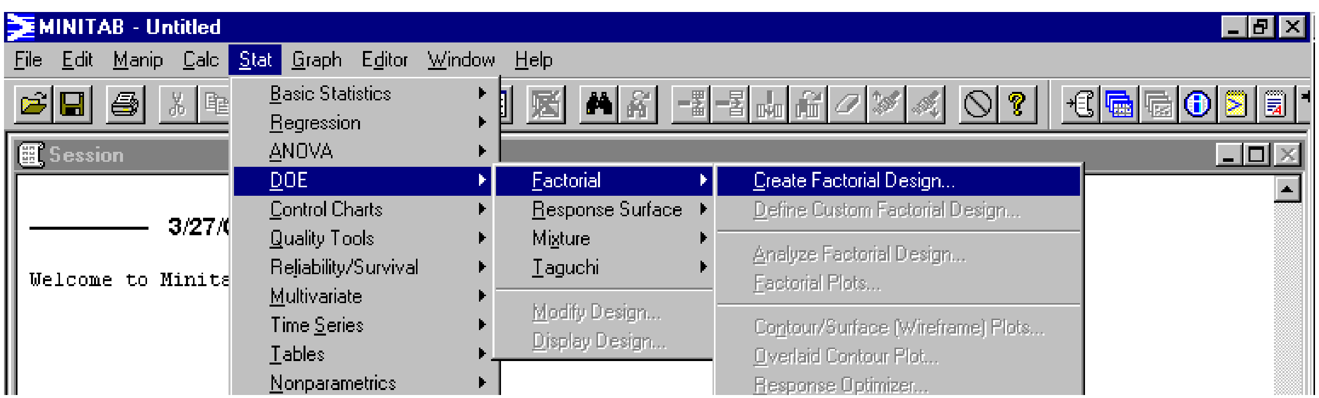
Exercise with Minitab

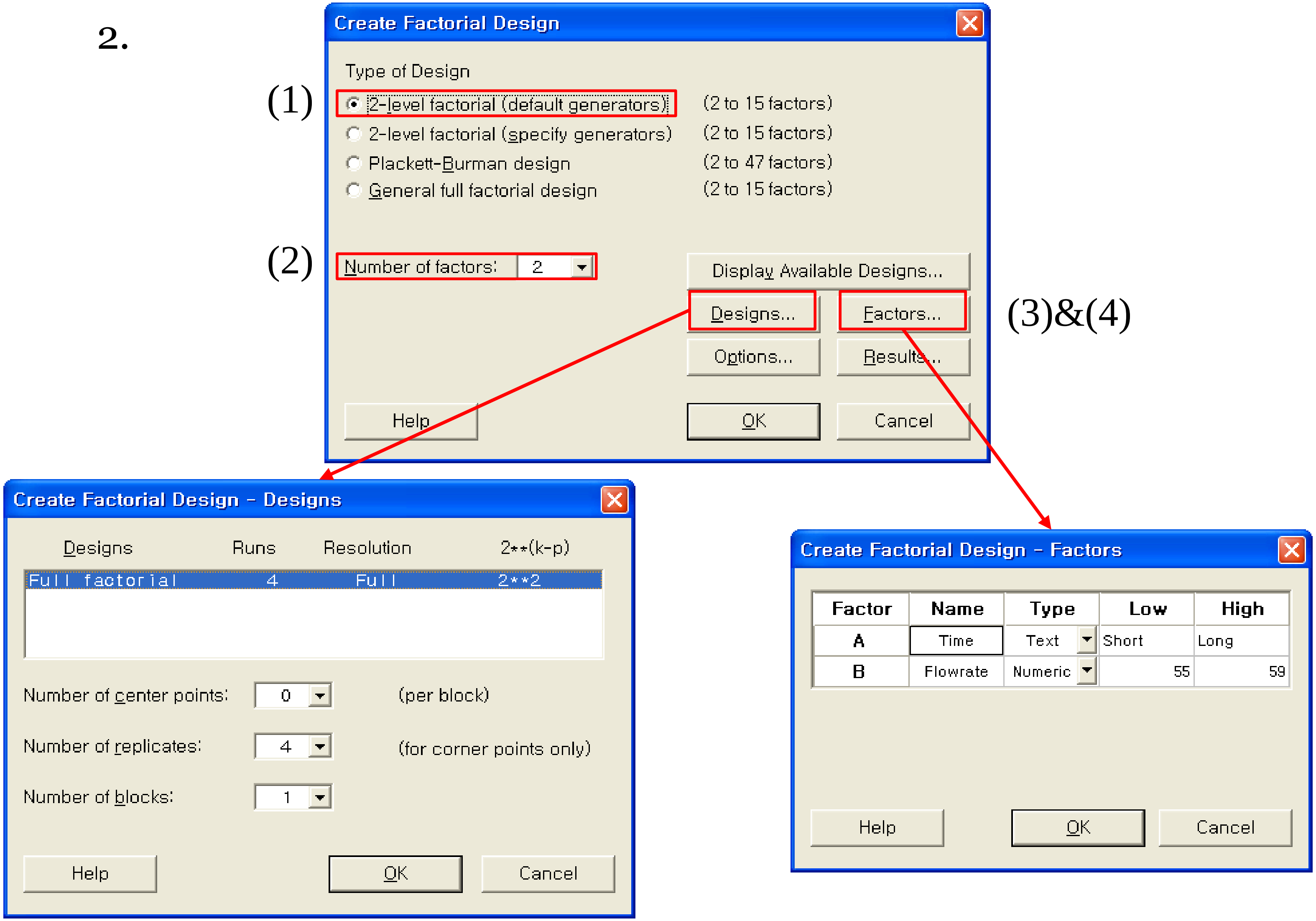
- When you have replicates.
  - You're a process engineer @ a semiconductor plant who wants to determine factors affecting thickness of epitaxial layer on silicon wafer. The main factors (or input variables) you think are (deposition) time and (arsenic) flowrate. Assume only linear relationship.
  - Solution
    1. 2^2 factorial design with 4 replicates @ corners

       stat>DOE>Factorial>create Factorial Design



Exercise with Minitab (cont.)

2.



Exercise with Minitab (cont.)

3. Run experiments according to design matrix

Worksheet 2 ***

↓	C1	C2	C3	C4	C5-T	C6
	StdOrder	RunOrder	CenterPt	Blocks	Time	Flowrate
1	11	1	1	1	Short	59
2	15	2	1	1	Short	59
3	3	3	1	1	Short	59
4	2	4	1	1	Long	55
5	9	5	1	1	Short	55
6	8	6	1	1	Long	59
7	7	7	1	1	Short	59
8	10	8	1	1	Long	55
9	1	9	1	1	Short	55
10	4	10	1	1	Long	59
11	12	11	1	1	Long	59
12	5	12	1	1	Short	55
13	16	13	1	1	Long	59
14	6	14	1	1	Long	55
15	13	15	1	1	Short	55
16	14	16	1	1	Long	55

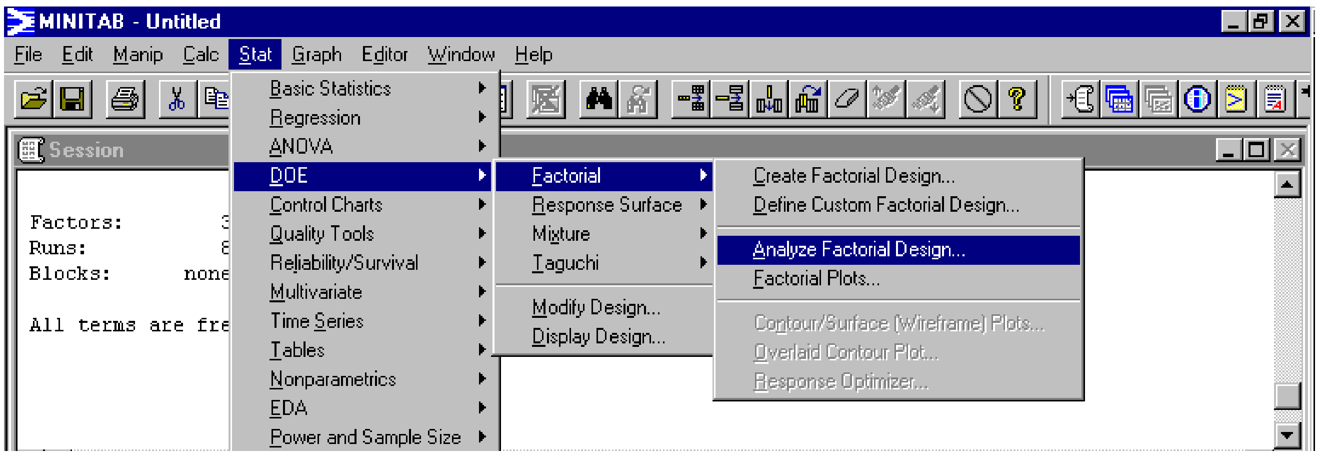
Why?

Exercise with Minitab (cont.)

4. Analysis of experimental results

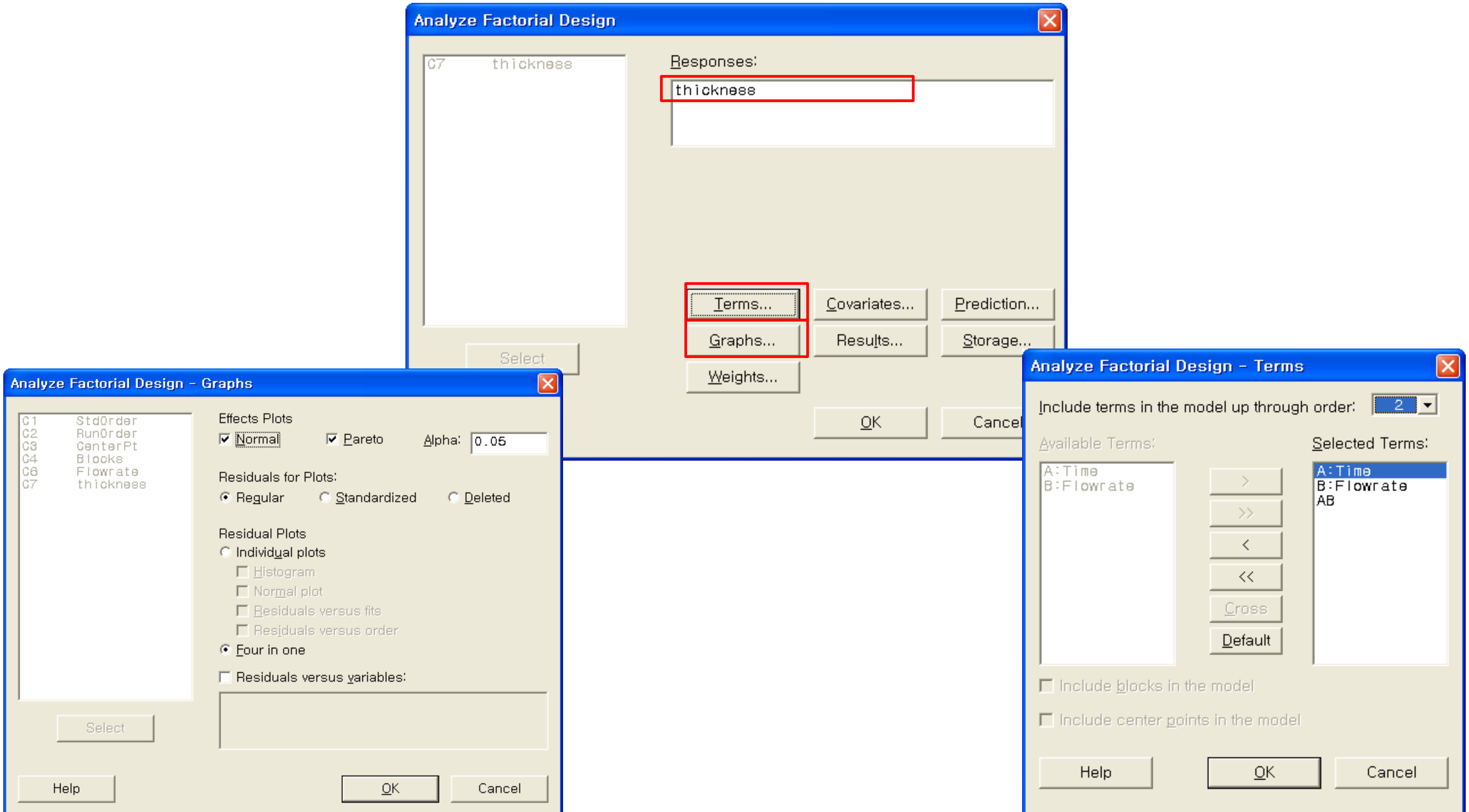
Using all analysis tools from least squares & main/interaction plots

DOE>Factorial>Analyze Factorial Design



Exercise with Minitab (cont.)

4. Analysis of experimental results (cont.)



Exercise with Minitab (cont.)

(a) ANOVA table (∵ we have replicates)

Estimated Effects and Coefficients for thickness (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		14.3884	0.03606	399.05	0.000
Time	0.8369	0.4184	0.03606	11.60	0.000
Flowrate	-0.0681	-0.0341	0.03606	-0.94	0.363
Time*Flowrate	0.0324	0.0162	0.03606	0.45	0.661

S = 0.144228   R-Sq = 91.88%   R-Sq(adj) = 89.85%

Analysis of Variance for thickness (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	2	2.82000	2.82000	1.41000	67.78	0.000
2-Way Interactions	1	0.00419	0.00419	0.00419	0.20	0.661
Residual Error	12	0.24962	0.24962	0.02080		
Pure Error	12	0.24962	0.24962	0.02080		
Total	15	3.07382				

Unusual Observations for thickness

Obs	StdOrder	thickness	Fit	SE Fit	Residual	St Resid
11	12	14.4150	14.7890	0.0721	-0.3740	-2.99R

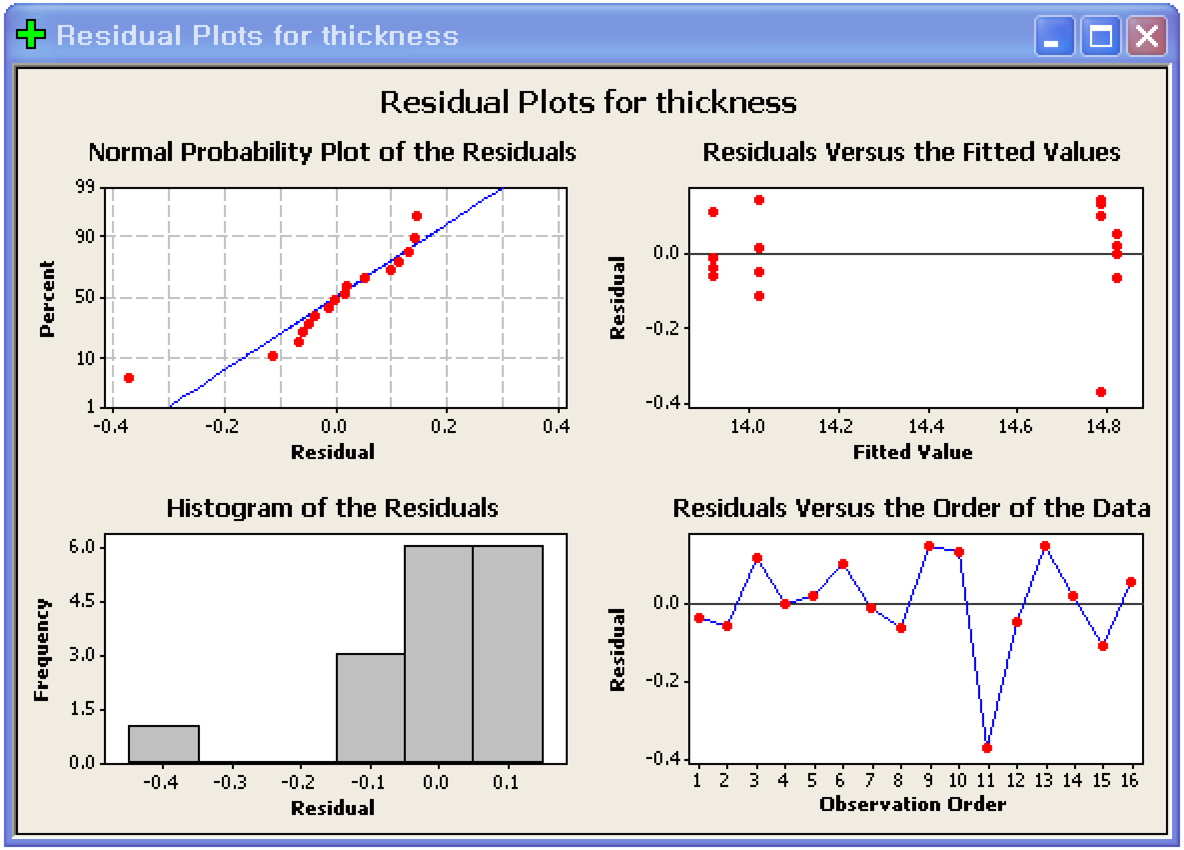
R denotes an observation with a large standardized residual.

Estimated Coefficients for thickness using data in uncoded units

Term	Coef
Constant	15.3592
Time	-0.04291
Flowrate	-0.0170313
Time*Flowrate	0.0080937

Exercise with Minitab (cont.)

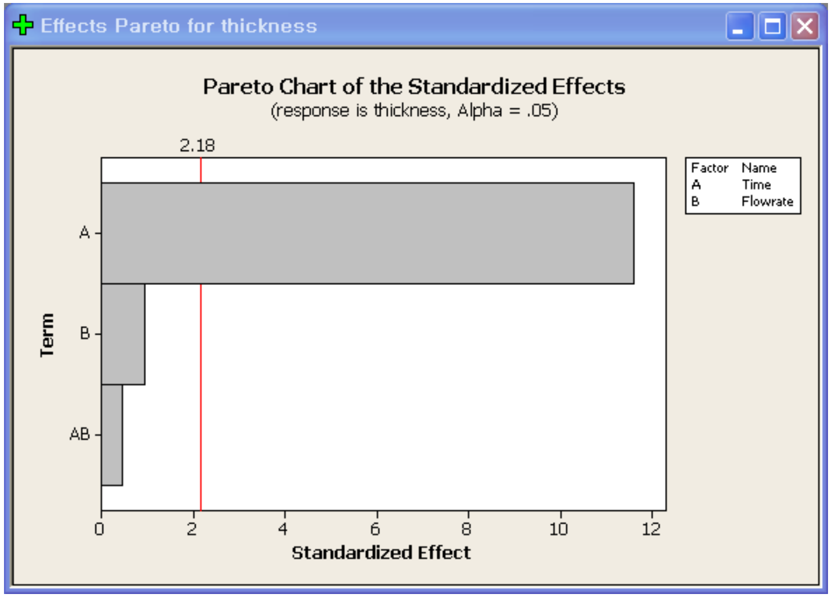
(b) Residual plots

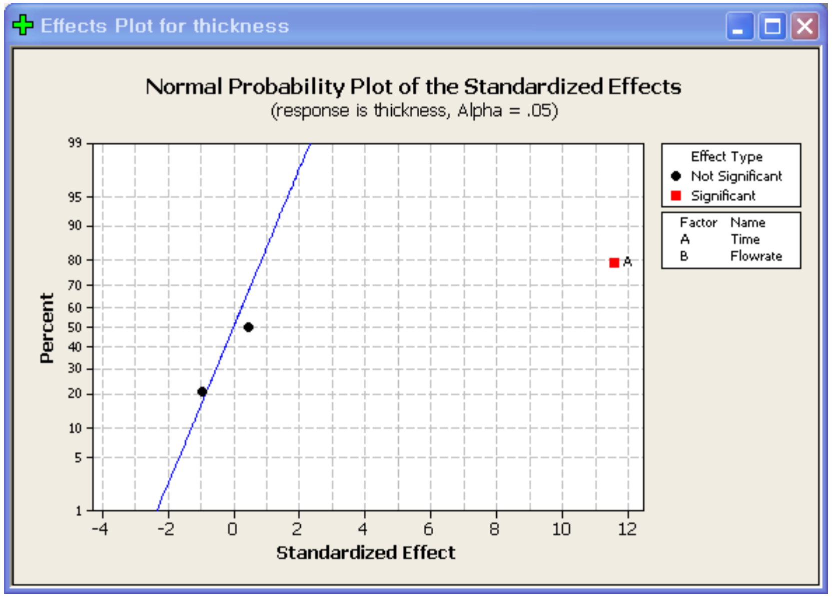


Exercise with Minitab (cont.)

(c) Plots for effects

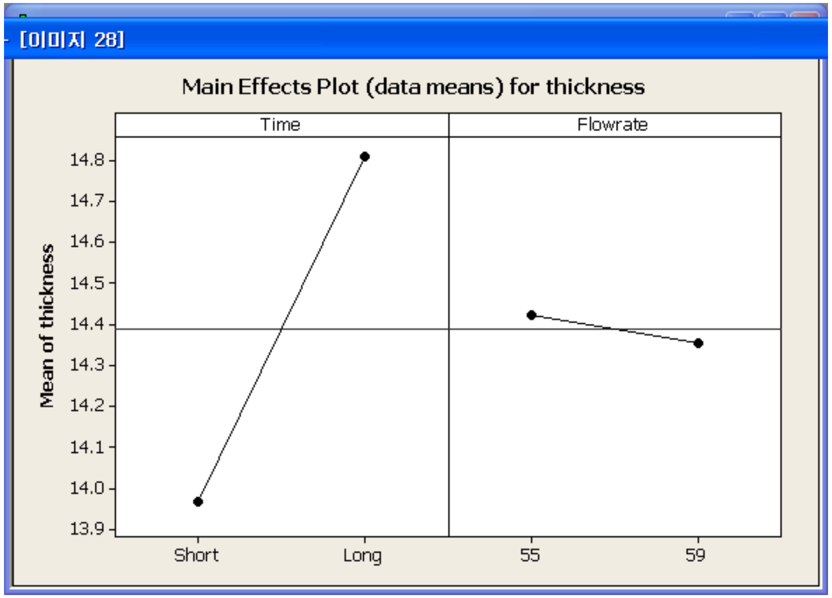
You can also determine which factors have significant effects.

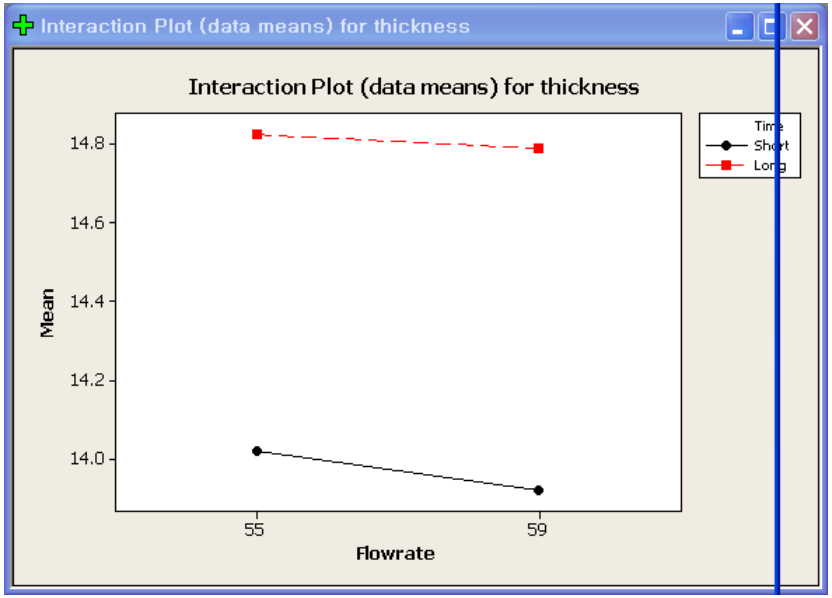




Exercise with Minitab (cont.)

Alternatively, main/interaction plot





Exercise with Minitab (cont.)

- Depending on your goal, you can refine a prediction model by selecting significant factors (variables) only.

  → less # of coefficients

  → more degree of freedom

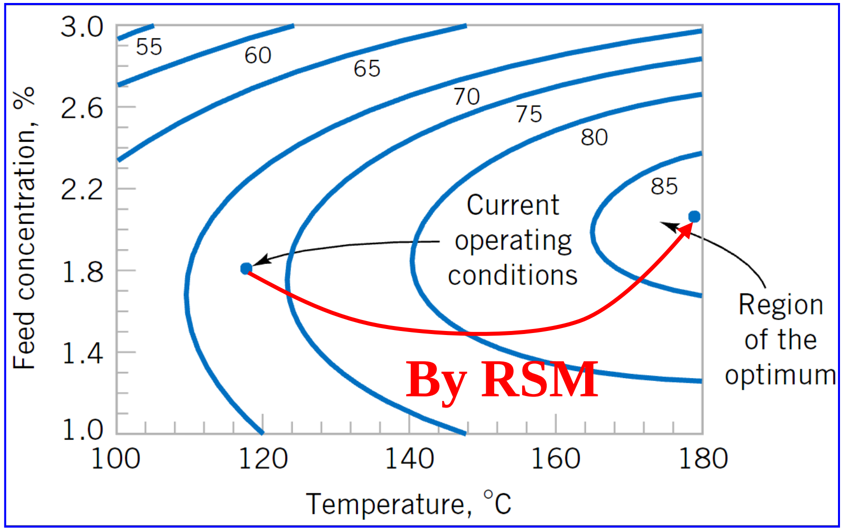
  → more accurate estimate of C.I ($S_{y/x}$ can decrease)

This is very useful even when you have ***many factors*** and ***no replicates***.

Principle of sparsity of effects: the system (process) is usually dominated by the main effects and low-order interactions. That is, the three factor and higher-order interactions are usually negligible.

Response Surface Methods (RSM)

- RSM
  - Objective: optimize a process (or system) using mathematical & statistical techniques.
  - But, the process is usually unknown. (i.e., relationships between x & y variables are unknown.)



➔ (1) The First step of RSM is to ***find a (approximate) model*** of the process using least squares (& DOE).

➔ (2) Next step is to ***improve process operation*** by moving to a better operating point using the model.

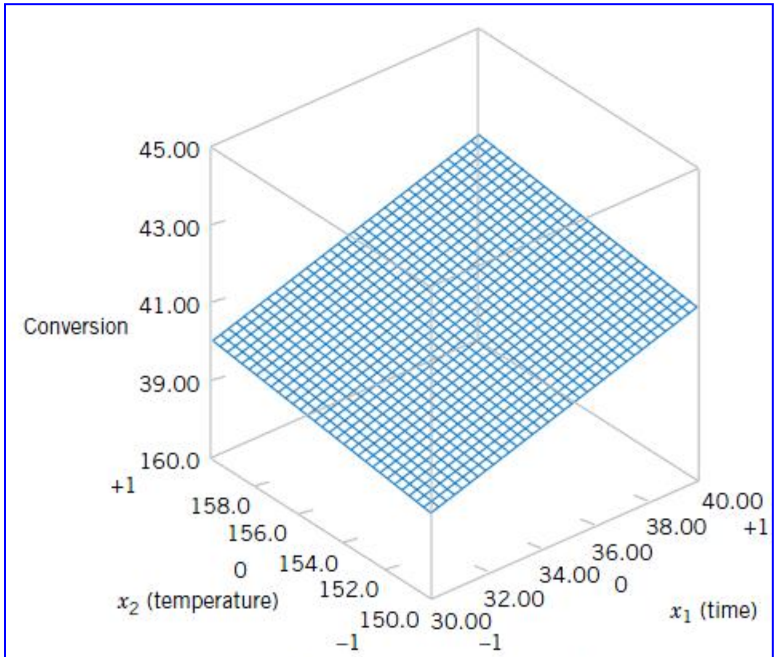
➔ (3) Repeat this until optimum is reached.

FYI (For Your Information)

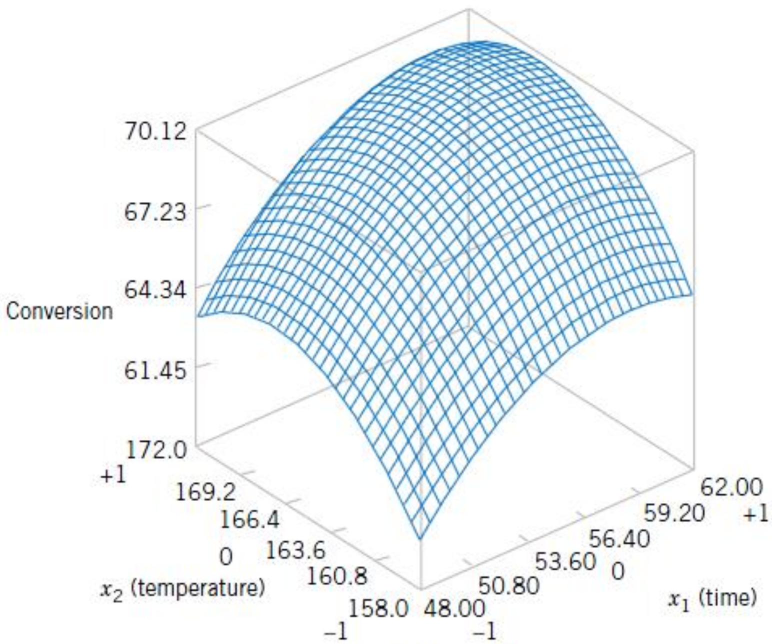
- Response surface?

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2$$

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + a_{11} x_1^2 + a_{22} x_2^2$$



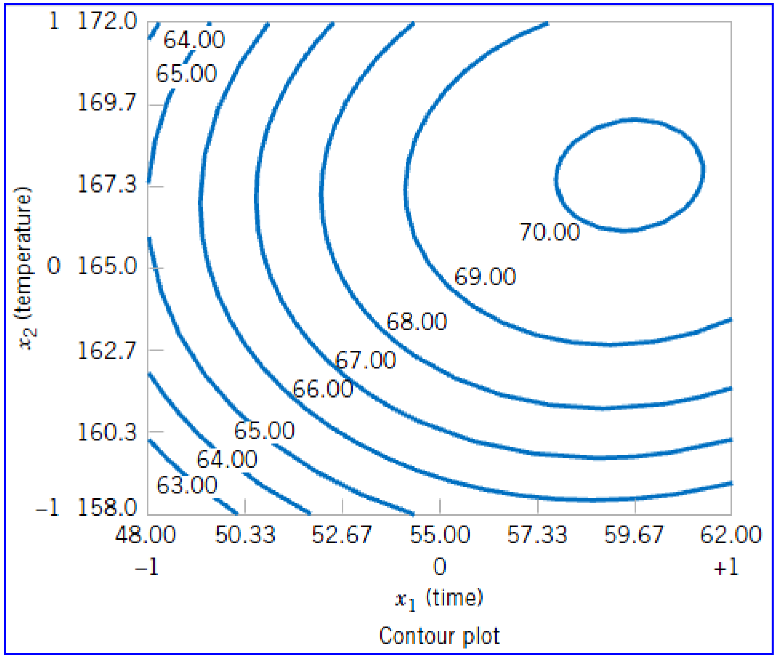
Three-dimensional surface plot

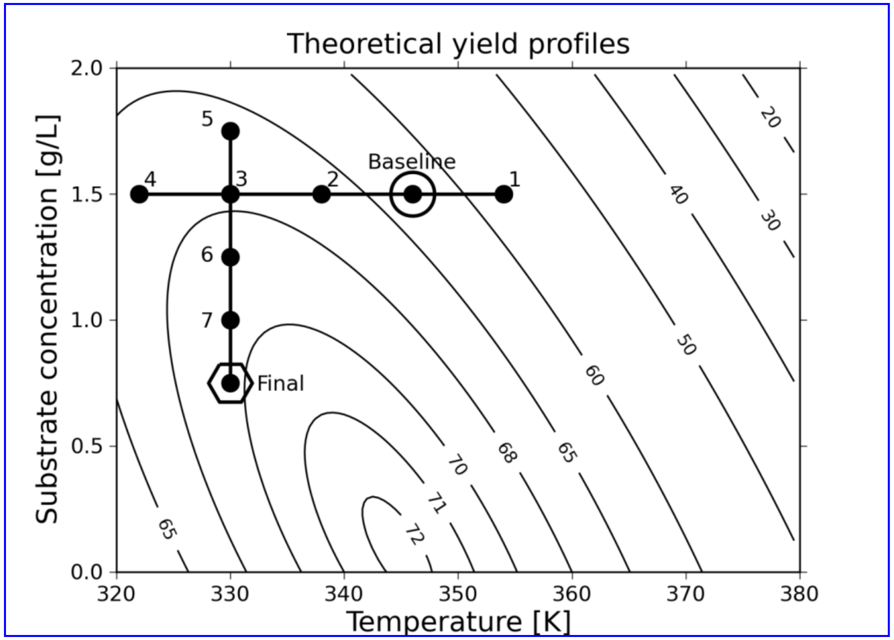


Surface plot

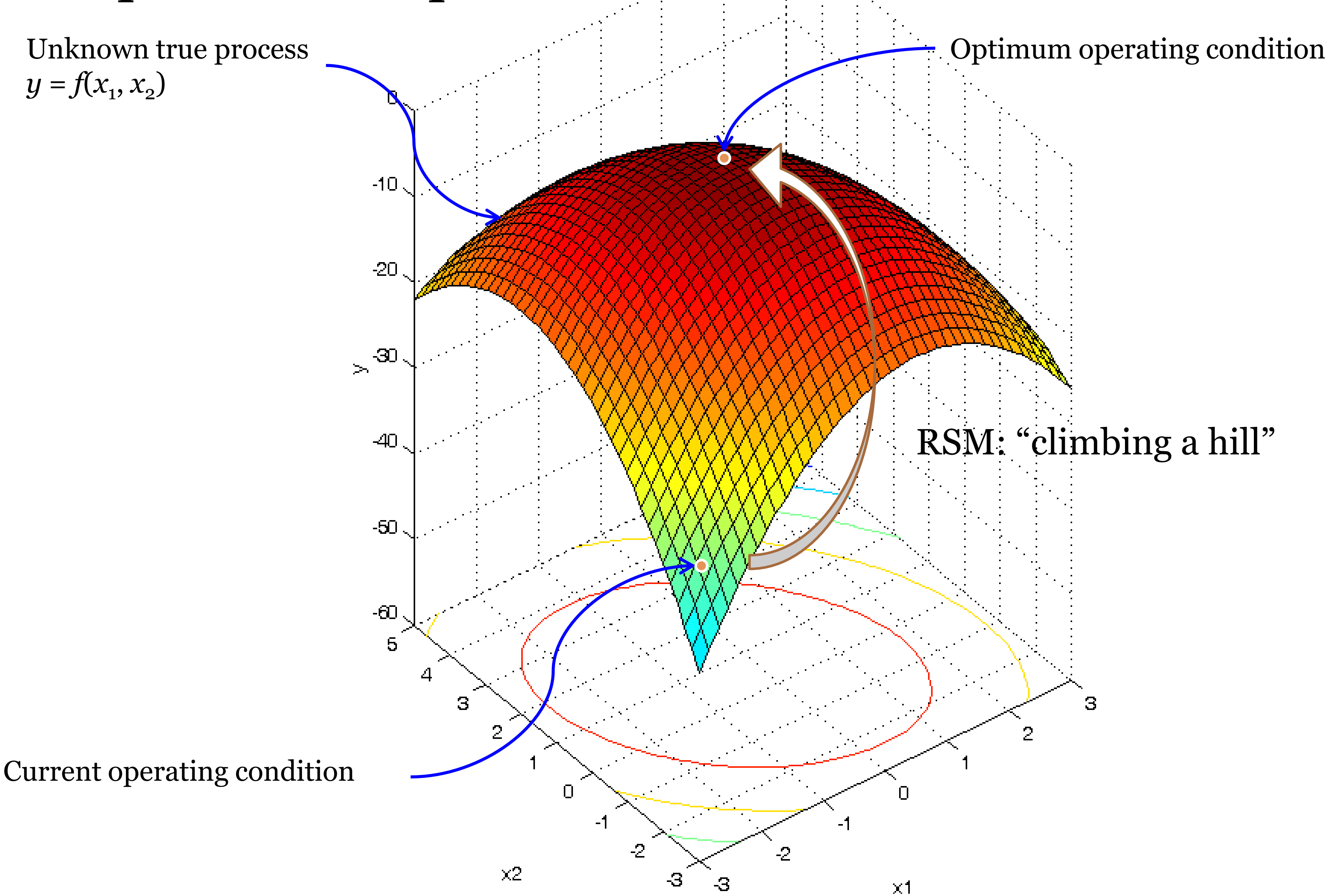
FYI (For Your Information)

- COST **costs too much** to find optimum when interaction exists.
  - Compare two cases





Graphical interpretation of RSM (1)

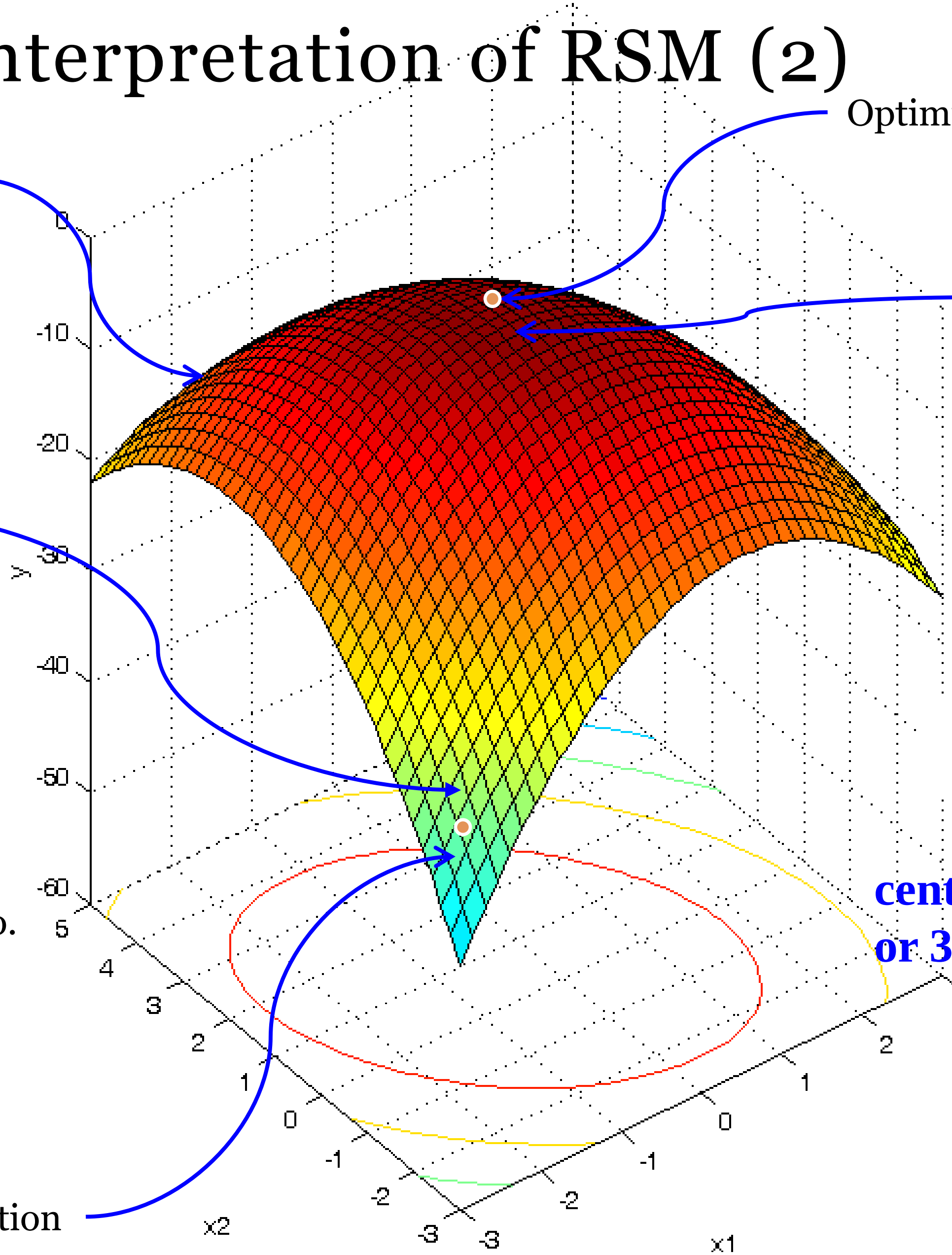


Graphical interpretation of RSM (2)

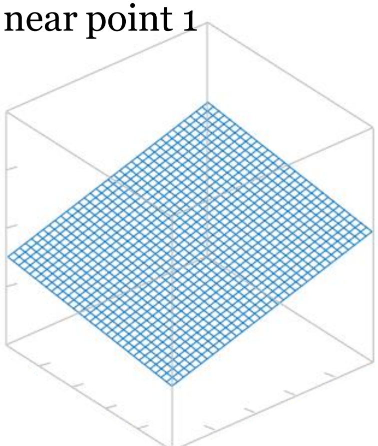
Unknown true process
$y = f(x_1, x_2)$

Optimum operating condition

Approximate model
near point 2

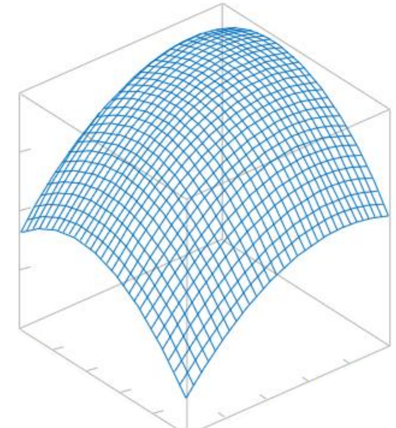


Approximate model
near point 1



A linear model form will do.

(full) factorial design



Quadratic model form
needed

**central composite design
or 3-level design**

Current operating condition

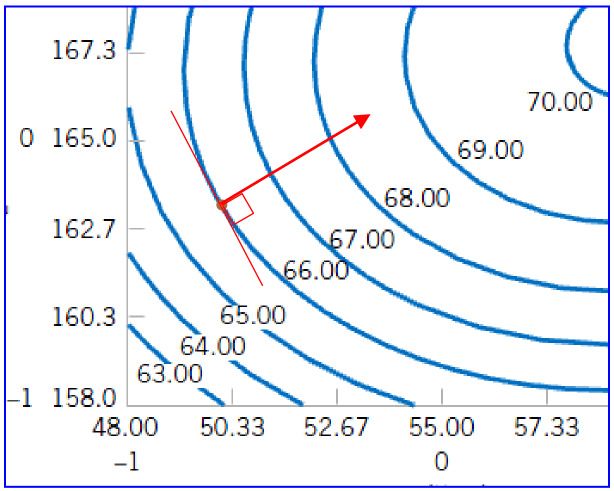
Graphical interpretation of RSM (3)

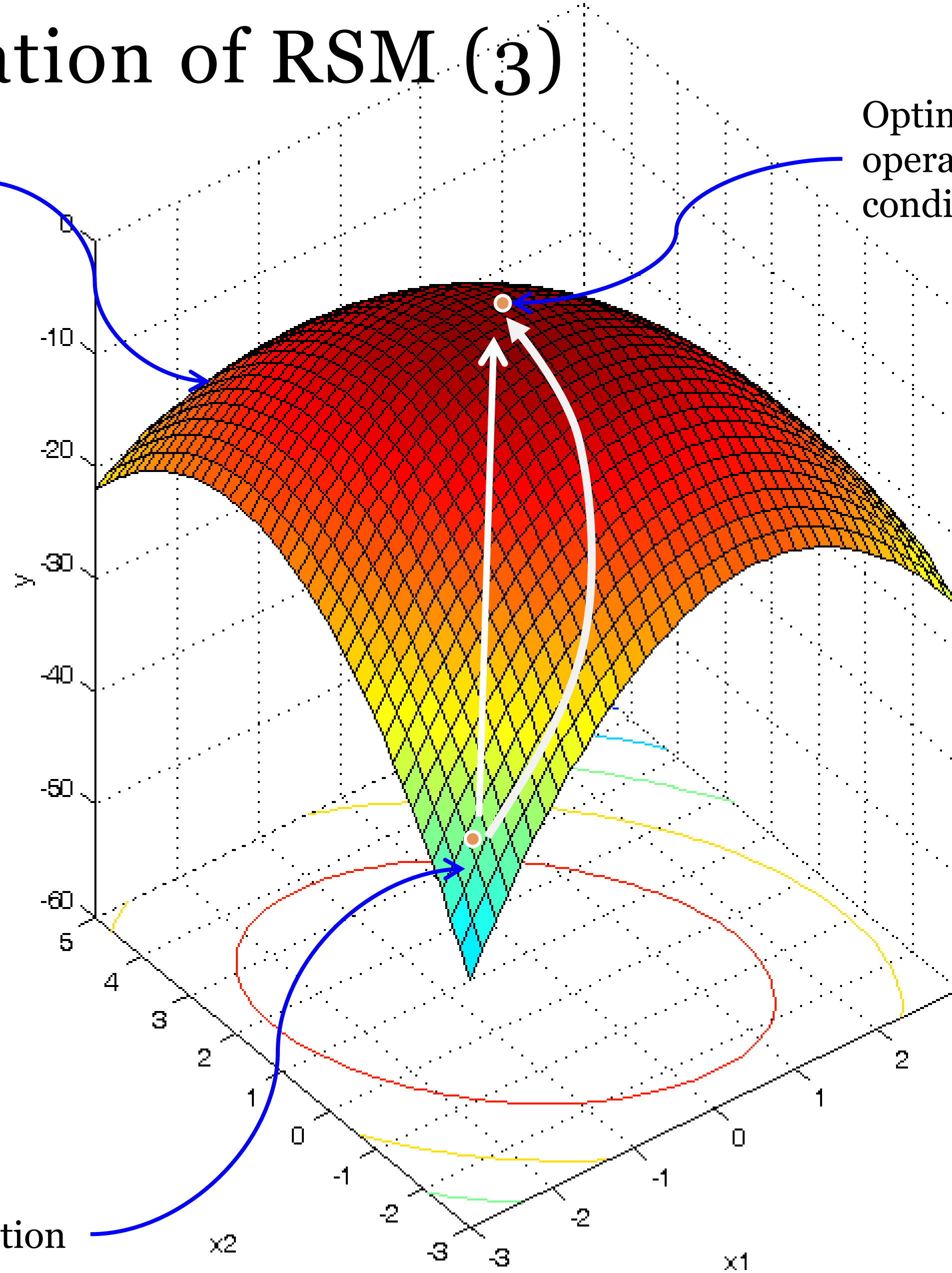
Unknown true process
$y = f(x_1, x_2)$

Optimum operating condition

The fastest way to climb a hill

Methods of steepest ascent





Direction of steepest ascent

$$= \begin{pmatrix} \dfrac{\partial y}{\partial x_1} & \dfrac{\partial y}{\partial x_2} \end{pmatrix}$$

$$\cong \begin{pmatrix} a_1 & a_2 \end{pmatrix}$$ when interaction is smaller than main effects

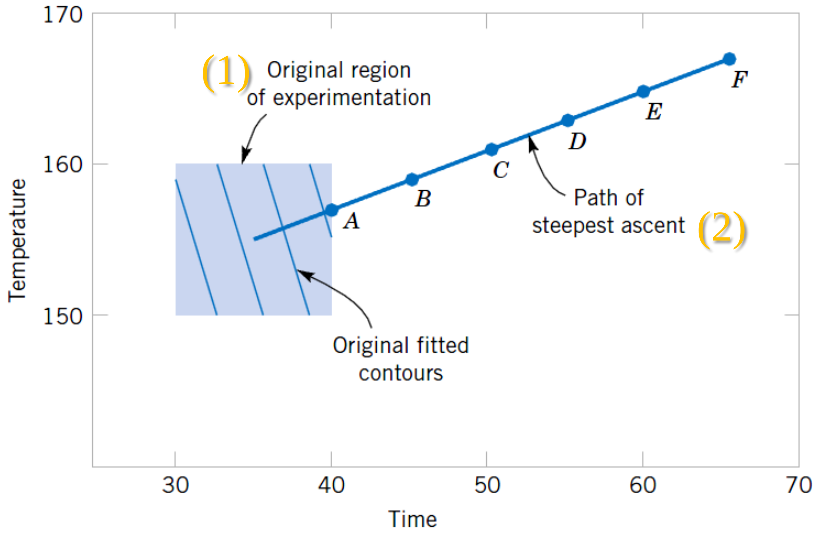
Current operating condition

RSM (cont.)

- General procedure
  1. Perform (fractional) factorial design around current operating conditions & fit a linear model form

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{123}x_1x_2x_3$$

  2. Calculate direction of S.A. & perform experiments along this direction until response doesn't improve. (step size to be determined carefully)



Point A: 40 minutes, 157°F, y = 40.5
Point B: 45 minutes, 159°F, y = 51.3
Point C: 50 minutes, 161°F, y = 59.6
$\boxed{\text{Point } D\text{: 55 minutes, 163°F, } y = 67.1}$
Point E: 60 minutes, 165°F, y = 63.6
Point F: 65 minutes, 167°F, y = 60.7

RSM (cont.)

3. Lay down a new factorial design.
4. Repeat steps 1 ~ 3 until linear model is insufficient.
   - Curvature shows up.
   - 2-factor interaction dominate main effects.
5. Estimate a quadratic model if curvature and/or interaction is large relative to main effects.
   - Add star points → central composite design
   - Or three-level design

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2$$

6. Plot response contour and move towards to best conditions

(most statistical software will do this)

RSM Exercise

Yield $y = f(T, S)$

Current operating conditions

- **T** = 325 K
- **S** = 0.75 g/L
- Profit = $407

Step 1

Experiment	T	S	Profit
1	−	−	193
2	+	−	310
3	−	+	468
4	+	+	571

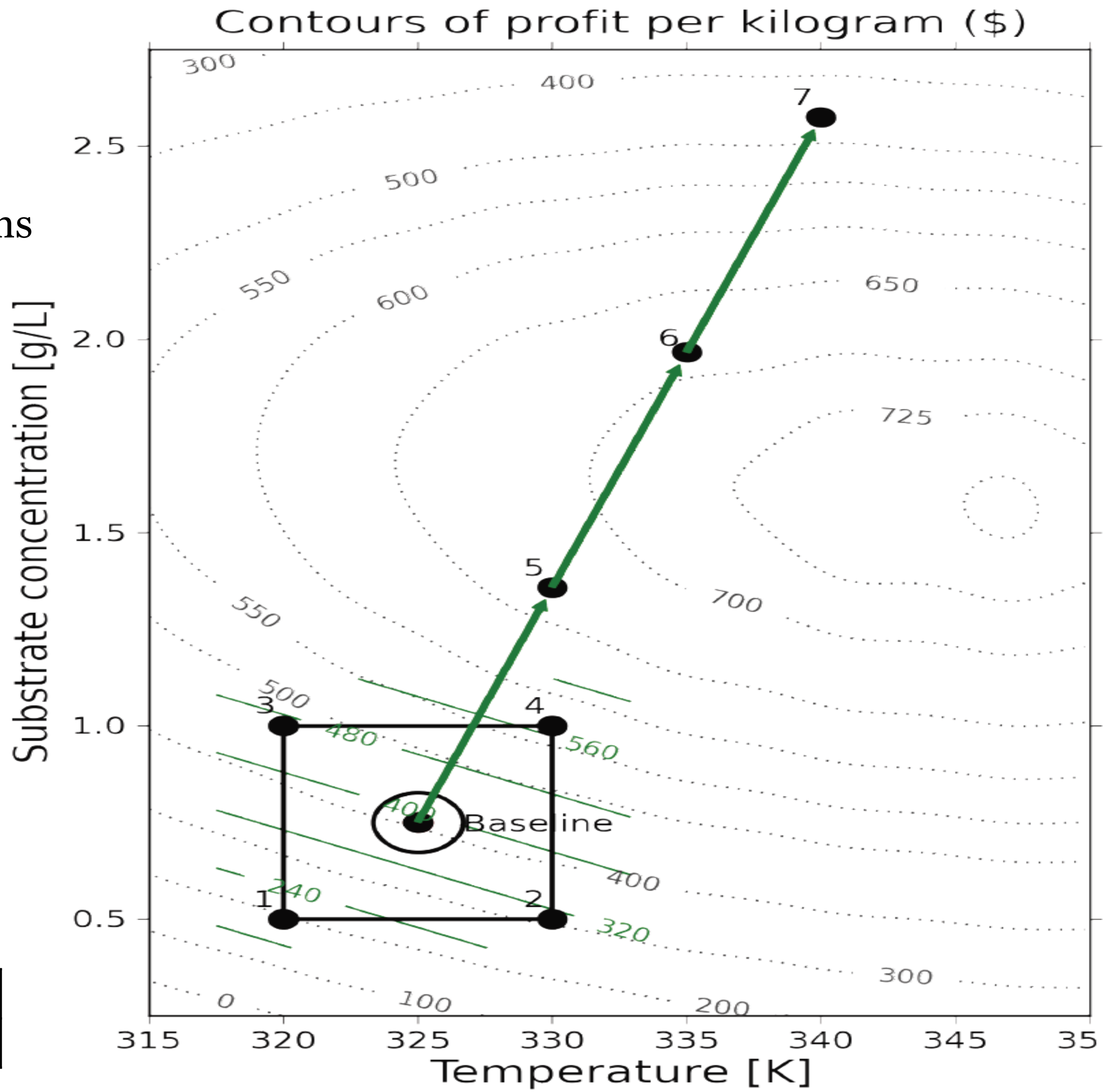
$$\hat{y} = 385.6 + 55x_T + 134x_S - 3.75x_T x_S$$

Step 2

Derection of S.A

$$= \begin{pmatrix} \dfrac{\partial y}{\partial x_T} & \dfrac{\partial y}{\partial x_S} \end{pmatrix} \cong \begin{pmatrix} 55 & 134 \end{pmatrix}$$

experiment	5	6	7
profit	$669	$688	$463



RSM Exercise

Step 3

Experiment	T	S	Profit
8	−	−	694
9	+	−	725
10	−	+	620
11	+	+	642
6	0 (335 K)	0 (1.97 g/L)	688

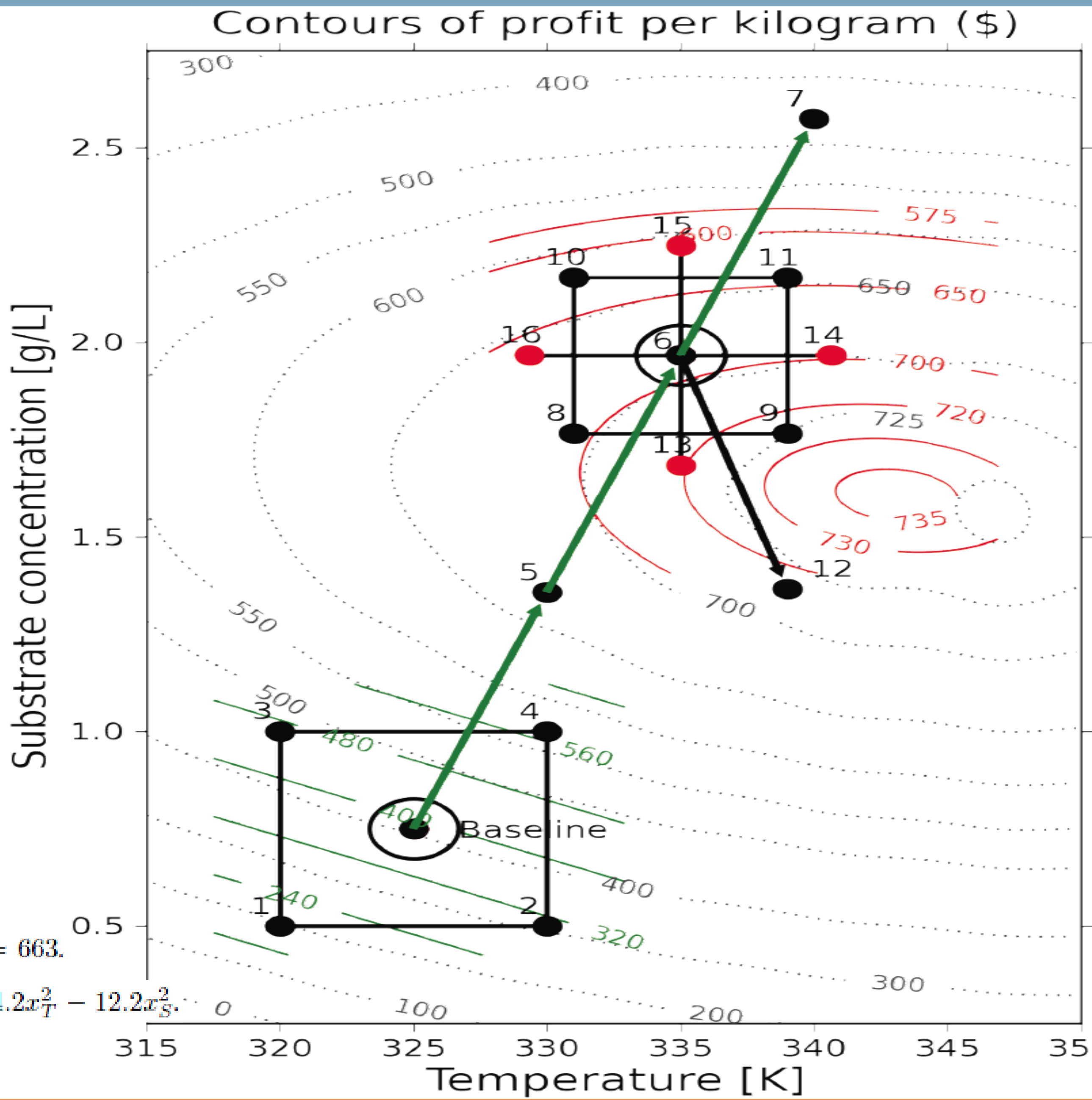
$$\hat{y} = 670 + 13x_T - 39x_S - 2.4x_T x_S$$

Derection of S.A

$$= \begin{pmatrix} \dfrac{\partial y}{\partial x_T} & \dfrac{\partial y}{\partial x_S} \end{pmatrix} \cong \begin{pmatrix} 13 & -39 \end{pmatrix}$$

Profit (12) = 716 < profit (9)

→ Strong interaction



Step 5

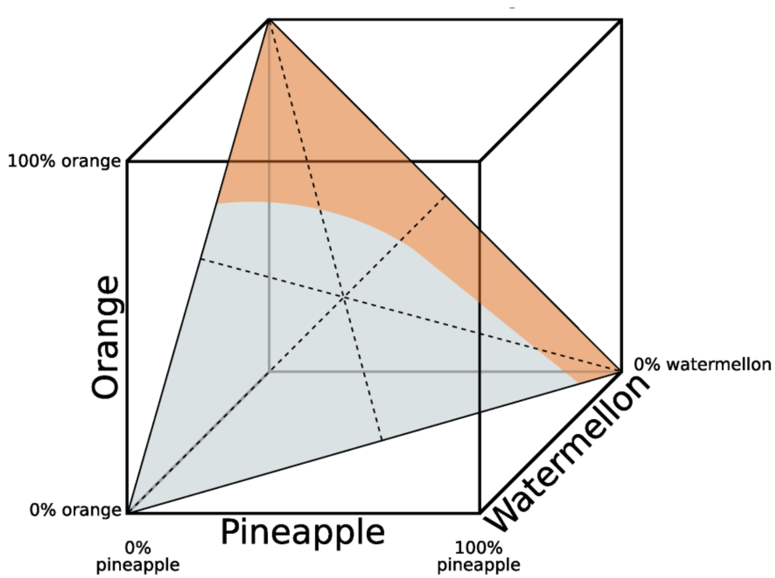
Star points

$y_{13} = 720$, $y_{14} = 699$, $y_{15} = 610$, and $y_{16} = 663$.

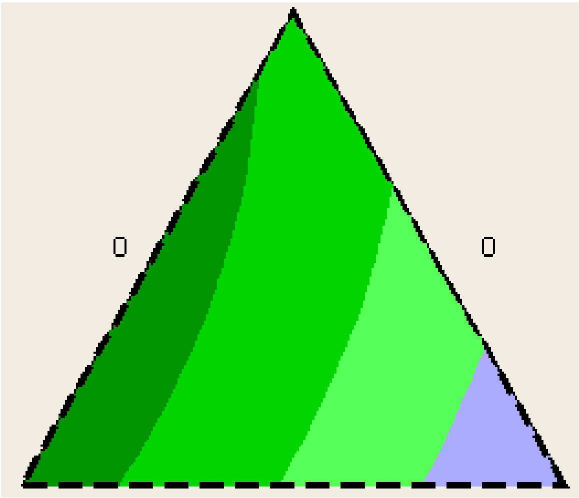
$$y = 688 + 12.9x_T - 39.1x_S - 2.4x_T x_S - 4.2x_T^2 - 12.2x_S^2.$$

Mixture design

- Mixture design
  - Ordinary factorial design with a constraint
    - $0 \le x_A, x_B, x_C \le 1,\ \ x_A + x_B + x_C = 1$
  - Of course, RSM can be used to determine best mixture.

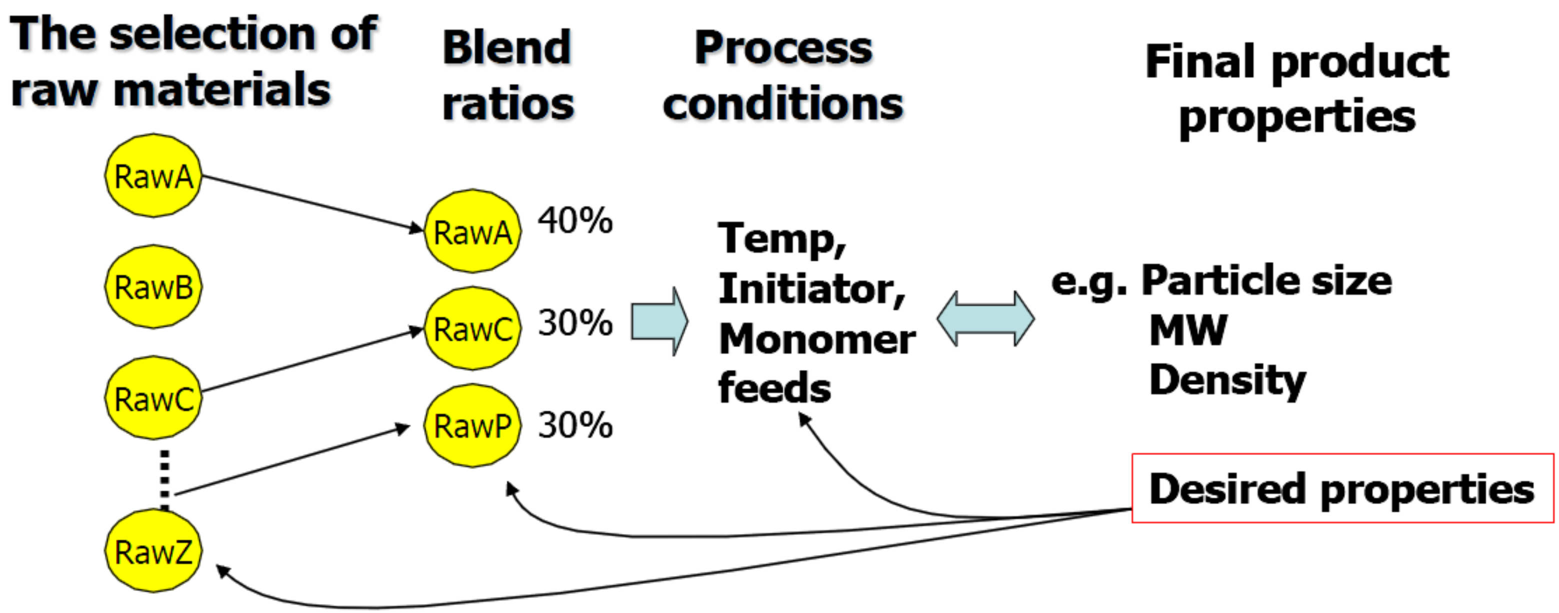


Mixture contour plot



Mixture design (cont.)

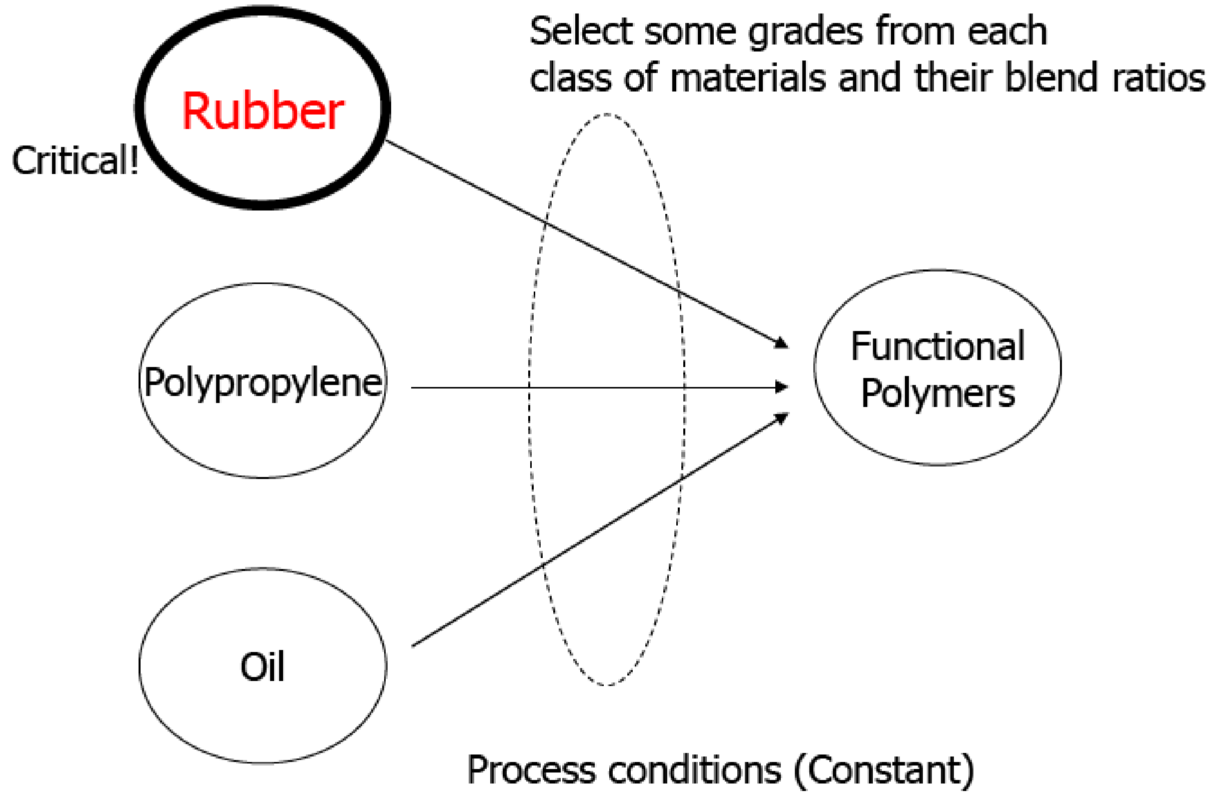
- Example: Product design (development)



Mixture design (cont.)

Example: Functional Polymer Development

Mitsubishi Chemicals



Mixture design (cont.)

- (Advanced) Mixture design example

Our approach has increased the resilience 1.7 times compared to previous products