Design and Analysis of Order-of-Addition Experiments

33rd QPRC

June 14–16, 2016

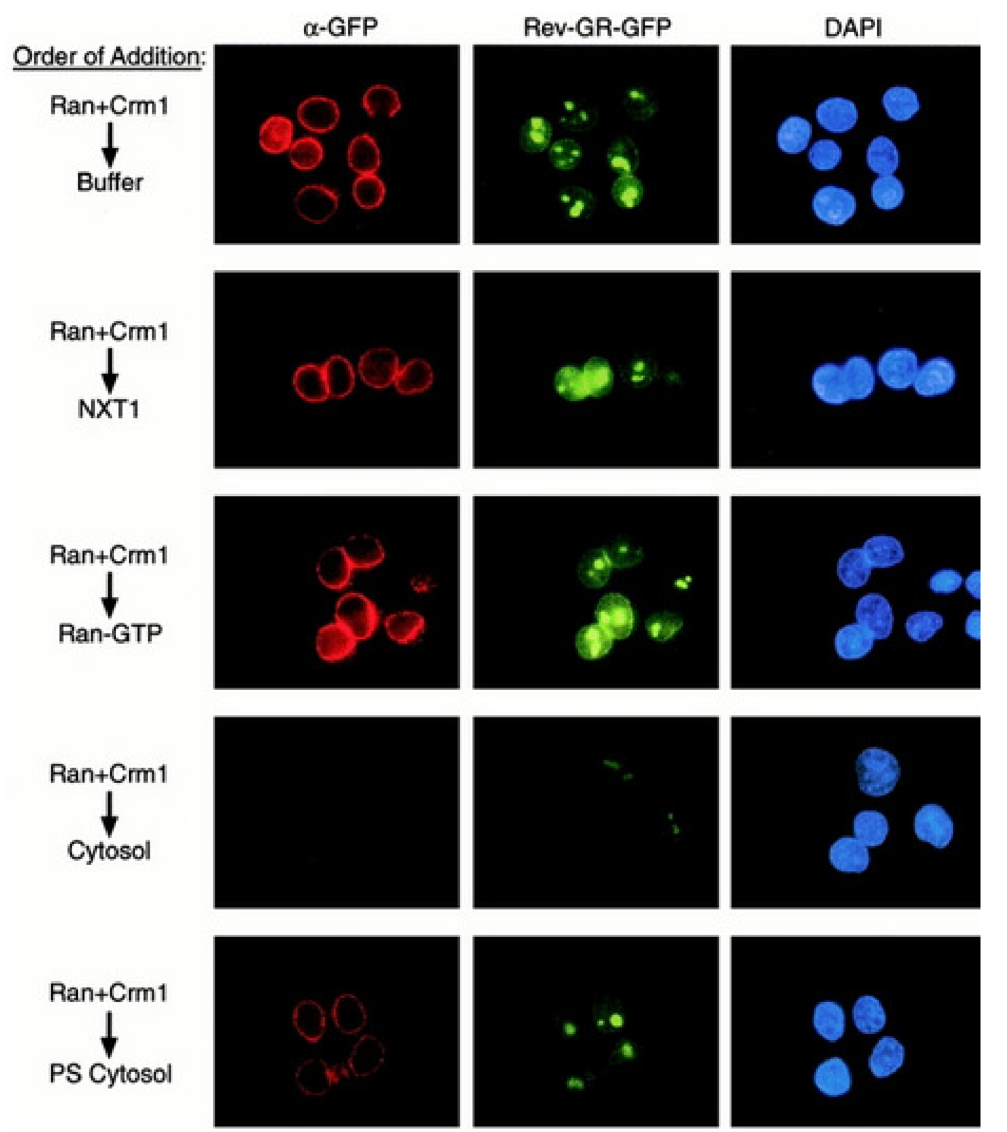
Joseph G. Voelkel, RIT


TEA
MILK
?

- Food
- Chemical
- Film

Order of addition experiment indicating the requirement for Ran, Crm1, and NXT1 early in the export pathway.

J of Cell Biology, 2001



short (1 hr) exposures to PDGF are sufficient to stimulate DNA synthesis, provided the cultures are incubated with forskolin after PDGF treatment (PF → F, P → F). Unlike PDGF, forskolin is required continually throughout the 20 hr lag period for growth to occur. Other permutations of the PDGF/forskolin order of addition (PF → P, F → P) do not result in Schwann cell growth.

Journal of Neuroscience, February 15, 2001

3.3 Order of addition

Order of addition can influence the signal generated to a large extent. The optimal order in which assay components interact should always be determined empirically. Keep in mind that some binding partners may interfere with the association of other binding partners if allowed to interact in the wrong order.

PerkinElmer
User's Guide to Alpha Assays:
Protein-Protein Interactions

Order-of-Addition Experiments

- Lady Tasting Tea: $m = 2$ components

$c_1 \rightarrow c_2$ 1: 1 2
$c_2 \rightarrow c_1$ 2: 2 1

- Consider $m = 3$ components
- $m!$ permutations
- *Order-of-Addition* (*OofA*) experiments.

1:	1	2	3
2:	1	3	2
3:	2	1	3
4:	2	3	1
5:	3	1	2
6:	3	2	1

1:	1 2 3
2:	1 3 2
3:	2 1 3
4:	2 3 1
5:	3 1 2
6:	3 2 1

Some Questions *(The Talk!)*

1. What are the factors for an m-component OofA experiment? Levels?
2. Say m = 4, 5, or 6 components: reasonable way to select a fraction of all $m!$ tc's?
3. Can we add process variables to the experiment in a natural way?
4. How can we analyze such experiments?

Earlier Work?

- Williams (1949, 1950)
  - Designs *not* created for OofA
    - But have been used as such.
  - Designs for cross-over studies—treatments in succession to the same e.u.
  - Interest: how to handle *residual* (carry-over) *effects* ?
    - For single preceding treatment
    - For any n of treatments, ignoring interaction effects
    - For two preceding treatments and their interaction

Williams

- Balance: "single preceding treatment" case
- m even: m runs ($m \times m$ LS)
  m odd: $2m$ runs (2 LS's)
- Questions:
  - Do these create "optimal" OofA designs?
  - "good" designs?...

1:	1	2	4	5	3
2:	2	3	5	1	4
3:	3	4	1	2	5
4:	4	5	2	3	1
5:	5	1	3	4	2
6:	1	3	2	5	4
7:	2	4	3	1	5
8:	3	5	4	2	1
9:	4	1	5	3	2
10:	5	2	1	4	3.

Factors? Fraction?

1. What are the factors for an m-factor OofA experiment? Levels?
2. Say m = 4, 5, or 6 components: reasonable way to select a fraction of all m! tc's?...

Van Nostrand (1995)

- Factors for the design?
- Idea: consider $\binom{m}{2}$ *Pseudo-factors (PF's)*
- $m = 3$ example, all runs

		F1<2	F1<3	F2<3
1:	1 2 3	1	1	1
2:	1 3 2	1	1	0
3:	2 1 3	0	1	1
4:	2 3 1	0	0	1
5:	3 1 2	1	0	0
6:	3 2 1	0	0	0

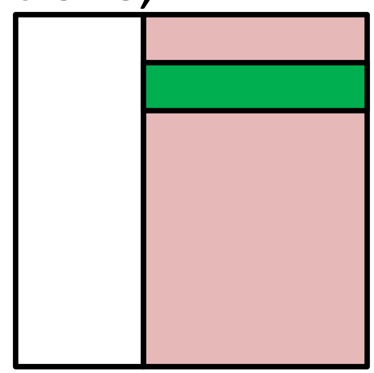
- Note: not all possible combinations of PF's are possible
- Transitive property
  If 1<2 & 2<3 ⇒ 1<3

	F1<2	**F1<3**	F2<3
1 2 3	1	**0**	1.

Van Nostrand's vs. our approach

2. Say m = 4, 5, or 6 components: reasonable way to select a fraction of all $m!$ tc's?

- Van Nostrand
  - Start with 2^{k-p} designs (PF's) with certain WL patterns
  - If $m = 5$, then $k = 10$
  - Note: $2^{10} = 1024$, but only $5! = 120$ valid PF comb'ns
  - For the fraction, only keep valid PF combinations, getting $\approx N$
  - Use mean VIF as goodness measure
- Our approach
  - Start with full $m!$ runs
  - Generate corresponding PF combinations
  - Find optimal N-run design using χ^2 (balance) or D-criterion as goodness measure
    - Balance??...



2. Say m = 4, 5, or 6 components: reasonable way to select a fraction of all $m!$ tc's?

Balance?

To do this, first consider *Orthogonal Arrays.*

OA of strength t

- An $N \times k$ array A with k factors each at s levels is an OA (Orthogonal Array) with strength t if every $N \times t$ sub-array of A contains all possible t-tuples the same number of times
- OA with $t = 2$ often simply called an OA

- 2^{5-2} : OA with $t = 2$
  – Res III
- 2^{4-1} : OA with $t = 3$
  – Res IV.

	A	B	C	D	E
1	0	0	0	0	0
2	1	0	0	1	1
3	0	1	0	1	0
4	1	1	0	0	1
5	0	0	1	0	1
6	1	0	1	1	0
7	0	1	1	1	1
8	1	1	1	0	0

A	D	Freq
0	0	2
1	0	2
0	1	2
1	1	2.

OofA OA of strength t?

- $m = 3$ example:

		F1<2	F1<3	F2<3
1:	1 2 3	1	1	1
2:	1 3 2	1	1	0
3:	2 1 3	0	1	1
4:	2 3 1	0	0	1
5:	3 1 2	1	0	0
6:	3 2 1	0	0	0

- Note: we will consider the *PF's as the factors* in the OofA design

- Problem: even *full* design not balanced

F1<2	F1<3	Freq
1	1	2
0	1	1
0	0	2
1	0	1

- So even full design not a (standard) OA of even strength 2.

OofA OA's of strength t

Definition

An $N \times k$ array A with $k = \binom{m}{2}$ PF's is an OofA OA with strength t if every $\mathrm{N} \times t$ sub-array of A contains all possible t-tuples *in the same proportions as the full array with* m! *runs*.

OofA patterns for $t = 2$

- Ex: $m = 4 \Rightarrow \binom{4}{2} = 6$ PF's

  $\Rightarrow \binom{6}{2} = 15$ pairs of PF's.

		F1<2	F1<3	F1<4	F2<3	F2<4	F3<4
1:	1 2 3 4	1	1	1	1	1	1
2:	1 2 4 3	1	1	1	1	1	0
3:	1 3 2 4	1	1	1	0	1	1
4:	1 3 4 2	1	1	1	0	0	1
...							
21:	4 2 1 3	0	1	0	1	0	0
22:	4 2 3 1	0	0	0	1	0	0
23:	4 3 1 2	1	0	0	0	0	0
24:	4 3 2 1	0	0	0	0	0	0

OofA patterns for $t = 2$

- Ex: $m = 4 \Rightarrow \binom{4}{2} = 6$ PF's $\Rightarrow \binom{6}{2} = 15$ pairs of PF's
- 2 non-isomorphic patterns ($m = 4$; 24 runs)

Synergistic: 12

		F1<2	
		0	1
F1<3	0	**8**	**4**
	1	**4**	**8**

Independent: 3

		F1<2	
		0	1
F3<4	0	**6**	**6**
	1	**6**	**6**

OofA patterns for $t = 3$

- Same idea—more complex…

Some simple OofA-OA results

- OofA OA of $t = 2$
  - can prove: for $m > 3$ need $N = 0 \; mod \; 12$
- OofA OA of $t = 3$ for $m > 3$ need $N = 0 \; mod \; 24$

$\Rightarrow$ Williams' design for $m = 5$ with $N = 10$ is not an OofA-OA.

Fraction?

2. Say m = 4, 5, or 6 components: reasonable way to select a fraction of all $m!$ tc's?

- How to *select* a good fraction? Consider $t = 2$ only
  - When $N = 0 \ mod \ 12$?
  - In other cases?...

Selecting Good Fractions

- When $N = 0\ mod\ 12$
  - Conjecture: there is a closed-form way …
- In general
  - Use exchange algorithms
  - Average χ^2 (balance…) or D-criterion.

Some results: $m = 5, N = 10$

Williams

- From earlier…

						F1<2	F2<4
1:	1	2	4	5	3	1	1
2:	2	3	5	1	4	0	1
3:	3	4	1	2	5	1	1
4:	4	5	2	3	1	0	0
5:	5	1	3	4	2	1	0
6:	1	3	2	5	4	1	1
7:	2	4	3	1	5	0	1
8:	3	5	4	2	1	0	0
9:	4	1	5	3	2	1	0
10:	5	2	1	4	3	0	0

OofA-OA attempt

- Best result, via χ^2

						F1<2	F1<3
1:	1	2	5	3	4	1	1
2:	1	4	3	5	2	1	1
3:	2	1	5	4	3	0	1
4:	2	3	4	5	1	0	0
5:	3	4	1	2	5	1	0
6:	3	5	2	1	4	0	0
7:	4	3	2	1	5	0	0
8:	4	5	1	2	3	1	1
9:	5	1	3	4	2	1	1
10:	5	4	2	3	1	0	0.

Some results: $m = 5, N = 10$

Williams

- Aver $\chi^2 = 0.8$ (45 tables)
- Worst $\chi^2 = 3.2$

Actual Worst: F1<2 vs F2<4

F1<2 \ F2<4	0	1
0	3	2
1	2	3

Expected

	0	1
0	1.7	3.3
1	3.3	1.7

OA OofA attempt

- Aver $\chi^2 = 0.4$
- Worst $\chi^2 = 0.8$

Actual Worst: F1<2 vs F1<3

F1<2 \ F1<3	0	1
0	4	1
1	1	4

Expected

	0	1
0	3.3	1.7
1	1.7	3.3.

From χ^2 to D

$\boldsymbol{\chi^2}$	$\boldsymbol{D}$
• Homebrewed algorithm – Difficulty due to restricted tc's	• Well-studied algorithms
• Slow-running	• Fast

But: will D give "good" designs??...

From χ^2 to D

- Theorem: An OofA OA of $t = 2$ exists ($\chi^2 = 0$ for its PF design matrix $\mathbf{P}$) in N runs if and only if its associated $\mathbf{X} = [\mathbf{1}|\mathbf{P}]$ (if of full rank) has D-efficiency = 1.
  - D-efficiency measured wrt the full OofA design
- Recall:
  - $\chi^2 = 0$ only if $N = 0 \; mod \; 12$.

Some $m = 4,5$ results

- $m = 4,5, N = 12$
  - D-eff'y 1
  - $m = 4$: two non-isomorphic designs found (≈ equal)
  - $m = 5$: only one design found
- $m = 4,5, N = 18$
  - D-eff'y of top 1 of 100 (SAS Optex) and of top 1 of 1000 (R AlgDesign) = 0.97
  - But not isomorphic...

Some $m = 5{,}6$ results

- $m = 5, N =$24
  - D-eff'y = 1 for top 12 of 100 (SAS Optex) and for top 1 (at least) of 20 (R AlgDesign)
  - But none of these 13 were isomorphic (but $\approx$ equal)
- $m = 6, N =$24
  - D-eff'y = 1 for top 1 of 100 (SAS Optex) and for top 1 (at least) of 100 (R AlgDesign)
  - Not isomorphic (but $\approx$ equal.)

Adding Process Variables

- Can be done naturally—and well
- Example: $m = 5$, N =24, with 2^4 (main effects)
  - D-eff'y = 1
- Example: $m = 5$, N =24, with 3×2^2 (main effects)
  - D-eff'y = 1 (4000 iterations…)

Analysis

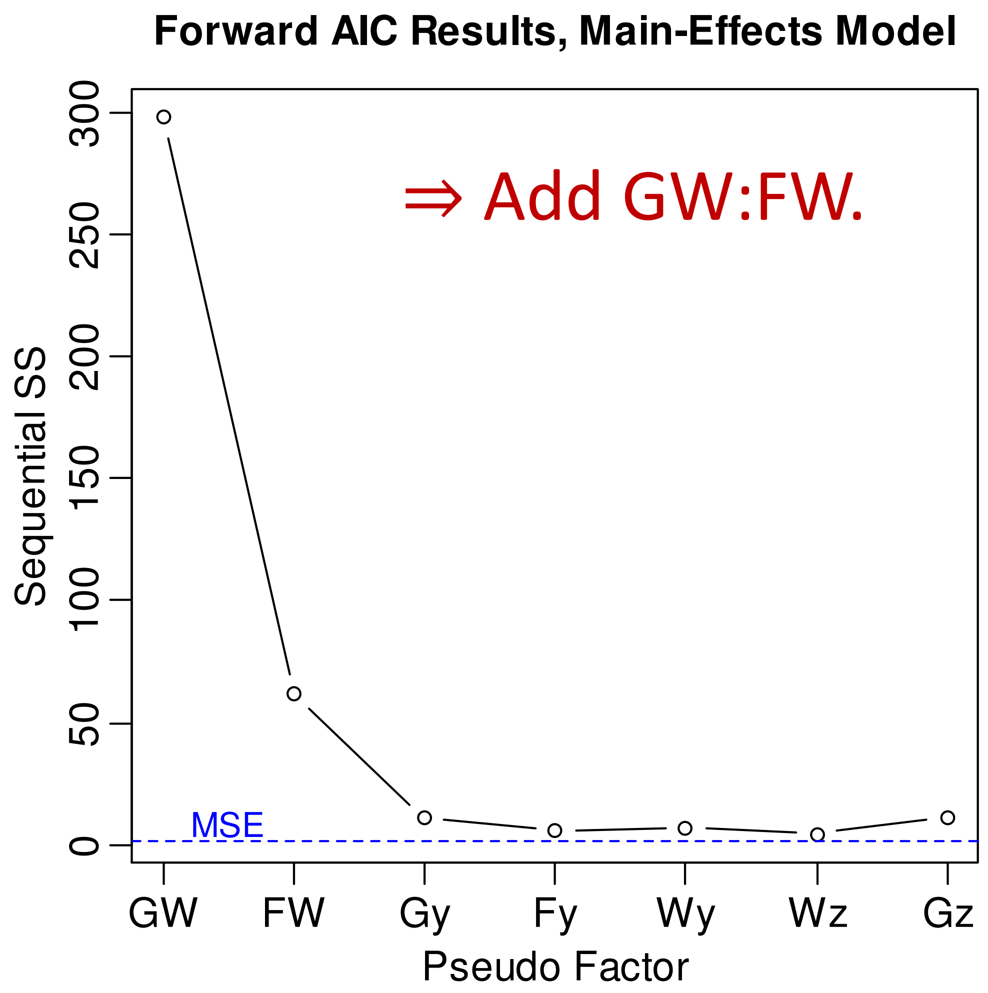
- Example (ala Van Nostrand)
- 6 components
  - G Gunpowder
  - F Flaming Rag
  - W Water
  - x,y,z
- Response: measure of exothermic reaction
- $m = 6, N = 24.$

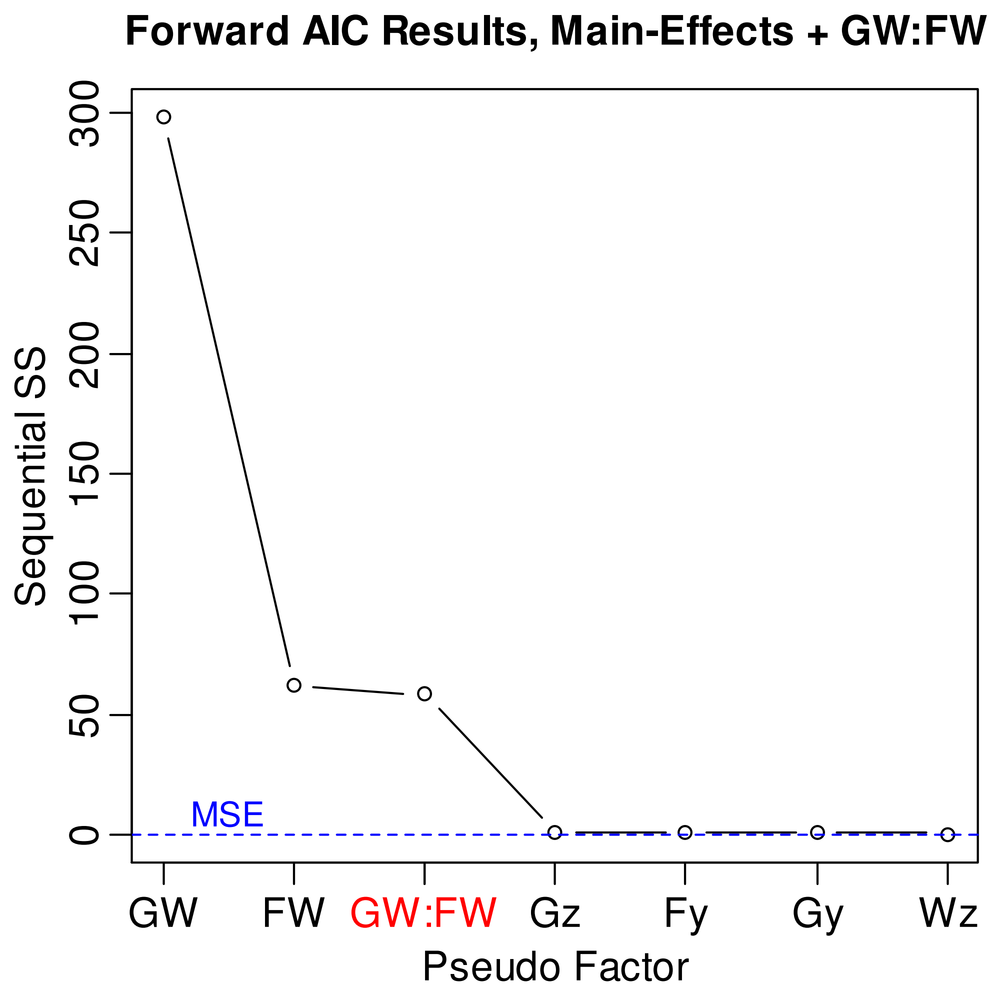
Caution!

- In many physical situations
  - Interactions may be expected
  - “Reaction works only if A<B and B<C”
  - So, not simply main effects, which is what $t = 2$ handles directly
- Idea: start with main effects; later, add interactions
- Two strategies to be shown…

Strategy 1: Regression

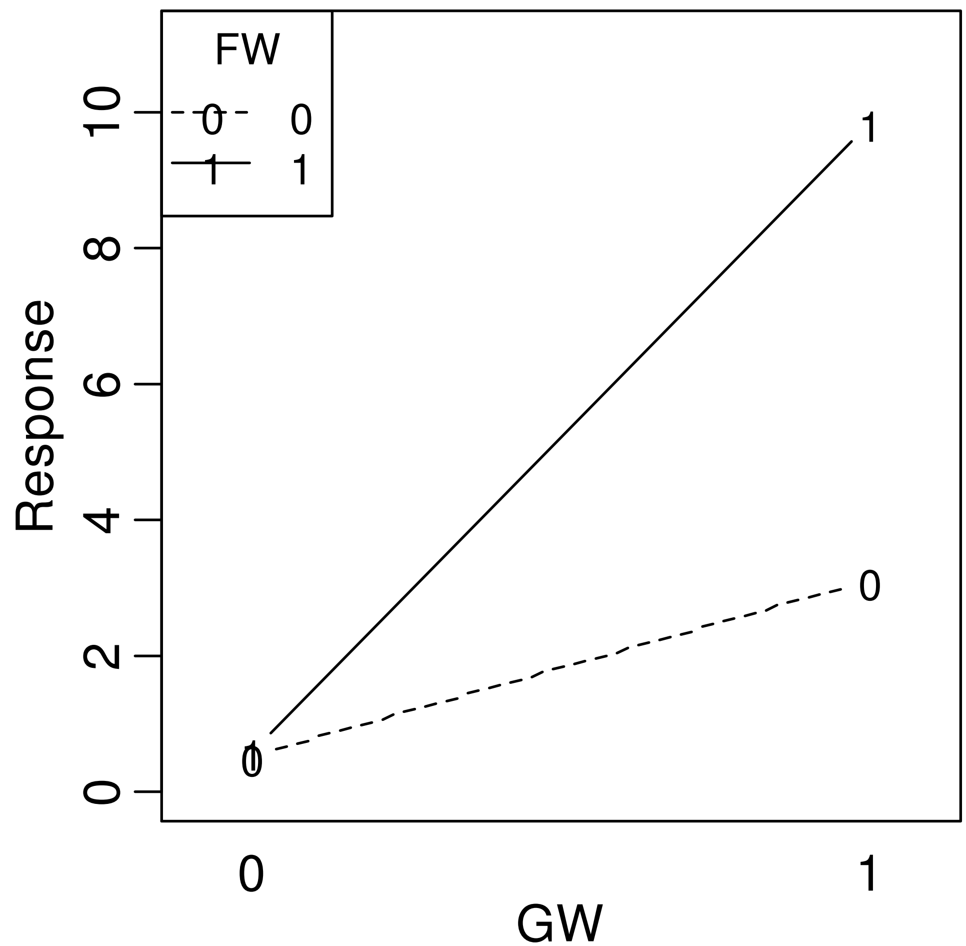
- Step 1. Forward selection: main-effects model in PF’s
  - GF GW Gx Gy Gz FW Fx Fy Fz Wx Wy Wz xy xz yz
- Step 2. Forward selection: reduced (?) model + 2fi(s).


Forward AIC Results, Main-Effects Model
⇒ Add GW:FW.
MSE
Sequential SS
0
50
100
150
200
250
300
GW
FW
Gy
Fy
Wy
Wz
Gz
Pseudo Factor


Forward AIC Results, Main-Effects + GW:FW
Sequential SS
0
50
100
150
200
250
300
MSE
GW
FW
GW:FW
Gz
Fy
Gy
Wz
Pseudo Factor

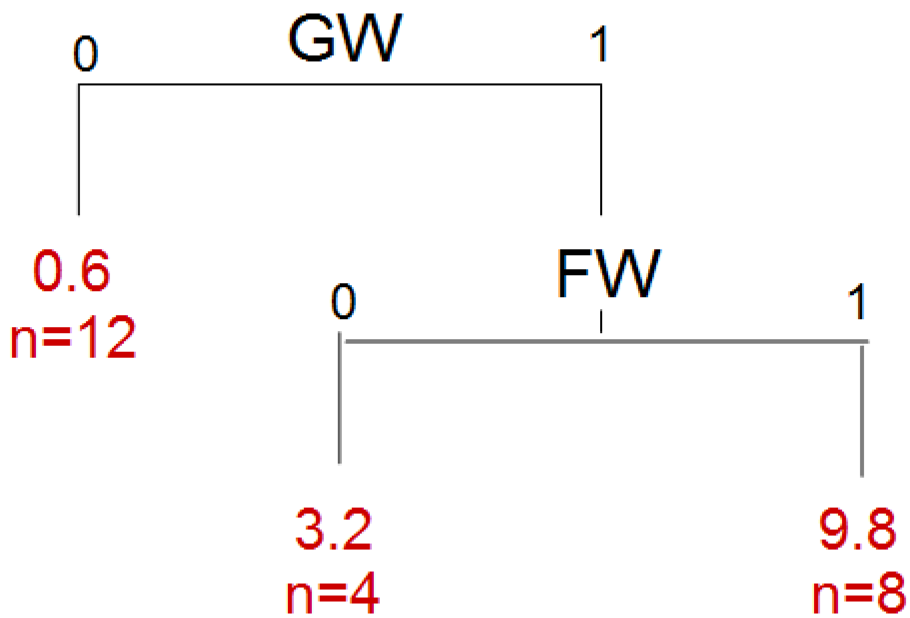
Model Summary

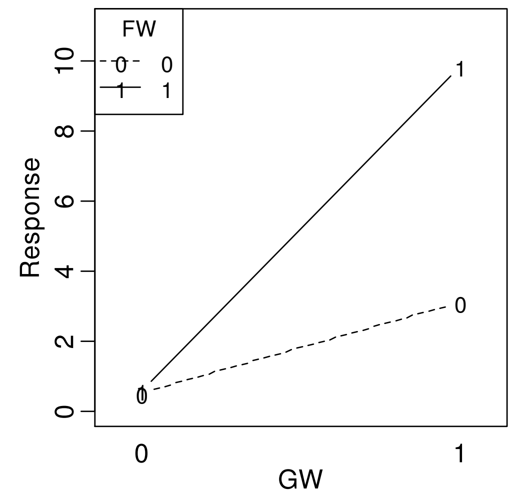
– G	Gunpowder
– F	Flaming Rag
– W	Water



Strategy 2: CART

- Automates search for (nested) interactions





Summary

- Order-of-Addition Experiments—not unusual
- Little information available to construct good designs
- The factors in an m-factor OofA experiment
- Optimal ways to select a fraction of all $m!$ tc's
- Addition of process variables to an OofA design
- The analysis of such experiments

Thank you!