KOREA

DISCOVERY SUMMIT

EXPLORING DATA
INSPIRING INNOVATION

jmp®

김광희, 옥창수, 정순우 (삼성전기)

DOE(Taguchi vs RSM)

Optimization Experimental Design

김광희 (Big Data 분석 전문가)
옥창수 (품질 관리 기술사)
정순우 (품질 관리 기술사)

DISCOVERY SUMMIT

jmp

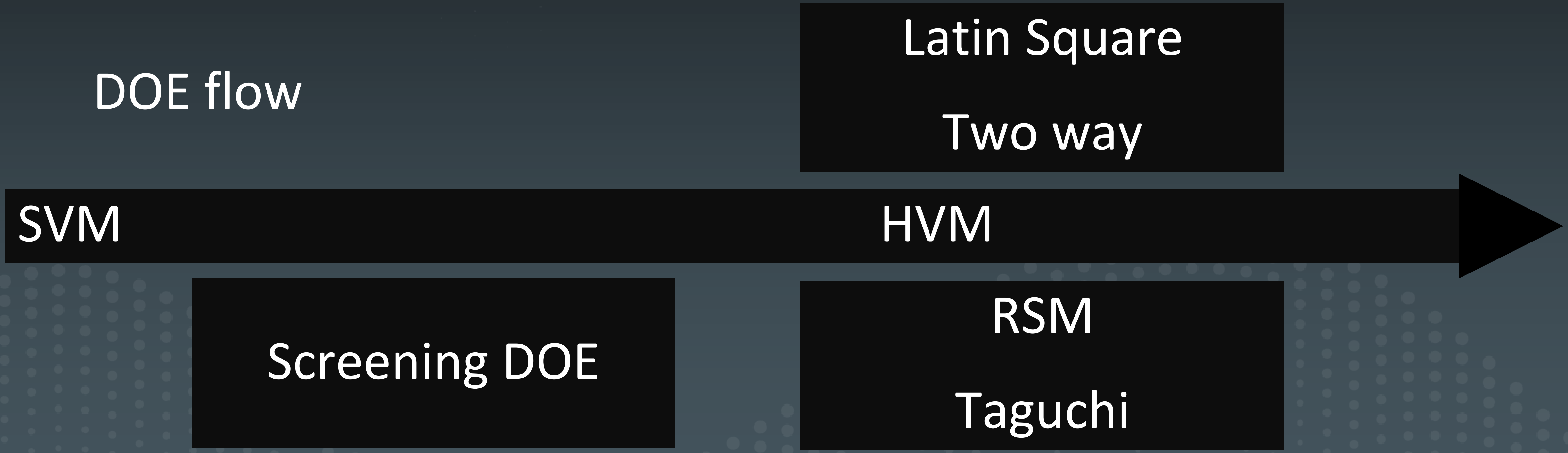
Content

1. What is DOE
2. DOE simulation
 - Taguchi
 - RSM
3. Result

Background

- Why Taguchi vs RSM ?

DOE flow



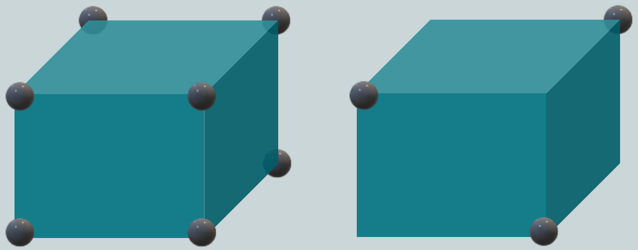
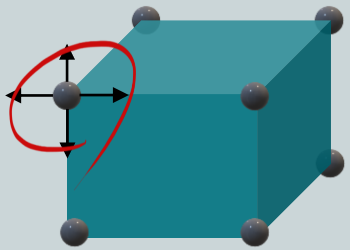
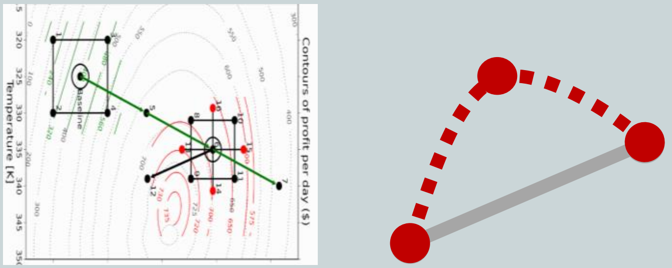
DOE 5 principals

- Randomization : Random assignment → Minimize bias
- Replication : Repeat → Reliability △ (Degrees of freedom for the error term increase.)
- Blocking : Grouping similar units → Controls for variables
- Orthogonality : Independent factor combinations → Reduce confounding
- Confounding : Overlap of factor effects → Affects clarity

DISCOVERY SUMMIT

jmp

Types of DOE designs

	Factorial Design	Taguchi	RSM
Definition			
Purpose	Study the effect of multiple factors	Minimize variation Robustness design	Identify optimal settings with fewer experiments
Remark	Liner effect	Find loss function	Optimum point can be estimated using curvature

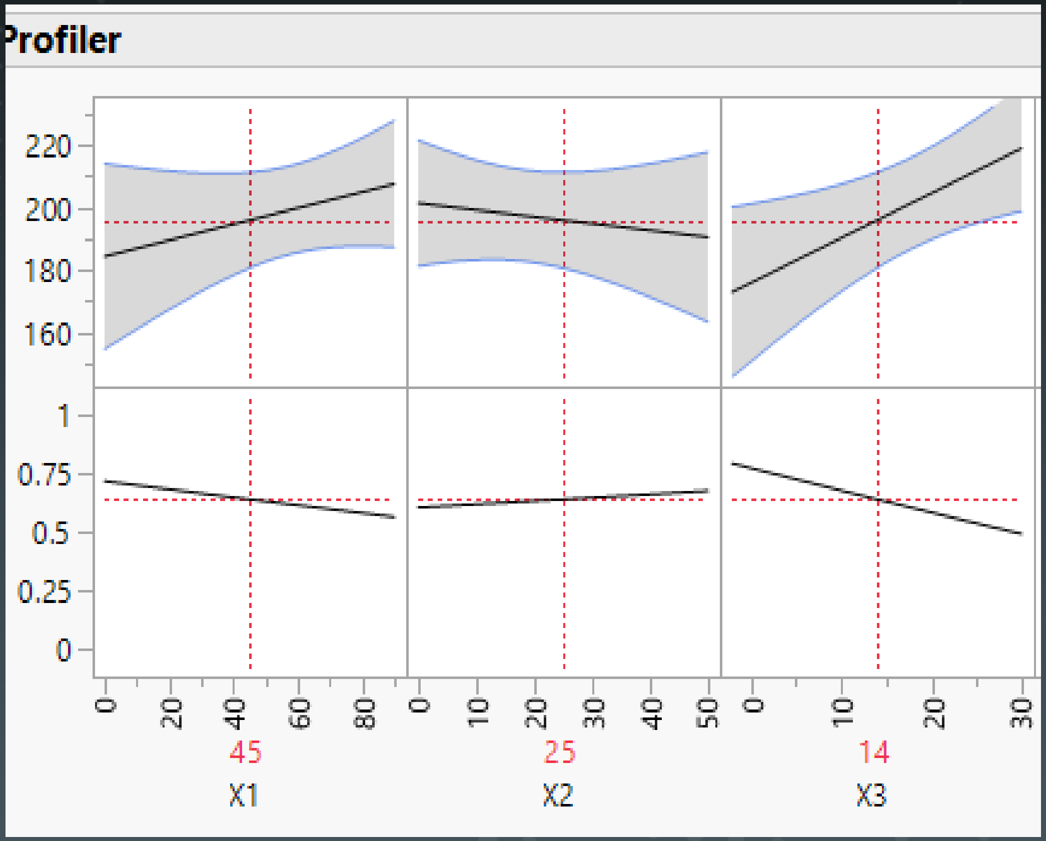
DISCOVERY SUMMIT

jmp

Factorial design

- Input & Out put

		A1	A2	A3
B1	C1			
	C2			
	C3			
B2	C1			
	C2			
	C3			
B3	C1			
	C2			
	C3			

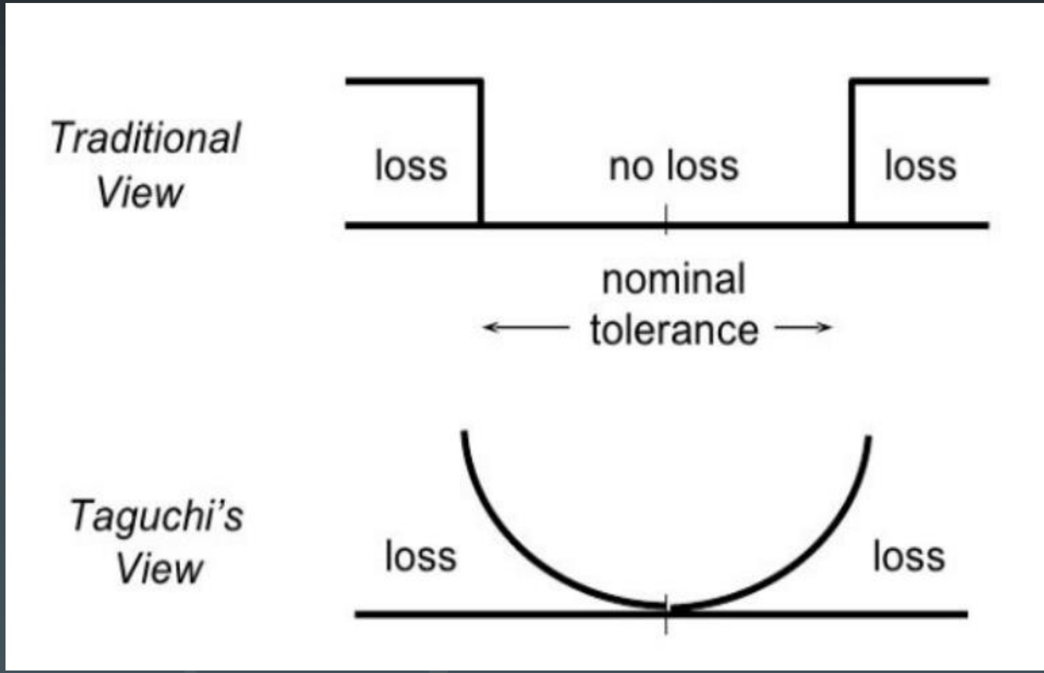


- Cost and Time
- Sensitivity to Noise
- Limited Optimization and Surface Analysis

※ 3^3 : 3-factor, 3-level design

Taguchi design

- Shift → Loss → On target



- Robust design against to noise → Preferred higher SNB

- Noise : Uncontrollable factors that affect the experiment

  (Ex : Environment condition,

  User variability and etc.)

Taguchi design

- SNB : Signal to Noise ratio → Preferred higher SNB

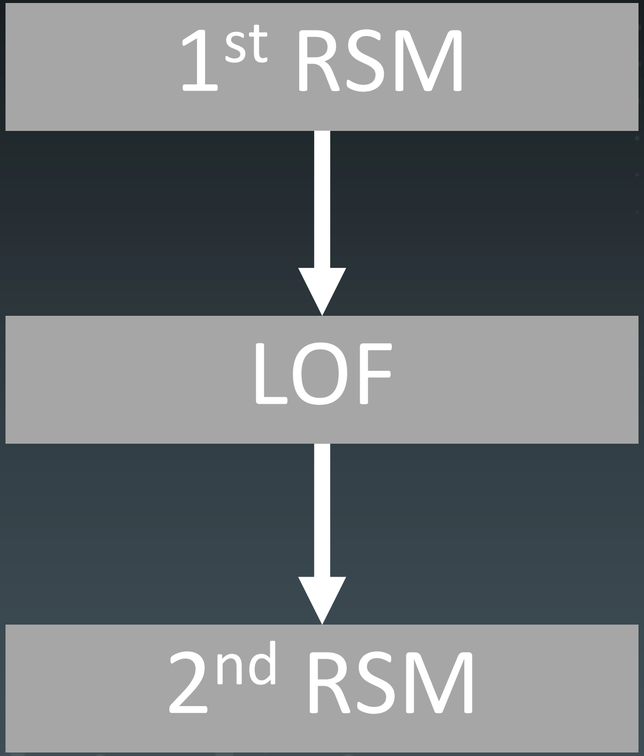
- Nominal is Best = $10log\left(\frac{\bar{y}^2}{s^2}\right)$

Signal

- Larger is Better = $-10log\left(\frac{\sum_{i=1}^{n}\frac{1}{y^2}}{n}\right)$

- Smaller is Better = $-10log\left(\frac{\sum_{i=1}^{n}y^2}{n}\right)$

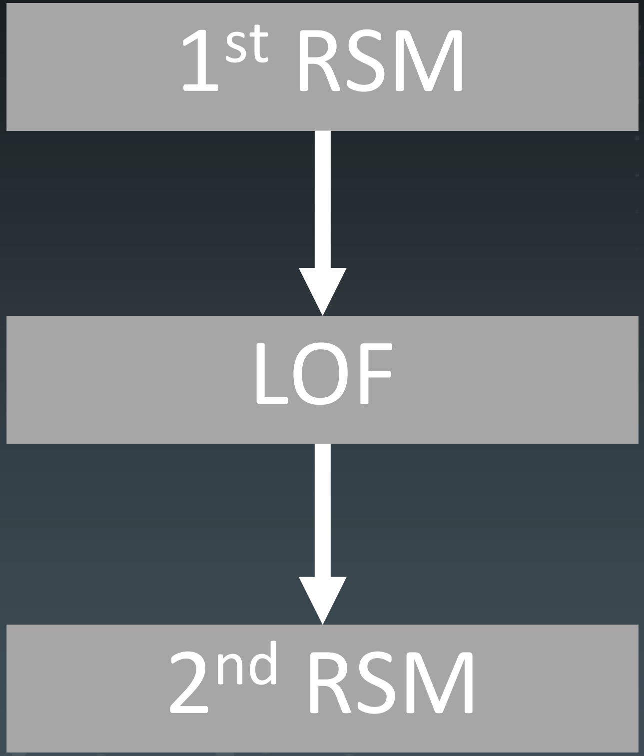
RSM design



- 1st order model : Assumes a linear relationship
- 2nd order model : Define optimized condition

  - 3Level : Accounts for a nonlinear relationship

RSM design



※ LOF : Lack of Fit

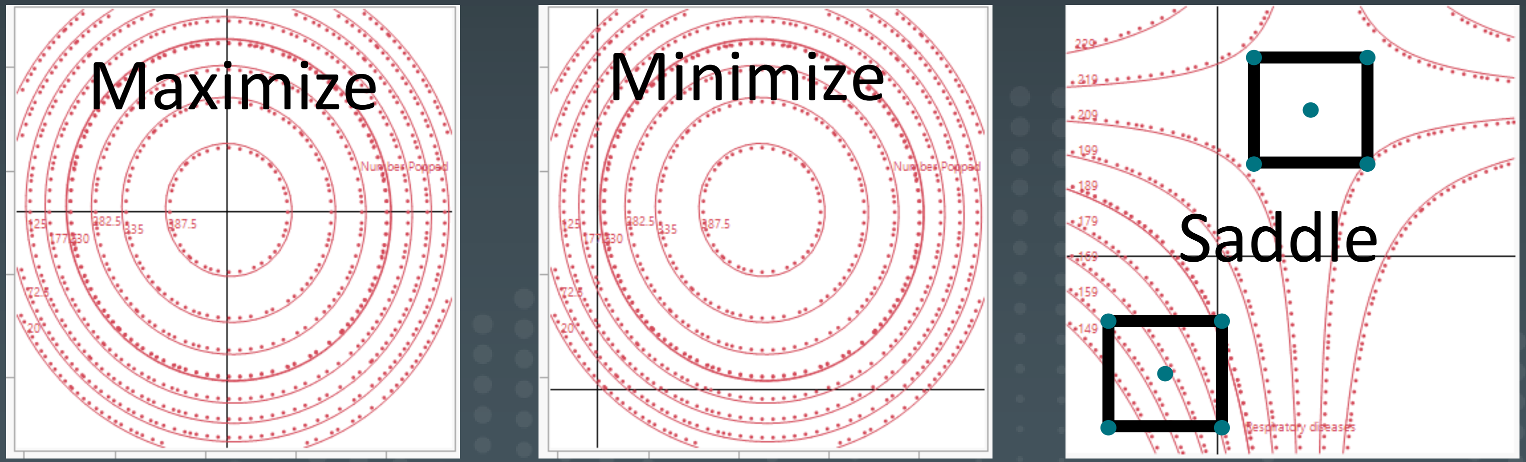
- The assumed regression model is validated through an F-test.

$H0$: No Lack of Fit (LOF).

$H1$: Lack of Fit (LOF) present.

As the F-value increases, and the p-value decreases, $H1$ is more likely to be accepted.

※ Contour plot (3types)



RSM

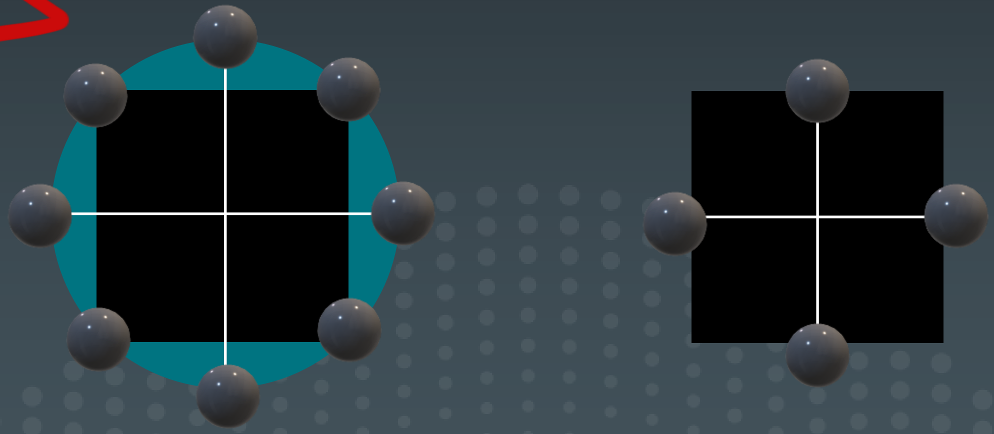
- CCD : Central composite Design

  → Rotatability : Equal variance

  → Corner : Extreme value

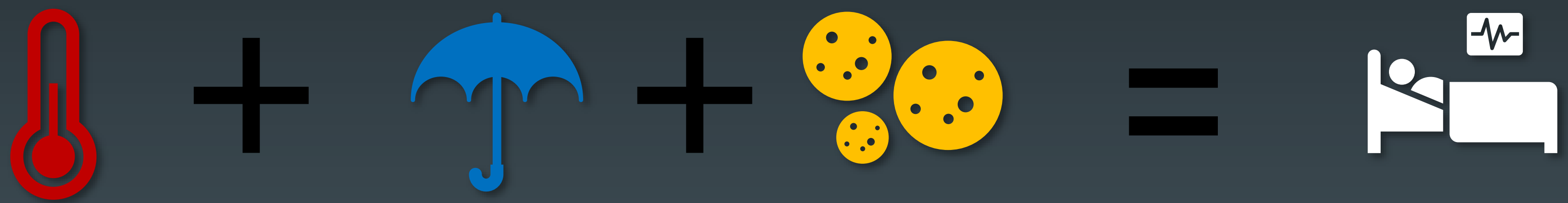
- Box benken

  → Edge : Realistic value



Condition

- Under what conditions of temperature, precipitation, and fine dust levels does the rate of respiratory diseases increase?

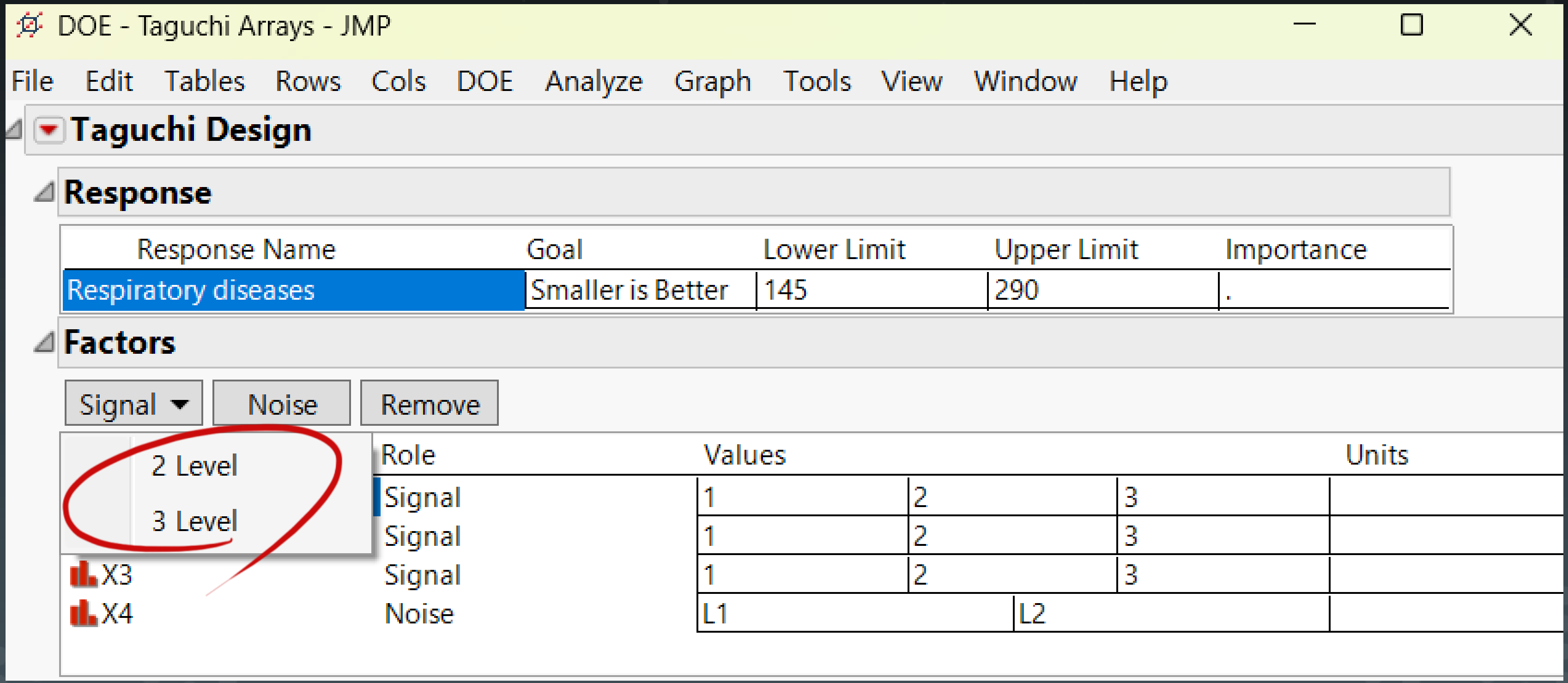


- Data reference

- https://www.weather.go.kr (기상청)

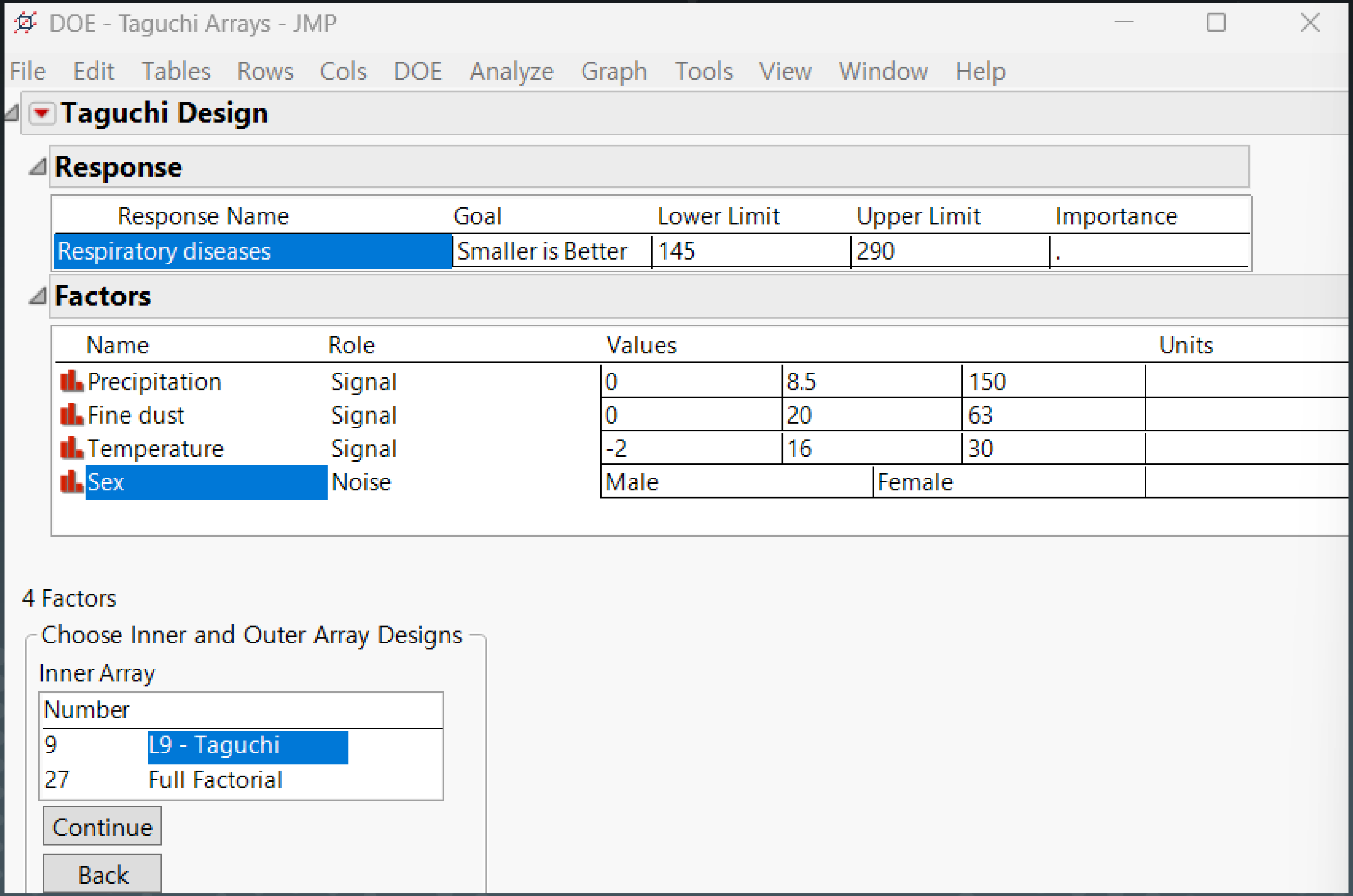
- https://ncv.kdca.go.kr(감염병 포털)

Taguchi



- Select Level
- Noise

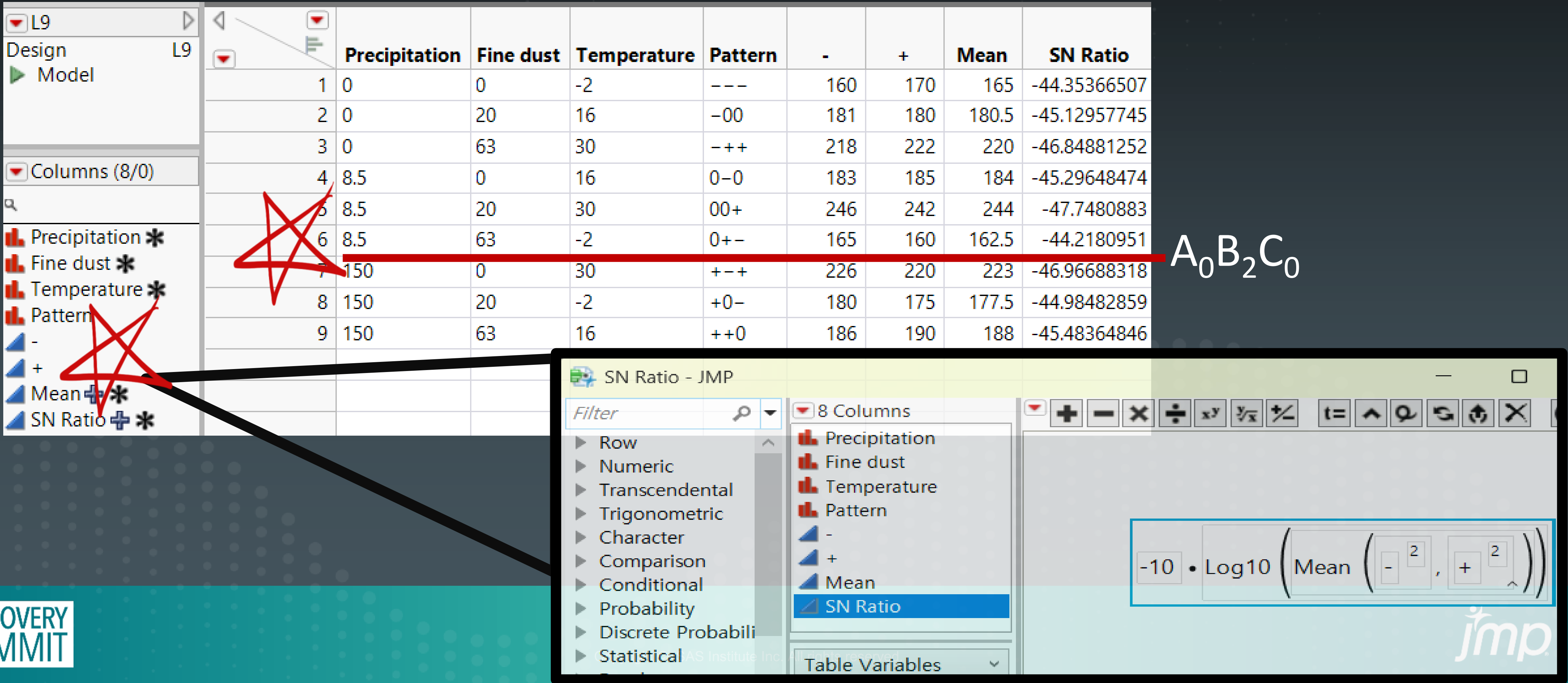
Taguchi



- Add Noise
- L9 : L9(3^4) Frequent orthogonal table in 3level, 4factors

Taguchi

- Mean, SNB (Smaller is Better = $-10log\left(\frac{\sum_{i=1}^{n} y^2}{n}\right)$)

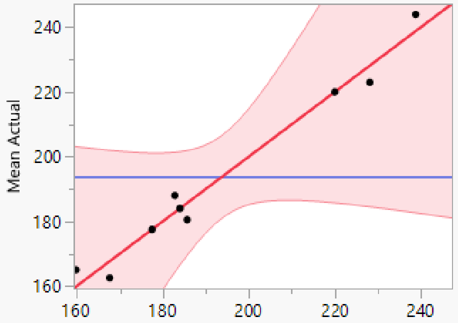


	Precipitation	Fine dust	Temperature	Pattern	-	+	Mean	SN Ratio
1	0	0	-2	−−−	160	170	165	-44.35366507
2	0	20	16	−00	181	180	180.5	-45.12957745
3	0	63	30	−++	218	222	220	-46.84881252
4	8.5	0	16	0−0	183	185	184	-45.29648474
5	8.5	20	30	00+	246	242	244	-47.7480883
6	8.5	63	-2	0+−	165	160	162.5	-44.2180951
7	150	0	30	+−+	226	220	223	-46.96688318
8	150	20	-2	+0−	180	175	177.5	-44.98482859
9	150	63	16	++0	186	190	188	-45.48364846

Taguchi

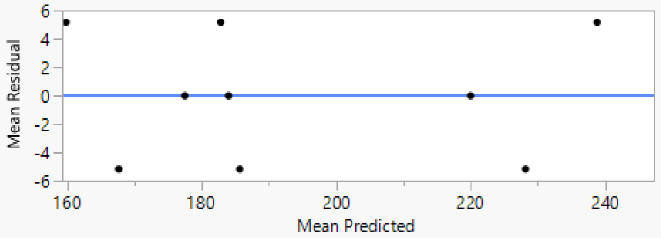
Response Mean

Actual by Predicted Plot

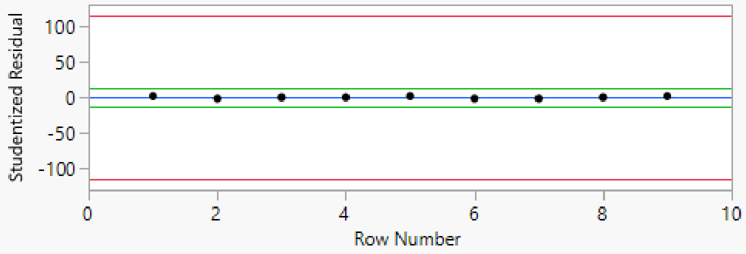


Mean Predicted RMSE=8.9489 RSq=0.98 PValue=0.0728

Residual by Predicted Plot



Studentized Residuals



Externally studentized residuals with 95% simultaneous limits (Bonferroni) in red, individual limits in green.

Scaled Estimates

Nominal factors expanded to all levels

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|
Intercept	193.83333	2.982976	64.98	0.0002*
Precipitation[0]	-5.333333	4.218566	-1.26	0.3335
Precipitation[8.5]	3	4.218566	0.71	0.5507
Precipitation[150]	2.3333333	4.218566	0.55	0.6358
Fine dust[0]	-3.166667	4.218566	-0.75	0.5312
Fine dust[20]	6.8333333	4.218566	1.62	0.2467
Fine dust[63]	-3.666667	4.218566	-0.87	0.4764
Temperature[-2]	-25.5	4.218566	-6.04	0.0263*
Temperature[16]	-9.666667	4.218566	-2.29	0.1490
Temperature[30]	35.166667	4.218566	8.34	0.0141*

- Mean

- C_0 : Mean ↓

- C_2 : Mean ↑

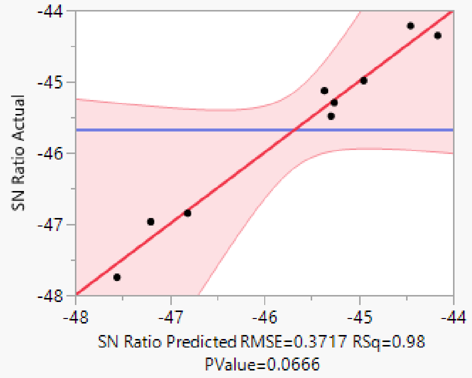
DISCOVERY SUMMIT

jmp

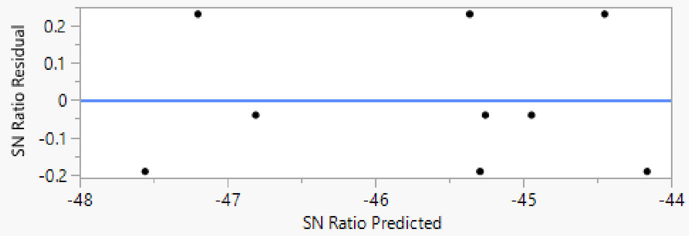
Taguchi

Response SN Ratio

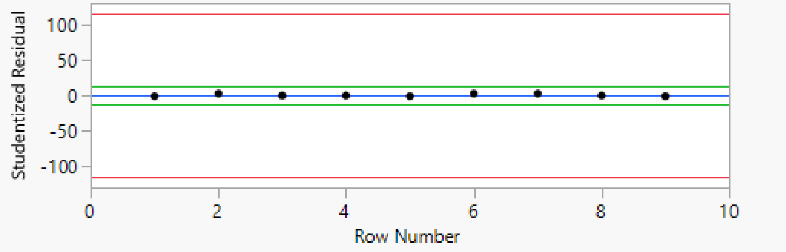
Actual by Predicted Plot



Residual by Predicted Plot



Studentized Residuals



Externally studentized residuals with 95% simultaneous limits (Bonferroni) in red, individual limits in green.

Scaled Estimates

Nominal factors expanded to all levels

Term	Scaled Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-45.67001	0.123915	-368.56	<.0001*
Precipitation[0]	0.2259909	0.175242	1.29	0.3262
Precipitation[8.5]	-0.084213	0.175242	-0.48	0.6783
Precipitation[150]	-0.141777	0.175242	-0.81	0.5034
Fine dust[0]	0.1309983	0.175242	0.75	0.5327
Fine dust[20]	-0.284156	0.175242	-1.62	0.2464
Fine dust[63]	0.1531572	0.175242	0.87	0.4743
Temperature[-2]	1.1511463	0.175242	6.57	0.0224*
Temperature[16]	0.3667724	0.175242	2.09	0.1714
Temperature[30]	-1.517919	0.175242	-8.66	0.0131*

- SN Ratio

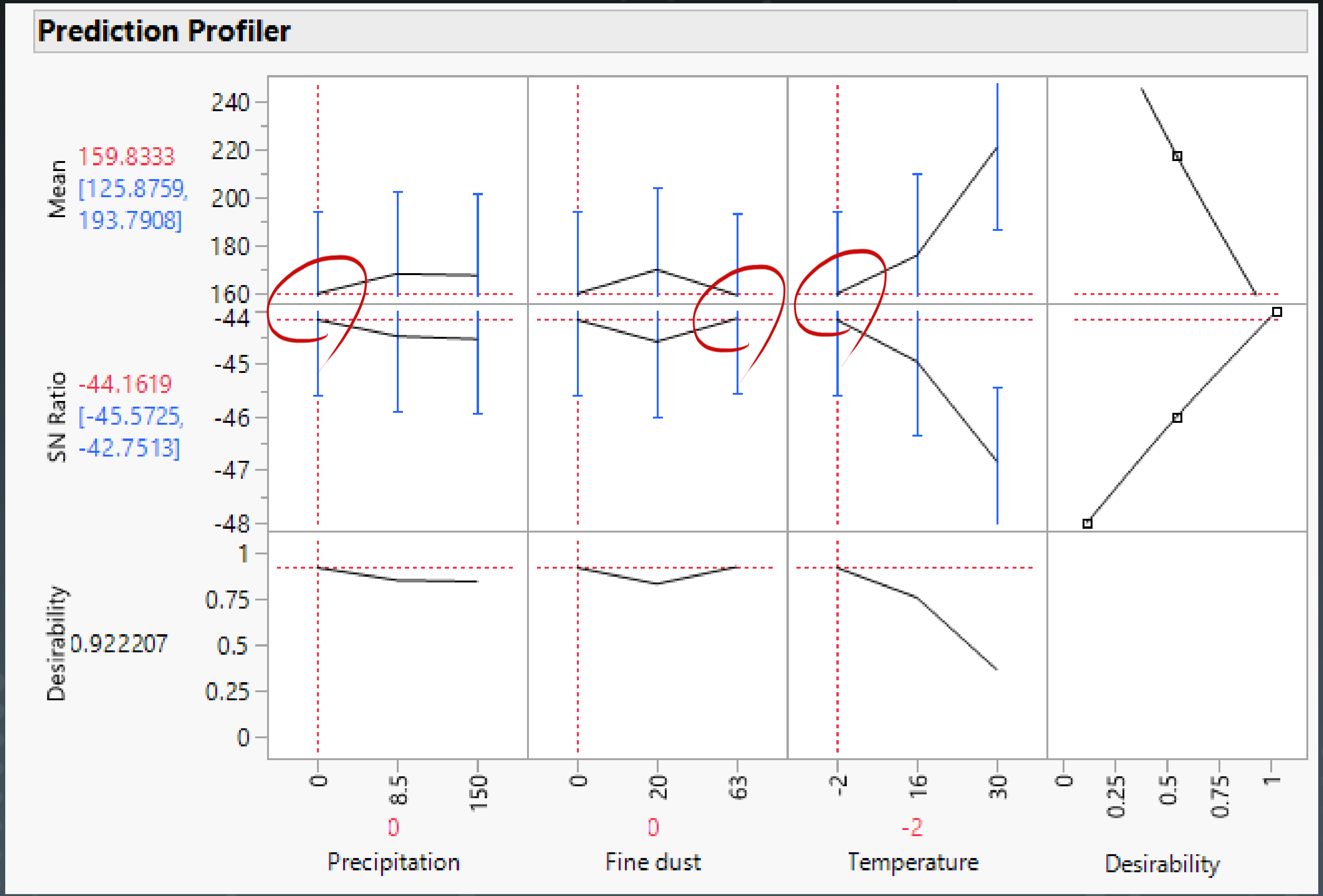
- C_0 : SN Ratio ↑

- C_2 : SN Ratio ↓

DISCOVERY SUMMIT

jmp

Taguchi



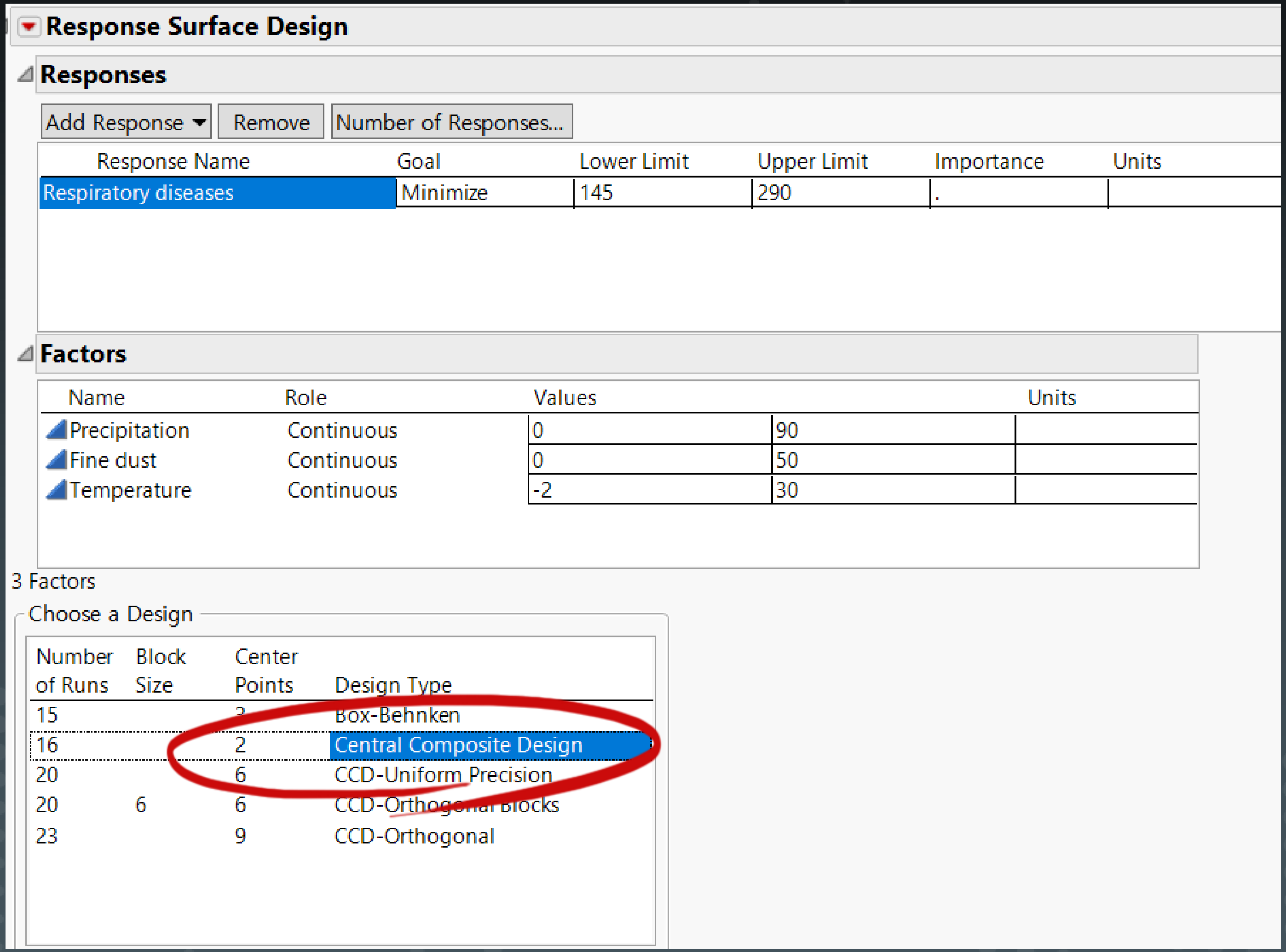
- Conclusion

- Mean : $A_0B_2C_0$ ↓

- SN Ratio : $A_0B_2C_0$ ↑

DISCOVERY SUMMIT

RSM



- 2nd RSM (Assume 1st model is not valid)
- Response : 1
- Factors : 3
- Design : CCD

RSM

	Pattern	Precipitation	Fine dust	Temperature	Respiratory diseases
1	−++	0	50	30	218
2	0A0	45	50	14	210
3	00a	45	25	-2	149
4	00A	45	25	30	223
5	+−−	90	0	-2	229
6	++−	90	50	-2	186
7	+−+	90	0	30	226
8	−−−	0	0	-2	163
9	A00	90	25	14	190
10	+++	90	50	30	210
11	a00	0	25	14	181
12	0a0	45	0	14	180
13	−−+	0	0	30	220
14	000	45	25	14	204
15	−+−	0	50	-2	165
16	000	45	25	14	206

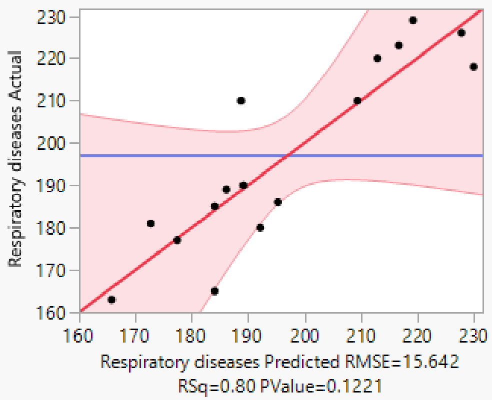
- DOE table
- 16 Legs

DISCOVERY SUMMIT

jmp

RSM

Actual by Predicted Plot



Effect Summary

Source	Logworth	PValue
Temperature(-2,30)	1.672	0.02130
Temperature*Temperature	0.917	0.12096
Precipitation*Temperature	0.878	0.13242
Precipitation(0,90)	0.828	0.14843 ^
Precipitation*Fine dust	0.797	0.15965
Fine dust*Fine dust	0.275	0.53136
Fine dust(0,50)	0.129	0.74280 ^
Precipitation*Precipitation	0.120	0.75826
Fine dust*Temperature	0.090	0.81193

Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	5	1267.9422	253.588	1.2679
Pure Error	1	200.0000	200.000	Prob > F
Total Error	6	1467.9422		0.5848
				Max RSq
				0.9728

- LOF : P value > 0.05

↓

- The lack of fit (LOF) is not significant.

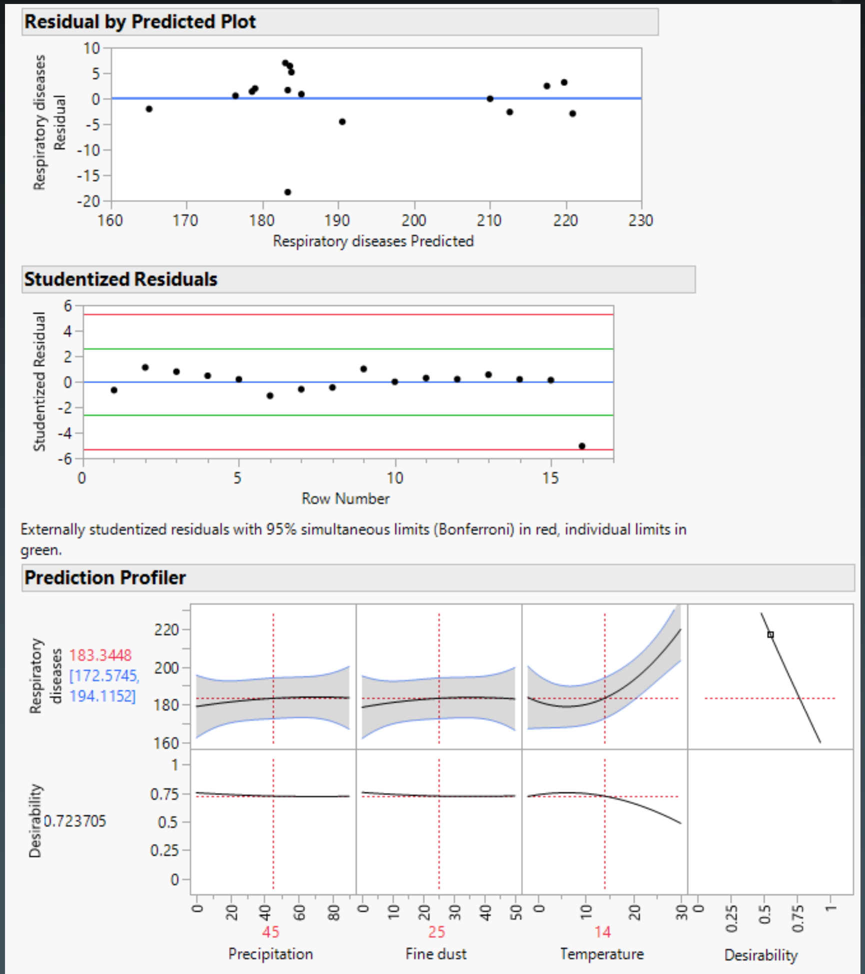
↓

- Accept the null hypothesis.

↓

- The model is valid.

RSM

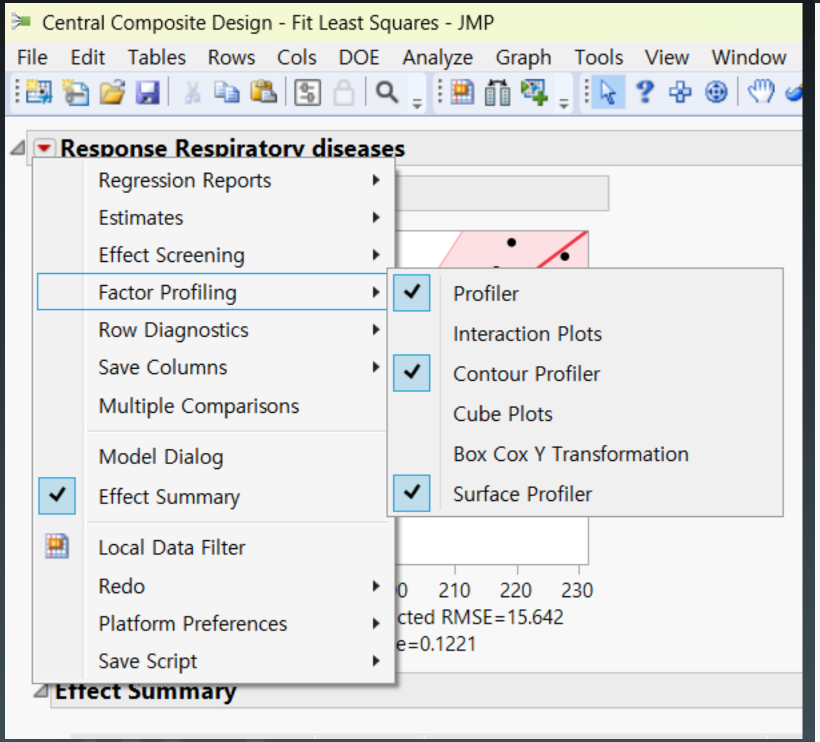


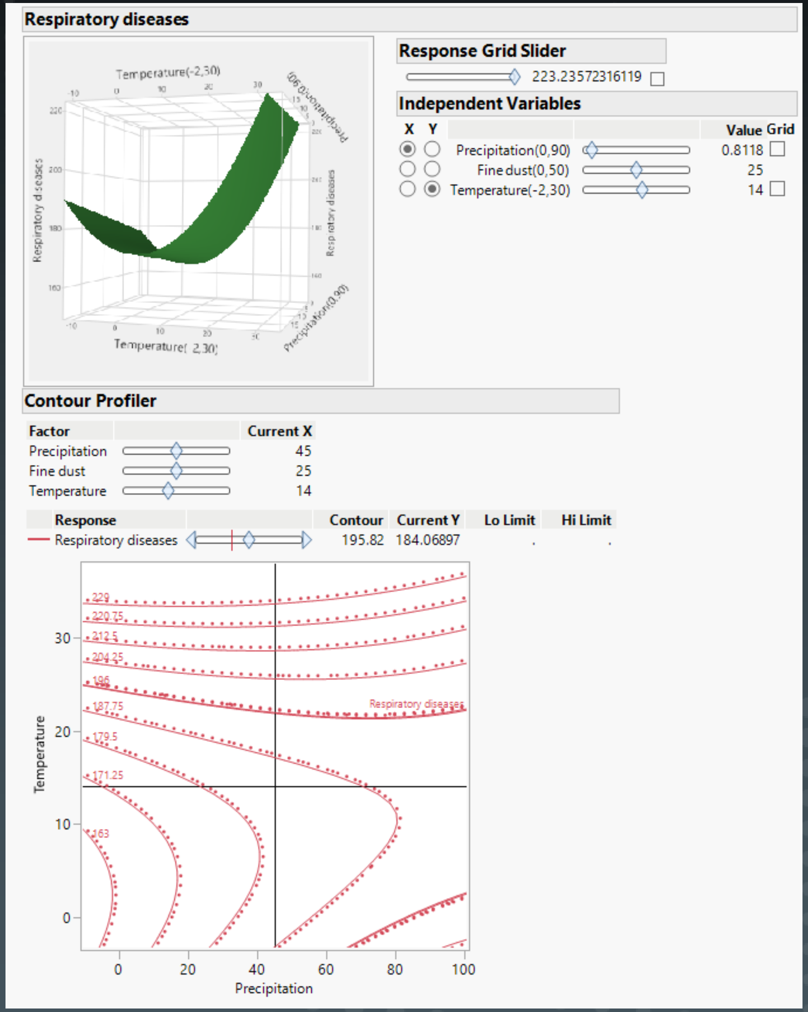
- Residuals : Random → No abnormality
- Studentized residual : < 2
- Curvature effect

DISCOVERY SUMMIT

jmp

RSM

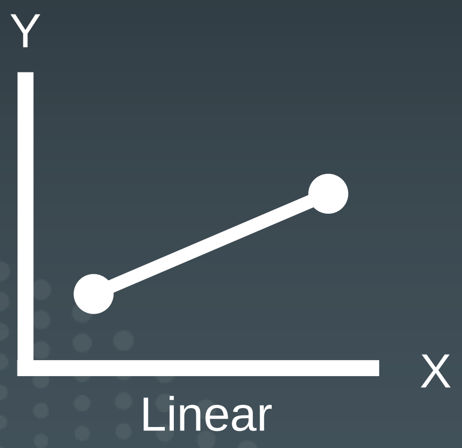


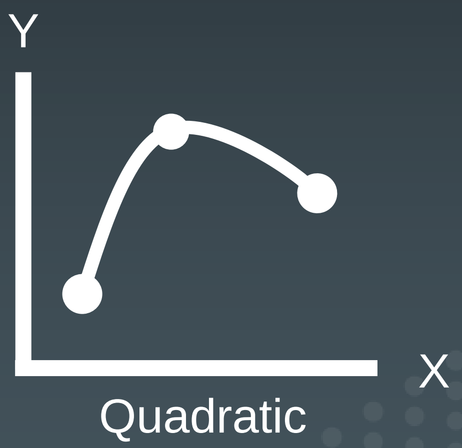


- Surface Profiler shows changes in rate of respiratory diseases
- Contour line shows rate of respiratory diseases

RSM

- If, LOF is not valid → Accept null hypothesis → Model is valid
- If LOF is valid, need to RSM analysis.





RSM vs Taguchi

	RSM	Taguchi
Purpose	Optimization of response values in the region of interest	Identify optimal conditions that are robust to the effects of noise
Input	Fewer test (using LOF), CCD/BBD	Orthogonal array, Loss function
Output	Response value behavior by input	Robust to noise

DISCOVERY SUMMIT

jmp

Thank you