

# Design of Experiments in Ecological Systems: some methods and issues

Jorge Luis Romeu, Ph.D.

Research Professor, Syracuse University

Email: [jlromeu@syr.edu](mailto:jlromeu@syr.edu);

Dept. Mechanical & Aerospace Engineering, and

EQS STAR Center for Environmental Quality

Seminar Series.

February 27, 2009

# Outline

- Problem statement
- Proposed solution
- Simulation Example
- Some applicable DOEs
- Implementation problems
  - and their consequences
- Other modeling issues
- Conclusions

# Problem Statement

- Given a network of water masses
  - E.g. Finger Lakes, Great Lakes, Gulf Ports
- Optimize some performance measures:
  - Wetland Preservation; water use; exports, etc
- Subject to the Key set of economic, social, labor, political, environmental, climatic, cultural, etc. problem variables/constraints
- Maintaining employment, health etc. levels

# Modeling Methods

- Theoretical (physics law, or relation)
  - But, can we come up with such equation?
- Empirical (regression)
  - But, can we find enough data to implement?
- Discrete Event Simulation
  - Don't need to relax model assumptions
  - Can include complex interactions
  - But run time can be very long!

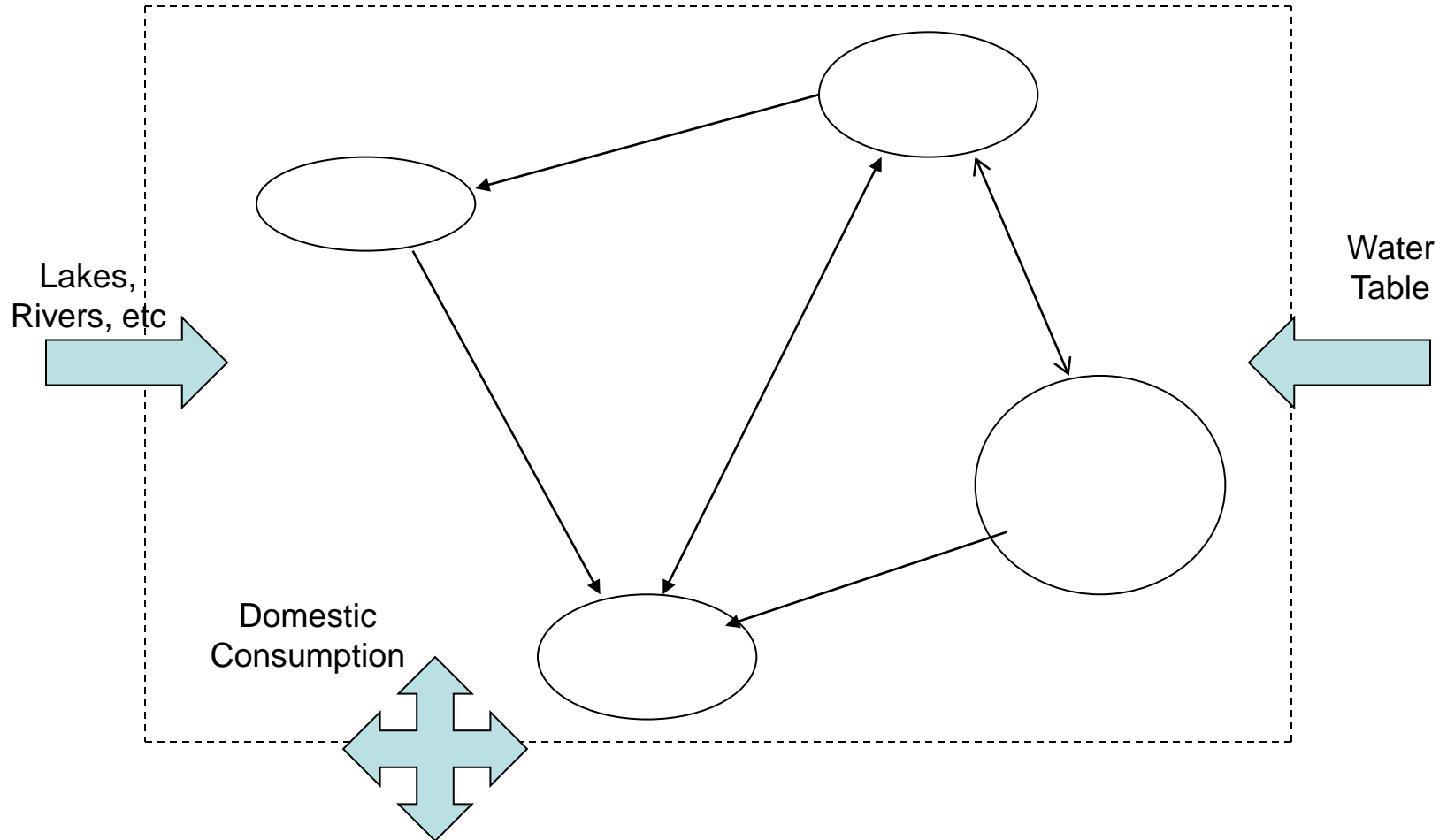
# Simulation Modeling Problem

- Consider a Complex Simulation Model
- Main Issues:
  - Too many variables to analyze
  - Complex Dynamical System structure
- Proposed solution: Find Key Variables
  - Via Design of Experiments (DOE)
  - Derive a set of simpler Meta Models
  - Use them as proxies for the Full Model

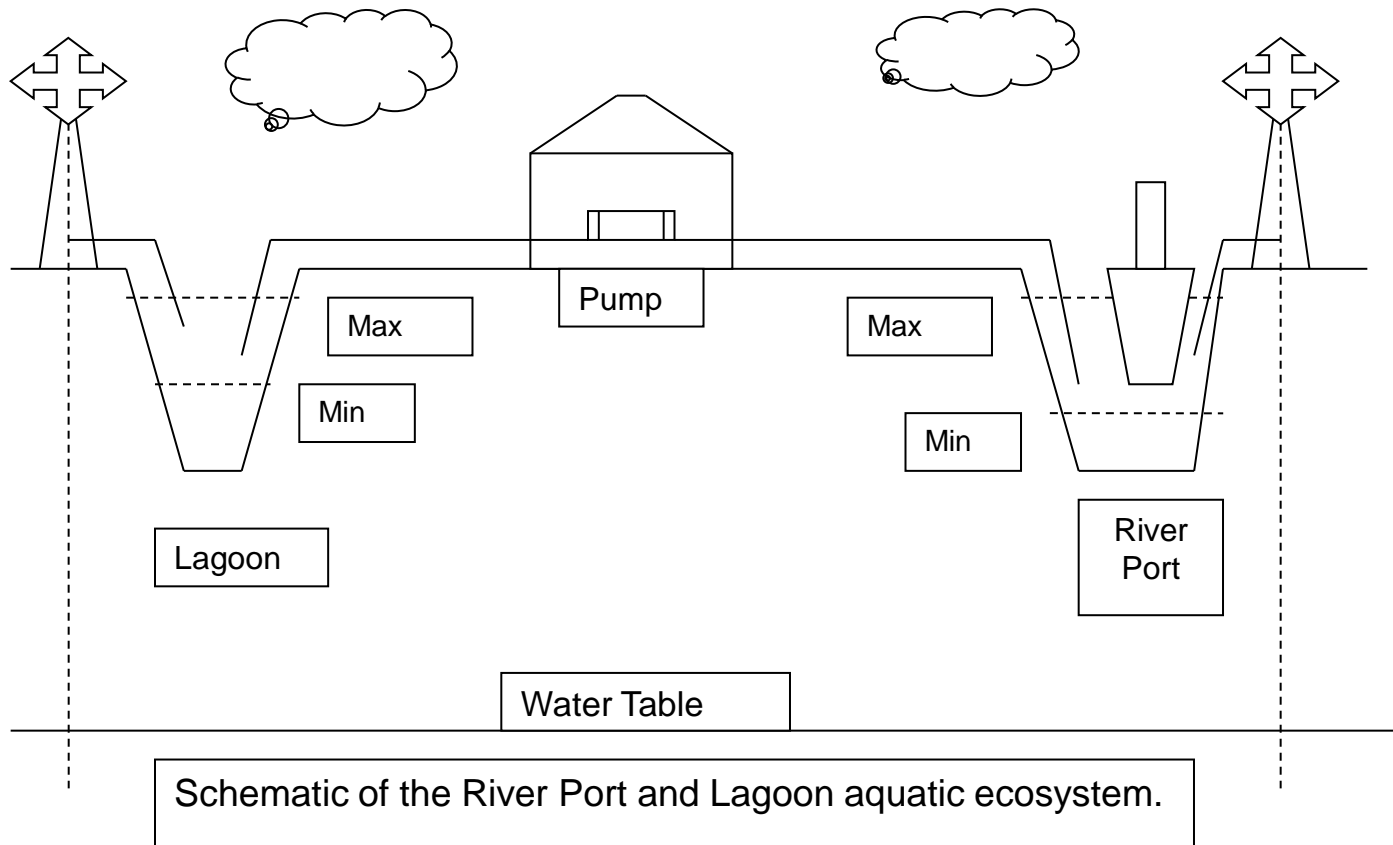
# Simulation Example

- Given a network of water masses
  - For both, civilian and industrial use
- Optimize some performance measures
  - e.g. operational, social, political, ecological
- Subject to a set of (conflicting) political, labor, socio-economic, etc. constraints
  - Maintaining levels of production, employment
  - Tax revenues, social services, economic, etc.

# A Network of Interconnected Water Masses



# Example: River Port w/Lagoon





# Controlled Variables: Economic

- Replenishing Levels (MIN)
- Reservoir Capacity (MAX)
  - Ordering Schedule
  - Transfer Policy
  - Usage Policy
  - Shortage Policy
    - Profitability
- System's Initial Conditions

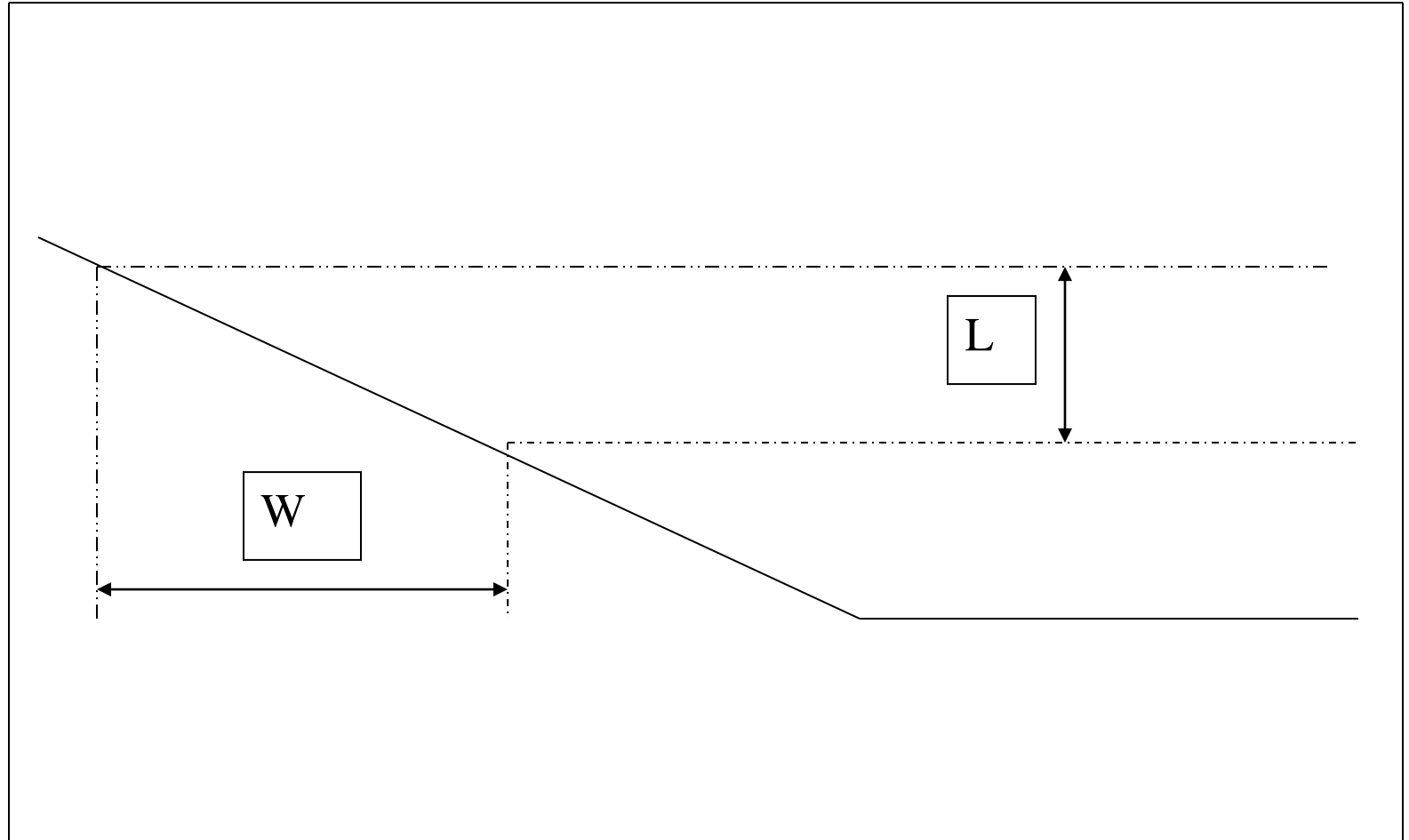
# Controlled Variables: Social

- Allocation to each sector
  - Size of the Reservoirs
    - Transfer Policy
- Generation of electricity
- Hospitals and schools
- Transportation uses
  - Recreation uses

# Controlled Variables: Ecologic

- Wetland Area
- Wetland Depth
- Transfer Speed
- Water Table Use
  - Pollution Level
- Fish/Foul Population

# WetLand v. Level



# Uncontrolled Variables

- ECONOMIC
- Political issues
- Labor issues
- Water Theft
- Water Leaks
- Markets
- Financial
- ECOLOGIC
- Evaporation
- Temperature
- Salinity
- Reproduction
- Weather
- Water Table

# And Associated Costs

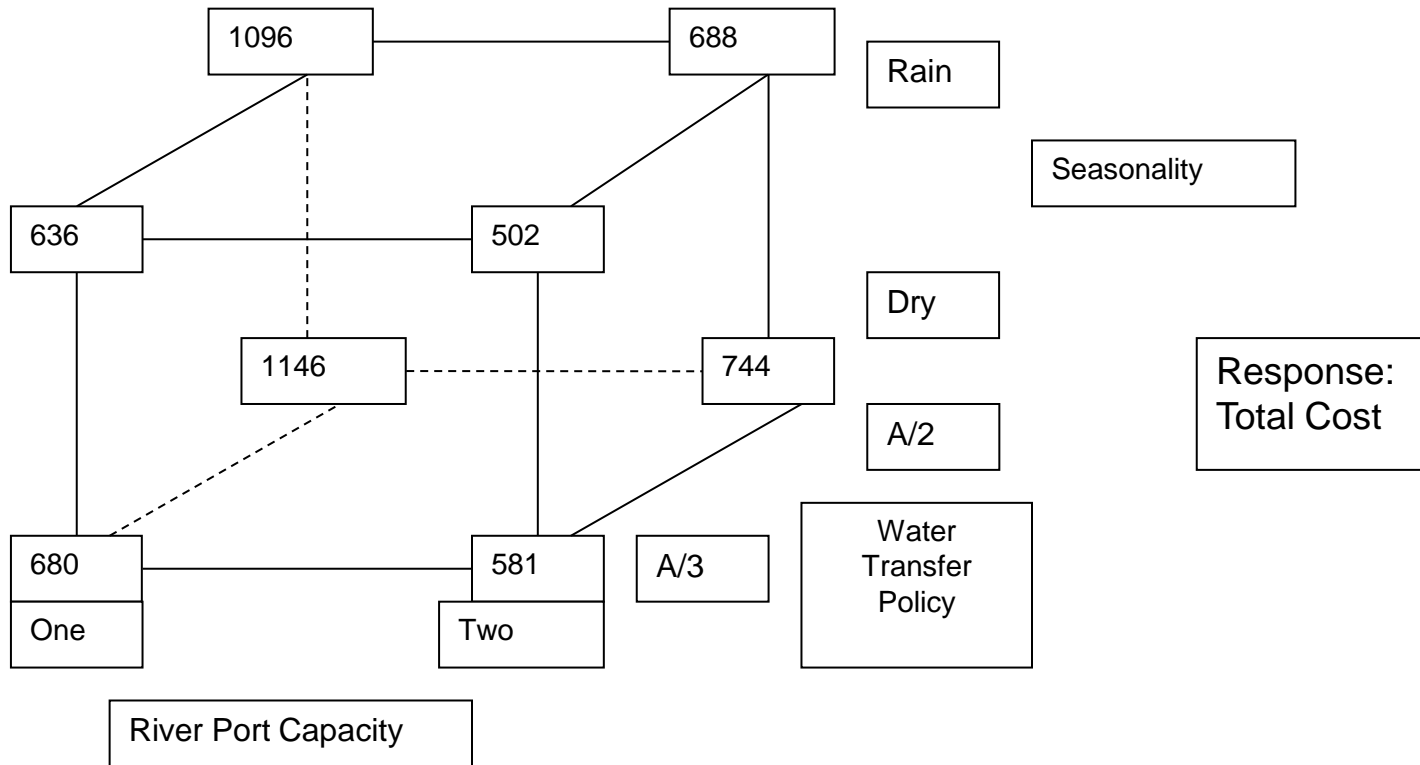
- Of Importing Water from other places
- Transferring from Social to Economic
- Allocation to various constituencies
- Of Water shortages and rationing
- Indirect costs (labor, political, social)
- Ecological costs (degradation, loss)
- Total costs (compound response)

# Additional Model Uses

Multi-criteria (ecological, social, economic, etc.) system responses (consolidating elements in the system) can be obtained, by combining (say  $k$ ) contrasting and competing individual responses into a single, complex one. The (linear) combinations formed quantify the contrasting policies and philosophies of the different constituencies. Comparisons of competing and contrasting policies, via the simulation model results, can help diverse constituencies to rationally discuss their differences, and better reach a consensus.

# Simple DOE Example

Complete Factorial Experiment for the Simulation





# DOE Model Limitations:

- Analyzes limited variables (here,  $k=3$ )
  - For,  $2^K$  Factors/Interacts are generated
- The Effect of Interaction, when  $k > 2$ 
  - Can affect results, if present and strong
- Need to find Robust Responses
  - Handling specific “noise variables”
- Need to Identify “significant few” variables
  - To reduce model Size, maintaining Info level.

# And their Consequences ...

- If large number of factors to analyze
  - Strong factor interaction may exist
  - Dependent on the model structure
  - Requires special methods for analysis
- Different objective of models derived:
  - To describe/study, forecast or control
- Robust Parameter analysis capability
  - To derive a response equation that is
  - Resilient to “noise” or uncontrolled factors

# Methods for Key Variable Id

- Full Factorial Designs
- Fractional Factorial Designs
- Plackett-Burnam Designs
- Controlled Sequential Bifurcation
- Latin Hypercube Sampling
- Other modeling approaches
  - Bayesian, Hierarchical, Taguchi, PCA, etc.

# Full Factorials

- Most expensive (in time and effort)
  - Prohibitive with current number of factors
- Most comprehensive information
  - Provides info on all factor interactions
- Two Examples with a  $2^3$  Full Factorial
  - First case: mild interaction (AB only)
  - Second: strong and complex interaction
  - Notice how the Model-Estimations vary

# 2<sup>3</sup> Full Factorial DOE: Variables Used

- A = Replenishing Levels (MIN)
- B = Reservoir Capacity (MAX)
  - C = Transfer Policy
- Mild interaction assumed
  - A\* B only

**Full Factorial Experiment 2<sup>3</sup>**

Run	A	B	C	AB	AC	BC	ABC	Avg.
1	-1	-1	-1	-1	1	1	1	-1.07
2	1	-1	-1	-1	-1	-1	1	3.72
3	-1	1	-1	-1	-1	1	-1	-0.58
4	1	1	-1	-1	1	-1	-1	12.04
5	-1	-1	1	1	1	-1	-1	7.75
6	1	-1	1	1	-1	1	-1	15.45
7	-1	1	1	1	-1	-1	1	11.09
8	1	1	1	1	1	1	1	18.31
TotSum								66.71
Effect	8.08	3.75	9.62	1.84	-0.62	-0.65	-2.08	

**Regression Estimations**

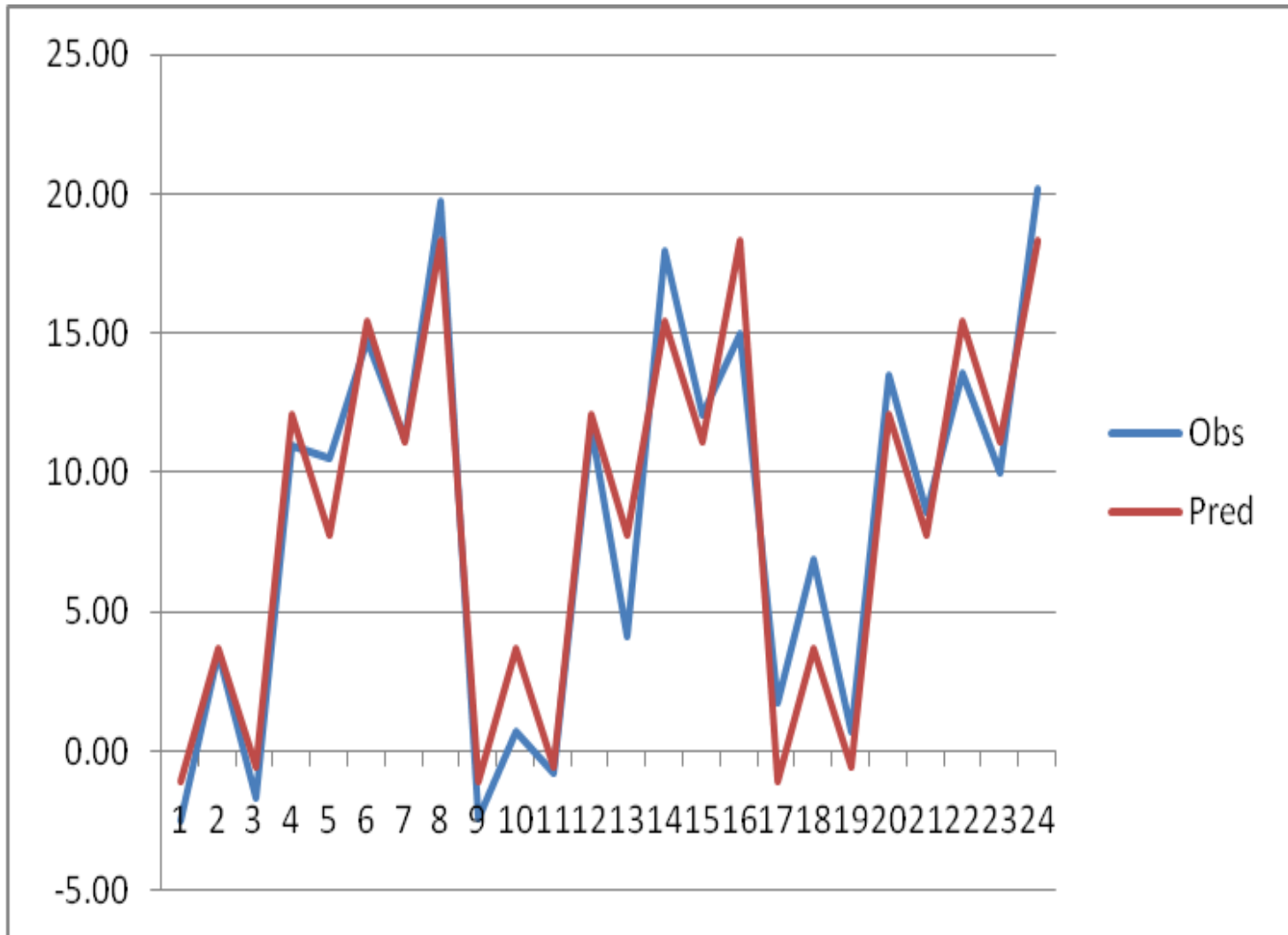
RegCoef	A	B	C	AB	b0
Estimat.	4.04	1.88	4.81	0.92	8.34
TRUE	4	2	5	1	10

**Meta Model:**  $Y_{ijkl} = 8.33 + 4.04A + 1.88B + 4.81C + 0.92AB$

**True Model:**  $Y = 10 + 4*A + 2*B + 5*C + AB + \epsilon$

Mild Interaction (AB only)

# Meta Model Re-creation ability: mild interaction.



# Fractional Factorials

- Analyzes only a Fraction of the Full
  - Reduces substantially time/effort
  - Confounding of Main Effects/Interactions
  - If Interactions present, this is a problem
  - Only for Powers of Two (no. of runs)
- Numerical Example: half fractions
  - Of the previous Full Factorial –and others
  - Assess Model-Estimation agreement



# Fractional Factorials

First Fraction: L1

Run	A	B	C=AB	Avg.
1	1	-1	-1	-0.33
2	-1	1	-1	-0.33
3	-1	-1	1	-0.33
4	1	1	1	1.00
<b>TotSum</b>				0.00
<b>Effect</b>	7.429	3.130	11.460	
<b>Signif.</b>	No	No	Yes	

$$Y_1 = 7.3 + 3.71A + 1.57B + 5.73C^*$$

C \*: Factor C is confounded with AB

Second Fraction: L2

Run	A	B	C=AB	Avg.
1	-1	-1	-1	-1.00
2	1	1	-1	0.33
3	1	-1	1	0.33
4	-1	1	1	0.33
<b>TotSum</b>				0.00
<b>Effect</b>	8.728	4.375	7.784	
<b>Signif.</b>	Yes	No	Yes	

$$Y_2 = 8.33 + 4.36A + 2.18B + 3.89C^*$$

Untangling the Confounded Structure

<b>(L1+L2)/2</b>	<b>8.079</b>	<b>3.753</b>	<b>9.622</b>
<b>(L1-L2)/2</b>	-0.649	-0.623	1.838
<b>Effects</b>	<b>8</b>	<b>4</b>	<b>10</b>

Notice how, by averaging both Half Fraction results, we obtain the Full Factorial results again.

$$\text{True Model: } Y = 10 + 4^*A + 2^*B + 5^*C + AB + \varepsilon$$

# Again $2^3$ Full Factorial: Same Variables Used, but now With Strong Interaction

- A = Replenishing Levels (MIN)
- B = Reservoir Capacity (MAX)
  - C = Transfer Policy
- Strong interaction assumed
  - A\*B, A\*C, B\*C
  - Overall: A\*B\*C

# Full Factorial: Complex, Strong Interaction

Model Parameters							
Variables	A	B	C	AB	AC	BC	ABC
RegCoef	3	-5	1	-12	8	-10	-15
RegEstim	1.94	-4.38	1.73	-12.14	7.34	-10.52	-15.26
MainEffEst	3.88	-8.76	3.47	-24.28	14.68	-21.05	-30.51
MainEffects	6	-10	2	-24	16	-20	-30
Var. of Model		12.5173		StdDv		3.53799	
Var. of Effect		2.0862		StdDv		1.44437	
Student T (0.025DF)		2.47287					
C.I. Half Width		3.57177					
Factor	A	B	C	AB	AC	BC	ABC
Signific.	Yes	Yes	No	Yes	Yes	Yes	Yes

## True Model and Estimated Meta Model:

$$Y = 3A - 5B + C - 12AB + 8AC - 10BC - 15ABC$$

<b>RegEstim</b>	1.94A	-4.38B	1.73C	-12.14AB	+7.34AC	-10.52BC	-15.26ABC
-----------------	-------	--------	-------	----------	---------	----------	-----------

# Corresponding Half Fractions

Run	Half			Analysis:						
	A	B	First C=AB	Half(a)			Avg.	Var	Model	
2	1	-1	-1	Y1	Y2	Y3	-15.87	0.59	-14	
3	-1	1	-1	7.18	9.21	5.28	7.22	3.87	6	
5	-1	-1	1	-16.75	-19.75	-22.02	-19.51	6.97	-22	
8	1	1	1	-31.61	-27.62	-33.04	-30.76	7.89	-30	
<b>TotSum</b>				-56.21	-54.7	-65.82	-58.91	19.32		
<b>Effect</b>	-17.17	5.92	-20.81		ModlVar.	4.83	StdDev=	2.2	EffVar	
<b>Signif.</b>	Yes	Yes	Yes		T(.975,df)	2.75	CI-HW=	3.49	StdDev	
Run	Second			Half(b)						
	A	B	C=-AB	Y1	Y2	Y3	Avg.	Var	Model	
1	-1	-1	-1	-5.64	-0.28	9.43	1.17	58.32	2	
4	1	1	-1	4	1.47	2.49	2.65	1.62	2	
6	1	-1	1	49.73	54.94	56.86	53.84	13.62	54	
7	-1	1	1	5.99	7.88	2.56	5.48	7.26	2	
<b>TotSum</b>				54.08	64.01	71.34	63.14	80.82		
<b>Effect</b>	24.92	-23.44	27.75		ModlVar.	20.2	StdDev=	4.49	EffVar	
<b>Signif.</b>	Yes	Yes	Yes		T(.975,df)	2.75	CI-HW=	7.14	StdDev	
<b>(a+b)/2</b>	3.88	-8.76	3.47	<b>MainEff</b>	<b>"C"</b>					
<b>(a-b)/2</b>	-21.05	14.68	-24.28	<b>Interact</b>	<b>C=AB</b>					
<b>Coefs</b>	6	-10	2							

**NOTE: FRACTIONAL FACTORIAL RESULTS, GIVEN THE STRONG INTERACTIONS, ARE POOR.**

# Plackett-Burnam (PB) DOEs

- A Fractional Factorial (FF) DOE
- Analyses “holes” between adjacent FFs
- Reduces time/effort, considerably
- Confounding of Main Effects/Interactions
- Numerical Example: 11 main effects
  - Compare PB to a  $2^{11}$  Full Factorial
  - Not all Interactions are strong/significant
- Counter Example: strong interactions

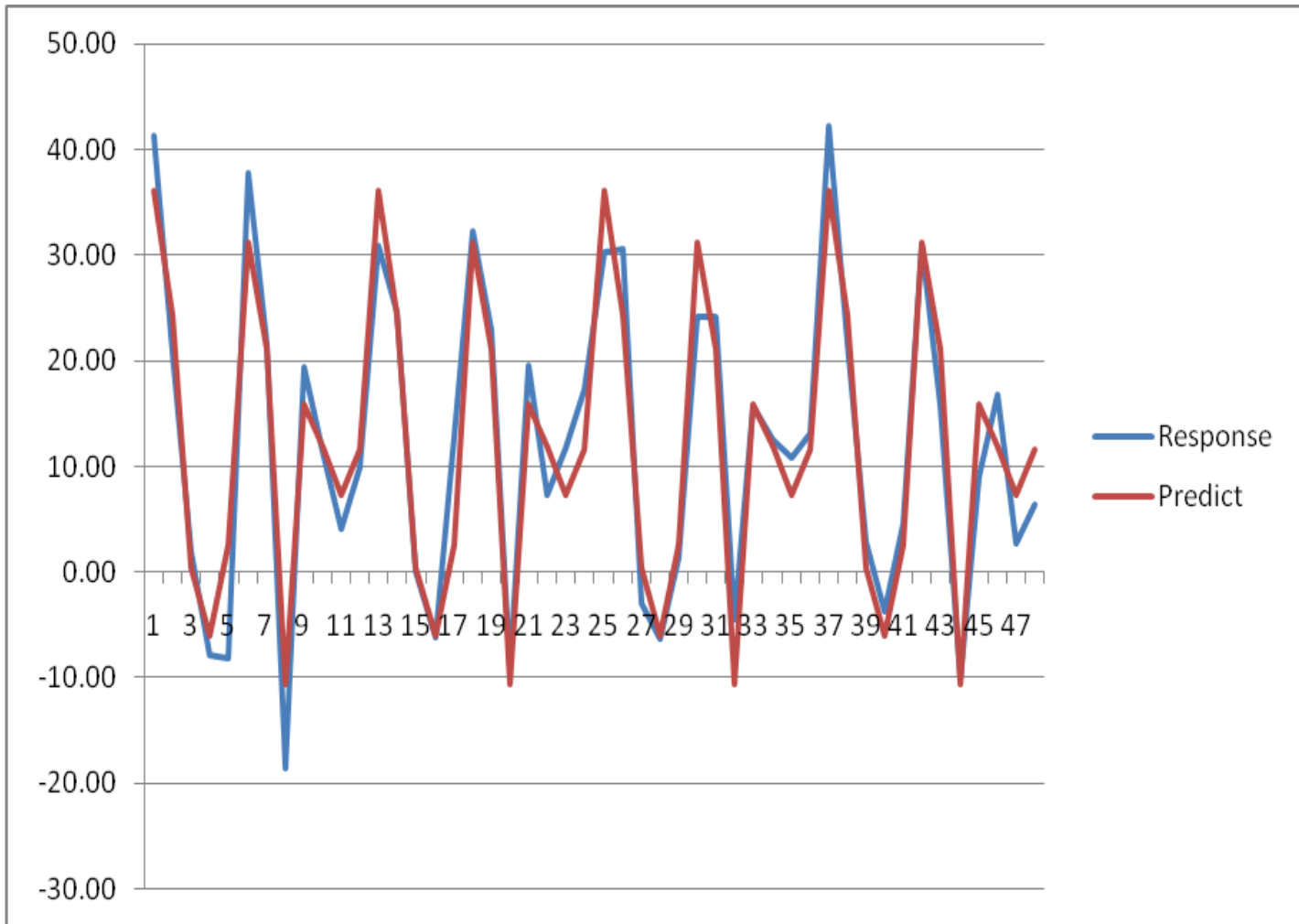
# Plackett-Burnam w/o Interaction

- A=Replenishing Levels (MIN)
- B=Reservoir Capacity (MAX)
  - C=Ordering Schedule
  - D=Transfer Policy
- E=Allocation to each sector
  - F=Size of the Reservoirs
- G=Generation of electricity
  - H=Hospitals and schools
    - I=Wetland size
    - J=Water Table
  - K=Fish/Foul Population

# Plackett-Burnam Design (no interaction)

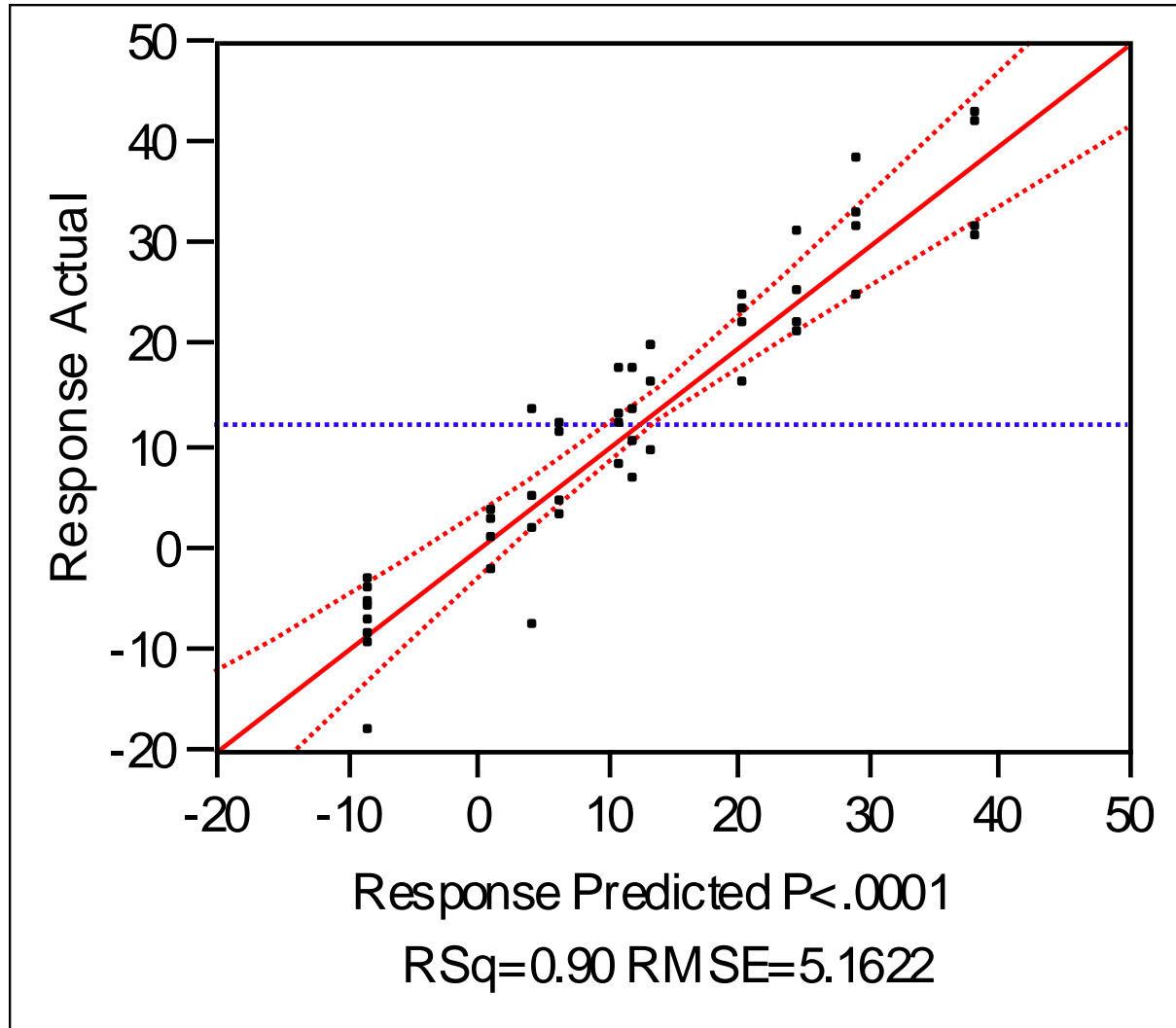
Run	A	B	C	D	E	F	G	H	I	J	K	Avg	
1	1	1	-1	1	-1	-1	-1	1	1	1	-1	1	36.14
2	1	1	1	-1	1	-1	-1	-1	1	1	1	-1	24.39
3	-1	1	1	1	-1	1	-1	-1	-1	1	1	1	0.5
4	1	-1	1	1	1	-1	1	-1	-1	-1	1	1	-5.96
5	1	1	1	-1	1	1	-1	1	-1	-1	-1	1	2.62
6	1	1	1	1	-1	1	1	-1	1	-1	-1	-1	31.26
7	-1	1	1	1	1	-1	1	1	-1	1	-1	-1	21.12
8	-1	-1	1	1	1	1	-1	1	1	-1	1	-1	-10.54
9	-1	-1	-1	-1	1	1	1	-1	1	1	-1	1	15.92
10	1	-1	-1	-1	-1	1	1	1	-1	1	1	-1	12.02
11	-1	1	-1	-1	-1	-1	1	1	1	-1	1	1	7.33
12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	11.66
<b>Factors</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>J</b>	<b>K</b>	<b>Bo</b>	
<b>RegCoef</b>	<b>6</b>	<b>2</b>	<b>0</b>	<b>-4</b>	<b>-6</b>	<b>0</b>	<b>-2</b>	<b>4</b>	<b>8</b>	<b>-8</b>	<b>0</b>	<b>12</b>	
<b>RegEst.</b>	<b>4.5</b>	<b>2.3</b>	<b>-0.1</b>	<b>-4.3</b>	<b>-3.6</b>	<b>1.4</b>	<b>-0.8</b>	<b>5.2</b>	<b>6.1</b>	<b>-7.6</b>	<b>-2.8</b>	<b>12.2</b>	
<b>MainEff</b>	12	4	0	-8	-12	0	-4	8	16	-16	0	n/a	
<b>EstimEff</b>	9.1	4.7	-0.2	-8.6	-7.2	2.8	-1.5	10.4	12.3	-15.2	-5.6	12.2	
<b>Signific.</b>	Yes	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	

# Meta Model Forecasting Ability





# Actual by Predicted Plot



# Plackett-Burnam with Interaction

- A=Replenishing Levels (MIN)
- B=Reservoir Capacity (MAX)
  - C=Ordering Schedule
  - D=Transfer Policy
- E=Allocation to each sector
  - F=Size of the Reservoirs
- G=Generation of electricity
- H=Hospitals and schools
  - I=Wetland size
  - J=Water Table
- K=Fish/Foul Population

# Model with Moderate Interaction structure:

<b>Factors</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>J</b>	<b>K</b>	<b>Bo</b>
<b>RegCoef</b>	6	2	0	-4	-6	0	-2	4	8	-8	0	12

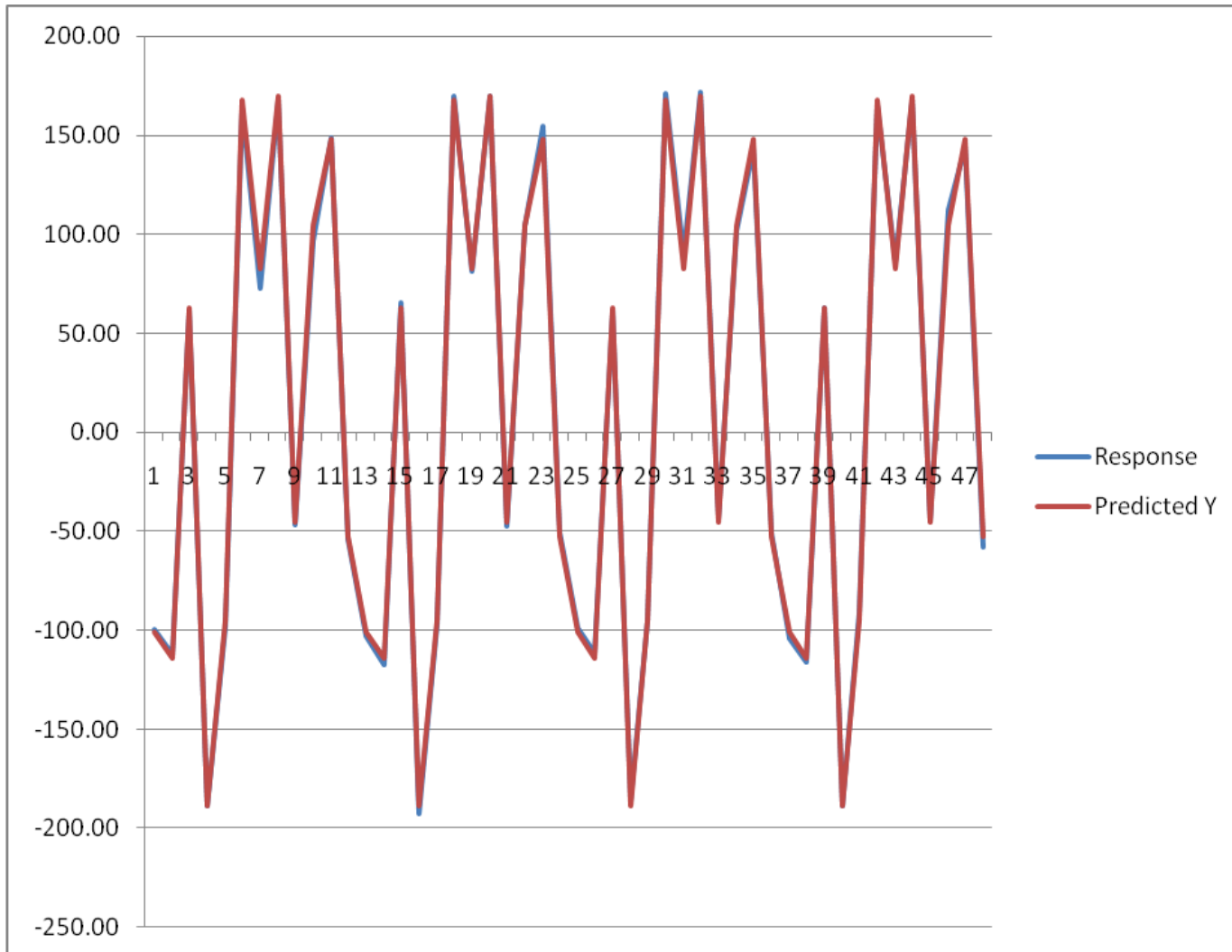
$$\text{Interaction: } 2*A*B - 4*H*I + G*J + D*E$$

Plackett-Burnam (n=12 rows) Analysis Results:

<b>Factors</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>J</b>	<b>K</b>
<b>MainEff</b>	12	4	0	-8	-12	0	-4	8	16	-16	0
<b>FacEstim</b>	-98.6	61.1	41.3	-86.5	98.4	66.4	79.7	51.8	-26.6	37.6	-96.0
<b>RegPar.</b>	6	2	0	-4	-6	0	-2	4	8	-8	0
<b>RegEstim</b>	-49.3	30.5	20.6	-43.2	49.2	33.2	39.8	25.9	-13.3	18.8	-48.0
<b>Signific.</b>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

**Results are seriously confounded and numerically erroneous.**

# Meta Model Forecasting Capabilities



# Controlled Sequential Bifurcation

- Method to Identify significant Main Effects
- Requires prior knowledge of Effect signs
  - To ensure all effects are in same direction
  - Requirement is unrealistic in most cases
- Branch and Bound-like approach
  - Top-Down approach most often
- Adaptive procedure to assess estimations
  - Using the approach but not the method

There are two groups of significant variables after Plackett-Burnam:

Positive: B, C, E, F, G, H, J;

and

Negative: A, D, I, K.

We Perform a Resolution IV FF  
To one of the two groups

# Plackett-Burnam Result

## Group of “Positive” Vars:

B, C, E, F, G, H, J;

- B=Reservoir Capacity (MAX)
  - C=Ordering Schedule
- E=Allocation to each sector
  - F=Size of the Reservoirs
- G=Generation of electricity
  - H=Hospitals and schools
    - J=Water Table

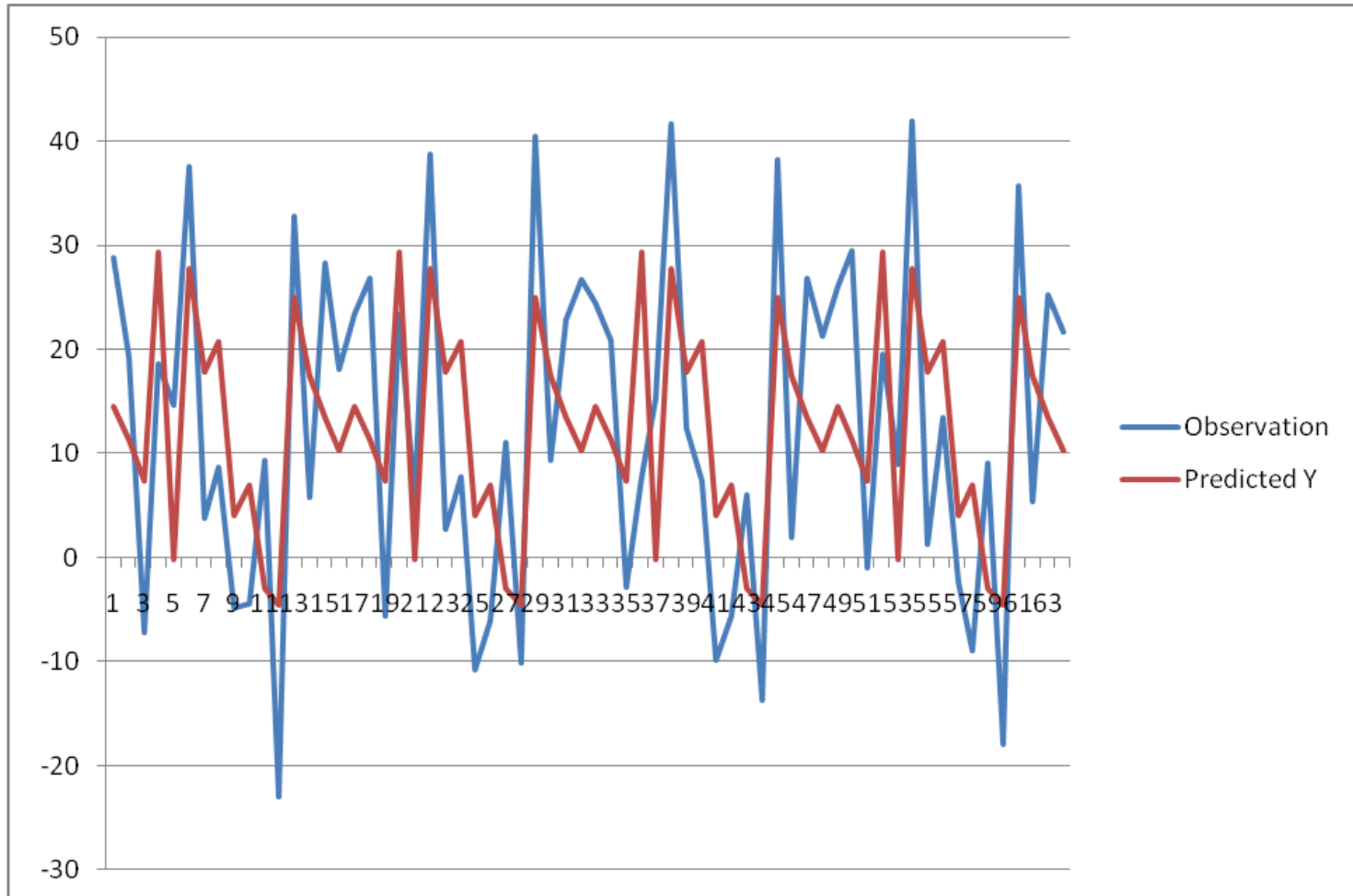
# Performing a Resolution IV FF to the “Positive” group: B, C, E, F, G, H, J

Factors	B	C	E	F	G	H	J	Bo
TRUE	12	4	0	-8	-12	0	-4	12
EffectEstim	12.14	2.53	1.17	-7.20	-11.82	0.39	-3.49	13.59
RegCoef	6	2	0	-4	-6	0	-2	12
RegEst.	6.07	1.26	0.59	-3.60	-5.91	0.19	-1.75	6.80
Signific.	Yes	Yes	No	Yes	Yes	No	Yes	

Notice how, once all the Plackett-Burnam (erroneously estimated) variables of the “same sign” were re-analyzed as a sub-group, estimations became closer to True values, both in sign and in magnitude.



Descriptive ability of the model improves;  
But its Forecasting capability deteriorates.



# Latin Hypercube Sampling

- Multiple regression analysis approach
  - Sampling at “best” points in sample space
- Regression selection methods
  - To obtain most efficient Meta Model set
- Provides a list of Alternative Meta Models
  - Some, not as efficient -but close enough
  - Their factors can be “controlled” by the user
- Very effective modeling approach.

# Latin Hypercube Example

Assume we have a three dimensional ( $p = 3$ ) problem in variables B, I, J (reservoir capacity; wetland size and water table use) and that these are respectively distributed Normal, Uniform and Exponential,. Assume that we want to draw a random sample of size  $n = 10$ . Divide each variable, according to its probability distribution, into ten equi-probable segments (Prob. =  $0.1 = 1/10$ ), identifying each segment with integers 1 through 10.

Then, draw a random variate (r.v.) from each of the ten segments, for each of the three variables B, I, J. Finally, obtain the  $10!$  permutations of integers 1 through 10. Randomly assign one of such permutations (e.g. segments 2,1,5,4,6,9,8,10,7 for B), to each of the variables, select the corresponding segment r.v., and form the vector sample, as below:

## Example of Latin Hypercube Sampling Segments

Sample	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
<b>B</b>	2	3	1	5	4	6	9	8	10	7
<b>I</b>	4	2	7	1	5	9	10	8	6	3
<b>J</b>	8	6	2	7	1	5	4	3	9	10

# Latin Hypercube Example

- Air Force Iraq Simulation Model
  - Fifty plus model variables
  - Two different responses of interest
  - Identify the Key or Relevant Few
  - Preserve as much Info as possible
- Analysis Results
  - Three Key Variables were identified
  - Ninety Percent of the Info ( $R^2 = 0.9$ )

# Regression Selection Analyses Results

## SUMMARY OUTPUT

### *Regression Statistics*

R Square 0.974347

Observations 450

### *Coeff* *P-value*

Intercept 0.034007 0.218928

X Variable 1 -3.00E-05 3.70E-193

X Variable 2 3.94E-05 0

## SUMMARY OUTPUT:

### *Regression Statistics*

R Square 0.854967

Observations 450

### *Coeff* *P-value*

Intercept 6.158105 1.26E-07

X Variable 1 0.000251 6.90E-75

X Variable 2 -0.00026 4.90E-143

X Variable 3 0.047911 1.96E-05

X Variable 4 -1.46362 1.09E-62

# SUMMARY OF IMPROVED META MODELS DERIVED:

RESPONSE	KEY VARIABLES	INDEX OF FIT	F-STATISTIC
ECONOMIC	AH, AK, AM, AR	97,6%	4723.7
ECONOMIC	AH, AK	97.4%	8488.9
VIOLENCE	AH, AK, AM, AR	85.5%	655.8
VIOLENCE	AH, AK, AR, AX	85.1%	637.5
VIOLENCE	AH, AK	72.4%	587.5
VIOLENCE	AH,AK,AM,AP,AR,AS,AT, AU,AV,AW	94.8%	800.4

# CORRELATION MATRIX FOR THE FOUR COMMON KEY VARIABLES AND THE TWO RESPONSES:

---

	<i>KEYVAR 1</i>	<i>KEYVAR 2</i>	<i>Column 3</i>	<i>Column 4</i>	<b>Violent</b>	<b>Econ</b>
KV 1	1					
KV 2	-0.07464	1				
Col 3	0.00345	0.088232	1			
Col 4	0.032984	0.0929	0.055228	1		
<b>Violence</b>	<b>0.44621</b>	<b>-0.7560</b>	-0.0018	<b>-0.4056</b>	1	
<b>Econ</b>	<b>-0.4637</b>	<b>0.90353</b>	0.028	0.08433	<b>-0.8671</b>	1

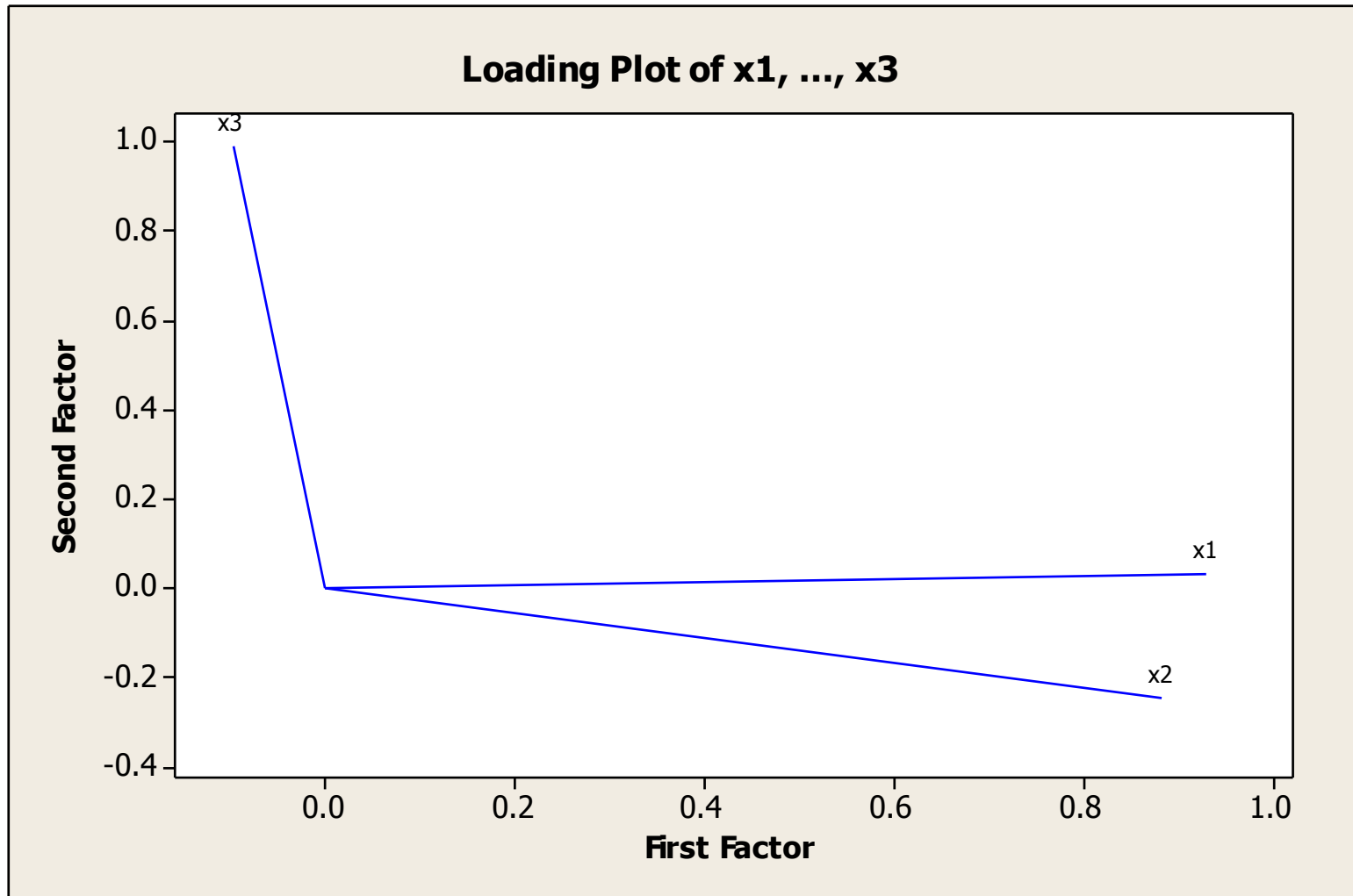
# Principal Components

- Can also be used with Latin Hypercube
  - When variables are strongly correlated
  - Alternative dimension reduction technique
- Main problem: how to interpret it:
  - To identify Key variables through loadings?
  - To use the PCA Main Factors, instead?
  - Alternative approaches?
- Needs evaluation and comparison w/DOE



Example of Varimax Factor Rotation :  
Project Variables X1 and X2 on F1  
Then, Project Variable X3 on Factor 2.

Variable	Factor1	Factor2
x1	0.930	0.030
x2	0.883	-0.249
x3	-0.097	0.989



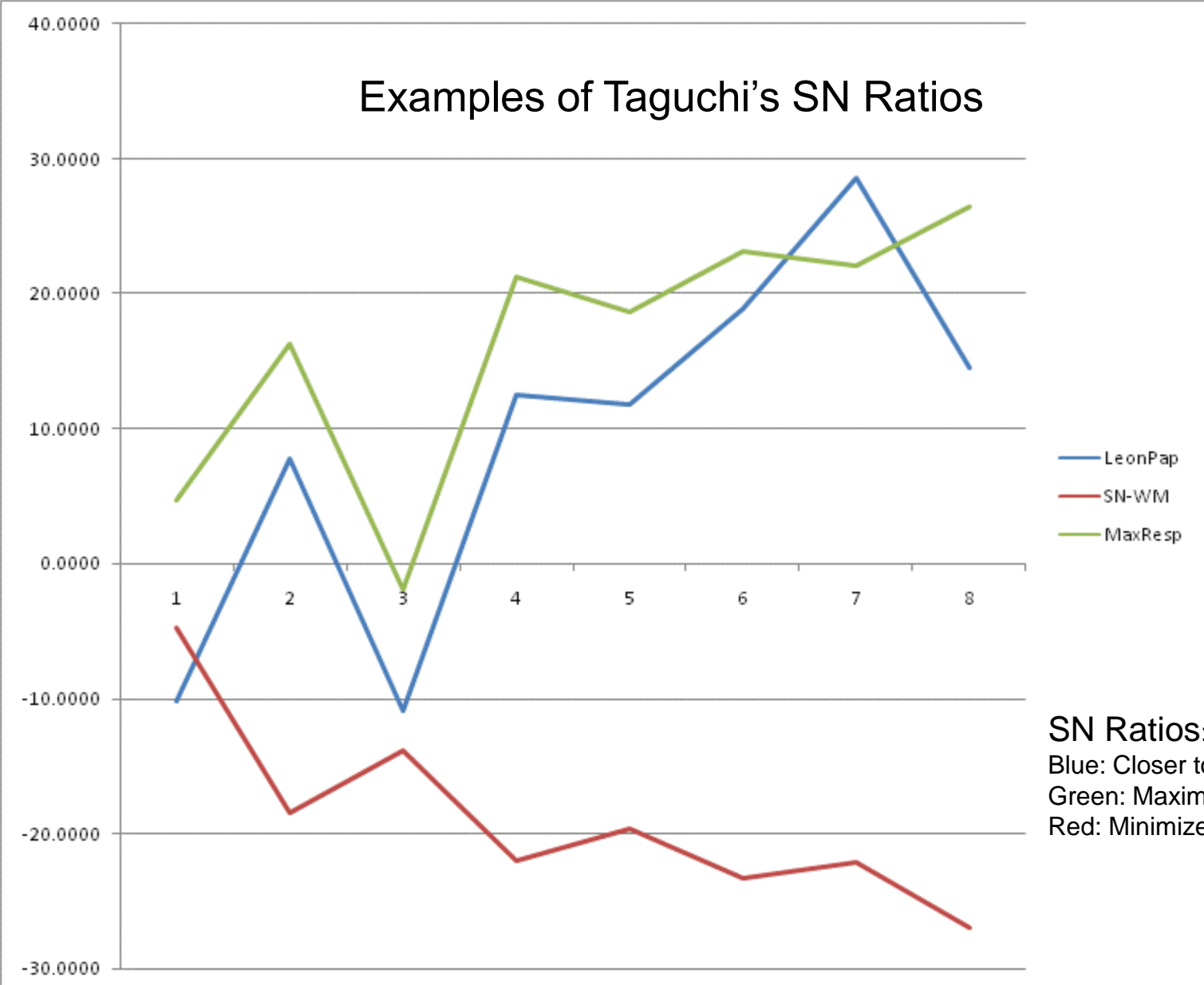
# Other Approaches

- Bayesian
  - Assume a prior on Meta Model terms
- Hierarchical
  - Sub-model output yields upper level input
- Taguchi
  - Derive results resilient to “noise” parameters
  - Parameters representing “uncontrolled” vars
  - Provides many conceptual DOE ideas.

# Taguchi Approach

- Analyzes both Location and Variation
  - Of the performance measure of interest
- Best combination of both these together
  - To obtain most efficient Meta Model
- Optimize Location, resilient to Variation
- Minimize Variation, resilient to Location
- Determine regions of joint optimality
- Determine Variation is Not an issue
- Can be equivalently implementing w/DOE

# Examples of Taguchi's SN Ratios



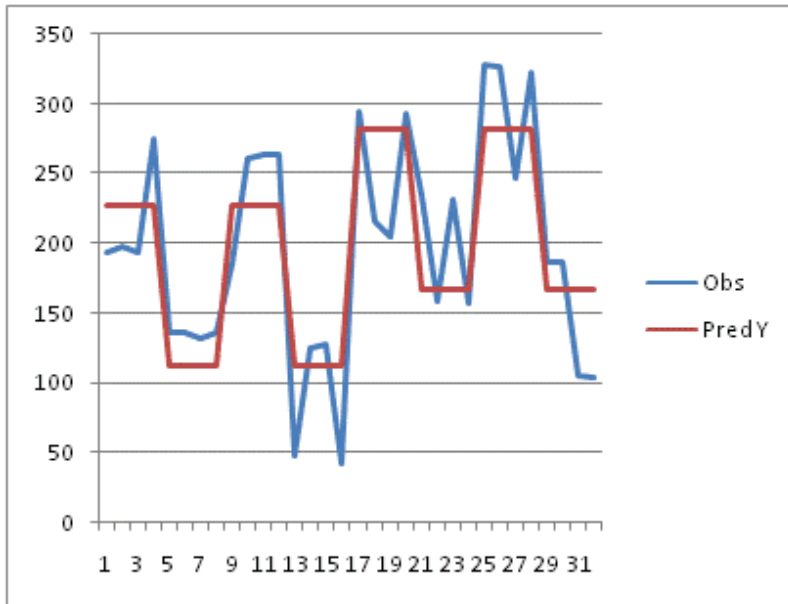
SN Ratios:  
Blue: Closer to Target  
Green: Maximize Yield  
Red: Minimize Yield

# Example of Taguchi

- Response: Wet Land Size
- X1=Reservoir Capacity (MAX)
- X2=Generation of electricity
  - X3=Hospital Capacity
  - X4=Social Services
- X5=Fish/Foul Population

## Comparison of *Combined* DOE and Taguchi's Approach

X1	X3	X2	X4	X5	1	2	3	4	Var	LnVar	Average	TagMinim
1	1	1	-1	-1	194	197	193	275	1616.25	7.39	214.75	-46.75
1	1	-1	1	1	136	136	132	136	4.00	1.39	135.00	-42.61
1	-1	1	-1	1	185	261	264	264	1523.00	7.33	243.50	-47.81
1	-1	-1	1	-1	47	125	127	42	2218.92	7.70	85.25	-39.51
-1	1	1	1	-1	295	216	204	293	2376.67	7.77	252.00	-48.15
-1	1	-1	-1	1	234	159	231	157	1852.25	7.52	195.25	-45.97
-1	-1	1	1	1	328	326	247	322	1540.25	7.34	305.75	-49.76
-1	-1	-1	-1	-1	186	187	105	104	2241.67	7.71	145.50	-43.59

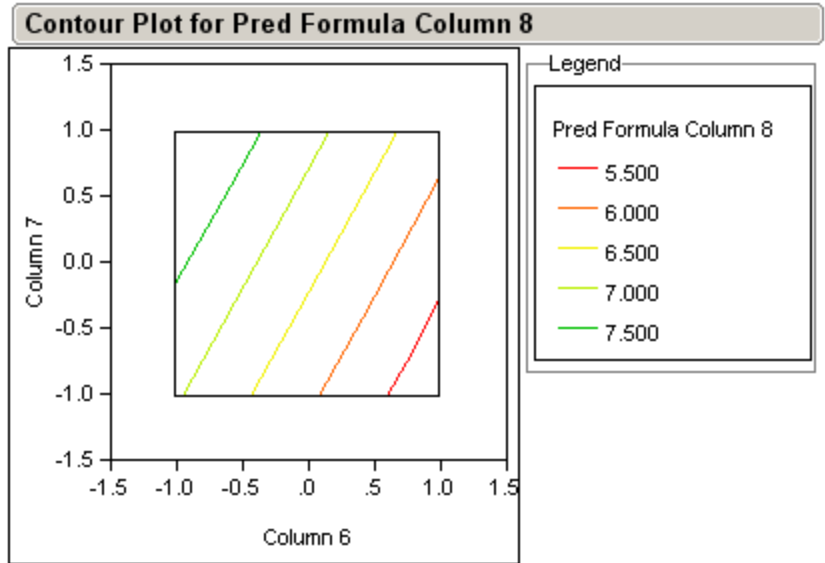
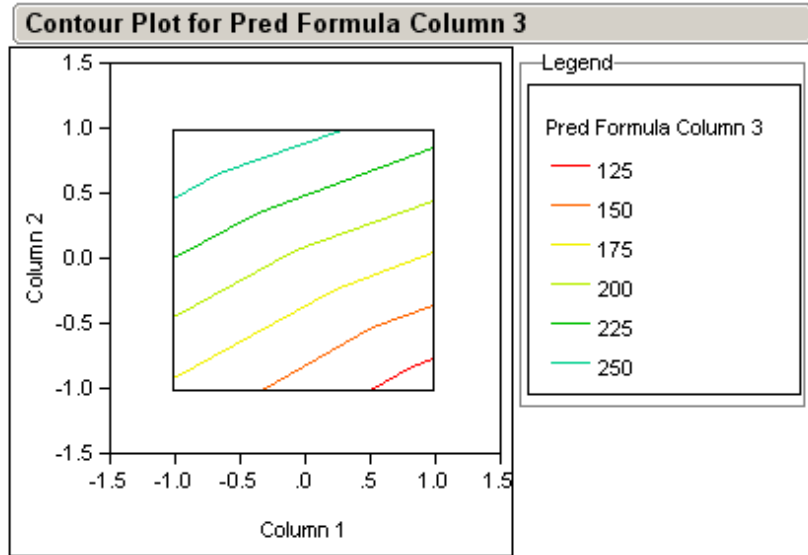


### Regression Analysis for the Main Effect influence

	Coef	Std Err	t Stat	P-value	Lower 95	Upper 95
Intrcpt	197.13	7.88	25.01	0.00	181.00	213.25
X Var 1	-27.50	7.88	-3.49	0.00	-43.62	-11.38
X Var 2	56.88	7.88	7.21	0.00	40.75	73.00

### Regression Analysis for the Variance Influence

	Coef	Std Err	t Stat	P-value	Lower 95	Upper 95
Intrcpt	6.77	0.78	8.70	0.00	4.77	8.77
X Var 1	-0.82	0.78	-1.05	0.34	-2.82	1.18
X Var 2	0.69	0.78	0.88	0.42	-1.31	2.69



**Optimal Solution:**

Overlaying both plots (for location and variation) we seek to Minimize both Yield (Errors) and Variation.

Jointly applying the two above (cols. 3 & 8).

**The Optimum is around (1, -1), yielding**

Estimated Minimum Output = 113; Min Variation = 5.3

**Estimated Yield:**

$$Y = 197.12 - 27.5X_1 + 56.9X_2$$

$$Y(1, -1) = 112.72$$

**Estimated Variation:**

$$Y = 6.77 - 0.82X_1 + 0.69X_2$$

$$Y(1, -1) = 5.26$$

## Alternative *Combined* DOE Approach

# Some Applications

- Model Size Reduction for:
- Evaluation of Decisions and Strategies
- Evaluation of Robust Strategies
- Trade-offs and Sensitivity analyses
- What-if, time to catastrophic fails, etc.
- Design and Optimization of Systems
- Study of key Factors on a System
- Arbitration and Conflict Resolution



# Composite Objective Functions

Ecologic:  $X_i$  is number of occurrences of  $i$ th item:

$$f(x_1, \dots, x_p) = \sum_i v_i x_i; .with : \sum v_i = 1$$

Economic:  $Y_i = a_i X_i$  is cost of No.  $i$ th item occurrences:

$$g(x_1, \dots, x_p) = \sum_i \lambda_i y_i; .with : \sum \lambda_i = 1$$

$$l(w_1, \dots, w_n) = \sum_i \delta_i w_i; .with : \sum w_i = 1$$

Arbitration and Trade-Off:  $\alpha$  is the preference or weight:

$$H(g, l) = \alpha g + (1 - \alpha)l; .with : 0 < \alpha < 1$$

# Example of approach use:

- Reduce Model to Key Variables to:
- Minimize Total Water Operations Cost
- Subject to:
  - Maintaining specified labor levels
  - Reducing pollution to specified levels
  - Maintaining specified social levels
  - Maintaining specified consumption levels
  - Increasing overall health indices

**Trade-Off Examples**

<b>Scenario</b>	<b>Ecologic</b>	<b>Health</b>	<b>Industry</b>	<b>Education</b>	<b>Recreation</b>	<b>Other</b>
<b>Best Ecologic</b>	<b>X1</b>	<b>Y1</b>	<b>Z1</b>	<b>W1</b>	<b>L1</b>	<b>M1</b>
<b>Best Health</b>	<b>X2</b>	<b>Y2</b>	<b>Z2</b>	<b>W2</b>	<b>L2</b>	<b>M2</b>
<b>Best Industry</b>	<b>X3</b>					
<b>Best Education</b>	<b>X4</b>					
<b>Best Recreation</b>	<b>X5</b>					
<b>Best Other</b>	<b>X6</b>					

Analyze Maxi-min and Mini-max results

# Conclusions

- A very complex problem
  - Size and interactions are serious issues
- Existing methods, not fully compliant
  - But a promise if worked around
- Meta Models extremely useful
  - For strategic and tactical decisions
  - In crisis, and to assess/avoid them
  - In theoretical studies.