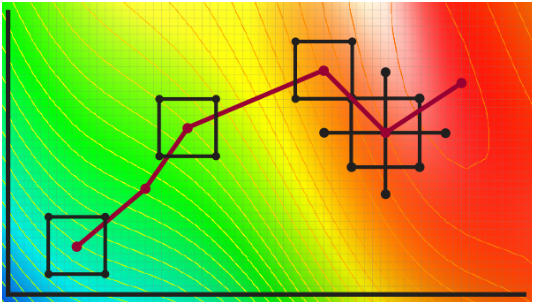
Experimentation for Improvement



http://learnche.org/

Design and Analysis of Experiments

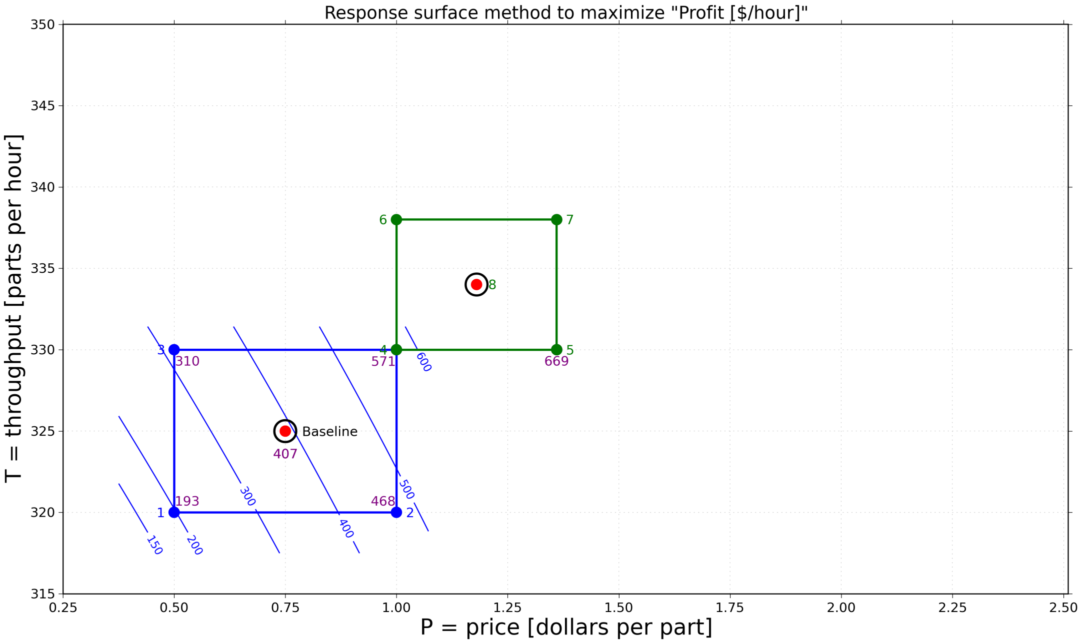
Copyright, sharing, and attribution notice

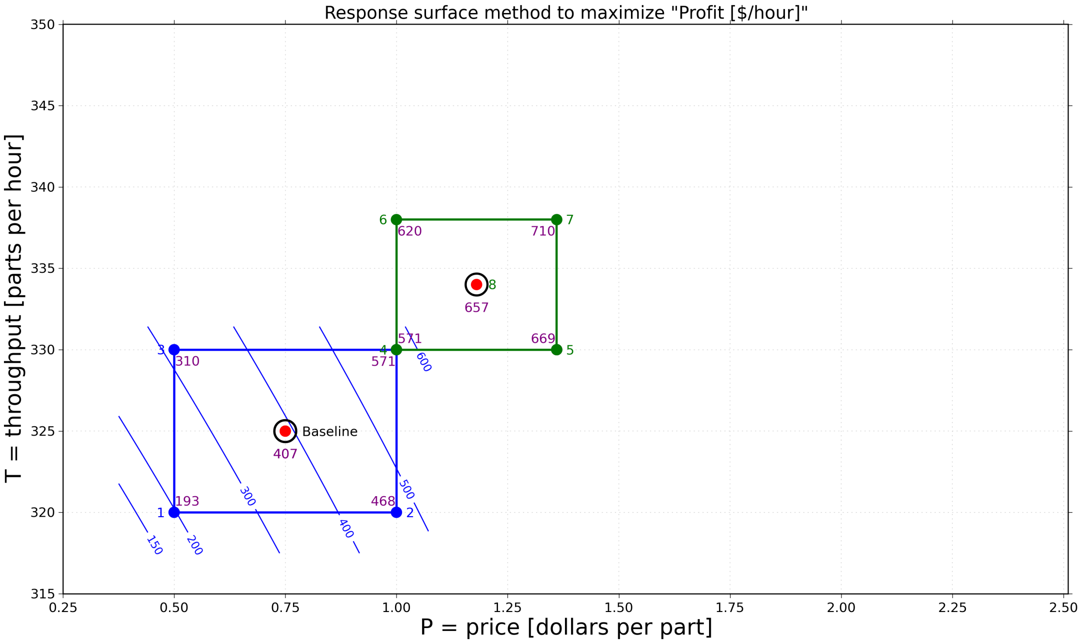
Outline of topics in this video

1. We keep climbing up the mountain ...
2. but, what happens if we encounter a constraint?
3. and, what happens when mistakes happen?

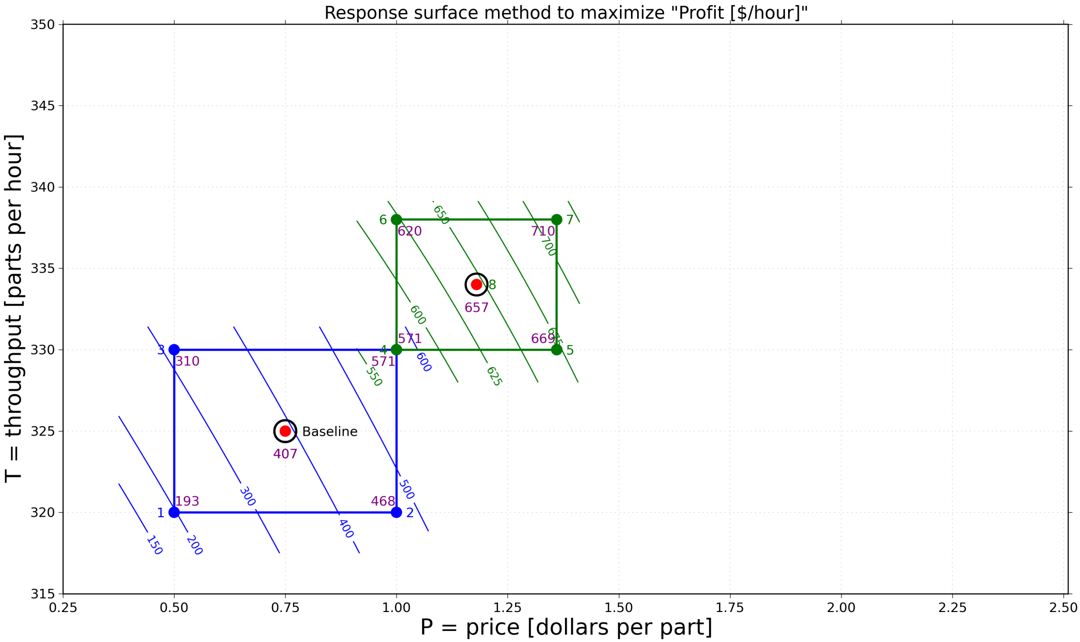
$$\text{coded value} = \frac{(\text{real value}) - (\text{center point})}{\frac{1}{2}\text{range}}$$

$$x_{\mathsf{T}} = \frac{337 - 339}{\frac{1}{2}(6)} = -\frac{2}{3}$$


Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
669
1
2
3
4
5
6
7
8
150
200
300
400
500
600



Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
350
345
340
335
330
325
320
315
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
1
2
3
4
5
6
7
8
193
468
310
571
571
669
620
710
657
407
Baseline
150
200
300
400
500
600


Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
350
345
340
335
330
325
320
315
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
1
2
3
4
5
6
7
8
193
468
310
571
571
669
620
710
657
407
Baseline
150
200
300
400
500
600
550
600
625
650
700

Taking the next step for experiment 9: fill in the blanks

1. Pick change in coded units in one factor.
2. Find the ratios for the other factor(s).
3. Calculate step size in coded units.
4. Convert these to real-world *changes*.
5. Get the real-world location of the next experiment.
6. Convert these back to coded-units.
7. Predict the next experiment's outcome.
8. Now run the next experiment, and record the values

Price	Throughput
$\Delta x_{\mathsf{P}} = 1.5$ (this was chosen)	
$\Delta x_{\mathsf{P}} = 1.5$	$\Delta x_{\mathsf{T}} =$ ____________
$\Delta \mathsf{P} =$ ____________	$\Delta \mathsf{T} =$ ____________ parts per hour
$\mathsf{P}^{(9)} =$ ____________	$\mathsf{T}^{(9)} =$ ____________ parts per hour
$x_{\mathsf{P}}^{(9)} =$ ____________	$x_{\mathsf{T}}^{(9)} =$ ____________

$\hat{y} =$ ____________ profit per hour

$y^{(9)} =$ ____________ profit per hour

Taking the next step for experiment 9

	Price	Throughput
1. Pick change in coded units in one factor.	$\Delta x_{\mathsf{P}} = 1.5$ (this was chosen)	
2. Find the ratios for the other factor(s).		$\Delta x_{\mathsf{T}} = \dfrac{b_{\mathsf{T}}}{b_{\mathsf{P}}} \cdot \Delta x_{\mathsf{P}}$ $\Delta x_{\mathsf{T}} = \dfrac{22.5}{47} \cdot 1.5$
3. Calculate step size in coded units.	$\Delta x_{\mathsf{P}} = 1.5$	$\Delta x_{\mathsf{T}} = 0.718$
4. Convert these to real-world *changes*.	$\Delta \mathsf{P} = \Delta x_{\mathsf{P}} \cdot \frac{1}{2}(0.36)$ $\Delta \mathsf{P} = \$0.27$	$\Delta \mathsf{T} = \Delta x_{\mathsf{T}} \cdot \frac{1}{2}(8)$ $\Delta \mathsf{T} = 2.87 \approx 3$ parts per hour
5. Finally, take a step from the baseline! Get the real-world location of the next experiment.	$\mathsf{P}^{(9)} = \mathsf{P}^{(8)} + \Delta \mathsf{P}$ $\mathsf{P}^{(9)} = \$1.18 + 0.27$ $\mathsf{P}^{(9)} = \$1.45$	$\mathsf{T}^{(9)} = \mathsf{T}^{(8)} + \Delta \mathsf{P}$ $\mathsf{T}^{(9)} = 334 + 3$ $\mathsf{T}^{(9)} = 337$ parts per hour

Taking the next step for experiment 9

5. Get the real-world location of the next experiment.
6. Convert these back to coded-units.
7. Predict the next experiment's outcome.
8. Now run the next experiment, and record the values

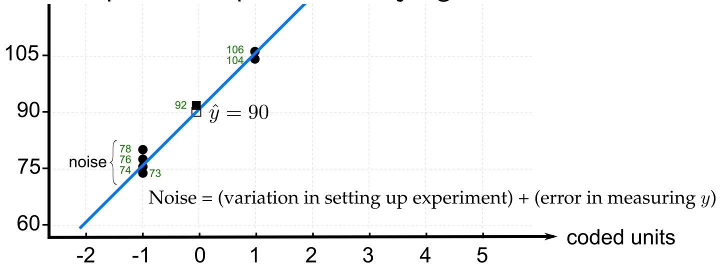
Price	**Throughput**
$\mathsf{P}^{(9)} = \$1.45$	$\mathsf{T}^{(9)} = 337$ parts per hour
$x_{\mathsf{P}}^{(9)} = 1.5$	$x_{\mathsf{T}}^{(9)} = \dfrac{337-334}{\frac{1}{2}(8)} = 0.75$ (not 0.718)

$$
\begin{aligned}
\hat{y} &= 645 + 47x_{\mathsf{P}} + 22.5x_{\mathsf{T}} - 2x_{\mathsf{P}}x_{\mathsf{T}}\\
\hat{y}^{(5)} &= 645 + 47(1.5) + 22.5(0.75) - 2(1.5)(0.75)\\
\hat{y}^{(5)} &\approx 645 + 71 + 17 - 2\\
\hat{y}^{(5)} &\approx 731 \text{ profit per hour}
\end{aligned}
$$

$$y^{(5)} = \$717 \text{ profit per hour}$$

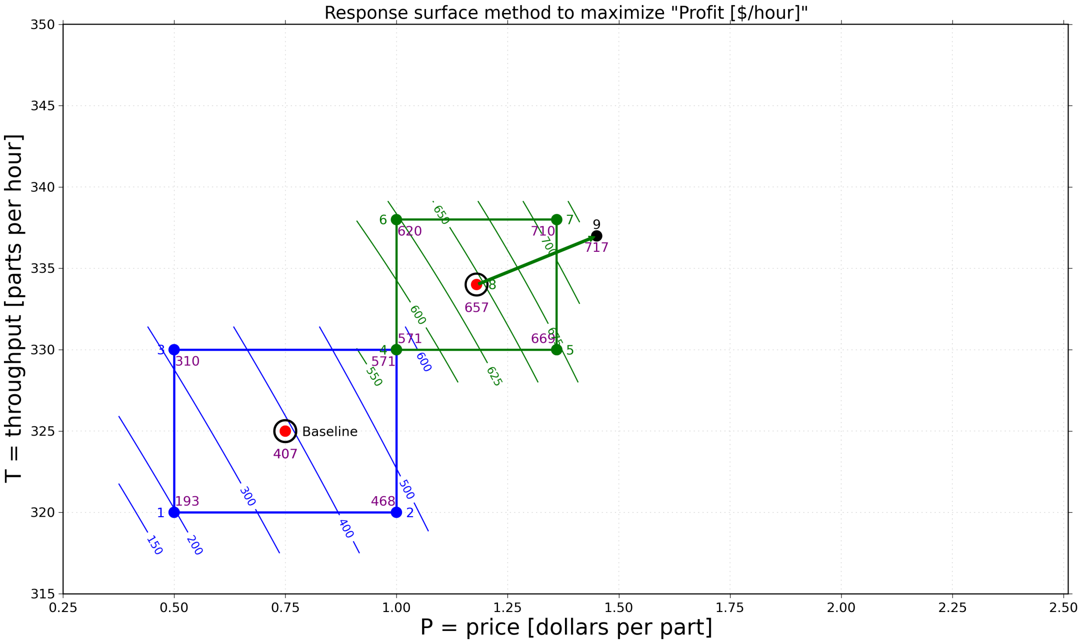
Judging the prediction error

1. We should have replicated experiments to judge the noise level

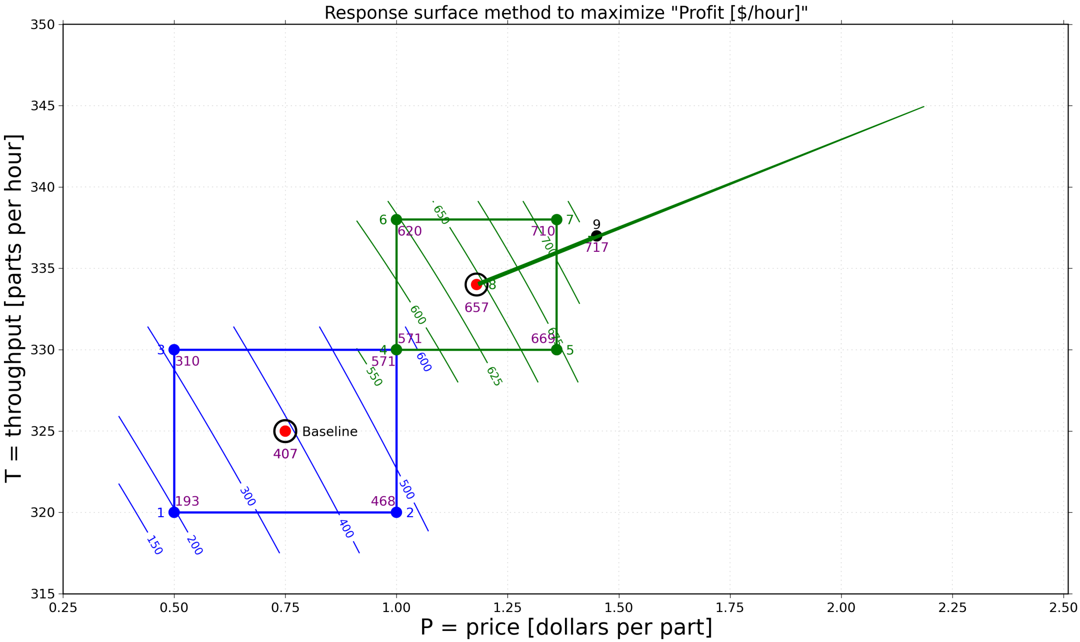


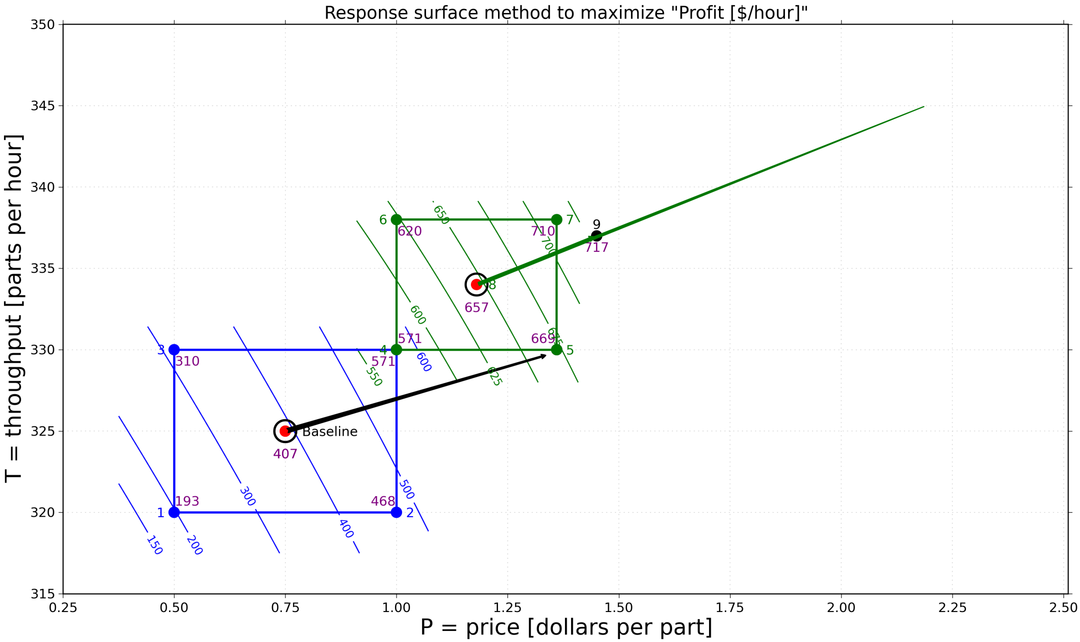
We will see more on replicated experiments in the next video

2. Contrast prediction error with the model's coefficients:
   - $\hat{y}^{(9)} = 645 + 47x_{\text{P}} + 22.5x_{\text{T}} - 2x_{\text{P}}x_{\text{T}} = \$731$
   - $y^{(9)}_{\text{actual}} = \$717$
   - error $= \hat{y} - y \approx \$13$ (ignoring the decimals)

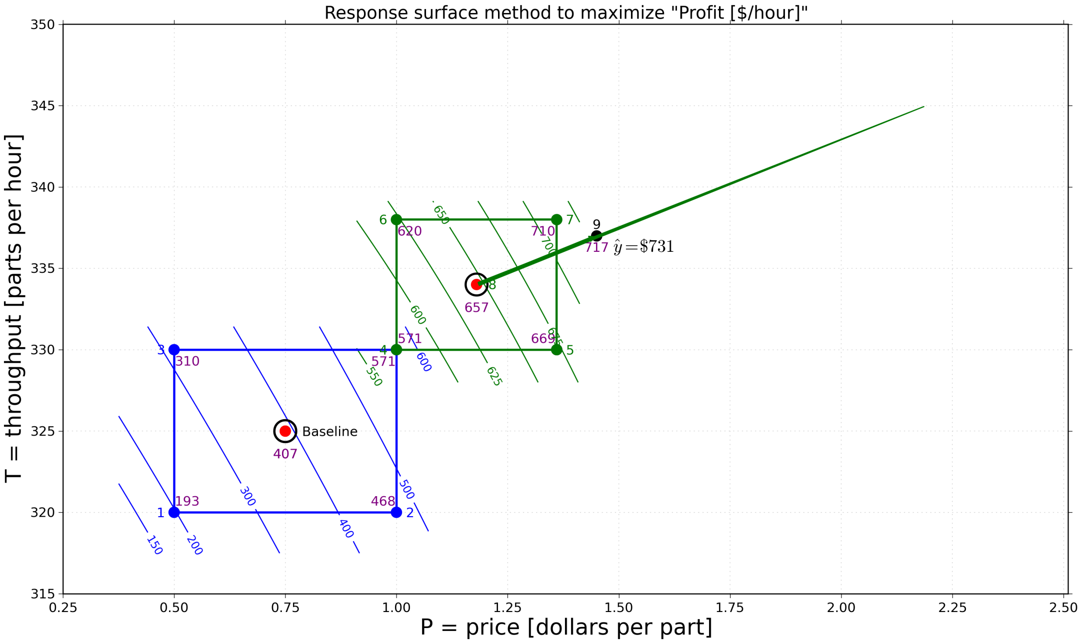


Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
620
710
669
657
717
1
2
3
4
5
6
7
8
9
150
200
300
400
500
600
550
625
650
675
700





Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
571
620
710
669
657
717
1
2
3
4
5
6
7
8
9
150
200
300
400
500
600
550
600
625
650
675
700



Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
620
710
669
657
717
$\hat{y} = \$731$

Taking the next step for experiment 10: fill in the blanks

1. Pick change in coded units in one factor.
2. Find the ratios for the other factor(s).
3. Calculate step size in coded units.
4. Convert these to real-world *changes*.
5. Get the real-world location of the next experiment.
6. Convert these back to coded-units.
7. Predict the next experiment's outcome.
8. Now run the next experiment, and record the values

Price	**Throughput**
$\Delta x_{\text{P}} = 2.5$ (this was chosen)	
$\Delta x_{\text{P}} = 2.5$	$\Delta x_{\text{T}} =$ __________
$\Delta \text{P} =$ __________	$\Delta \text{T} =$ __________ parts per hour
$\text{P}^{(10)} =$ __________	$\text{T}^{(10)} =$ __________ parts per hour
$x_{\text{P}}^{(10)} =$ __________	$x_{\text{T}}^{(10)} =$ __________

$\hat{y}^{(10)} =$ __________ profit per hour

$y^{(10)} =$ __________ profit per hour

Taking the next step for experiment 10: solution

	Price	Throughput
1. Pick change in coded units in one factor.	$\Delta x_{\mathsf{P}} = 2.5$ (this was chosen)	
2. Find the ratios for the other factor(s).		
3. Calculate step size in coded units.	$\Delta x_{\mathsf{P}} = 2.5$	$\Delta x_{\mathsf{T}} = 1.2 = \dfrac{b_{\mathsf{T}}}{b_{\mathsf{P}}} \times \Delta x_{\mathsf{P}} = \dfrac{22.5}{47} \times 2.5$
4. Convert these to real-world *changes*.	$\Delta \mathsf{P} = \$0.45$	$\Delta \mathsf{T} = 4.8 \approx 5$ parts per hour
5. Get the real-world location of the next experiment.	$\mathsf{P}^{(10)} = \$1.63$	$\mathsf{T}^{(10)} = 334 + 5 = 339$ parts per hour
6. Convert these back to coded-units.	$x_{\mathsf{P}}^{(10)} = 2.5$	$x_{\mathsf{T}}^{(10)} = \dfrac{339 - 334}{\frac{1}{2} \cdot (8)} = 1.25$

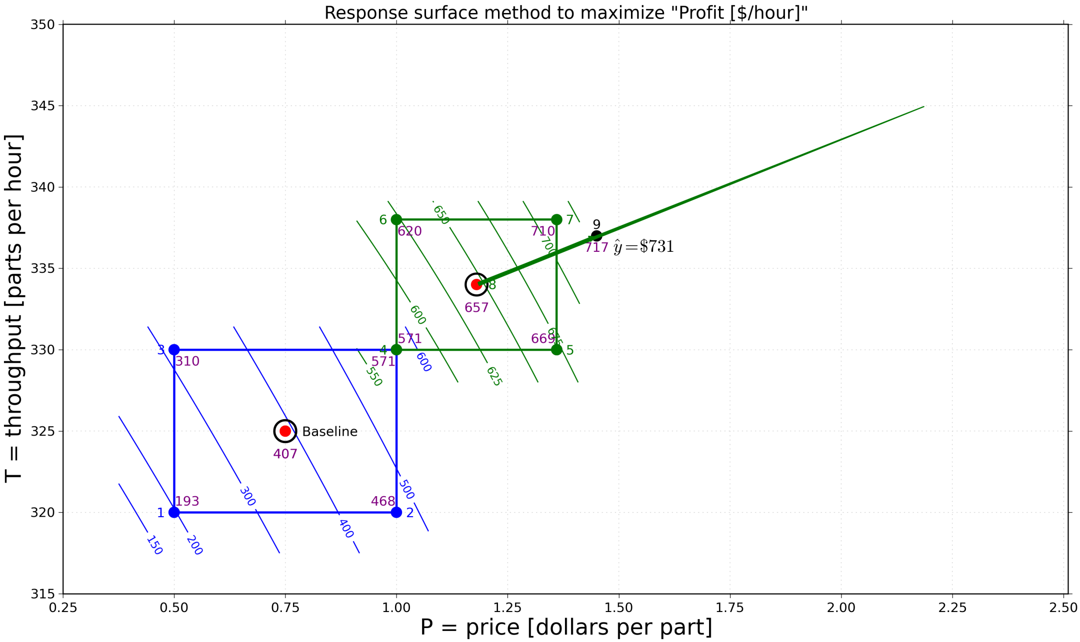
7. Predict the next experiment's outcome.

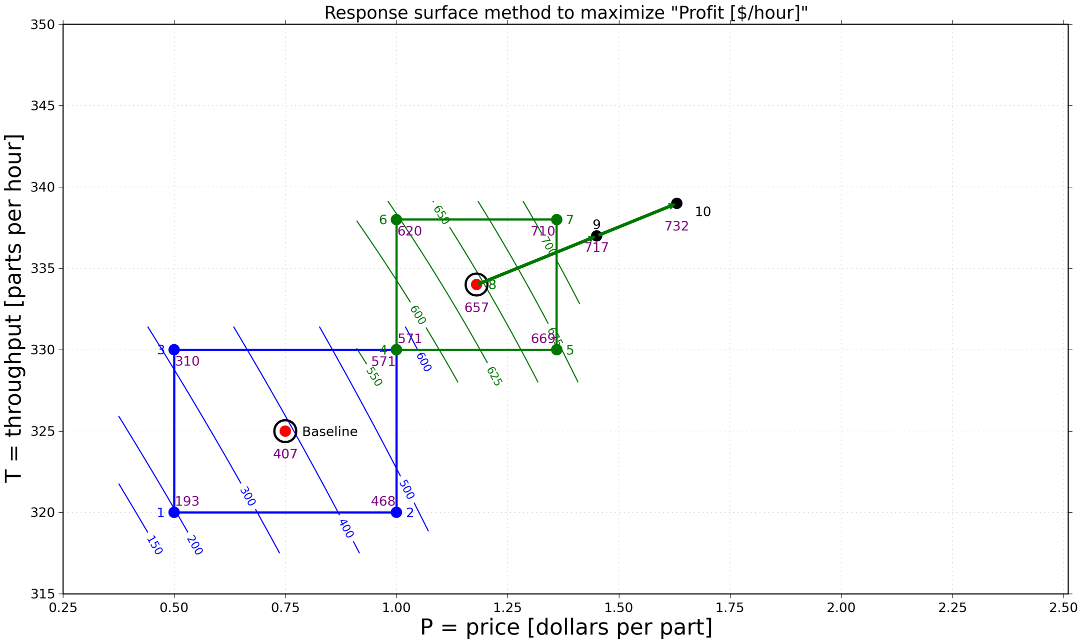
$$\hat{y} = 645 + 47x_{\mathsf{P}} + 22.5x_{\mathsf{T}} - 2x_{\mathsf{P}}x_{\mathsf{T}}$$

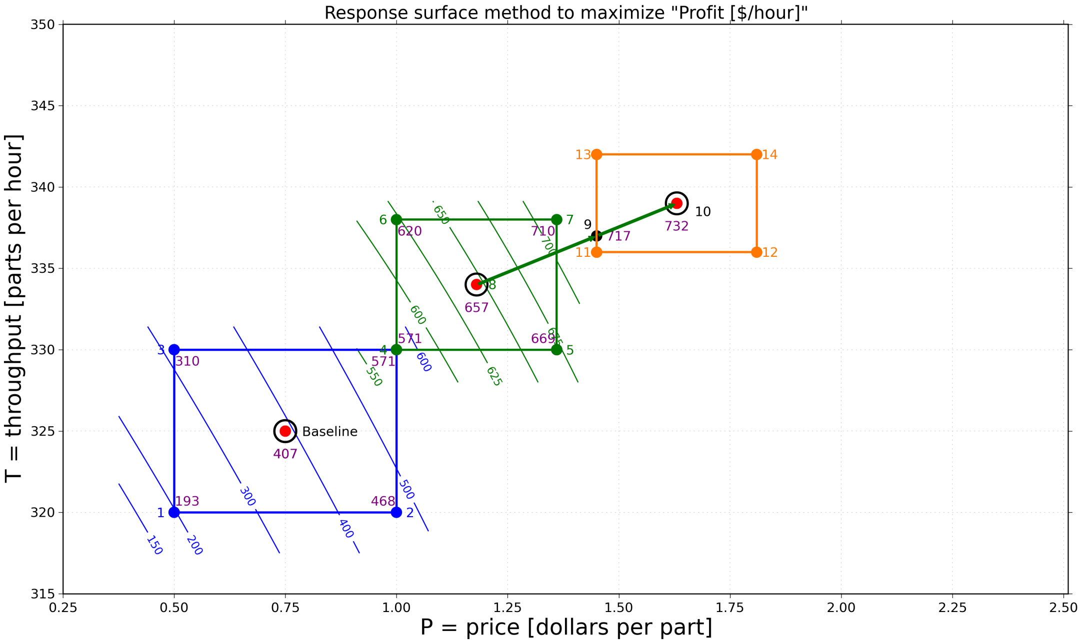
$$\hat{y}^{(10)} \approx \$785 \text{ profit per hour}$$

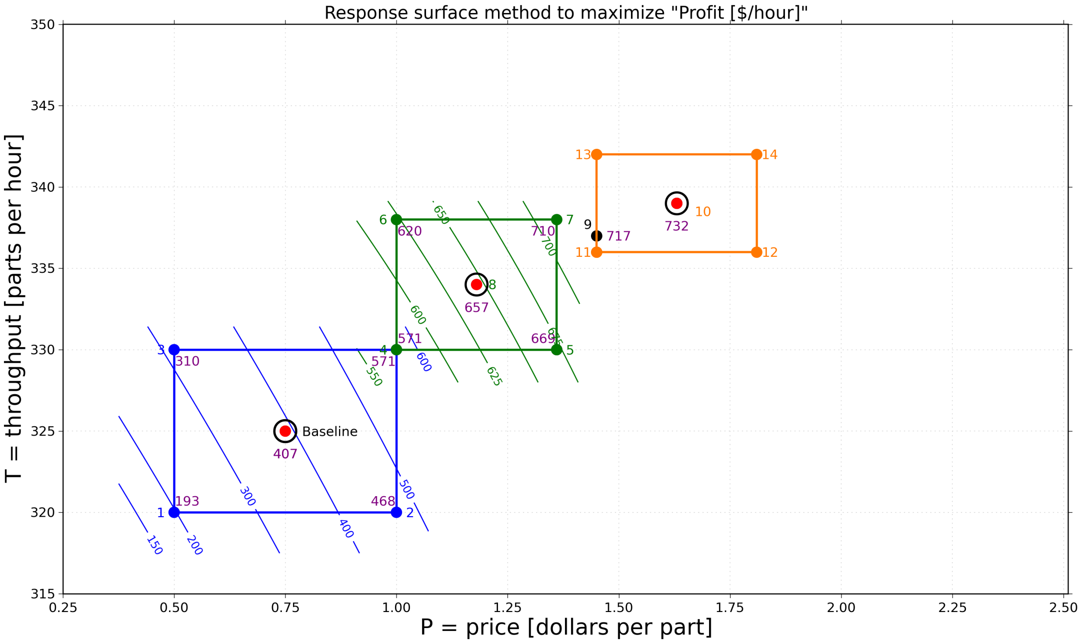
8. Now run the next experiment, and record the values

$$y^{(10)} = \$732 \text{ profit per hour}$$

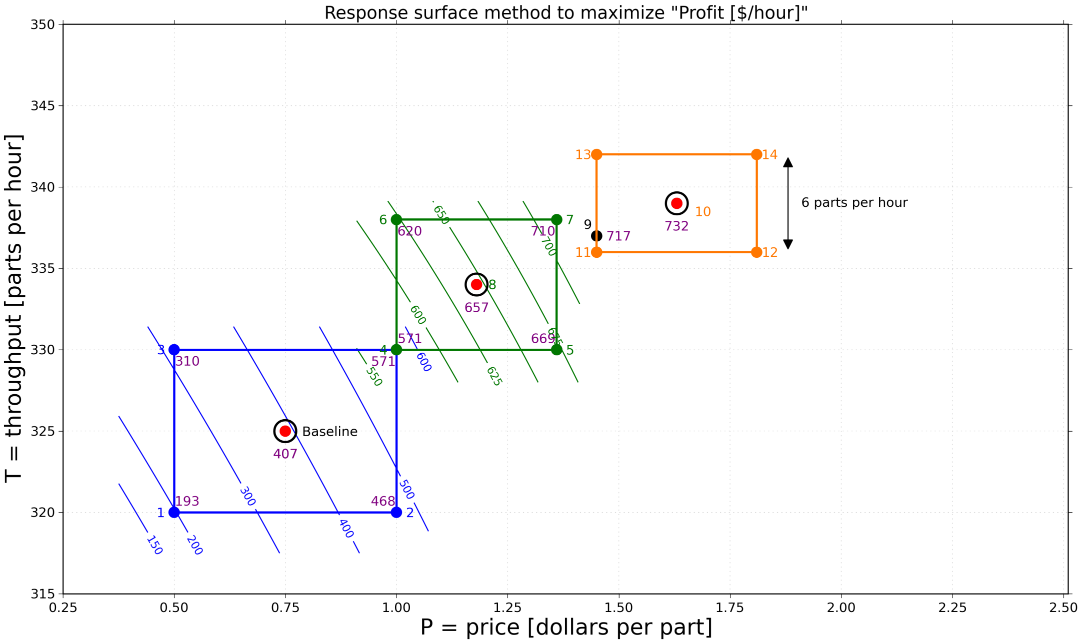

Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
620
710
669
657
717
$\hat{y} = \$731$
150
200
300
400
500
600
550
600
625
650
675
700
1
2
3
4
5
6
7
8
9
315
320
325
330
335
340
345
350
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50


Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
620
710
669
657
717
732
1
2
3
4
5
6
7
8
9
10
150
200
300
400
500
600
550
625
650
700
675
350
345
340
335
330
325
320
315
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50

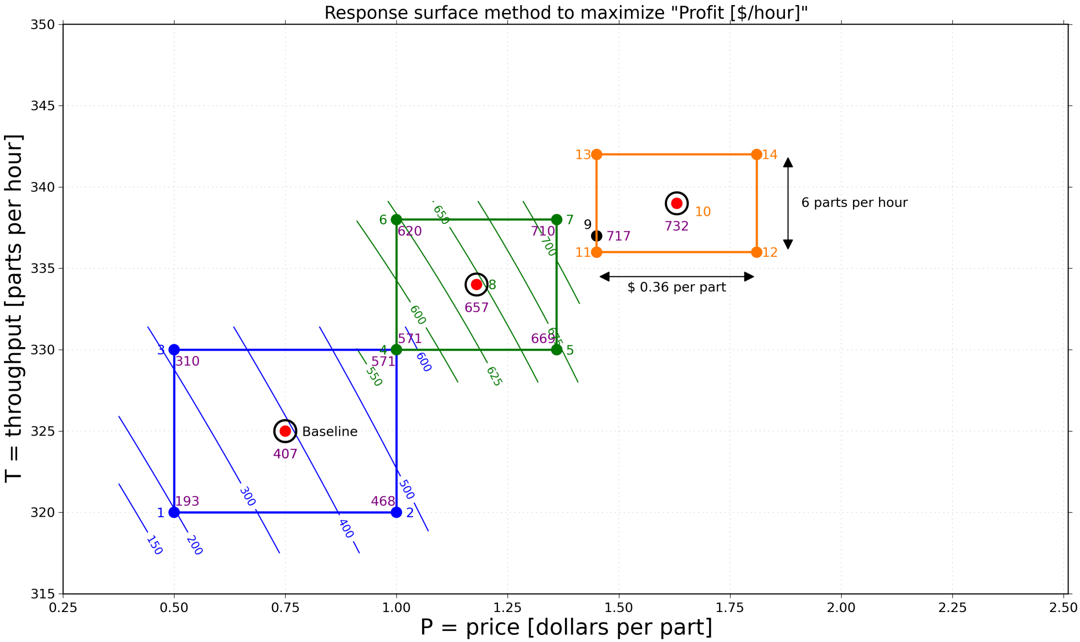

Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
620
710
669
657
717
732
1
2
3
4
5
6
7
8
9
10
11
12
13
14
150
200
300
400
500
550
600
625
650
700
675
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
315
320
325
330
335
340
345
350



Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
620
710
669
657
717
732
1
2
3
4
5
6
7
8
9
10
11
12
13
14
150
200
300
400
500
550
600
625
650
675
700



Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
620
710
669
657
717
732
6 parts per hour
1
2
3
4
5
6
7
8
9
10
11
12
13
14
150
200
300
400
500
600
550
625
650
675
700

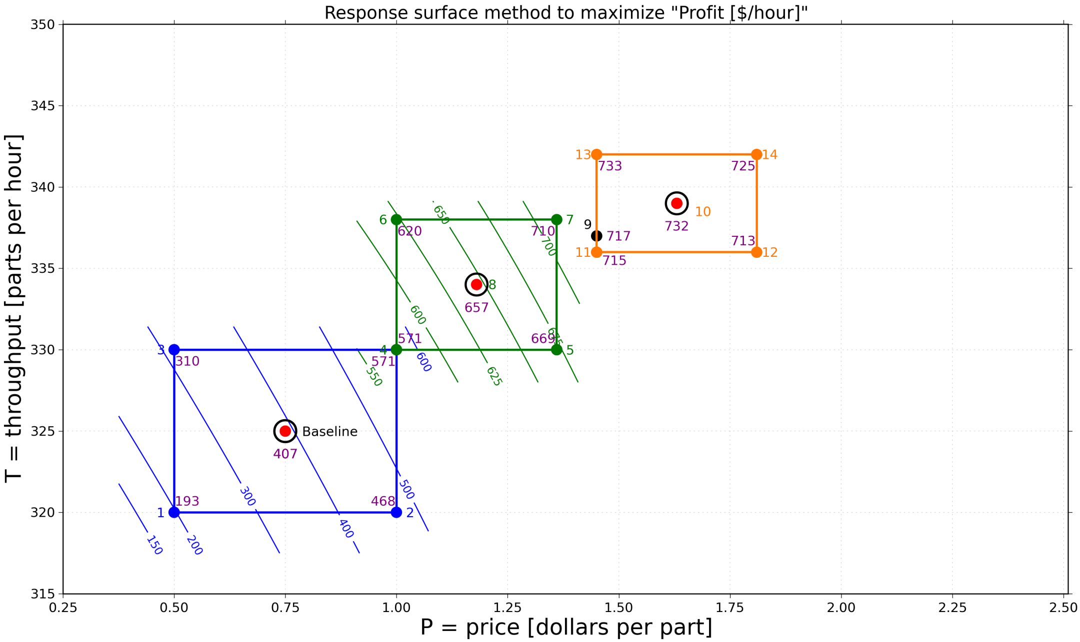


Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
350
345
340
335
330
325
320
315
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Baseline
407
193
468
310
571
620
657
710
669
717
732
1
2
3
4
5
6
7
8
9
10
11
12
13
14
150
200
300
400
500
550
600
625
650
675
700
6 parts per hour
$ 0.36 per part

Step size is a function of the factorial range

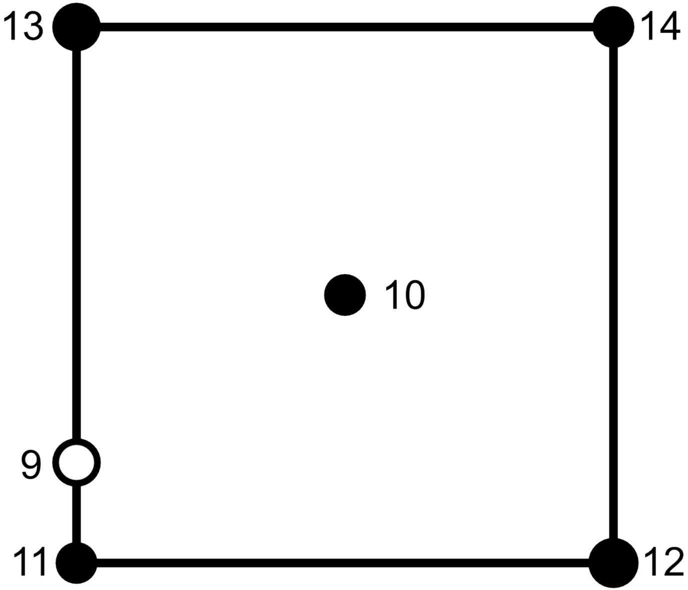
Let's work backwards:

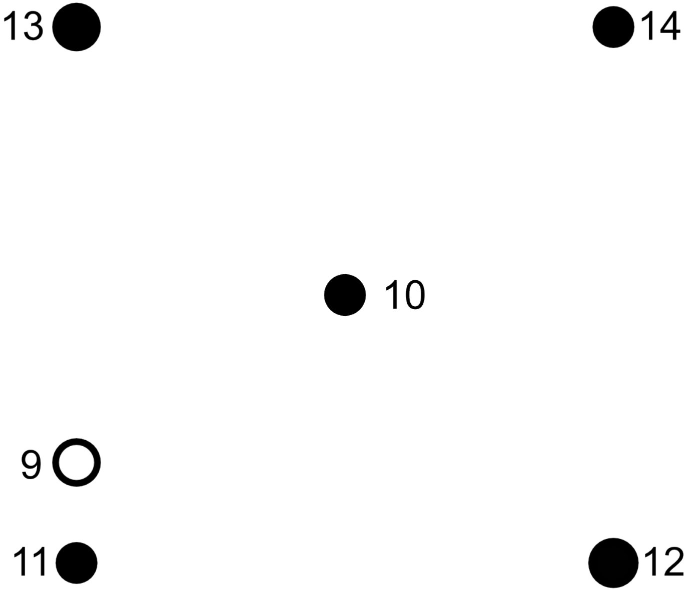
- You choose the step size, e.g. $\Delta x_{\mathsf{P}} = 2.5$
- Real-world step size $= \Delta \mathsf{P} = \Delta x_{\mathsf{P}} \cdot \frac{1}{2}\left(\text{range}_{\mathsf{P}}\right)$
- where the "range$_{\mathsf{P}}$" is picked by the experimenter

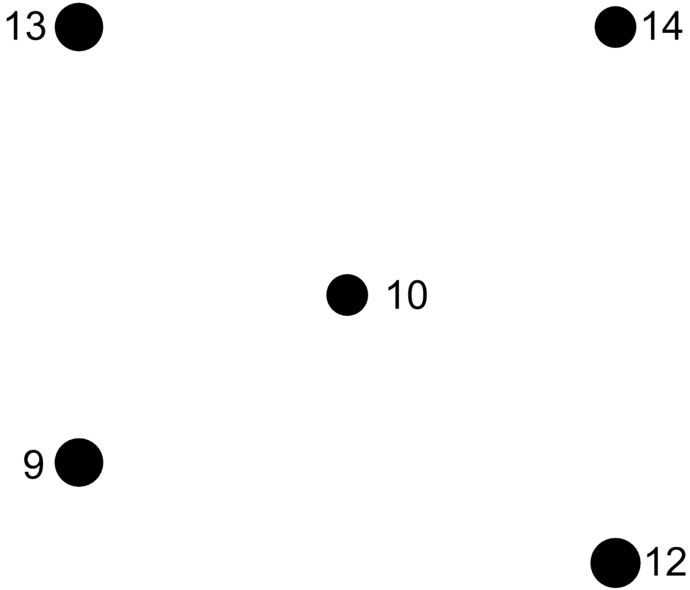

Response surface method to maximize "Profit [$/hour]"
T = throughput [parts per hour]
P = price [dollars per part]
Baseline
407
193
468
310
571
620
710
669
657
733
725
732
717
715
713
1
2
3
4
5
6
7
8
9
10
11
12
13
14
150
200
300
400
500
550
600
625
650
675
700

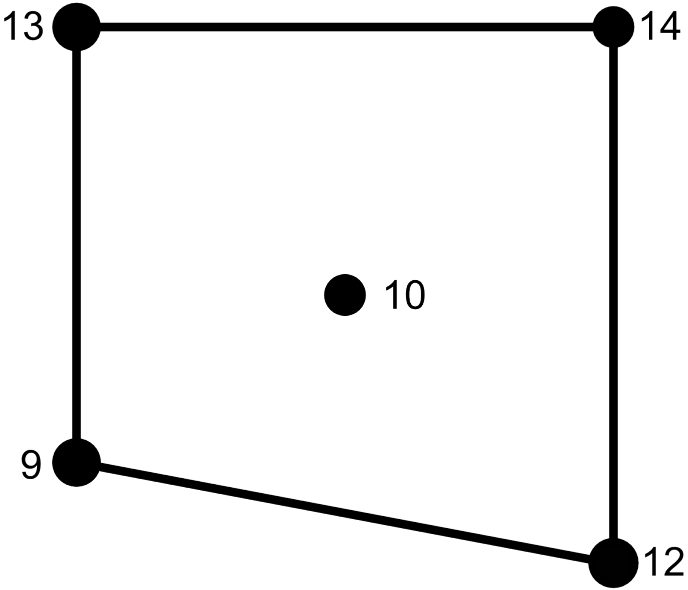
List of all experiments

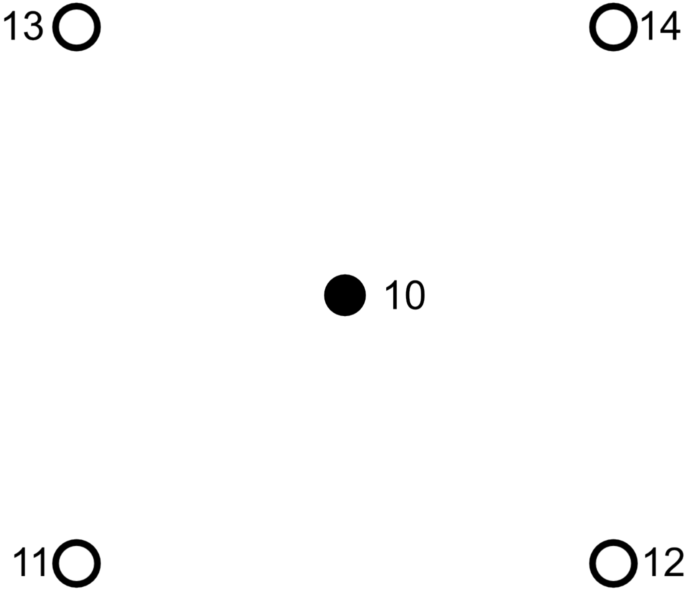
Experiment	P	T	$x_{\mathbf{P}}$	$x_{\mathbf{T}}$	Prediction $= \hat{y}$	Actual $= y$	Model
9	1.45	337	1.5	0.75		\$ 717	2
10	1.63	339	0	0		\$ 732	3
11	1.45	336	-1	-1		715	3
12	1.81	336	$+1$	-1		713	3
13	1.45	342	-1	$+1$		733	3
14	1.81	342	$+1$	$+1$		725	3

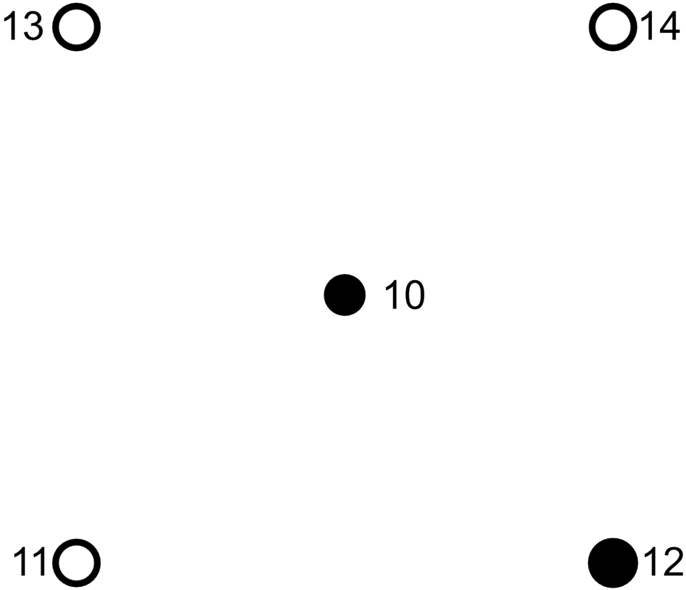

13
14
10
9
11
12

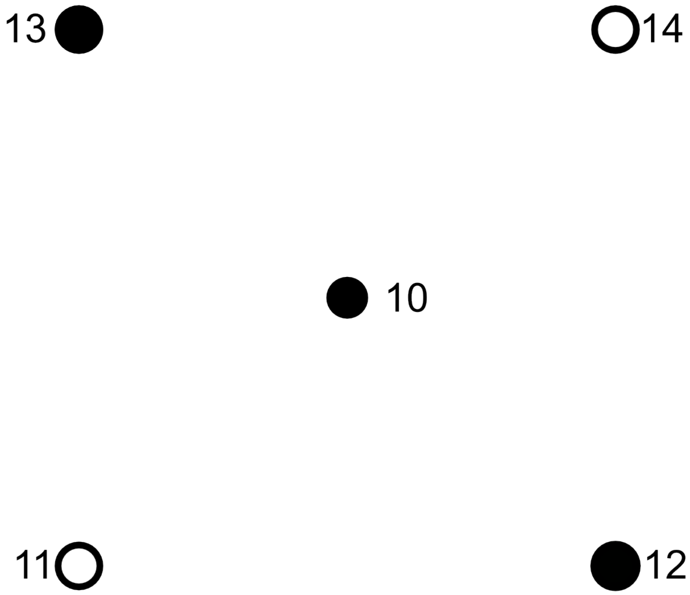

13
14
10
9
11
12

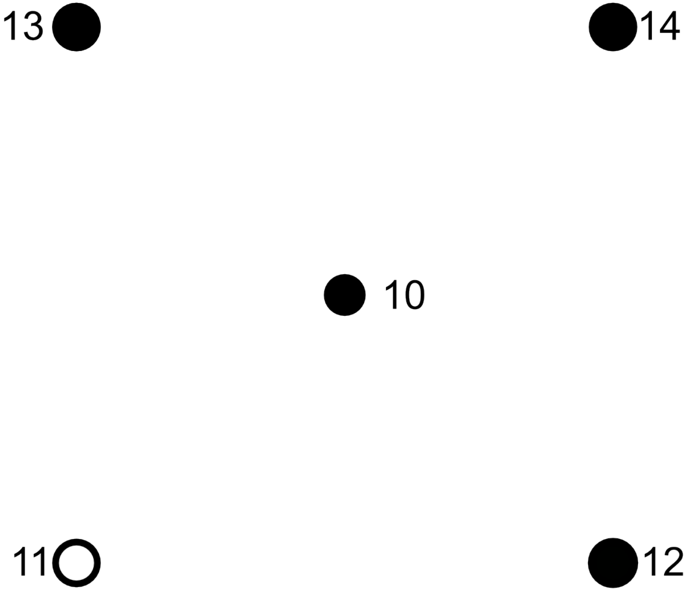

13
14
10
9
12

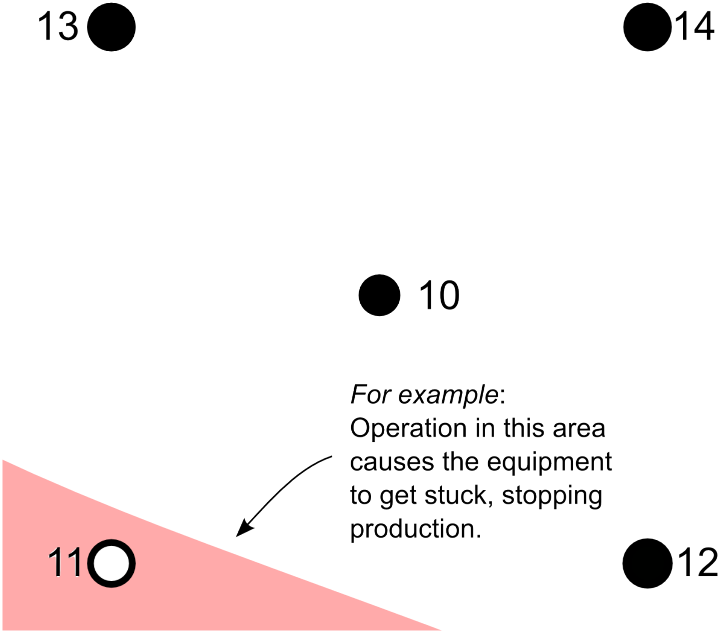

13
14
10
9
12

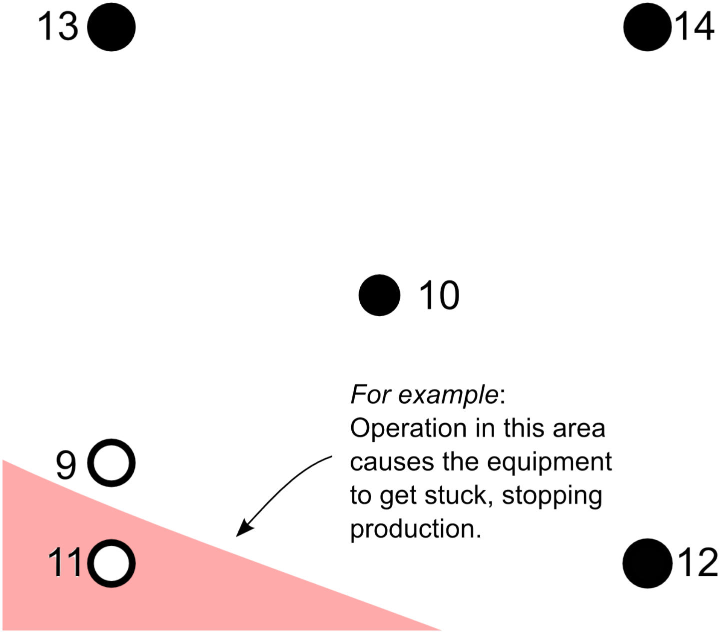

13
14
10
11
12

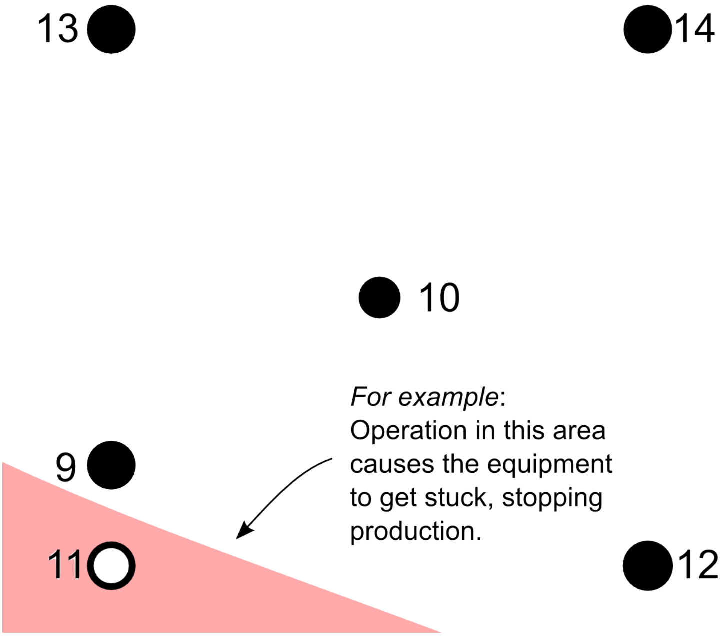

13
14
10
11
12

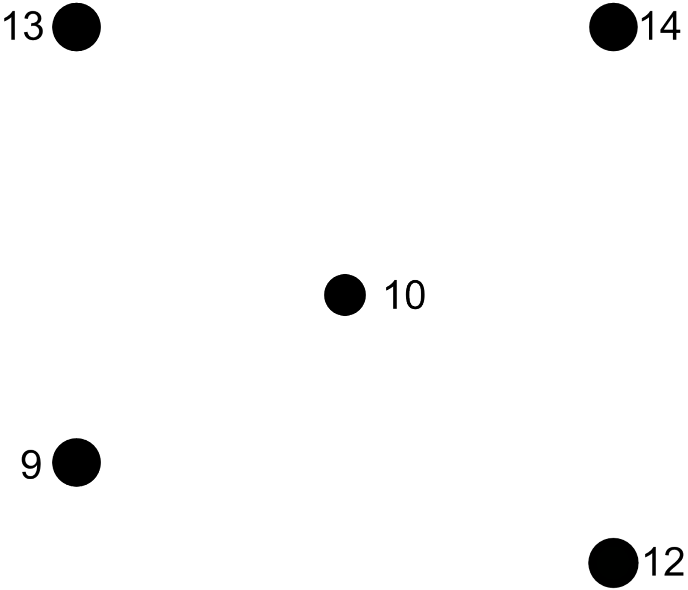

13
14
10
11
12

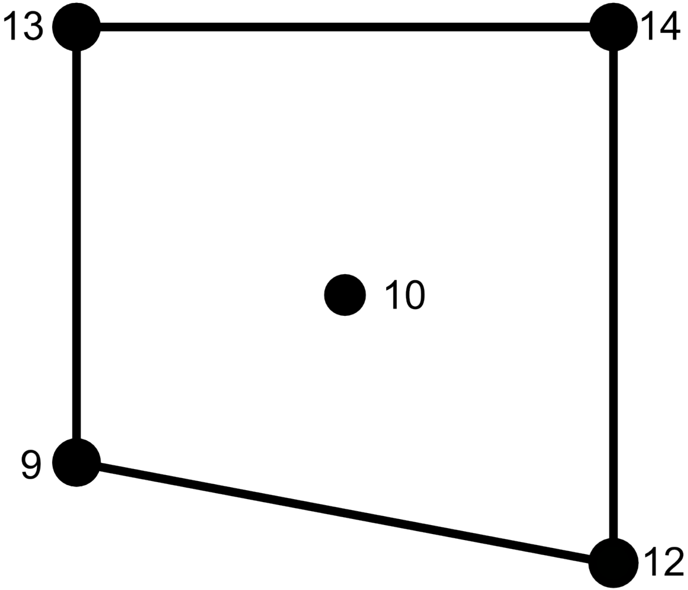

13
14
10
11
12

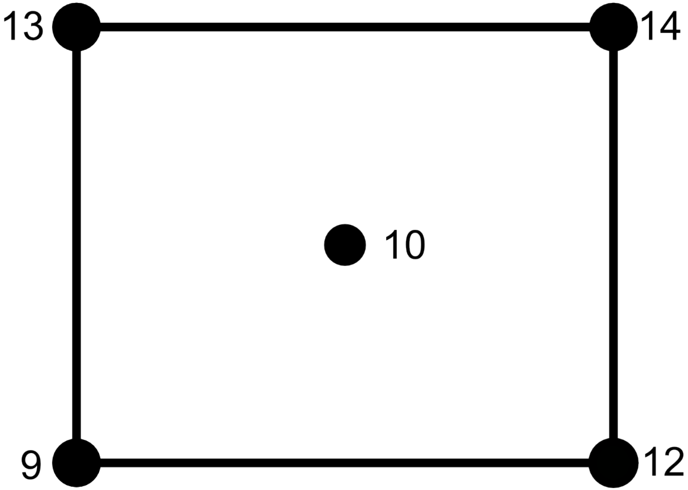

13
14
10
For example:
Operation in this area causes the equipment to get stuck, stopping production.
11
12

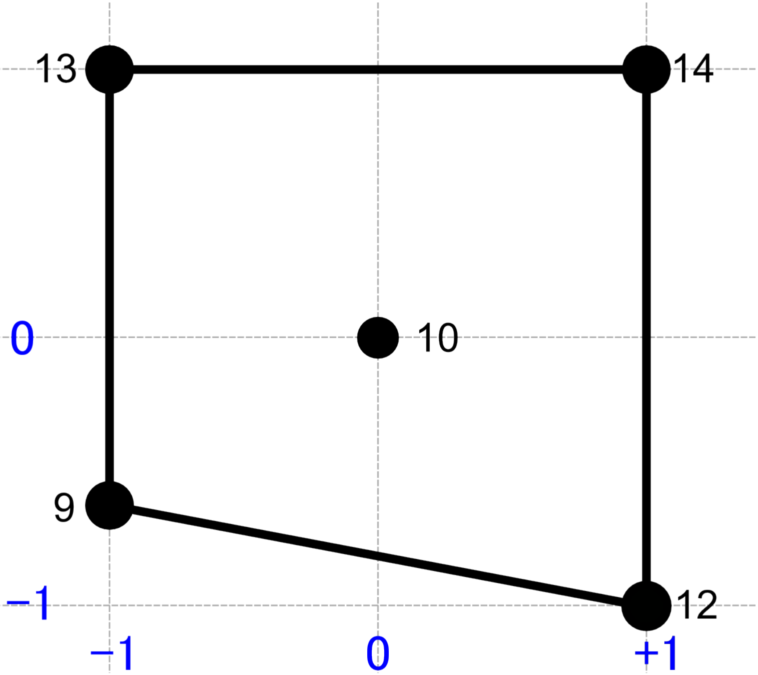

13
14
10
For example:
Operation in this area causes the equipment to get stuck, stopping production.
9
11
12


13
14
10
For example:
Operation in this area
causes the equipment
to get stuck, stopping
production.
9
11
12


13
14
10
9
12

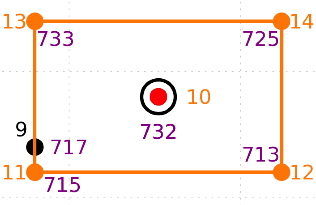

13
14
10
9
12


13
14
10
9
12


13
14
10
9
12
0
−1
−1
0
+1

Botched design

```r
P <- c(  0,   -1,  +1,  -1,  +1)
T <- c(  0, -2/3,  -1,  +1,  +1)
y <- c(732,  717, 713, 733, 725)
mod.base.4 <- lm(y ~ P*T)
summary(mod.base.4)
contourPlot(mod.base.4, P, T,
            title="Botched design")
```



Regular design

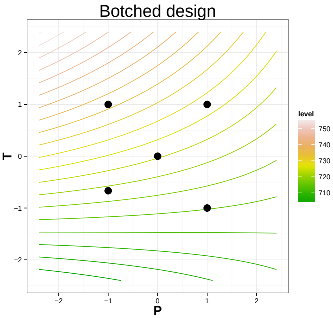
```r
P <- c(  0,  -1,  +1,  -1,  +1)
T <- c(  0,  -1,  -1,  +1,  +1)
y <- c(732, 715, 713, 733, 725)
mod.base.3 <- lm(y ~ P*T)
summary(mod.base.3)
contourPlot(mod.base.3, P, T,
            title="Regular design")
```

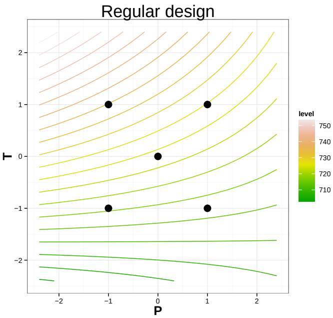
Data points used: 10, **9**, 12, 13, 14

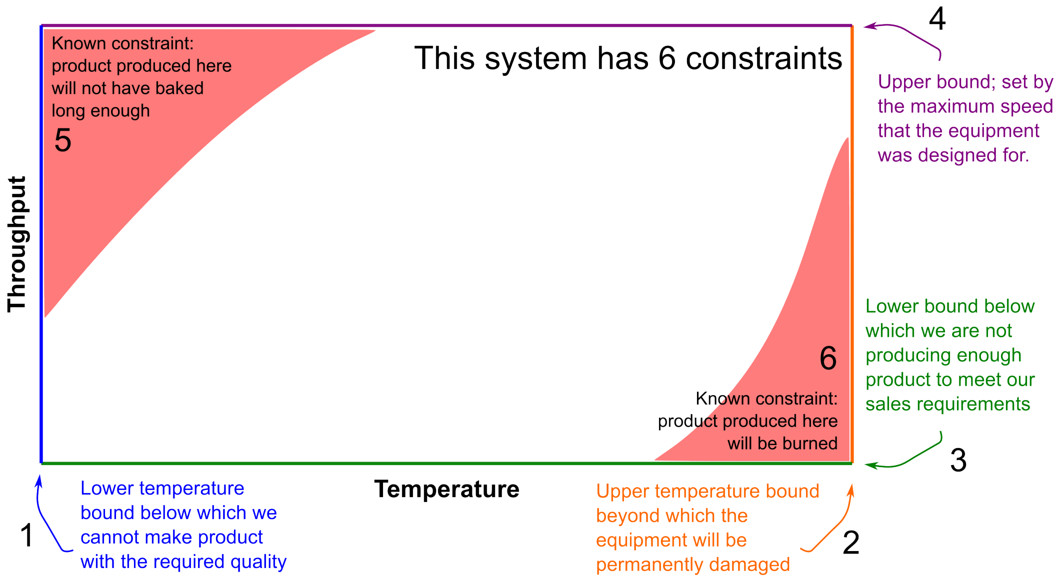
Data points used: 10, **11**, 12, 13, 14

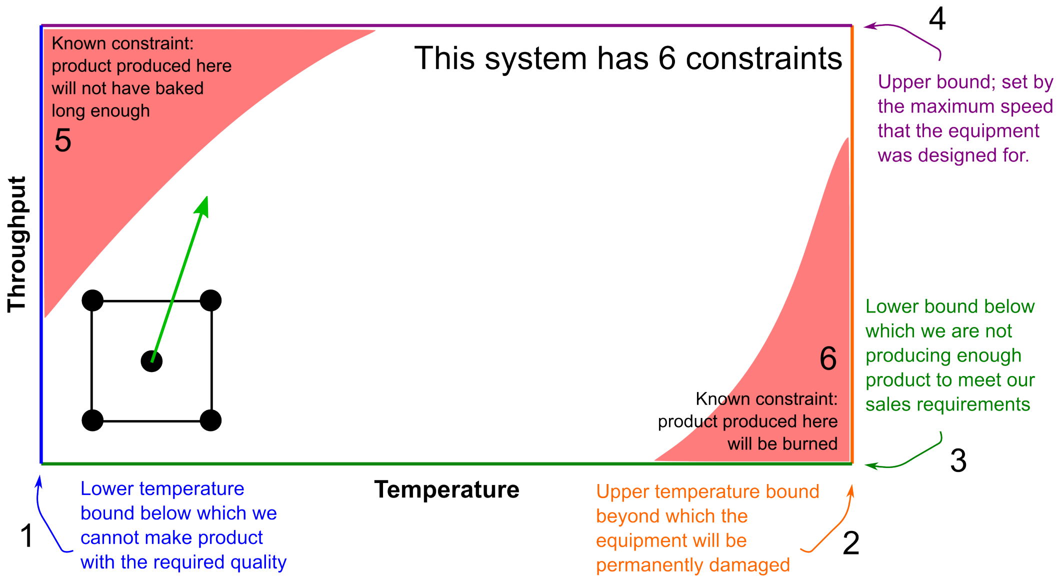
$$\hat{y} = 723.4 - 2.243 x_{\mathsf{P}} + 7.542 x_{\mathsf{T}} - 1.542 x_{\mathsf{P}} x_{\mathsf{T}}$$

$$\hat{y} = 723.6 - 2.500 x_{\mathsf{P}} + 7.500 x_{\mathsf{T}} - 1.500 x_{\mathsf{P}} x_{\mathsf{T}}$$

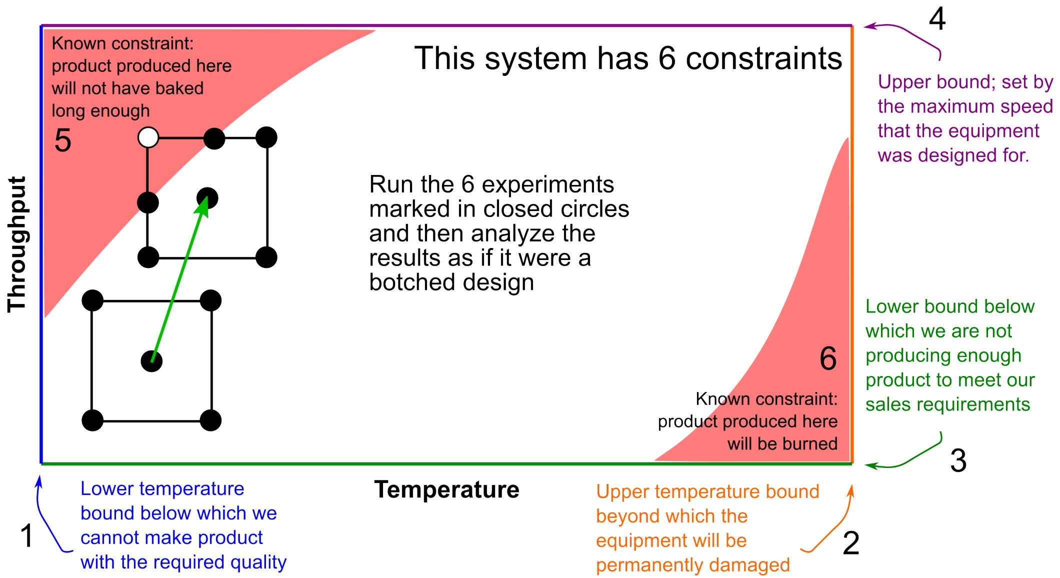




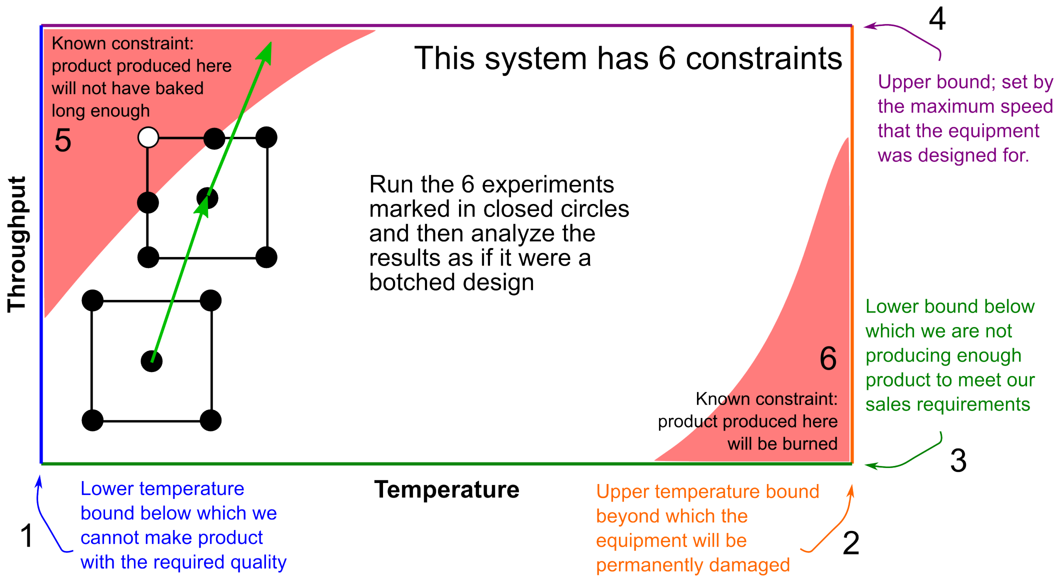

This system has 6 constraints
Known constraint: product produced here will not have baked long enough
5
6
Known constraint: product produced here will be burned
Throughput
Temperature
4
Upper bound; set by the maximum speed that the equipment was designed for.
Lower bound below which we are not producing enough product to meet our sales requirements
3
1
Lower temperature bound below which we cannot make product with the required quality
Upper temperature bound beyond which the equipment will be permanently damaged
2

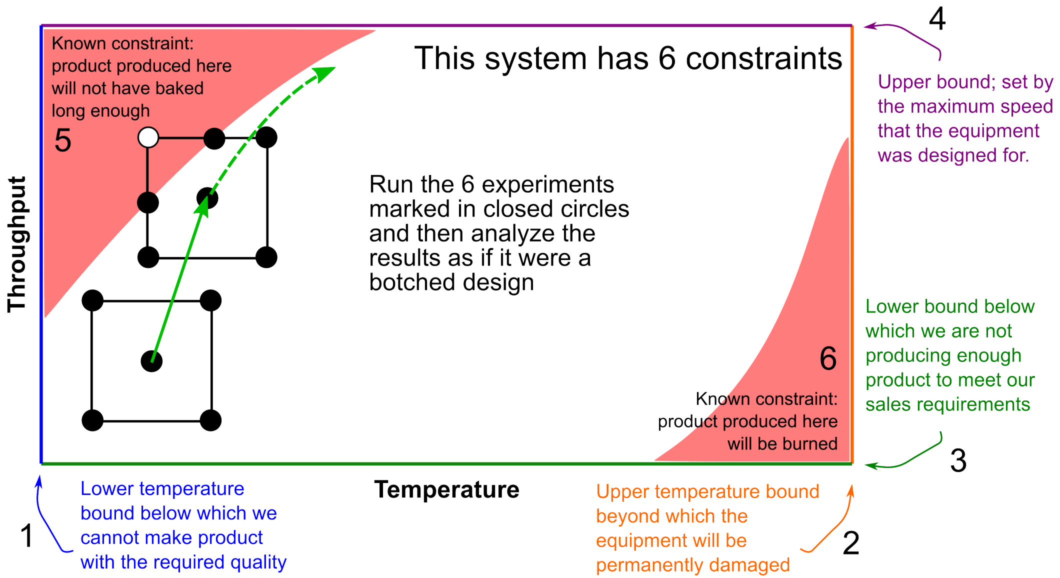


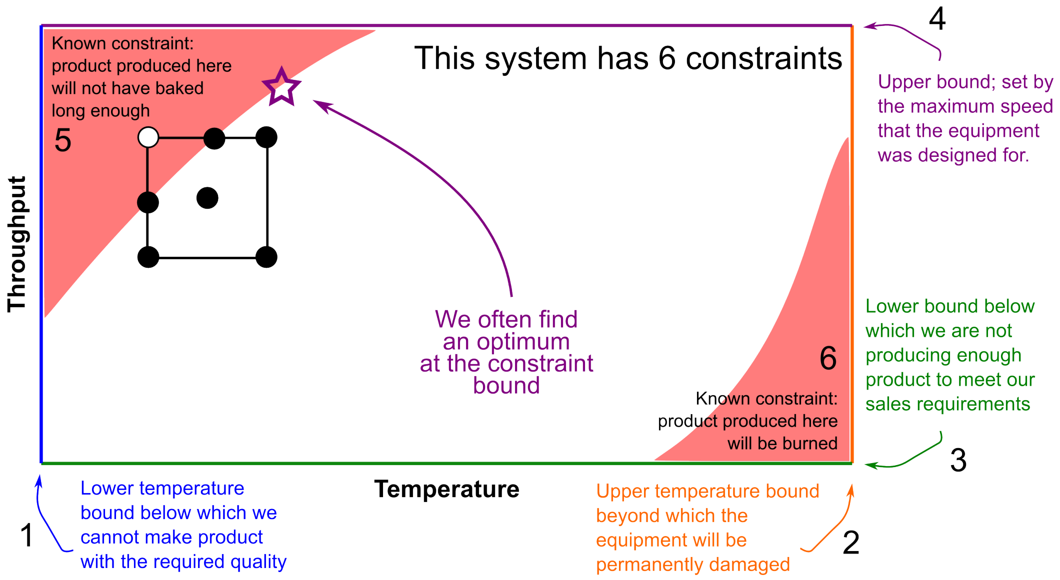
This system has 6 constraints
Known constraint: product produced here will not have baked long enough
5
Known constraint: product produced here will be burned
6
Throughput
Temperature
4
Upper bound; set by the maximum speed that the equipment was designed for.
Lower bound below which we are not producing enough product to meet our sales requirements
3
1
Lower temperature bound below which we cannot make product with the required quality
Upper temperature bound beyond which the equipment will be permanently damaged
2



This system has 6 constraints
Known constraint: product produced here will not have baked long enough
5
Run the 6 experiments marked in closed circles and then analyze the results as if it were a botched design
6
Known constraint: product produced here will be burned
Throughput
Temperature
4
Upper bound; set by the maximum speed that the equipment was designed for.
Lower bound below which we are not producing enough product to meet our sales requirements
3
1
Lower temperature bound below which we cannot make product with the required quality
Upper temperature bound beyond which the equipment will be permanently damaged
2


This system has 6 constraints
Known constraint: product produced here will not have baked long enough
5
Run the 6 experiments marked in closed circles and then analyze the results as if it were a botched design
6
Known constraint: product produced here will be burned
4
Upper bound; set by the maximum speed that the equipment was designed for.
Lower bound below which we are not producing enough product to meet our sales requirements
3
Throughput
Temperature
1
Lower temperature bound below which we cannot make product with the required quality
Upper temperature bound beyond which the equipment will be permanently damaged
2


This system has 6 constraints
Known constraint: product produced here will not have baked long enough
5
Run the 6 experiments marked in closed circles and then analyze the results as if it were a botched design
6
Known constraint: product produced here will be burned
Throughput
Temperature
4
Upper bound; set by the maximum speed that the equipment was designed for.
Lower bound below which we are not producing enough product to meet our sales requirements
3
1
Lower temperature bound below which we cannot make product with the required quality
Upper temperature bound beyond which the equipment will be permanently damaged
2


This system has 6 constraints
Known constraint: product produced here will not have baked long enough
5
Throughput
We often find an optimum at the constraint bound
6
Known constraint: product produced here will be burned
4
Upper bound; set by the maximum speed that the equipment was designed for.
Lower bound below which we are not producing enough product to meet our sales requirements
3
Temperature
1
Lower temperature bound below which we cannot make product with the required quality
Upper temperature bound beyond which the equipment will be permanently damaged
2