [Numeric]	FOR_primer	nM	Easy	Numeric	200	900
E [Numeric]	REV_primer	nM	Easy	Numeric	200	900
F [Numeric]	Probe	nM	Easy	Numeric	100	400
G [Numeric]	MG_C12	nM	Easy	Numeric	3	9
H [Numeric]	Tween	%	Easy	Numeric	0.005	0.02
J [Numeric]	Taq_Gold	U/ul	Easy	Numeric	0.01	0.04

3. Check the values for delta and sigma:

	Name	Units	Diff. to detect Delta("Signal")	Est. Subplot Sigma("Noise")	Delta/Sigma (Signal/Noise Ratio)
	Amplification		3	4	0.75

2^{9-1} Factorial Split-Plot
Power Calculation

Split-Plot Design Power

Recommended power is at least 80%.

	Name	Units	Delta (Signal)	Sigma (Noise)	Signal/Noise	Power for a	Power for b	Power for c	Power for D	Power for E	Power for F	Power for G	Power for H	Power for J
	Amplification		3	4	0.75	12.9%	12.9%	12.9%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%

Randomized Design Power

The power of an equivalent completely randomized design.

	Name	Units	Delta (Signal)	Sigma (Noise)*	Signal/Noise	Power for a	Power for b	Power for c	Power for D	Power for E	Power for F	Power for G	Power for H	Power for J
	Amplification		3	5.65685	0.53033	98.8%	98.8%	98.8%	98.8%	98.8%	98.8%	98.8%	98.8%	98.8%

The power for the whole-plot factors (a, b & c) falls way off due to fewer resets (only 8); the whole-plot factors form a 2^3 full factorial.

This is the cost associated with not randomizing all factors!

Note: *The whole-plot by subplot interactions have the error associated with subplot effects. If one of these interactions is selected, then the whole-plot term gets included for hierarchy, and its low power isn't as much of an issue.*

DNA Amplification
2^{9-1} Factorial run as Split-Plot

This split-plot design is really two designs:

1. The whole-plot factors (a, b, & c), for setting up the thermocycler to treat the whole plate.
2. The subplot factors (D, E, F, G, H & J), for creating the samples in each cell of the plate.

Since well-to-well, subplot variation occurs only within a plate, the plate-to-plate variation is not included in the subplot variance.

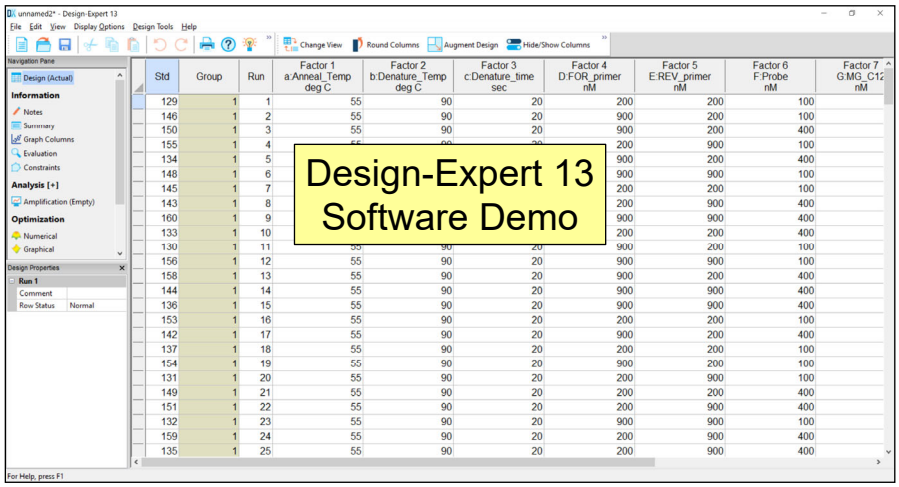
The error associated with resetting the thermocycler factors or, plate-to-plate variation, is included in the whole-plot variance.

DNA Amplification
2^{9-1} Factorial run as Split-Plot

Other considerations:

- Every run varies due to resetting factor levels.
- Whole-plot treatments only change by group.
- Subplot treatments may be altered on every run.
- To account for the differing errors, separate half-normal plots are created for the whole plot and subplot effects.
- The multiple error sources also influence the power calculations. The more the error and the fewer the changes the less the power.

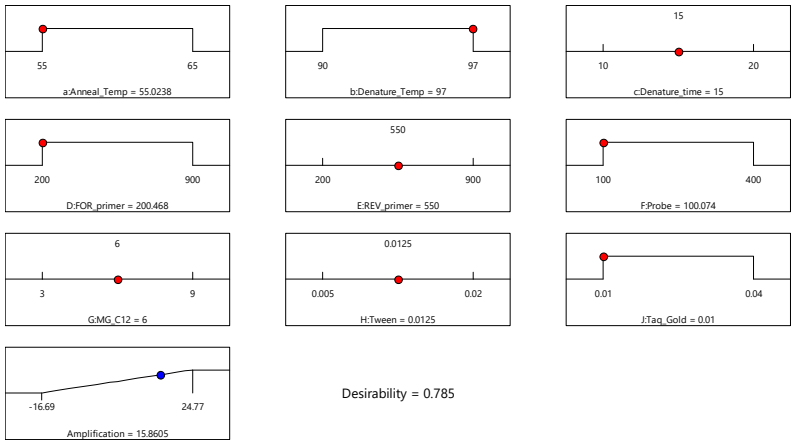
DNA Amplification
Analyze and Optimize



April 2021 Webinar 21

DNA Amplification
2^{9-1} Factorial run as Split-Plot

Recommended factor settings to maximize Amplification:



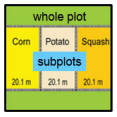
Desirability = 0.785

April 2021 Webinar 22

Agenda
Split Plot Pros and Cons

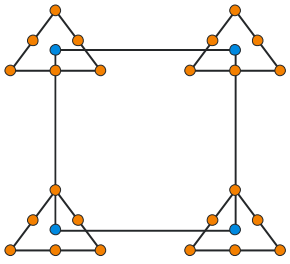
1. Restricting randomization
2. Factorial split plot
   (DNA Amplification)
3. **Combined mixture process split plot**
   (Reverse Phase HPLC)
4. Summary

Split-Plot Designs

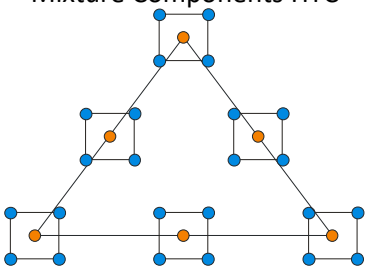


In combined designs it's common that either the process factors or the mixture components are hard to change (HTC).

Process Factors HTC



Mixture Components HTC

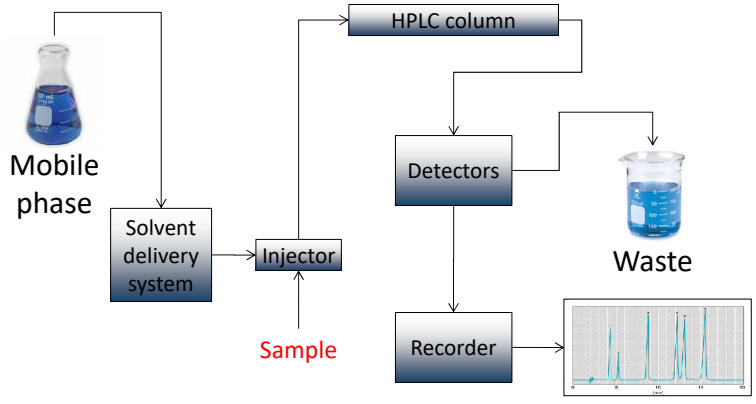


Combined Design
Reverse Phase HPLC

High performance liquid chromatography (or high-pressure liquid chromatography, HPLC) is a form of column chromatography used frequently in biochemistry and analytical chemistry to separate, identify, and quantify compounds based on their idiosyncratic polarities and interactions with the column's stationary phase.

Reversed phase HPLC (RP-HPLC) has a non-polar stationary phase and an aqueous, moderately polar mobile phase.

Reverse Phase HPLC
Apparatus



Reverse Phase HPLC
Four Mixture Components *(one fixed)*

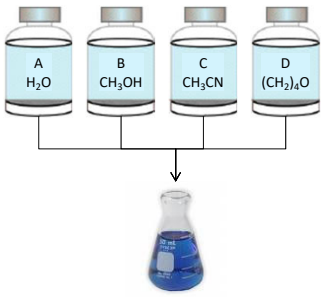
Mobile phase:

A. water
60 to 80 volume %

B. methanol
0 to 35 volume %

C. acetonitrile
0 to 20 volume %

D. tetrahydrofuran (THF)
fixed at 5 volume %

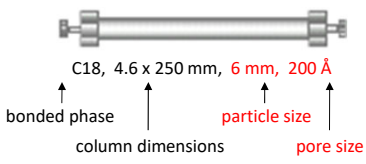


Mobile phase

Reverse Phase HPLC
Two Process Factors

Column factors:

- **particle size** (**3.5**, **6** or **10** mm)
- **pore size** (**100**, **200** or **300** Å)
- bonded phase is fixed (C18)
- column dimensions are fixed (4.6 x 250 mm)



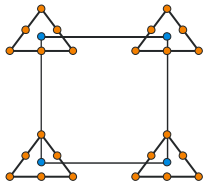
Reverse Phase HPLC
Two Process Factors

Design considerations:

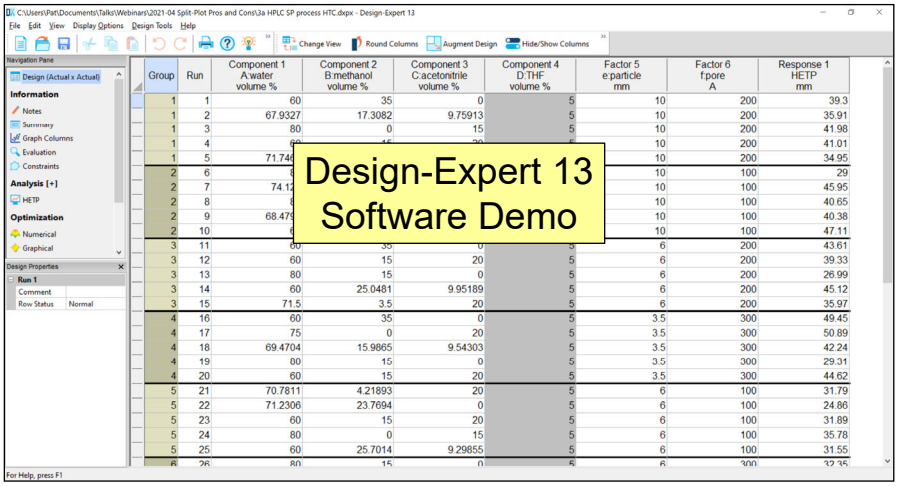
- The mobile phase is easy to mix and change.
- The column factors (particle and pore sizes) are difficult and time consuming to change.

The experiments will be easier, faster and less costly if the mixture components are specified as easy-to-change and the column factors are specified as hard-to-change.

Use a split-plot combined design with Mix components as ETC and Numeric factors as HTC.

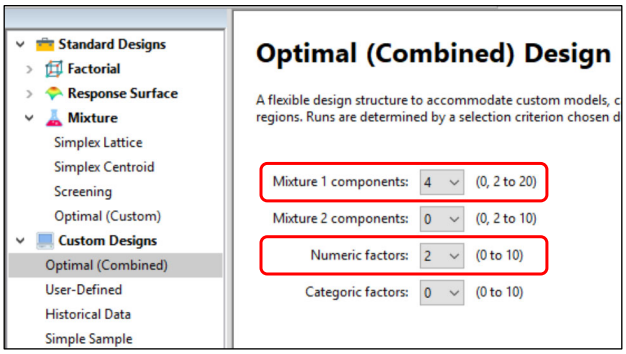


Reverse Phase HPLC
Build a Combined Design Split Plot



Build an Optimal Split-Plot Design
Process HTC *(page 1 of 4)*

1. Choose “File”, “New Design”, “**Custom Designs**” and “**Optimal (Combined)**” with “**4**” Mixture 1 components and “**2**” Numeric factors:



Next >>

Build an Optimal Split-Plot Design
Process HTC *(page 2 of 4)*

2. Leave the Mix components as “**Easy**” to Change and enter the names, low and high values:

Mixture 1 components: 4

Total: 100 ◉ Horizontal

Units: volume % ○ Vertical

	Name	Change	Low	High
A [Mixture]	water	Easy	60	80
B [Mixture]	methanol	Easy	0	35
C [Mixture]	acetonitrile	Easy	0	20
D [Mixture]	THF	Easy	5	5

Next >>

Build an Optimal Split-Plot Design
Process HTC *(page 3 of 4)*

3. Make the process factors "**Hard**" to change and enter their names, low and high values:

Numeric factors: 2 ○ Horizontal
Categoric factors: 0 ◉ Vertical

	e [Numeric]	f [Numeric]
Name	particle	pore
Units	mm	A
Change	Hard	Hard
Type	Discrete	Discrete
Levels	3	3
L[1]	3.5	100
L[2]	6	200
L[3]	10	300

Discrete factors are numeric factors with specific fixed levels; i.e. discrete levels.

Next >>

Build an Optimal Split-Plot Design
Process HTC *(page 4 of 4)*

4. Design for a quadratic by quadratic model:

Optimal (Combined) Design

Search: Best Optimality: I
Edit model... Quadratic x Quadratic / Scheffe
Blocks: 1
Variance ratio: 1 (0.0 to 1000.0)

Groups
Required groups: 6
Additional groups: 3
Center point groups: 0
Center point group size: 0
Total groups: 9

Runs
Required model points: 36
Additional model points: 9
Center points: 0
Total runs: 45

Next >>

5. Enter the response Height of an Equivalent Theoretical Plate (HETP):

	Name	Units
	HETP	MM

Finish

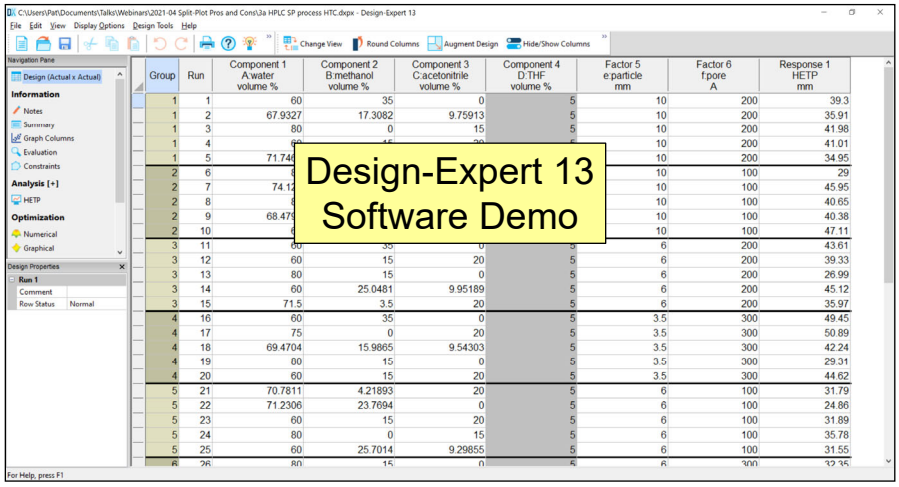
Reverse Phase HPLC
Split-Plot Design

Nine groups defined by combinations of the two HTC process factors.

The three ETC active mixture components vary within the groups.

Gro...	Run	Component 1 A:water volume %	Component 2 B:methanol volume %	Component 3 C:acetonitrile volume %	Component 4 D:THF volume %	Factor 5 e:particle mm	Factor 6 f:pore A	Response 1 HETP mm
1	1	60	35	0	5	10	200	
1	2	67.9327	17.3082	9.75913	5	10	200	
1	3	80	0	15	5	10	200	
1	4	60	15	20	5	10	200	
1	5	71.7468	23.2532	0	5	10	200	
2	6	80	15	0	5	10	100	
2	7	74.125	0.875	20	5	10	100	
2	8	80	0.875	14.125	5	10	100	
2	9	68.4792	17.5737	8.94711	5	10	100	
2	10	60	35	0	5	10	100	
3	11	60	35	0	5	6	200	
3	12	60	15	20	5	6	200	
3	13	80	15	0	5	6	200	
3	14	60	25.0481	9.95189	5	6	200	
3	15	71.5	3.5	20	5	6	200	
4	16	60	35	0	5	3.5	300	
4	17	75	0	20	5	3.5	300	
⋮						⋮		
7	35	60	35	0	5	10	300	
8	36	80	15	0	5	3.5	100	
8	37	80	0	15	5	3.5	100	
8	38	60	15	20	5	3.5	100	
8	39	69.6948	16.7867	8.51855	5	3.5	100	
8	40	60	35	0	5	3.5	100	
9	41	60	23.3935	11.6065	5	3.5	200	
9	42	80	15	0	5	3.5	200	
9	43	74.7203	0.279653	20	5	3.5	200	
9	44	80	0	15	5	3.5	200	
9	45	68.8637	26.1363	0	5	3.5	200	

Reverse Phase HPLC
Analyze and Optimize



Reverse Phase HPLC
Split-Plot Fixed Effects

After crossing the mixture model (ETC or subplot terms) with the process model (HTC or whole-plot terms), all terms have an ETC component, i.e., they are tested against the subplot error.

Reducing the model using Backward selection with the AICc criteria, and correcting for hierarchy yields: →

Fixed Effects [Type III]

Response 1: HETP

Mixture Component coding is **L_Pseudo**.
Sum of squares is **Type III - Partial**

REML (REstricted Maximum Likelihood) analysis
Kenward-Roger p-values

Source	Term df	Error df	F-value	p-value	
Subplot	14	23.20	16.03	< 0.0001	significant
Linear Mixture	2	27.51	3.24	0.0543	
AB	1	24.60	19.91	0.0002	
Ae	1	29.82	5.13	0.0309	
Af	1	23.58	2.01	0.1697	
BC	1	24.57	9.31	0.0054	
Be	1	12.40	0.1196	0.7352	
Bf	1	23.27	5.45	0.0285	
Ce	1	26.58	0.7773	0.3859	
ABe	1	24.35	7.35	0.0121	
Aef	1	29.70	8.42	0.0069	
Ae^2	1	28.94	10.23	0.0033	
Bf^2	1	29.99	3.89	0.0579	
Ce^2	1	29.05	12.33	0.0015	

Reverse Phase HPLC
Split-Plot Variance Components

More experimental error is generated when changing the process factors (the Group Variance) than when the mixture components are changed (the Residual Variance):

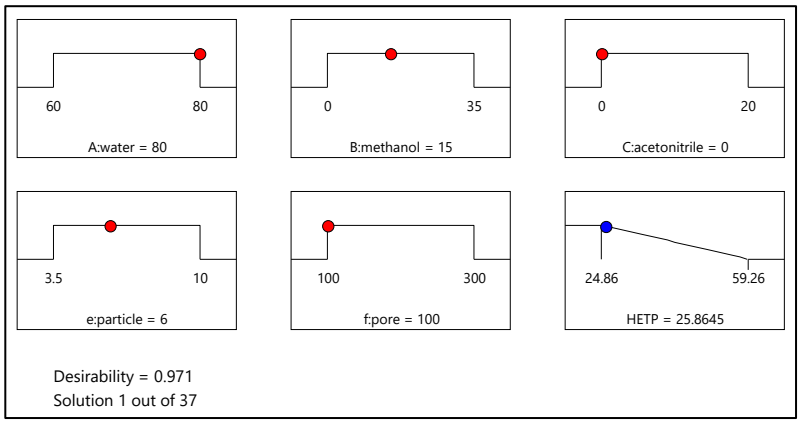
Transform | Model | ANOVA (REML) | Diagnostics | Model Graphs
Fixed Effects [Type III] | Variance Components | Model Selection Log

Variance Components

Source	Variance	Standard Error	95% CI Low	95% CI High
Group	12.33	8.11	-3.56	28.22
Residual	7.41	2.13	4.52	14.29
Total	19.74			

Reverse Phase HPLC
Split-Plot Optimization

Minimize the Height of an Equivalent Theoretical Plate (HETP):



Agenda
Split Plot Pros and Cons

1. Restricting randomization
2. Factorial split plot
   (DNA Amplification)
3. Combined mixture process split plot
   (Reverse Phase HPLC)
4. **Summary**

Split-Plot Design Summary
Advantages

Pros include:

- **Practical**: Randomizing hard-to-change (HTC) factors in groups, rather, than randomizing every run, is much less labor and time intensive. *(Carefully consider the cost savings of the split-plot design vs. the loss of power to detect HTC factor effects.)*
- **Flexible**: Factors that naturally have large experimental units can be easily combined with factors having smaller experimental units. *(Remember the DNA amplification: the plates are the "large units" and the 96 cells, on the plates, are the "small units".)*
- **More powerful**: Tests for the subplot effects from the easy-to-change (ETC) factors generally have higher power due to partitioning the variance sources.

Split-Plot Design Summary
Disadvantages

Cons include:

- **Unfamiliar**: They result in differing errors for whole plot versus subplot terms. REML should be used for a proper statistical test. If you apply a standard ANOVA, it may select too many whole-plot factors and too few sub-plot factors.
- **Less powerful**: Tests for the hard-to-change factors are less powerful, having a larger variance to test against and fewer changes to help overcome the larger error.
- **Different**: HTC and ETC factor effects are tested against different estimated noise. This can result in large HTC effects not being statistically significant, whereas small ETC effects are significant even though they may not be practically important.

Split-Plot Design Summary
Facts to Remember

Things to remember:

- Not randomizing the groups increases the risk of confounding the group effects with a lurking variable.
- When a group change occurs, factors that are not changing should be reset.
- If factors are not reset, then it is another restriction on the randomization.

Split Plots Pros and Cons:
Dealing with a Hard-to-Change Factor

Reminder, this presentation is posted at:
www.statease.com/webinar.html

If you have additional questions email them to:
stathelp@statease.com

Stat-Ease Training:
Sharpen Up Your DOE Skills

- Modern DOE for Process Optimization
- Mixture Design for Optimal Formulations
- Designed Experiments for Pharma, Food, etc. (teams)

Individuals	Teams (6+ people)
Improve your DOE skills	Choose your own date & time
Ideal for novice to advanced levels	Customize via company case studies

Learn more & then register:
www.statease.com

Contact:
workshops@statease.com

StatEase®
statistics made easy®

Make the most from every experiment!℠

Thank You for Attending

Pat Whitcomb