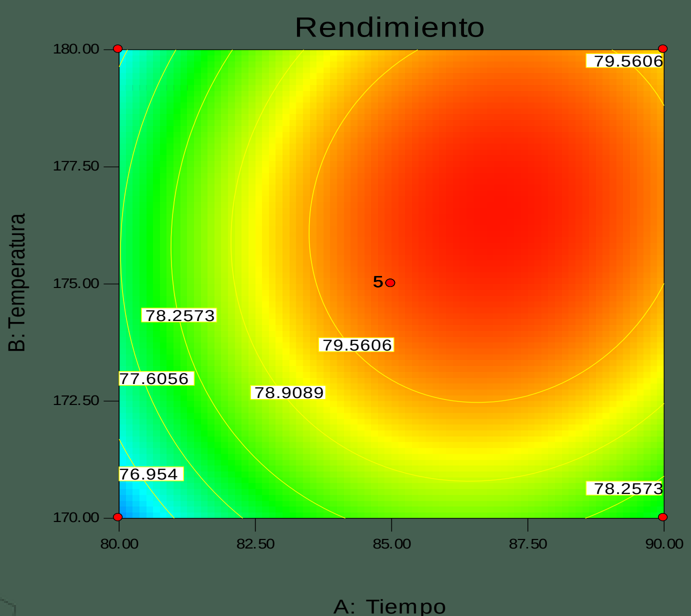
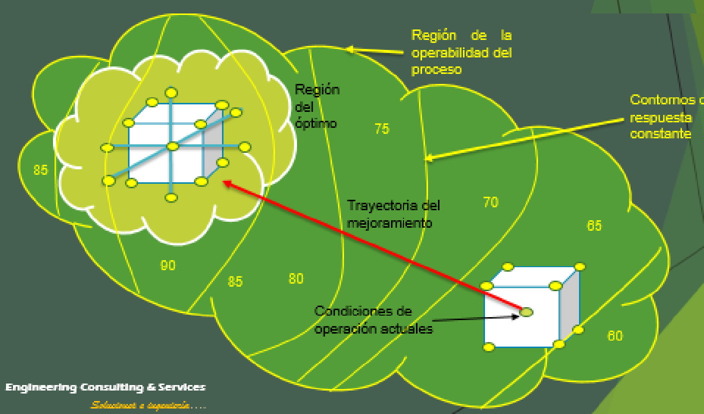
Diseño de experimentos usando el método de superficie de respuesta con diseño central compuesto rotacional.





Engineering Consulting & Services

Soluciones e ingeniería….

Ing. Jesús Alfredo Obregón Domínguez.

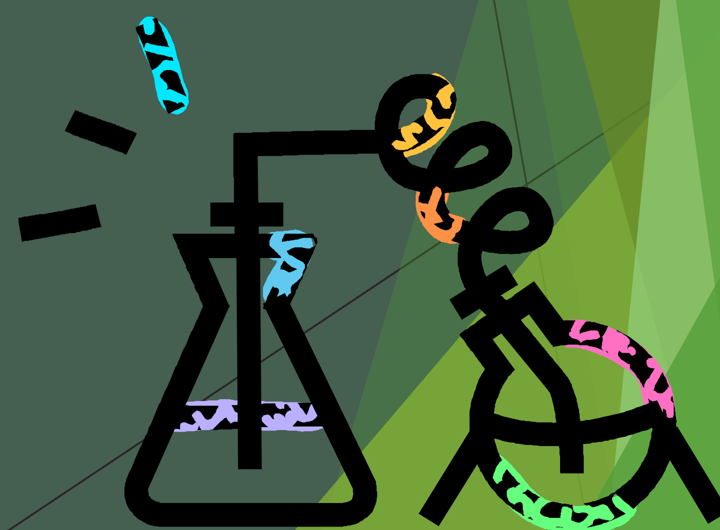
Estadística industrial

La estadística industrial es la rama de la estadística que busca implementar los procedimientos probabilísticos y estadísticos de análisis e interpretación de datos o características de un conjunto de elementos al entorno industrial, a efectos de ayudar en la toma de decisiones y en el control de los procesos industriales y organizacionales.

Diseño de experimentos

Un experimento es una prueba o ensayo.

El experimento diseñado es una prueba o serie de pruebas en las cuales se inducen cambios deliberados en la variable de entrada de un proceso o sistema, de manera que sea posible observar e identificar las causas de los cambios en la respuesta de salida (Montgomery, 2004).



Engineering Consulting & Services

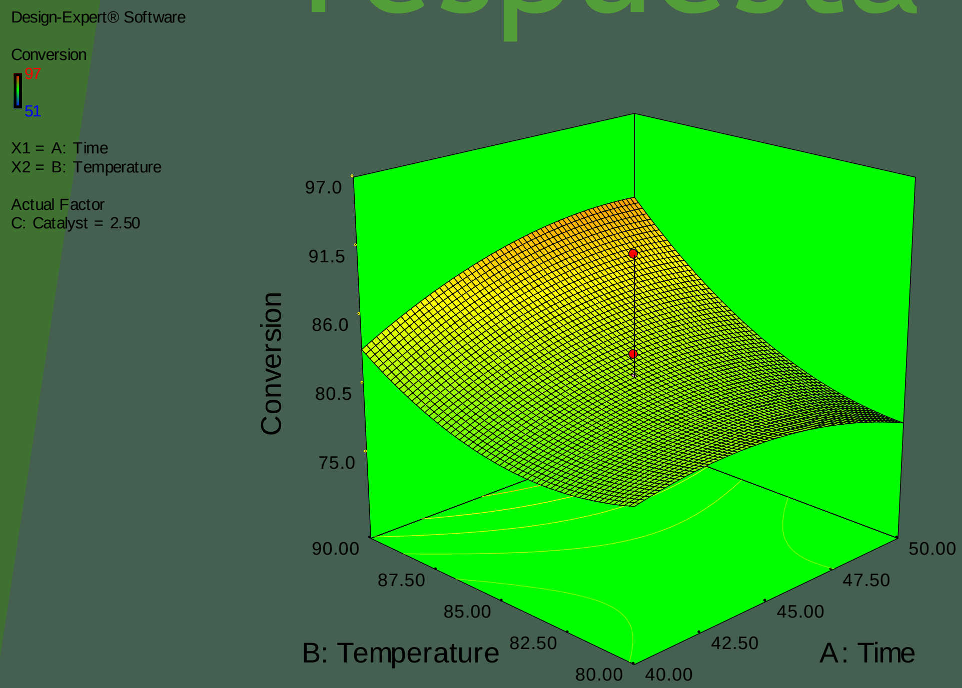
Soluciones e ingeniería....

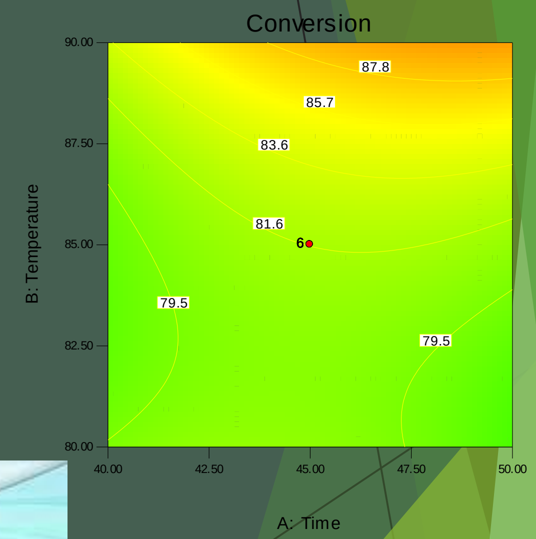
Diseño de experimentos

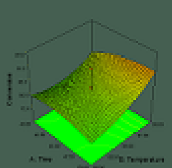
Los diseños de experimentos mas usados en estadística industrial son:

1. Plackett-Burman (screening).
2. Factoriales completos.
3. Factoriales fraccionados.
4. Diseño robusto de Taguchi.
5. Box-Behnken
6. **Superficie de respuesta de diseño compuesto central rotacional DCCR.**
7. Diseño de mezclas.
8. Diseño D-óptimo de parcelas divididas.

Superficie de respuesta DCCR.







Engineering Consulting & Services

Soluciones e ingeniería....

Superficie de respuesta DCCR

La metodología de superficie de respuesta (MSR) es una colección de técnicas matemáticas y estadísticas útiles en el modelado y el análisis de problemas en los que una respuesta de interés recibe la influencia de diversas variables y donde el objetivo es **optimizar** esta respuesta.

Ventajas

- La MSR tiene varias ventajas comparada con los experimentos clásicos y métodos de optimización, ya que nos permite obtener una gran cantidad de información a partir de un número pequeño de experimentos. Con el uso de los métodos clásicos, se consume mayor cantidad de tiempo y son necesarios mayor número de experimentos para explicar el comportamiento del sistema.

Superficie de respuesta DCCR

Ventajas

- Mejor calidad de información.
- El análisis multivariable permite verificar y cuantificar efectos sinérgicos y antagónicos entre variables estudiadas.
- Es posible optimizar mas de una respuesta al mismo tiempo.
- Es posible detectar el error experimental y evaluarlo.
- Utilizan operaciones matemáticas relativamente simples.

Superficie de respuesta DCCR

Optimizar

Buscar la mejor manera de realizar una actividad.

Diccionario de la real Academia Española

Optimizar (matemáticas)

Tomar una decisión para maximizar (ganancias, velocidad, eficiencia, tiempo, etc.) o minimizar un criterio determinado (costos, tiempo, riesgo, error, etc.)

Superficie de respuesta DCCR

¿Qué es un modelo?

Superficie de respuesta DCCR

¿Qué es un modelo?

Un modelo es producto de la abstracción de un sistema real, eliminando las complejidades, haciendo suposiciones pertinentes, y aplicando técnicas matemáticas se obtiene una representación simbólica del mismo.

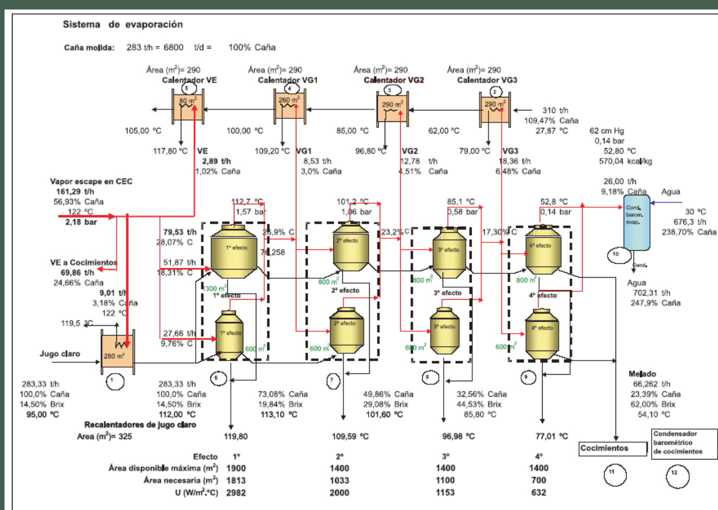


Figura 1. Caso base: resultados de los balances de masa y energía.

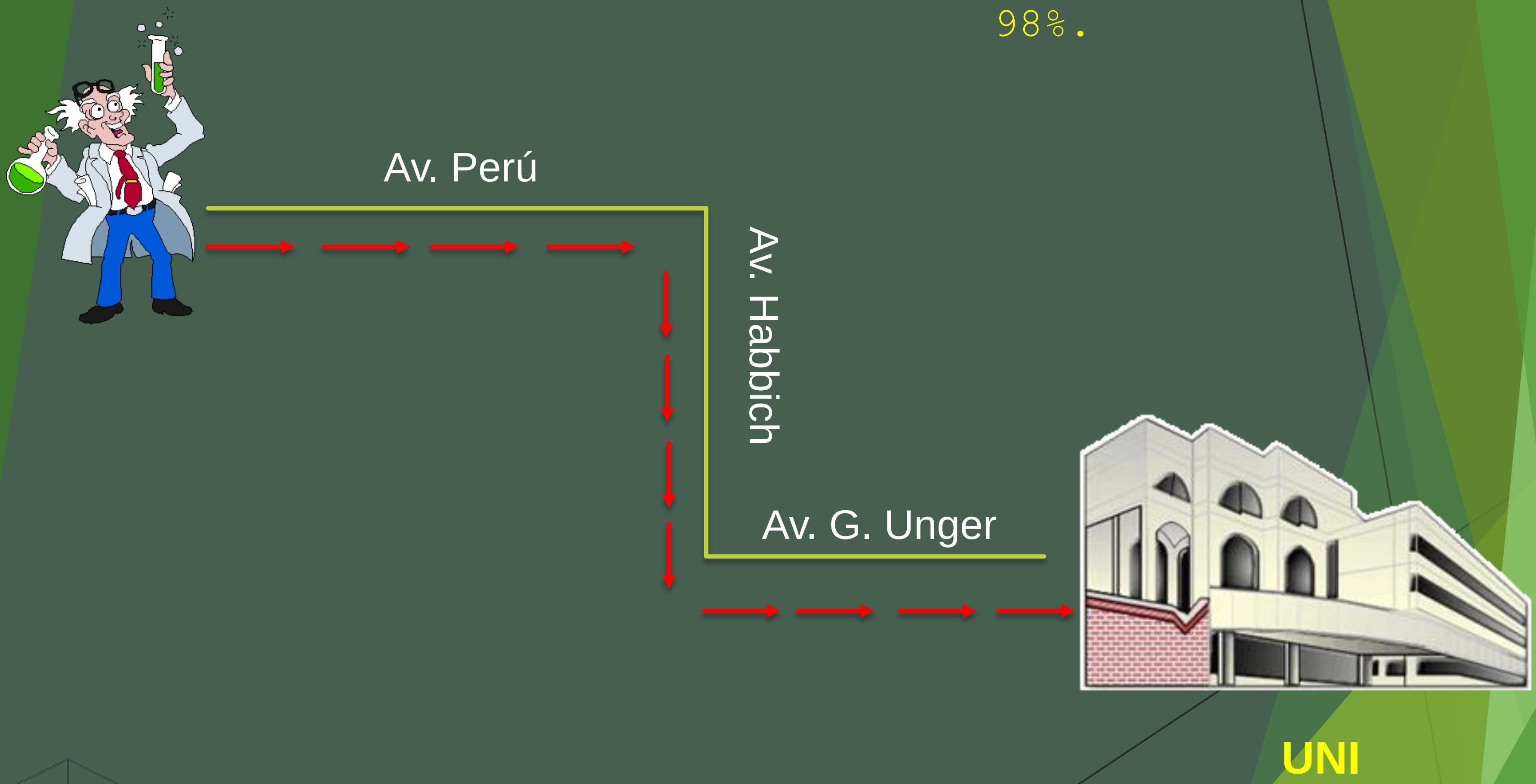
Engineering Consulting & Services

Soluciones e ingeniería....

Superficie de respuesta DCCR

¿Qué es un modelo?

Modelo describe el evento o fenómeno, aproximadamente un 98%.



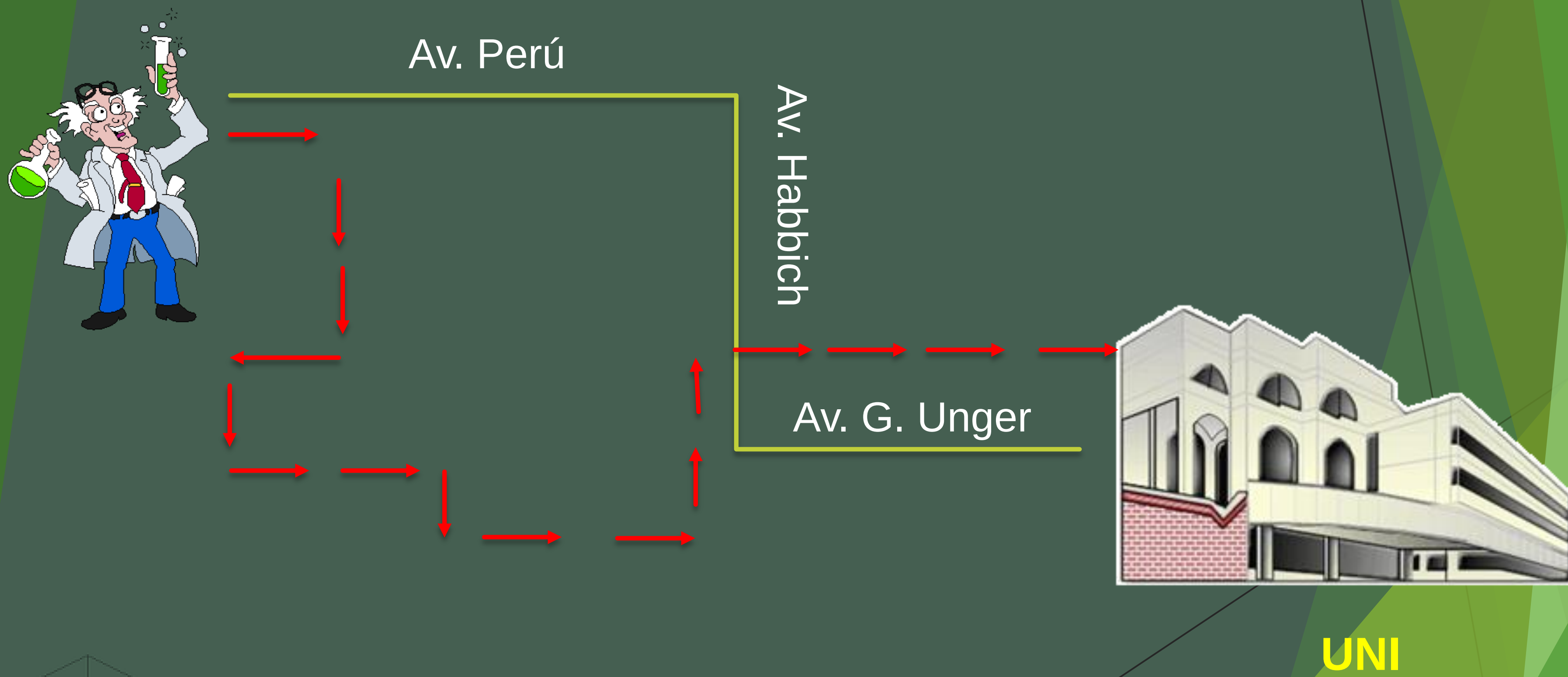
Engineering Consulting & Services

Soluciones e ingeniería....

Superficie de respuesta DCCR

¿Qué es un modelo?

Modelo describe el evento o fenómeno, aproximadamente un 35%.



Engineering Consulting & Services

Soluciones e ingeniería....

Superficie de respuesta DCCR

En la mayoría de los problemas de MSR, la forma de la relación entre la respuesta y las variables independientes es desconocida. Por lo tanto, el primer paso de la MSR es encontrar una aproximación adecuada de la verdadera relación funcional entre "*y*" y el conjunto de variables independientes. Por lo general, se emplea un polinomio de orden inferior en alguna región de las variables independientes. Si la respuesta está bien modelada por una función lineal de las variables independientes, entonces la función de aproximación es el modelo de primer orden:

$$y = \beta_0 + \beta_1 \chi_1 + \beta_2 \chi_2 + \ldots + \beta_\kappa \chi_\kappa + \varepsilon$$

Superficie de respuesta DCCR

Si hay curvatura en el sistema, entonces debe usarse un polinomio de orden superior, tal como el modelo de segundo orden:

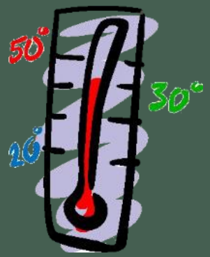
$$y = \beta_0 + \sum \beta_i x_i + \sum \beta_{ii} \beta_i^2 + \sum\sum \beta_{ij} x_i x_j + \varepsilon$$

En casi todos los problemas MSR se usa uno de estos modelos o ambos. Es probable que un modelo polinomial sea una aproximación razonable de la verdadera relación funcional en el espacio completo de las variables independientes, pero para una región relativamente pequeña suelen funcionar bastante bien.

Ejemplo interactivo

Un ingeniero químico quiere encontrar los niveles de tiempo (***x_1***) y temperatura (***x_2***) que maximicen el rendimiento (***y***) de un proceso. El rendimiento del proceso es una función de los niveles de tiempo y temperatura, por ejemplo:

$$y = f(\chi_1, \chi_2) + \varepsilon$$



Y = función (X_1, X_2)

El ingeniero opera actualmente el proceso con un **tiempo de reacción de 35 minutos** y una **temperatura de 155 °F**, que dan como resultado **rendimientos de cerca de 40%**.

Ejemplo interactivo

El ingeniero decide que la región de exploración para ajustar el modelo de primer orden deberá ser (30, 40) minutos de tiempo de reacción y (150, 160) °F. para simplificar los cálculos, las variables independientes se codificaron en el intervalo usual (-1, 1). Por lo tanto, si ξ_1 denota la variable natural tiempo y ξ_2 la variable natural temperatura, entonces las variables codificadas son.

$$X_1 = \frac{\xi_1 - 35}{5} \qquad X_2 = \frac{\xi_2 - 155}{5}$$

Planteamiento Factorial: 2^n + pc

N = numero de variables
2 = 2 niveles estudiados (-1, +1)
pc = puntos centrales (0)

Ejemplo:
2 variables (tiempo, temperatura)
2^2 = 4 ensayos + pc (5)

Ejemplo interactivo

Obsérvese que el diseño usado para recabar estos datos es un factorial 2^2 aumentado con cinco puntos centrales. Las replicas del centro del diseño se usan para estimar el error experimental y permitir la verificación de la adecuación del modelo de primer orden.

Variables naturales			Variables codificadas	
ε_1	ε_2		X_1	X_2
30	150		-1	-1
30	160		-1	1
40	150		1	-1
40	160		1	1
35	155		0	0
35	155		0	0
35	155		0	0
35	155		0	0
35	155		0	0

Ejemplo interactivo

Con este diseño el ingeniero realiza el experimento y recopila datos (respuesta).

Ejemplo interactivo

Datos del proceso para ajustar el modelo de primer orden

Variables naturales			Variables codificadas			Respuesta
ε_1	ε_2		X_1	X_2		y
30	150		-1	-1		39.3
30	160		-1	1		40.0
40	150		1	-1		40.9
40	160		1	1		41.5
35	155		0	0		40.3
35	155		0	0		40.5
35	155		0	0		40.7
35	155		0	0		40.2
35	155		0	0		40.6

Ejemplo interactivo

Efectos estimados

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	40.444	1	0.05728781	40.3042662	40.5846227	
A-Tiempo	0.775	1	0.08593171	0.56473268	0.98526732	1
B-Temperatu	0.325	1	0.08593171	0.11473268	0.53526732	1

Rendimiento = 40.44 +0.775*tiempo + 0.325*temperatura

Ejemplo interactivo

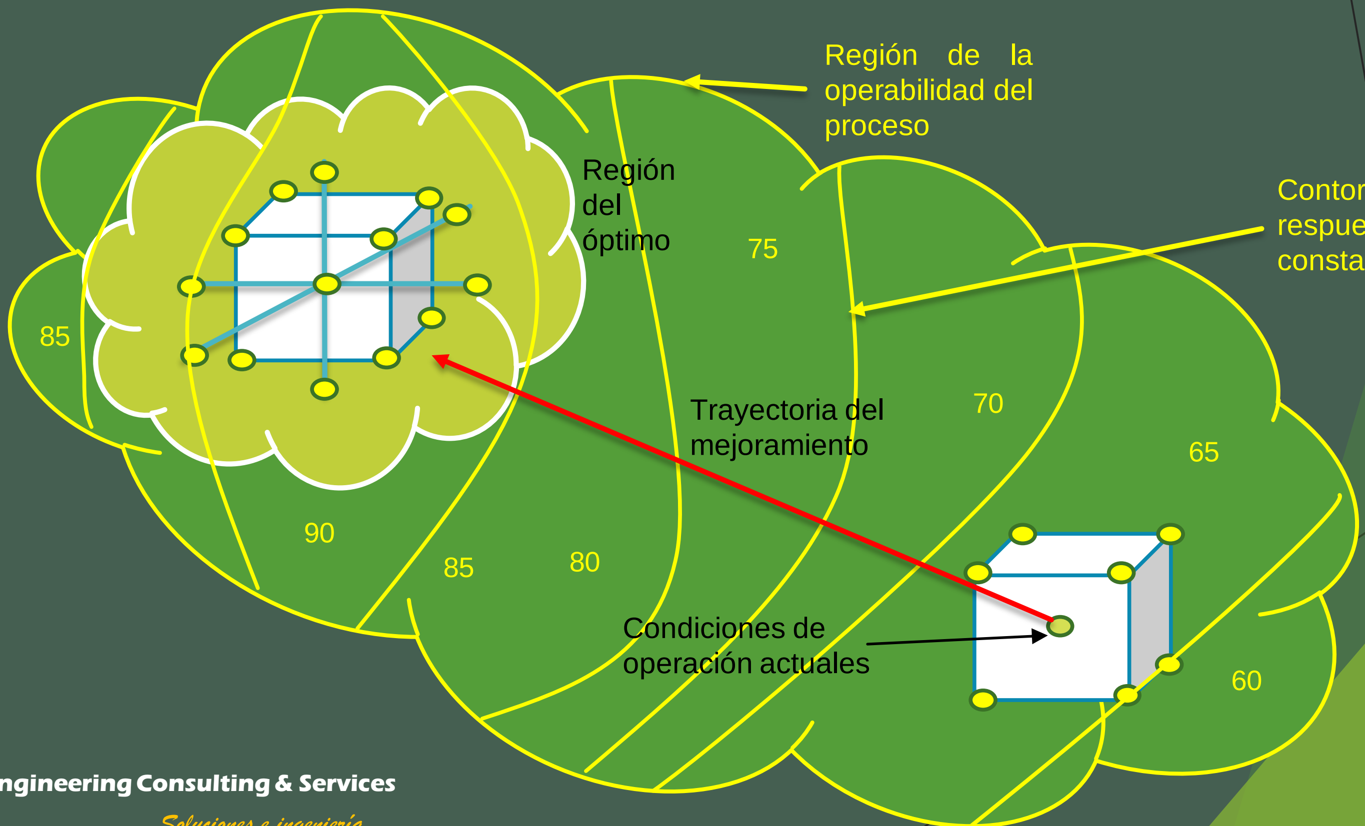
Análisis de varianza del modelo de primer orden

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Mean vs Total	14721.7778	1	14721.7778		
Linear vs Mean	2.8250	2	1.4125	47.821	0.0002
2FI vs Linear	0.0025	1	0.0025	0.072	0.7998
Quadratic vs 2FI	0.0027	1	0.0027	0.063	**0.8137**
Cubic vs Quadratic	0.0000	0			
Residual	0.1720	4	0.0430		
Total	14724.7800	9	1636.0867		

No hay indicios de un efecto cuadrático puro o curvatura

Ejemplo interactivo

El ingeniero opera actualmente el proceso con un **tiempo de reacción de 35 minutos** y una **temperatura de 155 °F**, que dan como resultado **rendimientos de cerca de 40%**, puesto que es improbable que esta región contenga el óptimo, el ingeniero ajusta un modelo de primer orden y aplica el **método de ascenso mas pronunciado**.



Ejemplo interactivo

El **método de ascenso rápido** es un procedimiento para moverse secuencialmente por la trayectoria de ascenso rápido, o sea, en la dirección del máximo incremento de la respuesta. Por supuesto, si lo que se busca es la minimización, entonces se utiliza **el método de descenso rápido**. El modelo ajustado de primer orden es:

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i \, x_i$$

Los experimentos se realizan a lo largo de la trayectoria de ascenso rápido hasta que ya no se observa incremento en la respuesta o hasta que la región de la respuesta deseada se alcanza. Entonces se usa un nuevo modelo de primer orden, se determina la dirección de una nueva trayectoria de ascenso rápido y de ser necesario, **se realizan experimentos adicionales** en esa dirección hasta que el experimentador sienta que está cerca del óptimo.

Ejemplo interactivo

Para la trayectoria de ascenso más rápido, se realiza:

a) Se elige el tamaño de paso de una de las variables del proceso Δx_j. La variable que tiene el coeficiente en valor absoluto más alto. En este caso se elige X_1.

b) El tamaño del paso para las otras variables es

$$\Delta x_i = \frac{\hat{\beta}_i}{\hat{\beta}_j / \Delta x_j} \ldots .i = 1, 2, .., k; -- para..i \neq j$$

En este caso:

$$\Delta x_2 = \frac{\hat{\beta}_2}{\hat{\beta}_1 / \Delta x_1} = \frac{0.325}{(0.775)/1.0} = 0.42$$

Ejemplo interactivo

Para convertir los tamaños de los pasos codificados (Delta X_1=1.0 y Delta X_2=0.42) a las unidades naturales de tiempo y temperatura se tiene:

$$\Delta x_1 = \frac{\Delta \xi_1}{5}$$

$$\Delta x_2 = \frac{\Delta \xi_2}{5}$$

$$\Delta \xi_1 = \Delta x_1(5) = 1 * 5 = 5 mi.$$

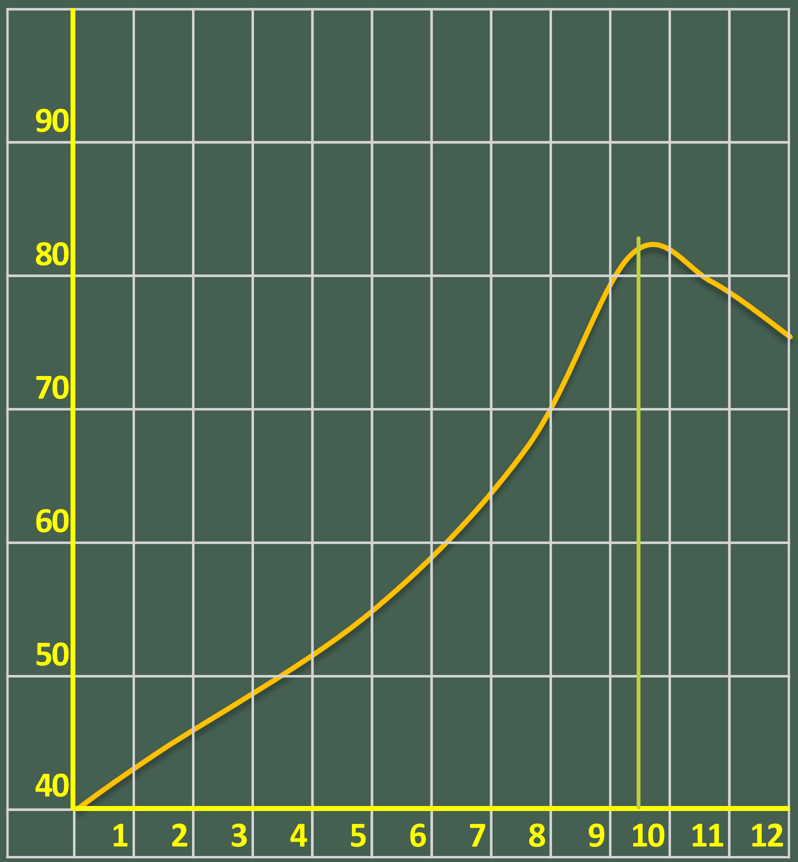
$$\Delta \xi_2 = \Delta x_2(5) = 0.42 * 5 = 2^{\circ} F$$

Tomando el punto correspondiente a (0,0) se realizan experimentos individuales adicionales, incrementando las variables en los pasos indicados arriba resultando en:

Ejemplo interactivo

	Variables codificadas		Variables naturales		Respuesta
Pasos	**X1**	**X2**	**ξ1**	**ξ2**	**Y**
Origen	0	0	35	155	
D	1	0.42	5	2	
Orig.+Δ	1	0.42	40	157	41.0
Orig.+2Δ	2	0.84	45	159	42.9
Orig.+3Δ	3	1.26	50	161	47.1
Orig.+4Δ	4	1.68	55	163	49.7
Orig.+5Δ	5	2.1	60	165	53.8
Orig.+6Δ	6	2.52	65	167	59.9
Orig.+7Δ	7	2.94	70	169	65.0
Orig.+8Δ	8	3.36	75	171	70.4
Orig.+9Δ	9	3.78	80	173	77.6
Orig.+10Δ	10	4.2	85	175	80.3
Orig.+11Δ	11	4.62	90	179	76.2
Orig.+12Δ	12	5.04	95	181	75.1

Ejemplo interactivo



Rendimiento contra pasos en la trayectoria de ascenso más pronunciado

Engineering Consulting & Services

Soluciones e ingeniería....

Superficie de respuesta DCCR

Se observa que el punto décimo representa el valor máximo de la trayectoria de experimentación por lo que ahora se tomará como nuevo punto central (0,0) el punto (85, 175) y la región de experimentación para ξ_1 es (80,90) y para ξ_2 es (170,180), con las variables codificadas X_1 y X_2 como sigue:

$$x_1 = \frac{\xi_1 - 85}{5}$$

$$x_2 = \frac{\xi_2 - 175}{5}$$

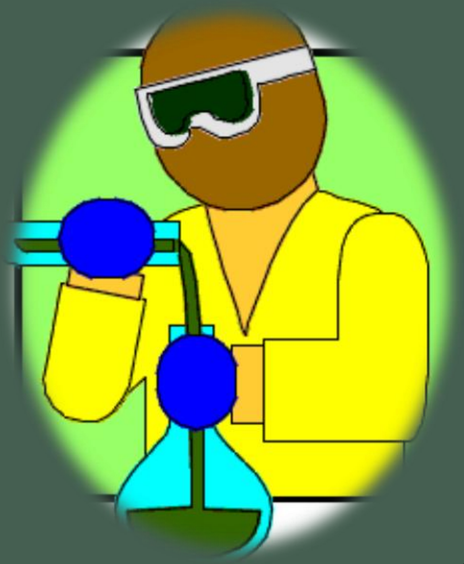
Superficie de respuesta DCCR

Haciendo nuevos experimentos alrededor del nuevo punto (0,0) se tiene:

Variables naturales			Variables codificadas	
ε_1	ε_2		X_1	X_2
80	170		-1	-1
80	180		-1	1
90	170		1	-1
90	180		1	1
85	175		0	0
85	175		0	0
85	175		0	0
85	175		0	0
85	175		0	0

Ejemplo interactivo

Con este nuevo diseño el ingeniero desarrolla el experimento y recopila datos (respuesta).



Superficie de respuesta DCCR

Datos del proceso para ajustar el modelo de primer orden

Variables naturales			Variables codificadas			Respuesta
ε_1	ε_2		X_1	X_2		y
80	170		-1	-1		76.5
80	180		-1	1		77.0
90	170		1	-1		78.0
90	180		1	1		79.5
85	175		0	0		79.9
85	175		0	0		80.3
85	175		0	0		80.0
85	175		0	0		79.7
85	175		0	0		79.8

Superficie de respuesta DCCR

Efectos estimados

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	78.967	1	0.45379062	77.8562811	80.0770523	
A-Tiempo	1	1	0.68068593	-0.66557841	2.66557841	1
B-Temperatu	0.5	1	0.68068593	-1.16557841	2.16557841	1

Rendimiento = 78.97 +1.0*tiempo + 0.5*temperatura

Superficie de respuesta DCCR

Datos del proceso para ajustar el modelo de primer orden

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Mean vs Total	56121.6100	1	56121.6100		
Linear vs Mean	5.0000	2	2.5000	1.349	0.3283
2FI vs Linear	0.2500	1	0.2500	0.115	0.7483
Quadratic vs 2FI	10.6580	1	10.6580	201.094	**0.0001**
Cubic vs Quadratic	0.0000	0			
Residual	0.2120	4	0.0530		
Total	56137.7300	9	6237.5256		

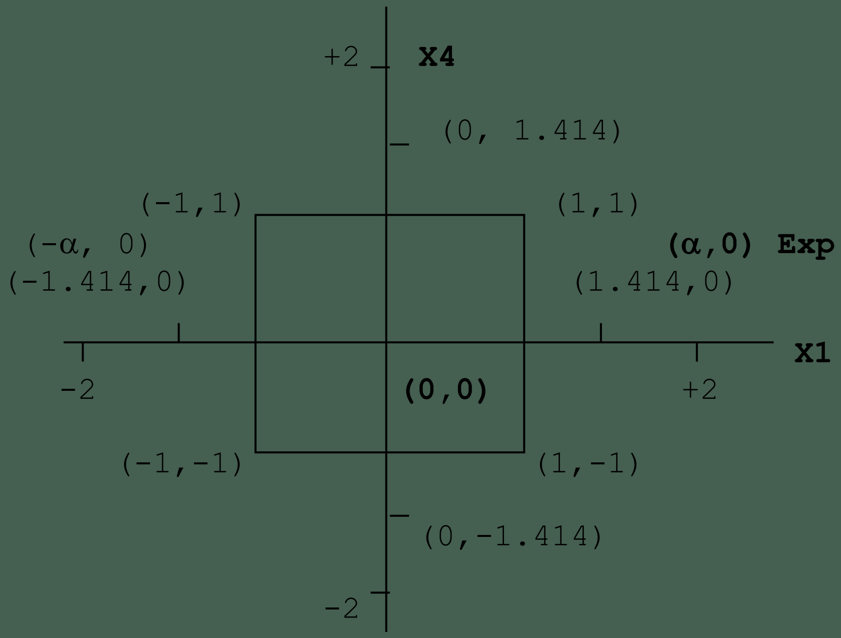
Hay indicios de un efecto cuadrático puro o curvatura

Engineering Consulting & Services

Soluciones e ingeniería….

Superficie de respuesta DCCR

Como la curvatura es significativa ahora aplicamos el modelo central compuesto que se muestra abajo para obtener un modelo de segundo orden.



Diseño Central Compuesto en las variables codificadas del ejemplo

Superficie de respuesta DCCR

La ecuación de segundo grado que nos dará tiene la forma siguiente:

$$y = \beta_0 + \sum \beta_i x_i + \sum \beta_{ii} \beta_i^2 + \sum\sum \beta_{ij} x_i x_j + \varepsilon$$

	Variables naturales			Variables codificadas			Respuesta
	ε_1	ε_2		X_1	X_2		y
Factoriales	80	170		-1	-1		76.5
	80	180		-1	1		77.0
	90	170		1	-1		78.0
	90	180		1	1		79.5
Centrales	85	175		0	0		79.9
	85	175		0	0		80.3
	85	175		0	0		80.0
	85	175		0	0		79.7
	85	175		0	0		79.8
Axiales	92.07	175		1.414	0		78.4
	77.93	175		-1.414	0		75.6
	85	182.07		0	1.414		78.5
	85	167.93		0	-1.414		77.0

Engineering Consulting & Services

Soluciones e ingeniería….

Superficie de respuesta DCCR

Modelo secuencial de suma de cuadrados

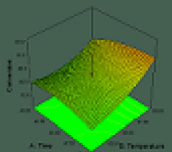
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Mean vs Total	80062.1569	1	80062.1569		
Linear vs Mean	10.0430	2	5.0215	2.685	0.1166
2FI vs Linear	0.2500	1	0.2500	0.122	0.7350
Quadratic vs 2FI	17.9548	2	8.9774	126.879	< 0.0001
Cubic vs Quadratic	0.0020	2	0.0010	0.010348979	0.9897
Residual	0.4933	5	0.0987		
Total	80090.9000	13	6160.8385		

$p<0.05$, existe efecto significativo

Prueba de falta de ajuste (Lack of Fit)

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Linear	18.4881188	6	3.08135313	58.13873829	0.0008
2FI	18.2381188	5	3.64762376	68.82308972	0.0006
Quadratic	0.28329185	3	0.09443062	1.781709771	0.2897
Cubic	0.28125	1	0.28125	5.306603774	0.0826
Pure Error	0.2120	4	0.0530		

$p>0.05$, existe buen ajuste



Superficie de respuesta DCCR

Resumen de modelos estadísticos

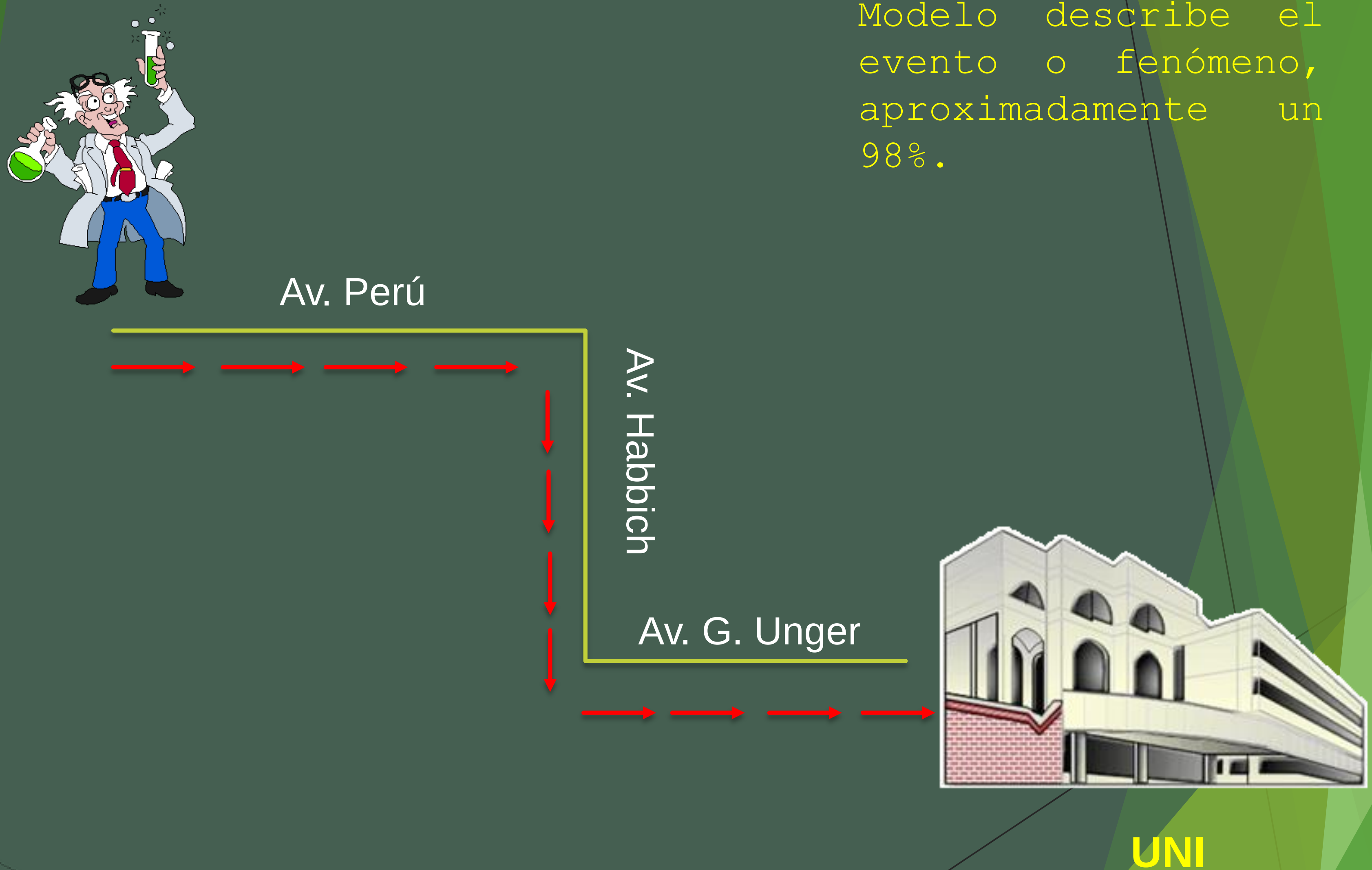
Source	Std. Dev.	R-Squared	Adjusted R-Squared	Predicted R-Squared	PRESS
Linear	1.3675	0.3494	0.2193	-0.0435	29.9945
2FI	1.4318	0.3581	0.1441	-0.2730	36.5891
Quadratic	0.2660	0.9828	0.9705	0.9184	2.3458
Cubic	0.3141	0.9828	0.9588	0.3622	18.3313

R^2 = 98.28%

R^2-Ajustado = 97.05%

Superficie de respuesta DCCR

Modelo describe el evento o fenómeno, aproximadamente un 98%.



Engineering Consulting & Services

Soluciones e ingeniería....

Superficie de respuesta DCCR

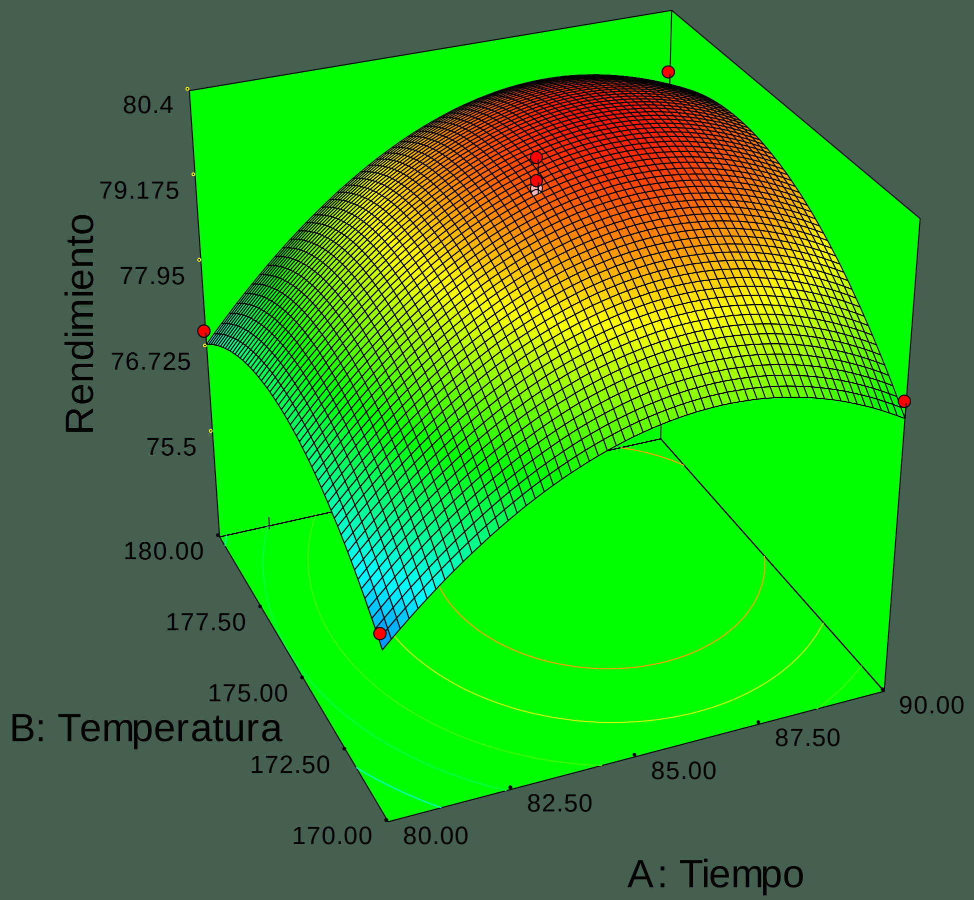


Grafico de superficie de respuesta

Engineering Consulting & Services

Soluciones e ingeniería….

Superficie de respuesta DCCR

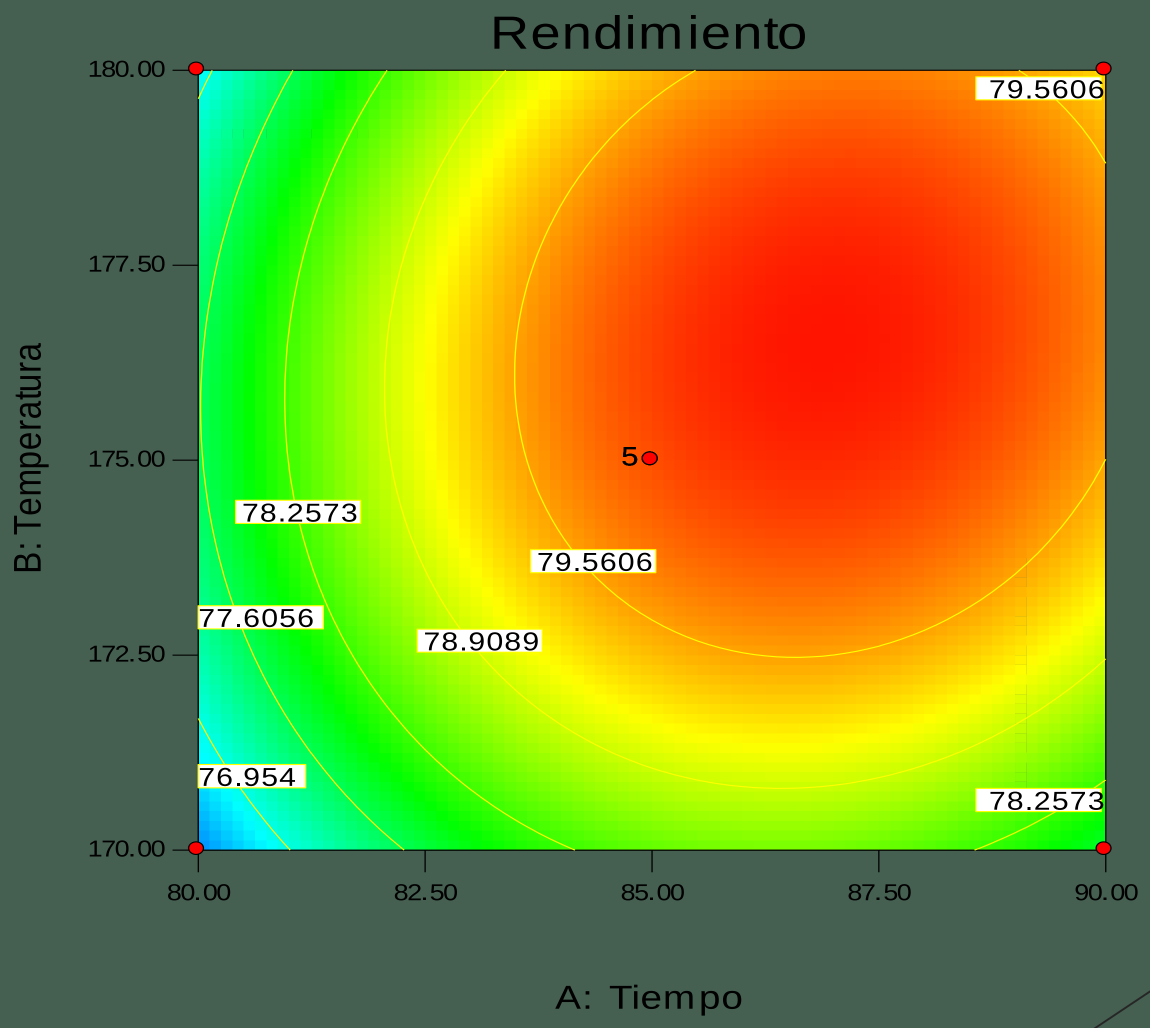


Gráfico de contorno

Engineering Consulting & Services

Soluciones e ingeniería....

Superficie de respuesta DCCR

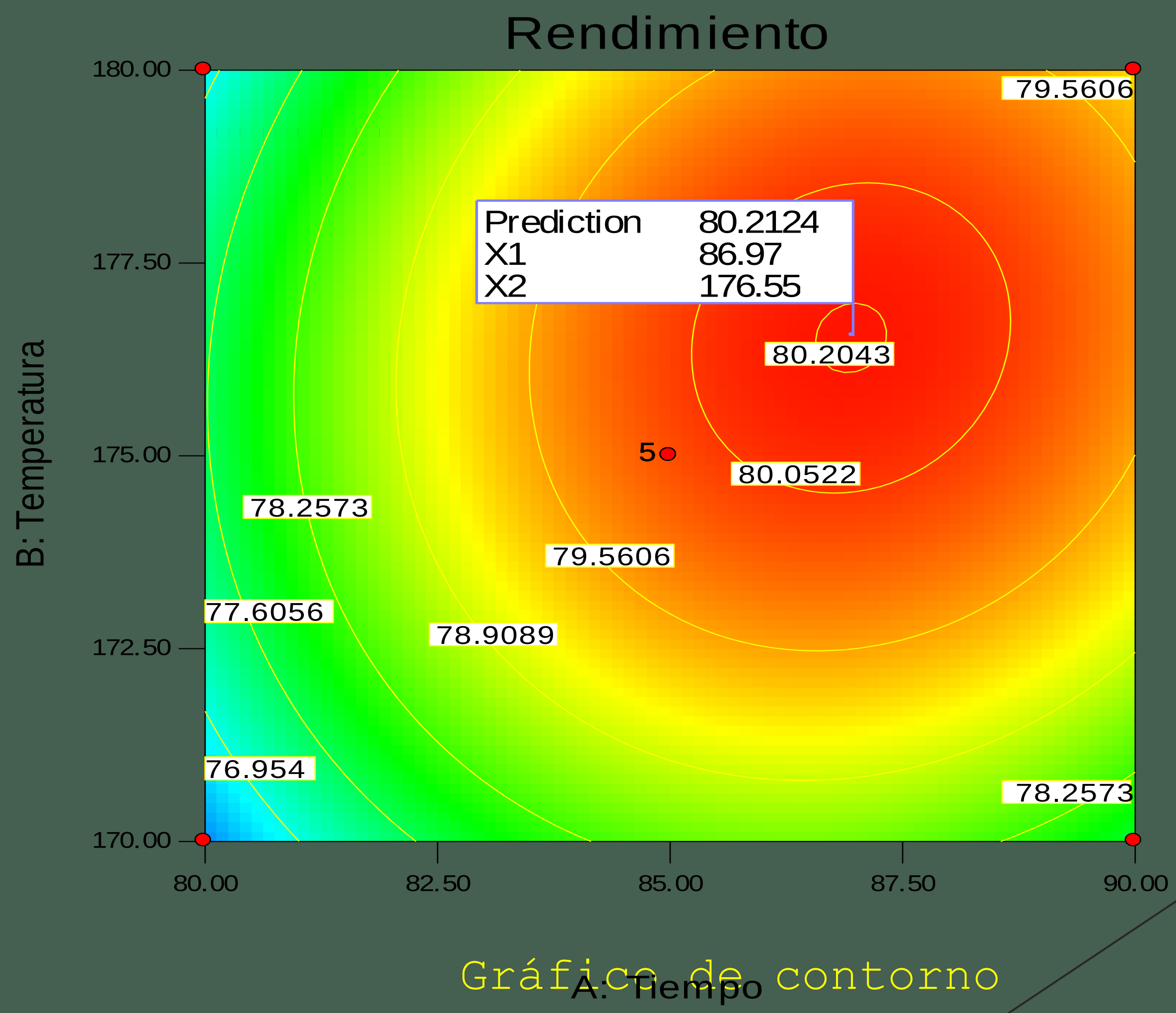
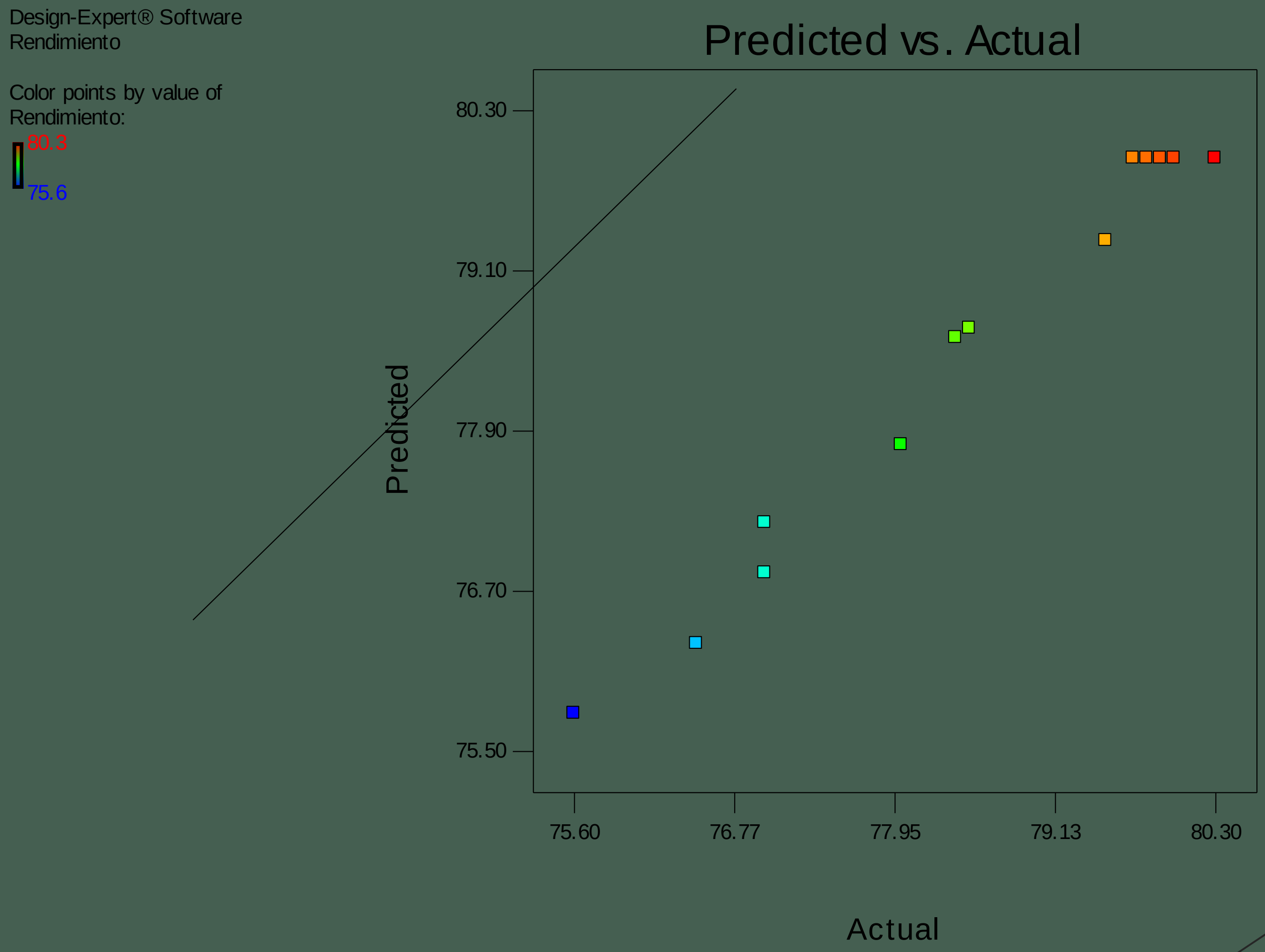


Gráfico de contorno

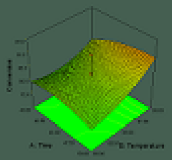
Engineering Consulting & Services

Soluciones e ingeniería....

Superficie de respuesta DCCR



Predichos ve. Experimentales u observados



Engineering Consulting & Services

Soluciones e ingeniería….

Superficie de respuesta DCCR

Análisis de varianza del modelo cuadrático de superficie de respuesta DCCR

	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	28.2477851	5	5.64955701	79.8456482	< 0.0001
A-Tiempo	7.91979797	1	7.91979797	111.9311481	< 0.0001
B-Temperatu	2.12316017	1	2.12316017	30.00679518	0.0009
AB	0.25	1	0.25	3.533270308	0.1022
A^2	13.1760978	1	13.1760978	186.2188609	< 0.0001
B^2	6.97392391	1	6.97392391	98.56303317	< 0.0001
Residual	0.49529185	7	0.07075598		
Lack of Fit	0.28329185	3	0.09443062	1.781709771	0.2897
Pure Error	0.212	4	0.053		
Cor Total	28.7431	12.0000			

Efectos estimados

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	79.940	1	0.12	79.66	80.22	
A-Tiempo	0.995	1	0.09	0.77	1.22	1.00
B-Temperatura	0.515	1	0.09	0.29	0.74	1.00
AB	0.250	1	0.13	-0.06	0.56	1.00
A^2	-1.376	1	0.10	-1.61	-1.14	1.02
B^2	-1.001	1	0.10	-1.24	-0.76	1.02

Rendimiento = 79.94 + 0.995*tiempo + 0.515*temperatura – 1.376*tiempo^2 – 1.001*temperatura^2 + 0.250*tiempo*temperatura.

Engineering Consulting & Services

Soluciones e ingeniería....

ESO ES TODO AMIGOS...........