Design of Experiments

Jayendran Venkateswaran
IE630 Simulation Modeling and Analysis
IIT Bombay

Example: Inventory Policy

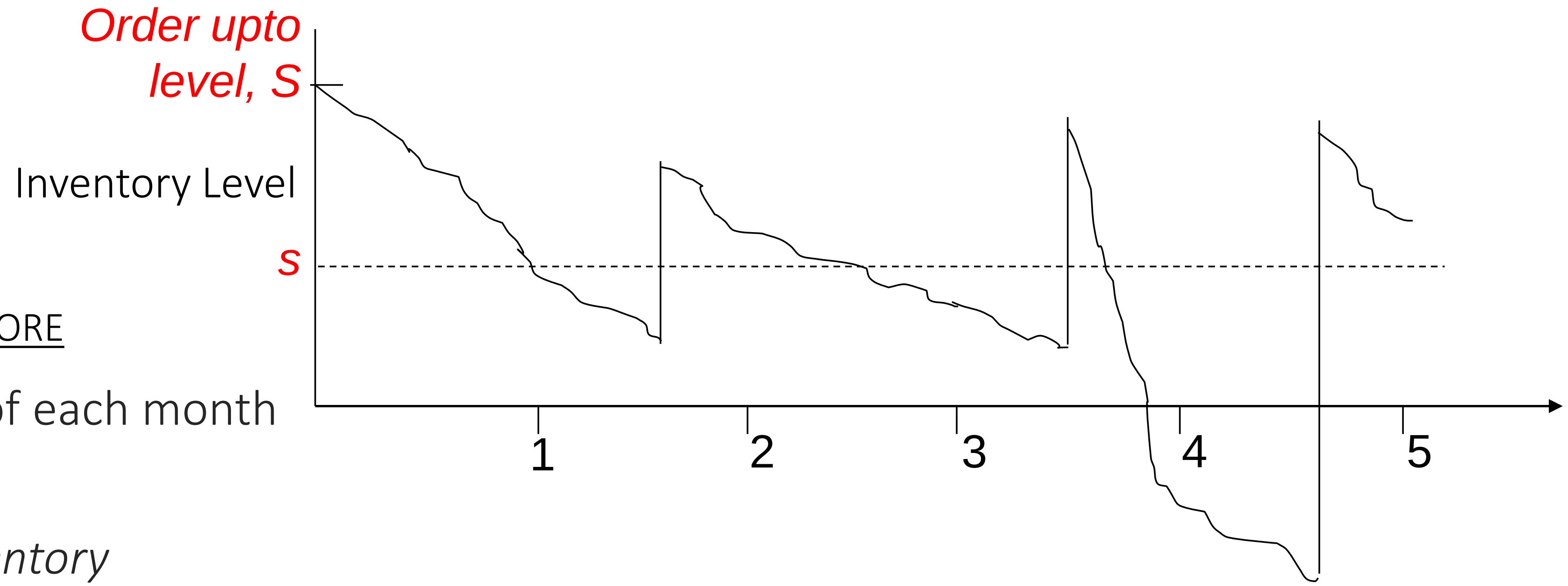
CUSTOMER VIEW

Customers Arrive
Exponential (λ=10)

Demand
1 unit w/ prob. 1/6;
2 units w/ prob.1/3;
3 units w/ prob. 1/3;
4 units w/prob. 1/6

Stores Inventory

Demand met from inventory
(if there)
Else, backlogged



BACK-END DECISION AT THE STORE

Inventory reviewed end of each month

If *Inventory* < *s* then

Order quantity = *S*-*Inventory*

The order quantity is then delivered after some days

Example: Inventory Policy

Demand, $D(t)$

- Inter-arrival times: EXPO(lambda=10) month
- Quantity: DISCRETE(1/6, 1, 3/6, 2, 5/6, 3, 6/6, 4)

$Inv\_level(t) = Inv\_level(t\text{-}1) - D(t) + OrderQuantityReceived^*$

Periodic review of inventory level every month

- Order Quantity = (S - *Inv_level*) IF *Inv_level* < s; 0 o.w.
- Delivery lag: UNIF(0.5, 1) month

Performance measure: Total Cost

= fixed cost (Rs. 32/order) + variable cost (Rs. 3 /item)

+ holding cost (Rs.1/unit/month) + backlog cost (Rs. 5/ unit/ month)

Source: Law and Kelton (2000)
Simulation Modeling & Analysis

Ex.: Compare Inventory Policies

Measure: Average total cost per month for the first 120 months of operations

Initial inventory level is 60

Suppose, given a combination of (s, S) → We can evaluate outcome.

For example: (s, S) = (20, 40) ; or (s, S)= (30, 60); etc…

How to systematically explore the design space such that we can find the best combination of (s, S)?

Source: Law and Kelton (2000)
Simulation Modeling & Analysis

Experiments

Experiments must be carefully planned and executed to be effective

An experiment is a deliberate manipulation of a process that intends to measure the effect of one or more experimental factors on some set of responses

Purpose

Which of many parameters or assumptions have the greatest effect on the response?

Which set of specifications appears to lead to optimal performance?

Factors, Levels & Responses

A **factor** is something you manipulate in your experiments

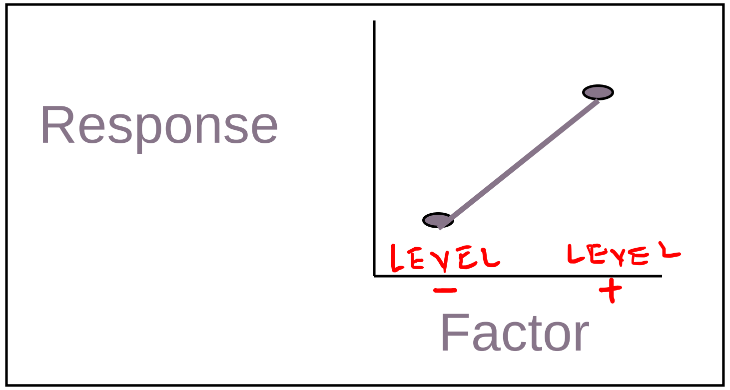
→ *Input parameters & assumptions* Factor has levels

Quantitative or Qualitative

A **response** is something you measure during or after the experiment

→ *Performance Measures*

A factor has an **effect** on a response if different **levels** (Settings) of the factor produce differing responses



Factorial Experiments

- Why do we need factorial experiments?
  - Many factors can be observed in action together
  - Interactions between factor, if any, can be observed
- INTUITION & INSIGHT
  - Include factors/ levels which you *feel* are unimportant
- Nomenclature & Design Matrix
  - 2^2 factorial design → 4 combinations
  - 2^3 factorial design → 8 combinations
  - 2^k factorial design
  - Full factorial design vs. Fractional factorial design

Design Matrix

2^2 factorial design

#	Factor 1	Factor 2	Response
1	–	–	R1
2	–	+	R2
3	+	–	R3
4	+	+	R4

2^3 factorial design

#	Factor 1	Factor 2	Factor 3	Response
1	–	–	–	R1
2	–	–	+	R2
3	–	+	–	R3
4	–	+	+	R4
5	+	–	–	R5
6	+	–	+	R6
7	+	+	–	R7
8	+	+	+	R8

Estimating Main Effects of factor *j*

Average **change** in response due to moving factor *j* from – level to + level, holding all other factors fixed.

(Trick: Sum the responses as per the sign of factor *j* & divide by half of the number of responses)

#	F1	F2	F3	Response
1	–	–	–	R1
2	–	–	+	R2
3	–	+	–	R3
4	–	+	+	R4
5	+	–	–	R5
6	+	–	+	R6
7	+	+	–	R7
8	+	+	+	R8

Estimating Interaction effect (two-way between j_1 & j_2)

Occurs when effect of one factor depends on settings of other factors

Half the difference between the

average effect of factor j_1 when factor j_2 is at + level and the average effect of factor j_1 when factor j_2 is at its – level (all other factors fixed)

Trick: Construct a truth table for (j_1 x j_2). Now Sum the responses as per the sign of (j_1 x j_2) & divide by half of the number of responses

#	F1	F2	F3	Response	$F_1 \times F_2$
1	–	–	–	R1	+
2	–	–	+	R2	+
3	–	+	–	R3	–
4	–	+	+	R4	–
5	+	–	–	R5	–
6	+	–	+	R6	–
7	+	+	–	R7	+
8	+	+	+	R8	+

$$e_{1\times 2} = e_{12} = \frac{1}{2}\left[\frac{(R_7 - R_5) + (R_8 - R_6)}{2} - \frac{(R_3 - R_1) + (R_4 - R_2)}{2}\right]$$

$$e_{21} = e_{12} = (R_1 + R_2 - R_3 - R_4 - R_5 - R_6 + R_7 + R_8)/4$$

|||ly $e_{23} = \ldots$

$e_{13} = \ldots$

Example 1 (Inventory)

Evaluate the effects of different combination of (s, d) on the response average monthly operating cost

Note: S = s + d

Factors	–	+
s	20	60
d	10	50

Design Matrix:

#	s	d	sxd	Response
1	–	–	+	142
2	–	+	–	119.81
3	+	–	–	145.83
4	+	+	+	148.31

Main effects, Interaction effects

Interpreting the results

Example 1 (Inventory)

How to use multiple replications data?

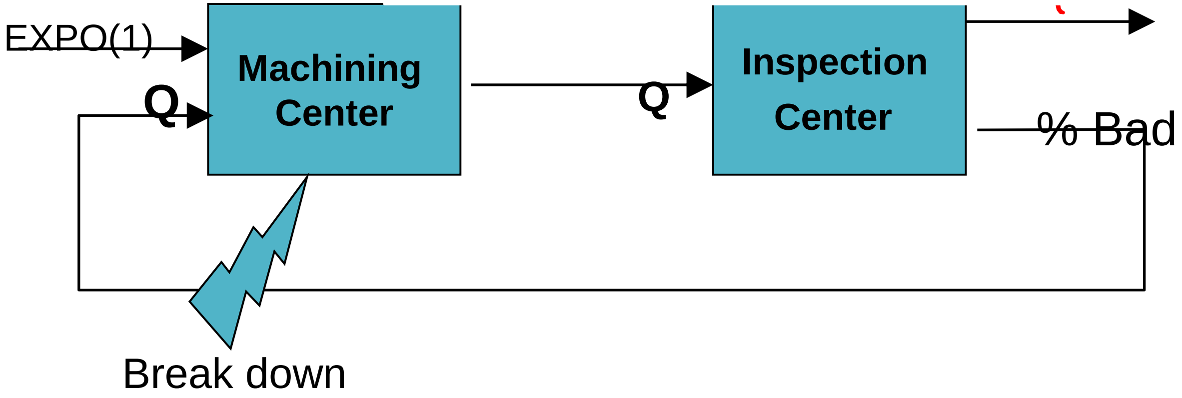
Example 1: Comparing inventory policies

With multiple replications, say 5

s	d=S-s	S=s+d	Rep1	Rep2	Rep3	Rep4	Rep5
20	10	30	142	135	143.31	149.1	139.31
20	50	70	119.81	117.18	128.02	123.13	124.54
60	10	70	145.83	144.82	143.66	146.54	144.31
60	50	110	148.31	145.28	145.66	150.4	149.67

e_s	16.165	18.96	8.995	12.355	15.065
e_d	-9.855	-8.68	-6.645	-11.055	-4.705
e_{sd}	12.335	9.14	8.645	14.915	10.065

Example 2



Factor	- (current)	+ (improved)
Machining times	U[0.64, 0.70]	U[0.585, 0.63]
Inspection times	U[0.75, 0.80]	U[0.675, 0.72]
Machine uptimes	Expo(360)	Expo(396)
Machine repair times	U[8, 12]	U[7.2, 10.8]
Scrap %	10%	9%
Queue rule	FIFO	SPT

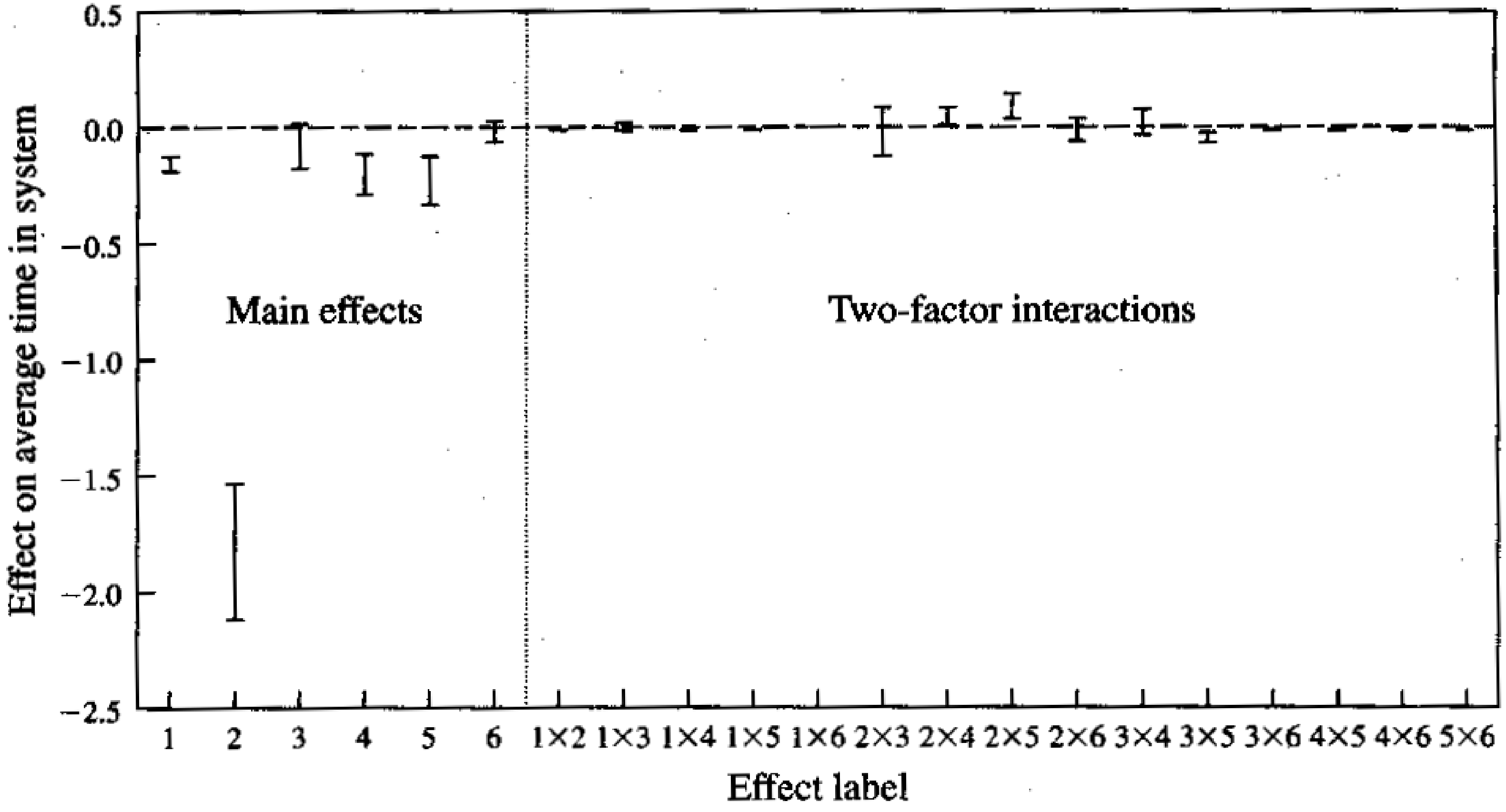
Design Matrix $2^6 = 64$ Combination

↓

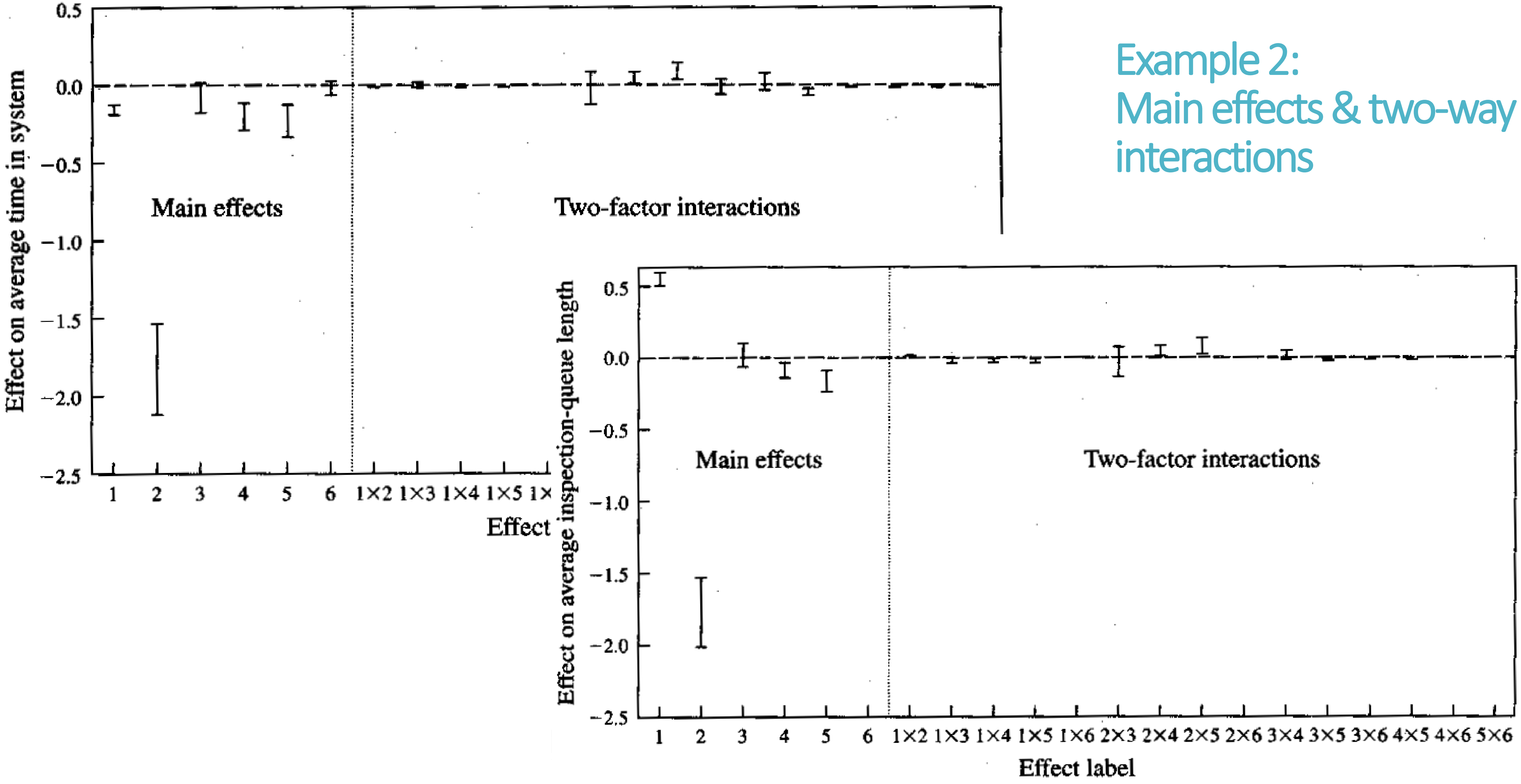
640 Simulation runs..

90% CI of the effects using 10 replications

- Suppose we can improve some process by 15%, where all do we invest? Will it improve time in system response?



Example 2: Main effects & two-way interactions



Example 2: Main effects & two-way interactions

FIGURE 12.2
Experimental design for small factory: Main effects and two-factor interactions.

Analysis of Variance (ANOVA)

Why?

- To identify the factors that are statistically significant
- Will NOT say which combination is the best

A division of the overall variability in data values in order to compare means.

Overall (or "**total**") variability is divided into two components:

- the variability "**between**" groups, and
- the variability "**within**" groups

Summarized in an "ANOVA" table

- Observe the p-value corresponding each factor

One factor experiments

Check how change in ONE factor affects performance

Which queue rule improves throughput?

Q rule	N	MEAN	SD
FIFO	10	188.20	3.88
LIFO	10	195.20	9.02
LPT	10	191.20	5.55
SPT	10	200.50	5.44

Similar to comparing alternatives!

Null hypothesis

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

Alternative hypothesis

H_A: at least one μ_i differs from the others

Analysis of Variance

Comparing 4 queue rules

Source	DF	SS	MS	F	P
QRule	3	1174.8	293.7	7.95	0.000
Error	35	1661.7	36.9		
Total	39	2836.5			

The P-value is small → reject H_0. There is sufficient evidence to conclude that *at least one queue rule is different* from the others.

Source	DF	SS	MS	F	P
QRule	3	80.1	40.1	0.46	0.643
Error	35	1050.8	87.6		
Total	39	1130.9			

The P-value is large → cannot reject H_0.

There is NOT sufficient evidence.

NOTE: Above data is for illustration only

Two way ANOVA Table

Table 5-3 The Analysis of Variance Table for the Two-Factor Factorial, Fixed Effects Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A treatments	SS_A	$a-1$	$MS_A = \frac{SS_A}{a-1}$	$F_0 = \frac{MS_A}{MS_E}$
B treatments	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$F_0 = \frac{MS_B}{MS_E}$
Interaction	SS_{AB}	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F_0 = \frac{MS_{AB}}{MS_E}$
Error	SS_E	$ab(n-1)$	$MS_E = \frac{SS_E}{ab(n-1)}$	
Total	SS_T	$abn-1$		

Variability due to the factor of interest

Unexplained random variations

Total variation from the grand mean

Further uses of Factorial Design

Simulation can be thought of as a mechanism that turns inputs parameters to outputs measures

Simulation is just a function!

Can we come up with a approximate formula that can be used as a proxy for full blown simulation?

This *meta-model* could be used in order to get at least a rough idea of what will happen for different of input-combinations

Regression model

Response Surface & Contour plots

Regression Model

General regression model for two-factor experiment

$$\underbrace{\hat{y}}_{\text{Estimated output measure}} = \beta_0 + \beta_1 \underbrace{x_1}_{\text{Variable for factor A}} + \beta_2 \underbrace{x_2}_{\text{Variable for factor B}} + \beta_{12} x_1 x_2 + \underbrace{\varepsilon}_{\text{Random error term}}$$

x_1 and x_2 are in coded scale -1 to +1

β_0 : average of responses of all combinations

β_1 : ½ * (Main effect of A) β_2 : ½ * (Main effect of B)

β_{12} : ½ * (interaction effect between A & B)

Regression model for Example-1 (inventory)

Compare inventory policies example

$\overline{e_s} =$ $\overline{e_d} =$ $\overline{e_{sd}} =$ $\overline{R} =$

Regression model, based on 5 replication data

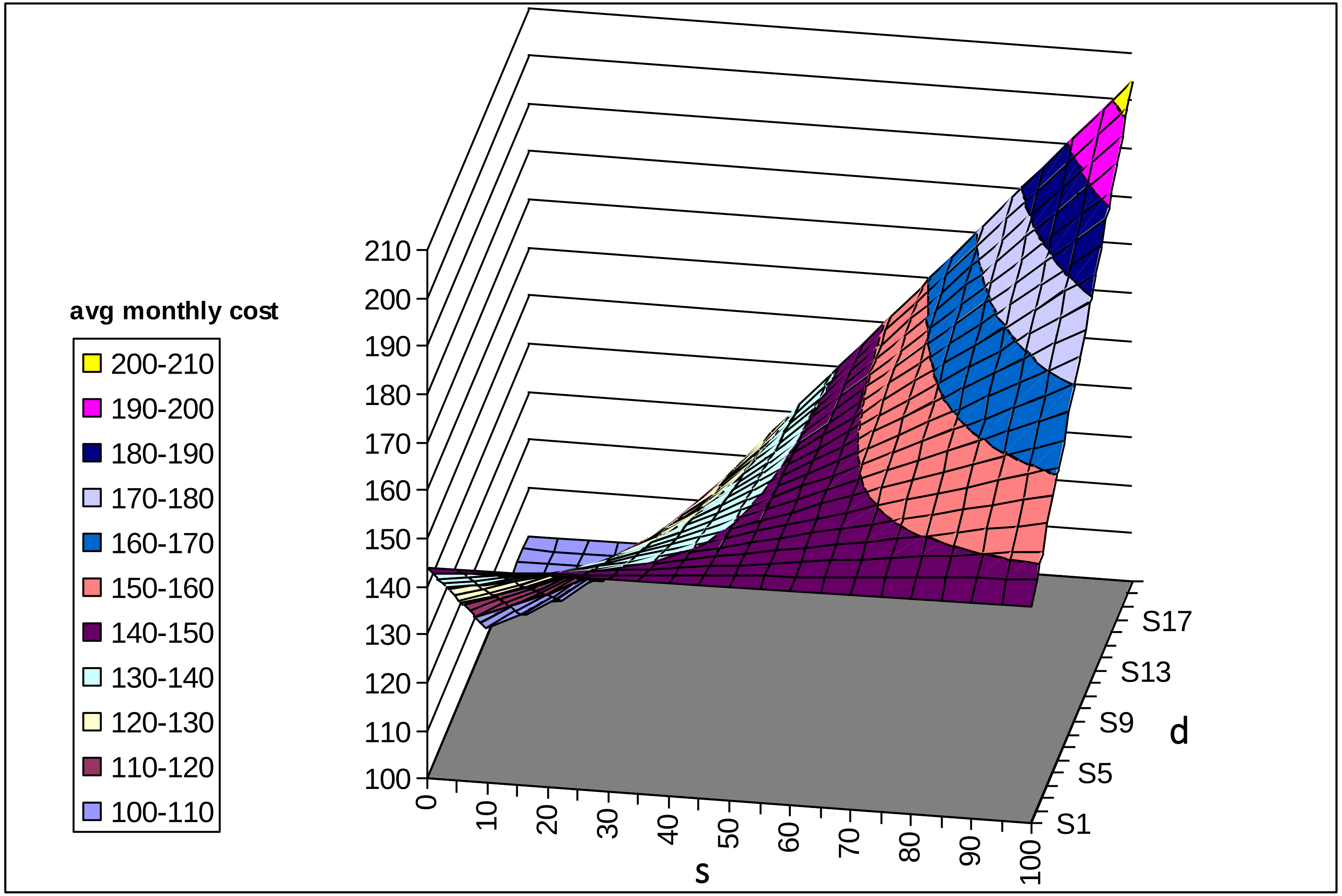
In terms of transformation (x_s, x_d) vs. Directly in terms of input parameters (s, d)

Compare inventory policies example (contd)

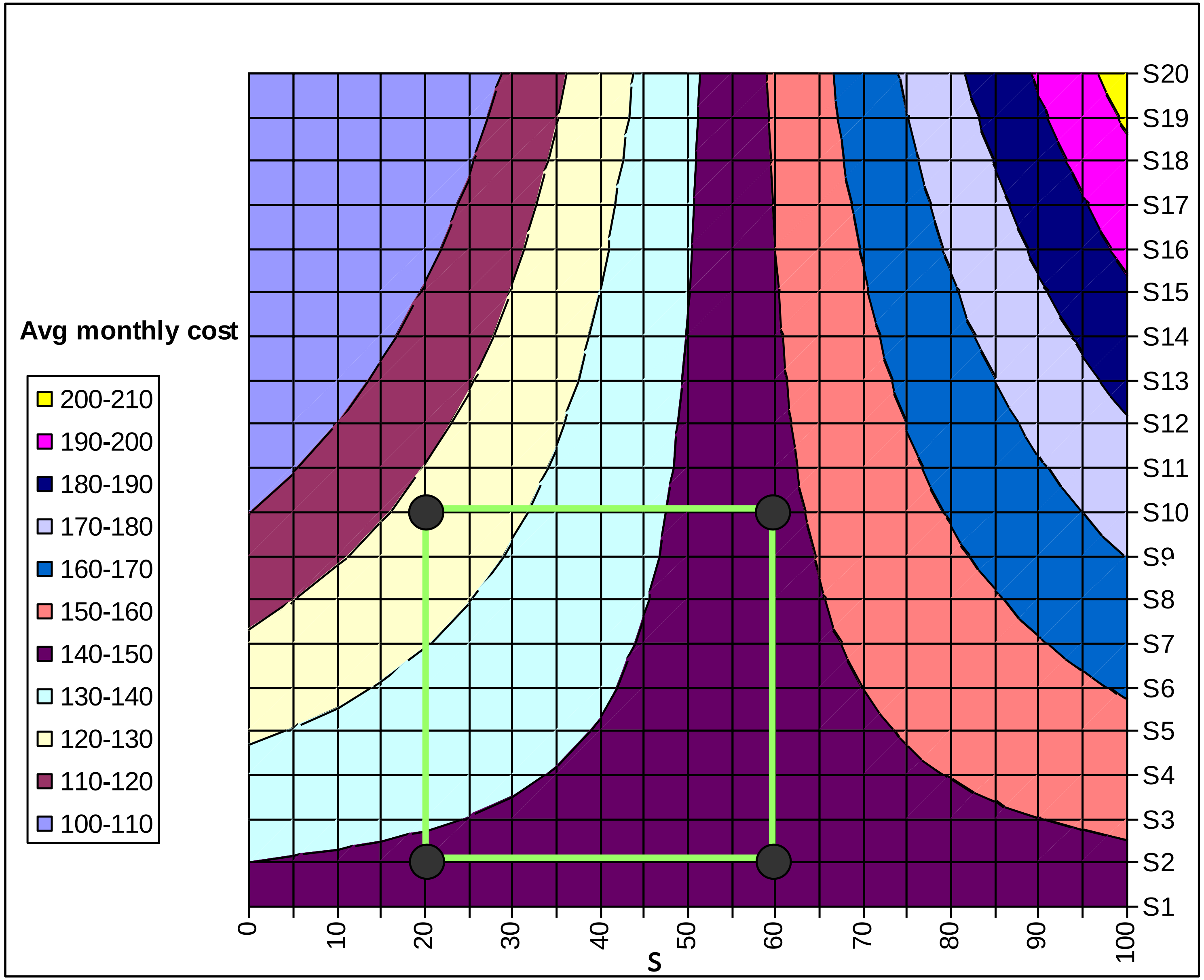
Based on regression model from 2^2 design, compute response for all 420 combinations of $s = 0, 5, 10, \ldots, 100$ and $d = 5, 10, \ldots, 100$.

- Use Excel spreadsheet to generate response (*doe-ex1.xls*, *Sheet: 'doe2x2'*)
- Response Surface: 3-D plot, with response on *z* axis & inputs (s, d) in x-y axis
- Contour plot: top view of response surface!

Response surface from regression model based on 2^2 design



Contour plot from regression model based on 2^2 design



Comparing inventory policies example (contd)

Let's make 10 replications of 16 combinations of s = 20, 40, 60, 80 and d = 20, 40, 60, 80

Using average of 10 replications

Fit a full quadratic model to the 16 data points

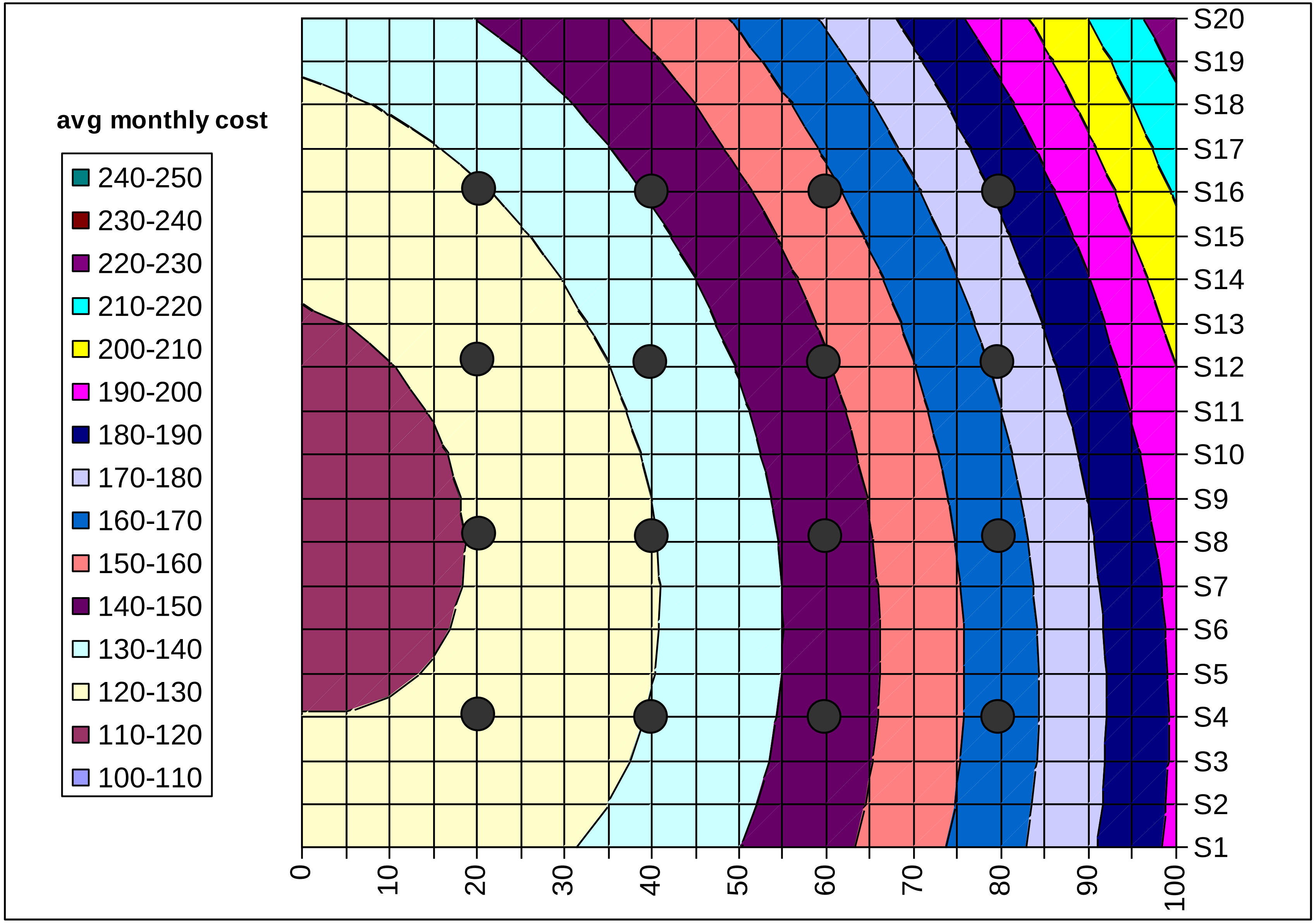
Use "Regression" from Tools>>Data Analysis

$$R(s,d) = 127.38 - 0.09s - 0.467d + 0.002sd + 0.007s^2 + 0.005d^2$$

Use regression model to generate response for all 420 combinations! → Create contour plot

Refer Excel (*doe-ex1.xls, Sheet: 'doe medium'*)

Contour plot from regression based on 4^2 design



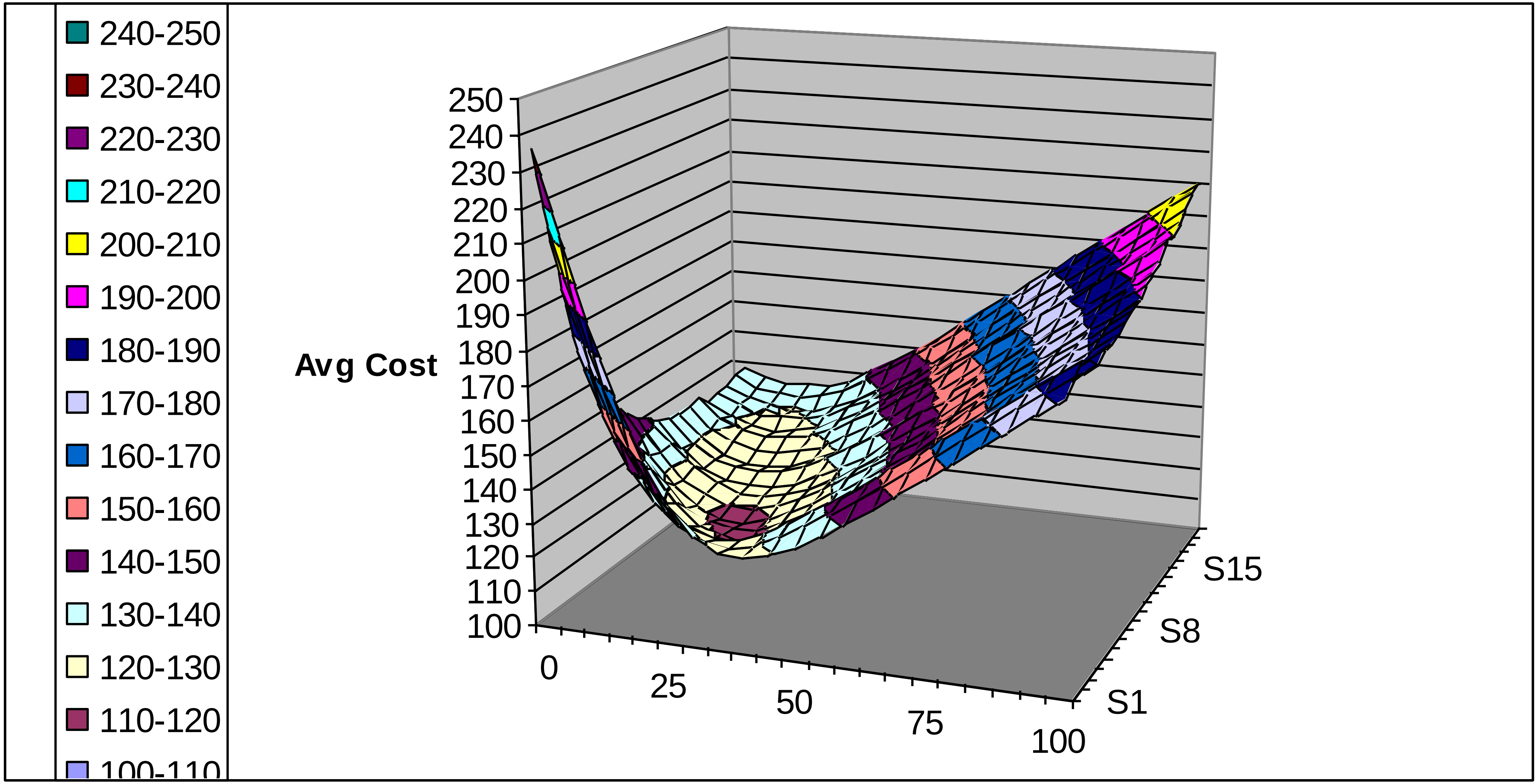
Comparing inventory policies example (contd)

Simulation generated response

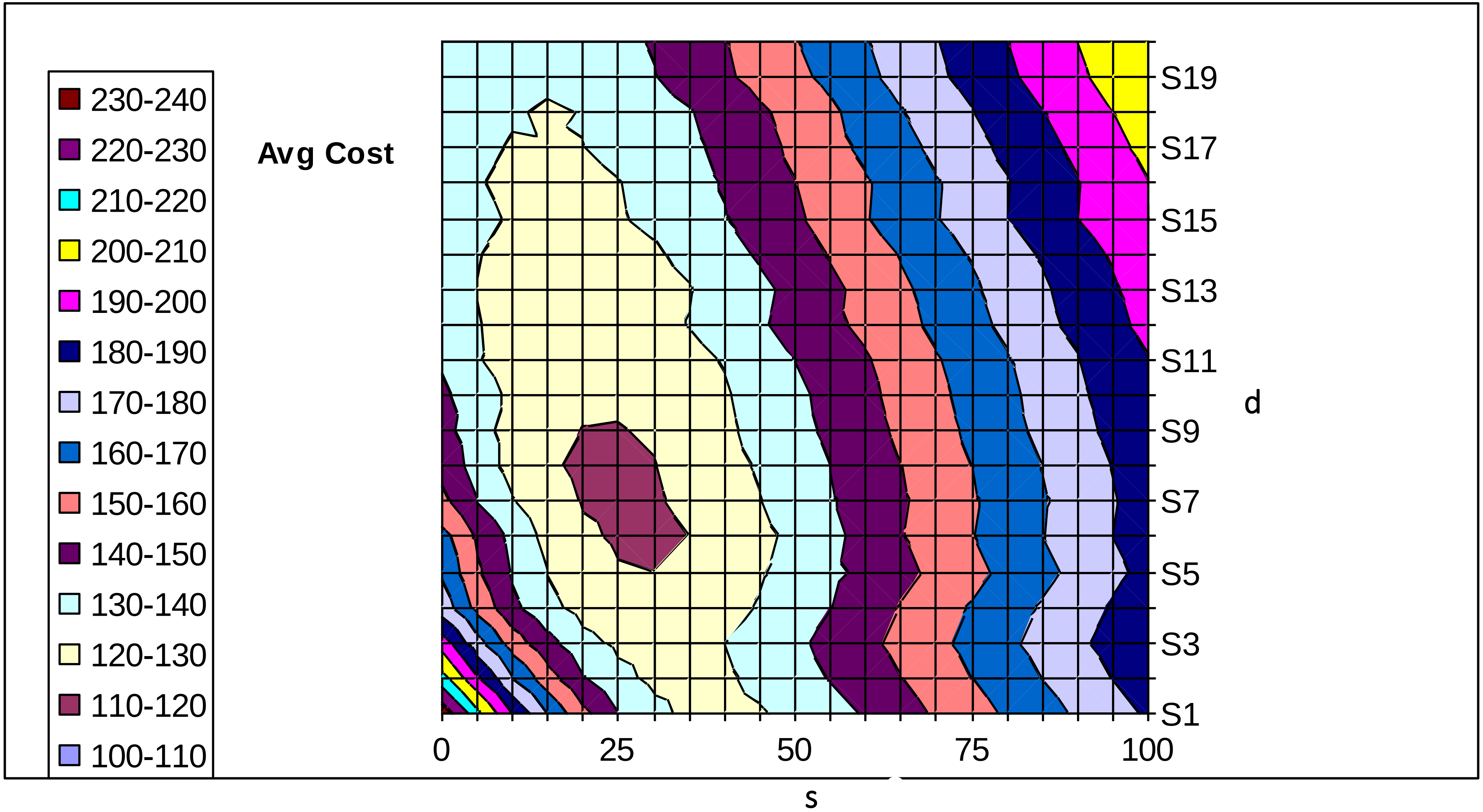
Let's run 10 replications of all 420 combinations of
$s = 0, 5, 10, \ldots, 100$ and $d = 5, 10, \ldots, 100$

Plot response surface and contour from simulation results
Using average of 10 replications

Response Surface from direct SIMULATION of all 420 combos



Contour from direct SIMULATION of all 420 combos



Summary

Brief overview of Design of Experiments

Further Reading

- Full vs partial factorial designs
- ANOVA
- Other meta-modelling techniques
- Handling multiple outputs
- Sensitivity analysis
- Gradient estimation
- Optimum seeking