DISEÑO DE PROCESOS ROBUSTOS

DICIEMBRE 2014

Índice general

1. INTRODUCCIÓN

Como introducción al **Diseño de Procesos Robustos** hemos considerado conveniente la exposición de algunos conceptos previos. Estos conceptos son:

- Interacción entre variables de control y de ruido.
- Valor medio y Varianza de una combinación lineal de variables aleatorias.
- Modelos sencillos de Superficie de Respuesta.

1.1. Interacción entre variables de control y de ruido

La existencia de interacción significativa entre las variables de control y las de ruido es fundamental para el diseño de Procesos Robustos. En las figuras 1 y 2 se presentan los gráficos de interacción entre una variable de control (x) y una variable de ruido (z) para distintas situaciones. Se supone que la variable x tiene dos niveles: -1 y +1 mientras que la variable de ruido (z) es una variable continua que varía de forma aleatoria entre -1 y +1.

El gráfico de la figura 1 señala la presencia de interacción. Observamos la falta de paralelismo de las respuestas. **Nótese que cuando la variable de control x está en el nivel -1 la respuesta apenas experimenta variación.** Ajustando la variable x en el citado nivel x=-1 podremos conseguir un Diseño Robusto frente a la acción de la variable de ruido z.

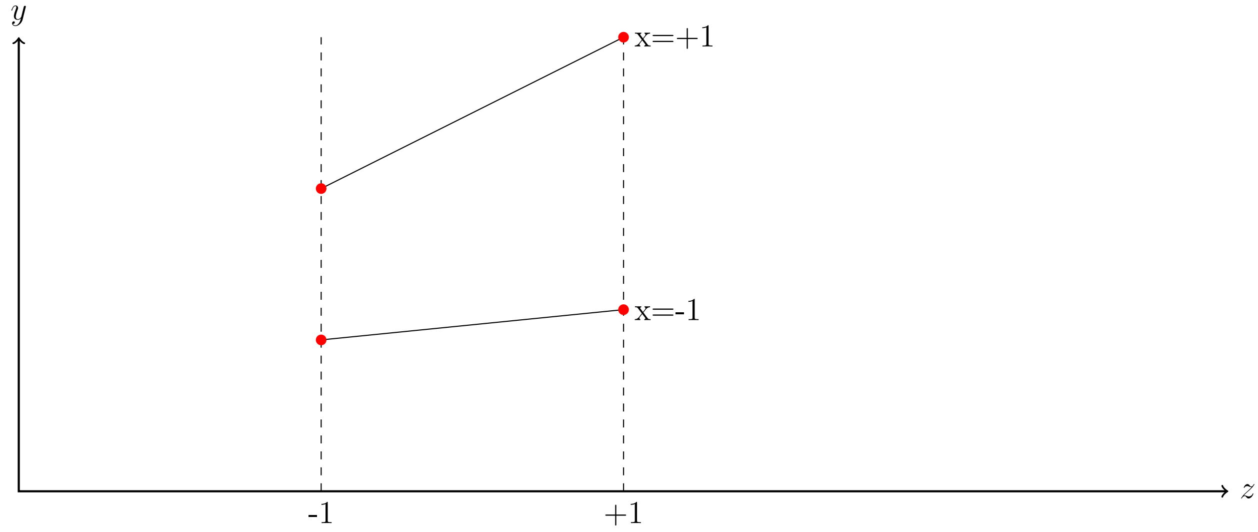


Figura 1: Presencia de Interacción

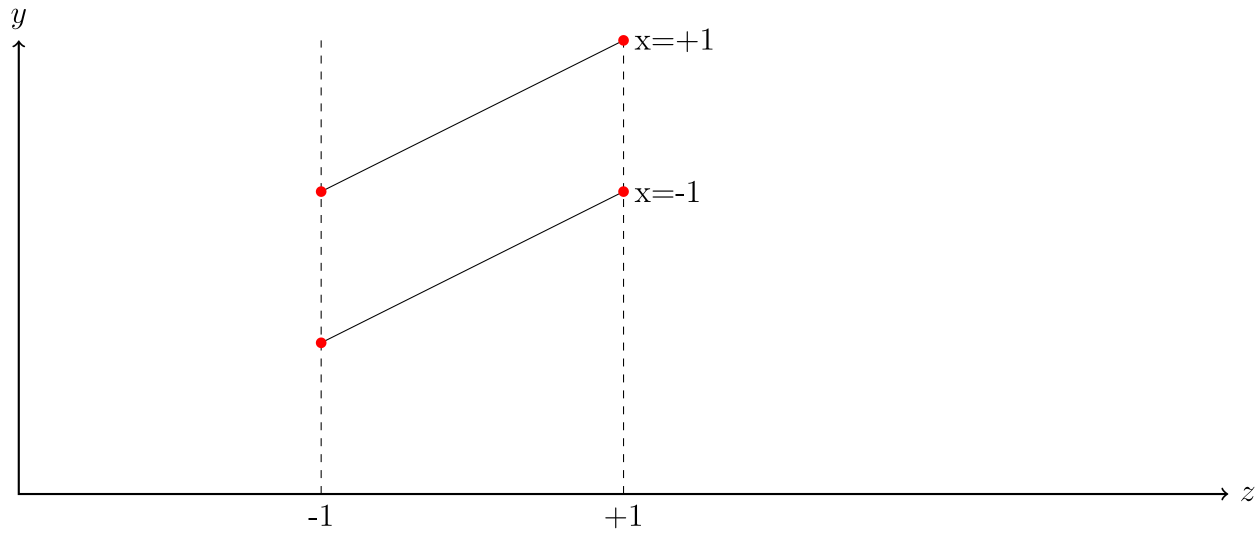


Figura 2: Ausencia de Interacción

La figura 2, por el contrario, no señala presencia de interacción al ser paralelas las lineas de respuesta para los niveles x=-1 y x=+1 de la variable de control. En consecuencia, **no es posible evitar la influencia de la variable de ruido z sobre la variabilidad** y no podremos obtener un diseño robusto.

En el caso general, que contempla la presencia de varias variables de control y varias variables de ruido, la existencia de interacciones es también necesaria para que sea posible el diseño de un Proceso Robusto.

1.2. Valor medio y Varianza de una combinación lineal

Sea una combinación lineal:

$$y = a_0 + a_1 y_1 + a_2 y_2 + a_3 y_3$$

Las variables y_1, y_2, y_3 son **variables aleatorias** con medias μ_1, μ_2, μ_3 y varianzas σ_1^2, σ_2^2, σ_3^2. Los coeficientes a_0, a_1, a_2, a_3 son valores constantes.

En el caso que y_1, y_2, y_3 no estén correlacionadas entre si, el valor medio y la varianza de la variable "y" serán:

$$E(y) = \mu = a_0 + a_1\mu_1 + a_2\mu_2 + a_3\mu_3 \tag{1}$$

$$V(y) = \sigma_y^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + a_3^2\sigma_3^2 \tag{2}$$

La aplicación de los operadores E(y) y V(y), como veremos, será fundamental en el diseño de Procesos Robustos.

1.3. Modelos sencillos de Superficie de Respuesta

El método de Superficie de Respuesta, aplicado al diseño de Procesos Robustos, se fundamenta en el **ajuste de un modelo lineal** polinomial que incluye tanto a las variables de control como a las variables de ruido. Este modelo lineal será habitualmente un modelo sencillo de primer o segundo orden en las variables de control incluyendo, además, los terminos de interación de dos variables. El término linealidad se refiere a los paramétros β_0, β_1, β_2 ... y no a las variables.

Consideremos un proceso en el que las variables de control sean, por ejemplo, x_1 y x_2 y la variable de ruido z_1. Un modelo sencillo que puede ajustarse para la variable de respuesta "y", teniendo en cuenta que necesariamente debe contener los términos de interacción entre las variables de control y la de ruido será:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \gamma_1 z_1 + \delta_{11} x_1 z_1 + \delta_{21} x_2 z_1 + \varepsilon$$

Ordenando los términos resulta:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + (\gamma_1 + \delta_{11} x_1 + \delta_{21} x_2) z_1 + \varepsilon \qquad (3)$$

El modelo contempla, también, la interación entre las dos variables de control y prescinde de otros posibles factores de entrada y de sus interacciones que quedan englobados en el término de error ε. El término de error ε se supone conforme con una distribución Normal de valor medio cero y varianza σ^2, es decir, del tipo $N(0, \sigma^2)$. Los coeficientes del modelo y la varianza residual σ^2 se estiman mediante **Análisis de Regresión** (ver nuestro capítulo de Regresión Lineal Múltiple de Junio 2011).

Otro modelo, ampliamente utilizado para realizar ajustes en una región en la que se presume existe una mayor curvatura, es un modelo que contiene, también, términos de segundo orden para las variables de control:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \gamma_1 z_1 + \delta_{11} x_1 z_1 + \delta_{21} x_2 z_1 + \varepsilon$$

Ordenando términos:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + (\gamma_1 + \delta_{11} x_1 + \delta_{21} x_2) z_1 + \varepsilon \quad (4)$$

Debe tenerse en cuenta que estos modelos son una aproximación a la verdadera función que relaciona la respuesta con las variables de entrada y que es desconocida. Recordemos que, mediante el desarrollo de Taylor, cualquier función puede desarrollarse en forma polinomial. La aproximación será más exacta cuanto mayor sea el orden del polinomio.

Los modelos, tales como (3) y (4), se usan con las variables codificadas. Como veremos en los ejemplos de aplicación, se supone que las variables de control son, en principio, cuantitativas y de tipo continuo, aunque se puede generalizar el análisis al caso de variables cualitativas. Para las variables de ruido se suponen ya conocidos por el investigador sus valores medios y sus varianzas. El ajuste de los modelos se realiza seleccionando la adecuada matriz de diseño y ejecutando los correspondientes experimentos. Las respuestas obtenidas serán la base de datos para el ajuste.

La codificación de las variables expresadas en sus unidades originales de longitud, tiempo, temperatura, etc, se realiza aplicando la oportuna transformación lineal. Tras la codificación, las variables se transformarán en variables adimensionales y serán ya aplicables las distintas matrices de diseño.

Sean x_1', x_2', z_1' las variables con sus valores originales y x_1, x_2, z_1 las variables codificadas. La codificación vendrá dada por las transformaciones:

Variables de control:

$$x_1 = \frac{x_1^{'} - \mu}{(R/2)}; \quad x_2 = \frac{x_2^{'} - \mu}{(R/2)}$$

donde μ son los respectivos valores medios y R los rangos de cada variable expresados en valores originales. Los valores máximo y mínimo que definen el rango se corresponden con los valores codificados -1 y +1.

Variables de ruido:

$$z_1 = \frac{z_1^{'} - \mu}{\sigma_{z_1^{'}}}$$

donde μ es el valor medio y $\sigma_{z_1^{'}}$ la desviación típica de la variable de ruido sin codificar. De esta forma, los valores medios de las variables codificadas resultarán siempre nulos y las varianzas iguales a 1.

Al aplicar los operadores E(y) y V(y) presentados en el punto 1.2 al modelo de respuesta (3) y dado que el término de error es de la forma $N(0, \sigma^2)$ y que la única variable aleatoria es z_1 tendremos:

$$E(y) = \mu = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 \quad (5)$$

$$V(y) = \sigma_y^2 = (\gamma_1 + \delta_{11} x_1 + \delta_{21} x_2)^2 \sigma_{z_1}^2 + \sigma^2 \quad (6)$$

Al estar codificadas las variables de forma que $\sigma_{z_1}^2 = 1$ resulta que **los modelos del valor medio y la varianza de la respuesta sólo dependerán de las variables de control y son independientes de las de ruido.** El valor de σ^2 se deduce del modelo ajustado para la respuesta (3).

En definitiva, a partir del modelo de respuesta obtendremos una

respuesta dual (valor medio y varianza) que optimizaremos actuando, solamente, sobre las variables de control. La optimización se realiza situando, a la vez, el valor medio expresado por (5) dentro de los valores especificados y minimizando la varianza dada por (6). Los coeficientes que afectan a las ecuaciones (5) y (6) así como el valor de la varianza σ^2 se estiman ajustando previamente el modelo de respuesta dado por (3).

Si los operadores E(y), V(y) se hubieran aplicado al modelo de respuesta dado por (4) tendríamos:

$$E(y) = \mu = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 \tag{7}$$

$$V(y) = \sigma_y^2 = (\gamma_1 + \delta_{11} x_1 + \delta_{21} x_2)^2 \sigma_{z_1}^2 + \sigma^2 \tag{8}$$

Los resultados expresados en (3), (4), (5), (6), (7) y (8) pueden generalizarse a cualquier número de variables.

2. MATRICES PARA DISEÑOS ROBUSTOS

La particularidad del método de Superficie de Respuesta para el diseño de Procesos Robustos consiste en que **incluye en un solo diseño** tanto a las variables de control como a las variables de ruido. La matriz del diseño dependerá del tipo de modelo selecionado para la Superficie de Respuesta, basicamente, modelos polinomiales de primer y segundo orden con términos de interacción. Definida la matriz de diseño, se ejecutan los correspondientes experimentos y a los resultados obtenidos se les ajusta el modelo.

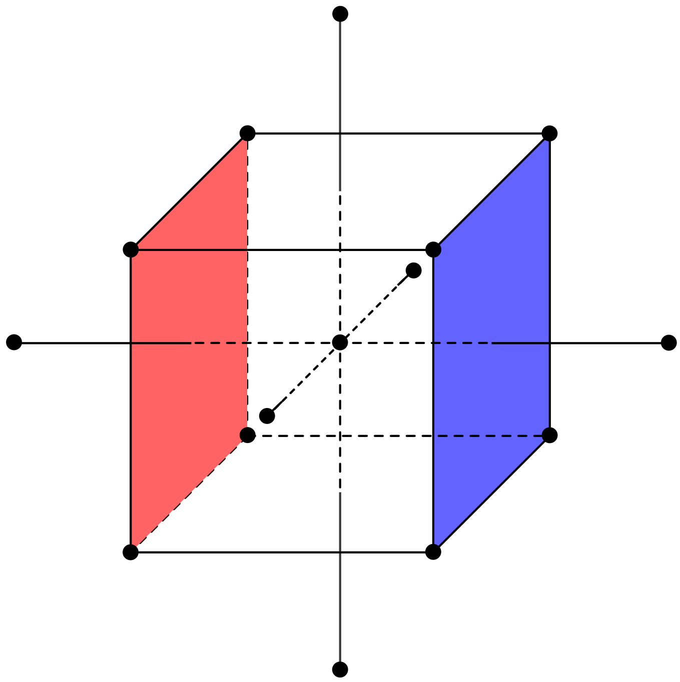
El número de filas en la matriz de diseño habrá de ser el suficiente para estimar todos los coeficientes del modelo y la varianza del error experimental. Las matrices utilizadas son siempre **ortogonales**, es decir, que dos columnas cualquiera de la matriz: a_{ij}, b_{ij} deben cumplir las condiciones: $\sum a_{ij} = 0$; $\sum b_{ij} = 0$, y $\sum\sum a_{ij} b_{ij} = 0$. El uso de matrices ortogonales minimiza la varianza de los coeficientes β_0, β_1, β_2 ... obtenidos por Análisis de Regresión.

Dependiendo del orden del modelo a ajustar, los diseños más utilizados son diseños factoriales del tipo 2^k, 3^k y diseños factoriales fraccionales del tipo 2^{k-p}, 3^{k-p}. Además, se han desarrollado diseños específicos que introducen en los diseños tradicionales puntos adicionales dando lugar a diseños con puntos centrales y axiales.

Presentamos a continuación, para el caso de k=3 factores, dos de estos diseños: Central compuesto y Box-Behnken. Los diseños factoriales completos y fraccionales ya han sido tratados con detalle en artículos anteriores.

2.1. Diseño Central Compuesto (DCC)

El DCC añade a los 8 puntos del diseño factorial 2^3, **varios puntos centrales y seís puntos axiales**. Su reprentación gráfica es:



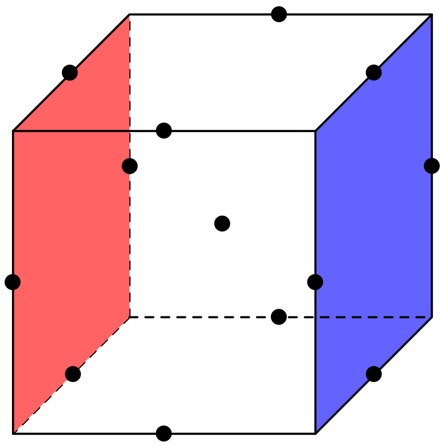
El diseño comprende un total de $14{+}n_c$ puntos (siendo n_c el número de puntos centrales). Los puntos axiales se situan a una distancia α del centro del diseño a fijar por el investigador. Dado que **un modelo de segundo grado** tal como el (4) tiene un total de 9 coeficientes a estimar, el CCD será suficiente para la estimación de todos lo coeficientes y de la varianza σ^2 del error experimental ε. El número de puntos centrales n_c, suele oscilar entre 3 y 5. La selección de n_c y de α aportan al investigador flexibilidad en los diseños. Con $\alpha = 1$, por ejemplo,tendríamos un diseño de caras centradas.

La matriz de diseño para 3 factores y $n_c = 3$ sería:

$$M = \begin{bmatrix} -1 & -1 & -1 \\ +1 & -1 & -1 \\ -1 & +1 & -1 \\ +1 & +1 & -1 \\ -1 & -1 & +1 \\ +1 & -1 & +1 \\ -1 & +1 & +1 \\ +1 & +1 & +1 \\ -\alpha & 0 & 0 \\ +\alpha & 0 & 0 \\ 0 & -\alpha & 0 \\ 0 & +\alpha & 0 \\ 0 & 0 & -\alpha \\ 0 & 0 & +\alpha \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

2.2. Diseño Box-Behnken (DBB)

El DBB tienen una representación geométrica bastante similar a una esfera en cuanto que la distancia de sus puntos al centro es la misma. El punto central suele ser triple. Su representación gráfica para k=3 variables es:



El diseño con $n_c = 3$ comprende un total de n=15 puntos capaces de sustentar, también, **un modelo de segundo grado** como el (4). La matriz de diseño con 3 puntos centrales sería:

$$M = \begin{bmatrix} -1 & -1 & 0 \\ +1 & -1 & 0 \\ -1 & +1 & 0 \\ +1 & +1 & 0 \\ -1 & 0 & -1 \\ +1 & 0 & -1 \\ -1 & 0 & +1 \\ +1 & 0 & +1 \\ 0 & -1 & -1 \\ 0 & +1 & -1 \\ 0 & -1 & +1 \\ 0 & +1 & +1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

El DBB es un diseño esférico, es decir, todos los puntos están a la misma distancia $r = \sqrt{2}$ del centro.

3. EJEMPLOS DE APLICACIÓN

Los ejemplos que vamos a exponer a continuación se desarrollan conforme al siguiente procedimiento:

a) **Selección del modelo**

Definidas las variables de control y de ruido se selecciona el modelo de ajuste para la **Superficie de Respuesta**. Los modelos, según lo **indicado en el punto 1.3**, serán modelos sencillos de primer y segundo orden que incluyan, necesariamente, términos de interacción entre las variables de control y de ruido. Los modelos serán de los tipos siguientes:

Modelo de primer orden en las variables de control

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + (\nu_1 + \delta_{11} x_1 + \delta_{21} x_2) z_1 + (\nu_2 + \delta_{12} x_1 + \delta_{22} x_2) z_2 + \varepsilon \quad (9)$$

Modelo de segundo orden en las variables de control

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + (\nu_1 + \delta_{11} x_1 + \delta_{21} x_2) z_1 + \varepsilon \quad (10)$$

b) **Selección de la matriz de diseño**

Para la selección de la matriz se procede de acuerdo con lo señalado en el punto 2. Para un modelo de primer orden como el (9) será suficiente una matriz de **diseño factorial a dos niveles** como el diseño 2^4 que aplicaremos en el ejemplo 3.1. En el caso de un modelo de segundo orden como el (10), utilizaremos un **diseño Central Compuesto** como el que aplicaremos en el ejemplo 3.2.

c) **Ajuste del modelo**

A los valores obtenidos para la variable de respuesta se les ajustan los modelos de superficie de Respuesta mediante el método de regresión por mínimos cuadrados.

d) **Optimización**

En la parte final de los ejemplos se optimiza la respuesta dada por expresiones similares para el valor medio (μ) y la varianza (σ_y^2) a las indicadas en las ecuaciones (5) y (6) para el modelo de primer orden (ejemplo 3.1) y a las ecuaciones (7) y (8) del modelo de segundo orden (ejemplo 3.2).

Tengamos en cuenta que la expresión general de la varianza es de la forma:

$$\sigma_y^2 = \sum (\nu_1 + \delta_{11} x_1 + \delta_{21} x_2)^2 \sigma_{z_j}^2 + \sigma^2$$

siendo necesario que los coeficientes "δ" correspondientes a los término de interacción $x_i z_j$ de la superficie de respuesta, sean significativos. La optimización simultánea de μ y σ_y^2 se realiza por superposición de las líneas de nivel respectivas.

3.1. Ejemplo con una matriz de diseño factorial 2^4

Definición de variables y objetivo

En un proceso industrial con 4 variables independientes, dos de ellas son variables de control: $x_1^{'}$ (temperatura) y $x_2^{'}$ (tiempo). Las otras dos son variables de ruido: $z_1^{'}$ (reflectividad) y $z_2^{'}$ (densidad), aleatorias y dependientes de las características de la materia prima utilizada. La variable dependiente "y" es la dureza del producto obtenido siendo el objetivo definir la región de variación de las variables de control para conseguir, con independencia de los valores aleatorios de las variables de ruido, alcanzar un **indice de dureza** ≥ 80 **con una varianza** ≤ 20.

Las variables independientes consideradas son variables continuas y tienen los siguientes rangos de variación:

$x_1^{'}$: 1100 a 1300ºC, con un valor medio de 1200ºC.

$x_2^{'}$: 18 a 22 horas con un valor medio de 20 horas.

Las variables de ruido, con rangos de variación $1{,}1 < z_1^{'} < 1{,}3$ y $600 < z_2^{'} < 800\frac{kg}{m^3}$, se consideran conformes con distribuciones normales:

$$z_1^{'} : N(\mu_{z_1^{'}}; \sigma_{z_1^{'}}) = N(1{,}2; 0{,}1) \text{ con } \sigma_{z_1^{'}} = 0{,}1.$$

$$z_2^{'} : N(\mu_{z_2^{'}}; \sigma_{z_2^{'}}) = N(700; 100) \text{ con } \sigma_{z_2^{'}} = 100\frac{kg}{m^3}$$

Las variables originales **se codifican** según las siguientes transformaciones:

$$x_1 = \frac{x_1^{'} - 1200}{100};\ x_2 = \frac{x_2^{'} - 200}{2};\ z_1 = \frac{z_1^{'} - 1,2}{0,1} \text{ y } z_2 = \frac{z_2^{'} - 700}{100}.$$

Con ello, resultan las siguientes equivalencias entre las variables independientes antes y después de la codificación:

Variables originales				Variables codificadas			
X_1'	X_2'	Z_1'	Z_2'	X_1	X_2	Z_1	Z_2
1100	18	1,1	600	-1	-1	-1	-1
1300	22	1,3	800	+1	+1	+1	+1

Modelo y matriz de Diseño

El modelo a ajustar será un modelo como el (9):

$$y = \beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{12}x_1x_2+(\nu_1+\delta_{11}x_1+\delta_{21}x_2)z_1+(\nu_2+\delta_{12}x_1+\delta_{22}x_2)z_2+\varepsilon$$

La matriz de diseño es la correspondiente a un diseño factorial completo a dos niveles del tipo 2^4 consistente en 16 tratamientos:

X_1	X_2	Z_1	Z_2
-1	-1	-1	-1
+1	-1	-1	-1
-1	+1	-1	-1
+1	+1	-1	-1
-1	-1	+1	-1
+1	-1	+1	-1
-1	+1	+1	-1
+1	+1	+1	-1
-1	-1	-1	+1
+1	-1	-1	+1
-1	+1	-1	+1
+1	+1	-1	+1
-1	-1	+1	+1
+1	-1	+1	+1
-1	+1	+1	+1
+1	+1	+1	+1

Los resultados obtenidos junto a la matriz de diseño se recogen en la página 15. En las pags 12 y 13 se realiza la discusión de resultados y en las pags. 16 a 21 se presentan los cálculos y gráficos efectuados utilizando el programa estadístico STATGRAPHICS.

Discusión de Resultados

- El modelo final de Superficie de Respuesta, ajustado según el método de regresión por mínimos cuadrados, se recoge al final de la pag 16 siendo su ecuación:

$$y = 81{,}06 + 1{,}68x_1 + 3{,}33x_2 + (3{,}82 + 4{,}82x_1 + 4{,}74x_2)z_1 + (4{,}44 + 2{,}52x_1 + 6{,}04x_2)z_2 + \varepsilon$$

En la tabla ANOVA de la citada página 16 se observa que todos los coeficientes incluídos son significativos con valores p <0.05 y que el modelo ajustado explica el 98.76 % de la variabilidad total. La varianza del término de error ε es de $\sigma^2 = 3{,}941$. El modelo final carece del término de interacción $\beta_{12}x_1x_2$ entre las variables de control ya que en un modelo inicialmente ajustado β_{12} no era significativo (pag 17).

- Los gráficos de interacción del final de la pag 18 muestran la clara interacción existente entre las variables de control y las variables de ruido que hacen, en definitiva, posible la aplicación del método.

- Al aplicar los operadores E(y) y V(y) al modelo de Superficie de Respuesta resulta:

  $E(y) = 81{,}06 + 1{,}68x_1 + 3{,}33x_2$

  $V(y) = (3{,}82 + 4{,}82x_1 + 4{,}74x_2)^2\sigma^2_{z_1} + (4{,}44 + 2{,}52x_1 + 6{,}04x_2)^2\sigma^2_{z_2} + \sigma^2$

  Teniendo en cuenta que las variables de ruido, en sus valores originales sin codificar, tienen por varianzas $\sigma^2_{z_i'} = 0{,}1^2$ y $\sigma^2_{z_2'} = 100^2$, las varianzas de las variables codificadas serán:

  $\sigma^2_{z_1} = \dfrac{0{,}1^2}{0{,}1^2} = 1$; $\sigma^2_{z_2} = \dfrac{100^2}{100^2} = 1$

  Por otra parte, $\sigma^2 = 3{,}941$. Sustituyendo los valores de las citadas varianzas en la expresión de V(y) resulta:

  $V(y) = \sigma^2_y = (3{,}82 + 4{,}82x_1 + 4{,}74x_2)^2 + (4{,}44 + 2{,}52x_1 + 6{,}04x_2)^2 + 3{,}941$ y finalmente.

  $V(y) = \sigma^2_y = 38{,}25 + 59{,}20x_1 + 89{,}85x_2 + 76{,}14x_1x_2 + 29{,}58x_1^2 + 58{,}95x_2^2$

- La pag 19 incluye los gráficos tridimensionales y de lineas de nivel para el valor medio $E(y) = 81{,}06 + 1{,}68x_1 + 3{,}33x_2$

- La pag 20 contiene el gráfico tridimensional y el de lineas de nivel de la varianza

  $V(y) = \sigma^2_y = 38{,}25 + 59{,}20x_1 + 89{,}85x_2 + 76{,}14x_1x_2 + 29{,}58x_1^2 + 58{,}95x_2^2$

- Finalmente, en la pag 21 se presentan los gráficos de lineas de nivel del valor medio E(y) y de la varianza V(y) **cuyas respuestas hay que optimizar** de forma simultánea.

- **Dado que el objetivo es obtener un valor medio $\mu = E(y) \geq 80$ con una varianza** $\sigma^2_y \leq 20$, se infiere **que la región acotada en el gráfico final de la pag 21 será la zona óptima** al ser $\mu \geq 80$ y situarse, simultaneamente, la varianza σ_y^2 en valores ≤ 20. Por ejemplo, a un punto P de esta región con $x_1 = -0{,}4$ y $x_2 = 0$ le corresponderá un valor medio $\mu = 80{,}4$ y una varianza $\sigma_y^2 = 19{,}3$. Al citado punto P, en sus unidades originales, le corresponden unos valores $x_1^{'} = 1160^{\underline{o}}\mathrm{C}$ y $x_2^{'} = 20h$. para las variables inepencientes de control $x_1^{'}$ (temperatura) y $x_2^{'}$ (tiempo). Estos valores se deducen de las ecuaciones de transformación puestas en forma inversa:

$$\begin{cases} x_1^{'} & = 1200 + 100 \cdot x_1 = 1160^{\underline{o}}\mathrm{C} \\ x_2^{'} & 20 + 2 \cdot x_2 = 20h \end{cases}$$

CONCLUSIÓN

Mediante el procedimiento de Superficie de Respuesa se ha acotado una **zona óptima de funcionamiento dependiente, unicamente, de las variables de control tiempo y temperatura.** En la zona óptima, la respuesta se caracteriza por tener un valor medio $\mu \geq 80$ y una varianza $\sigma_y^2 \leq 20$. Se ha conseguido, en definitiva, un diseño robusto frente a las variables de ruido reflectividad y densidad de carga.

Ejemplo 3.1 PROCESOS ROBUSTOS

	BLOCK	X1	X2	Z1	Z2	Y
1	1	-1	-1	-1	-1	84
2	1	1	-1	-1	-1	74.6
3	1	-1	1	-1	-1	73.1
4	1	1	1	-1	-1	59.2
5	1	-1	-1	1	-1	74.7
6	1	1	-1	1	-1	84
7	1	-1	1	1	-1	78
8	1	1	1	1	-1	85.3
9	1	-1	-1	-1	1	79.9
10	1	1	-1	-1	1	76.1
11	1	-1	1	-1	1	84.5
12	1	1	1	-1	1	86.5
13	1	-1	-1	1	1	65.2
14	1	1	-1	1	1	83.3
15	1	-1	1	1	1	95.6
16	1	1	1	1	1	112.9

MODELO FINAL

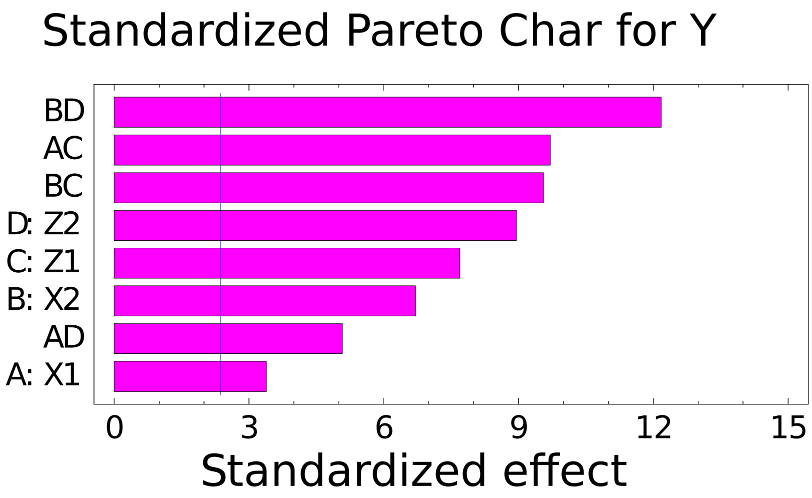
```
Analysis Summary
----------------
Ejemplo 3.1 PROCESOS ROBUSTOS

Estimated effects for Y
----------------------------------------------------------------------
average = 81.0562 +/- 0.49632
A:X1    = 3.3625  +/- 0.99264
B:X2    = 6.6625  +/- 0.99264
C:Z1    = 7.6375  +/- 0.99264
D:Z2    = 8.8875  +/- 0.99264
AC      = 9.6375  +/- 0.9926
AD      = 5.0375  +/- 0.99264
BC      = 9.4875  +/- 0.99264
BD      = 12.0875 +/- 0.99264
----------------------------------------------------------------------
Standard errors are based on total error with 7 d.f.
```

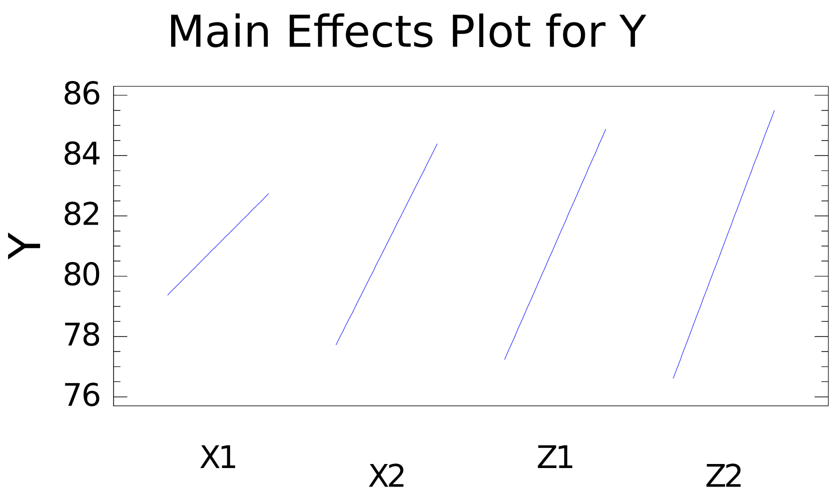
Analysis of Variance for Y

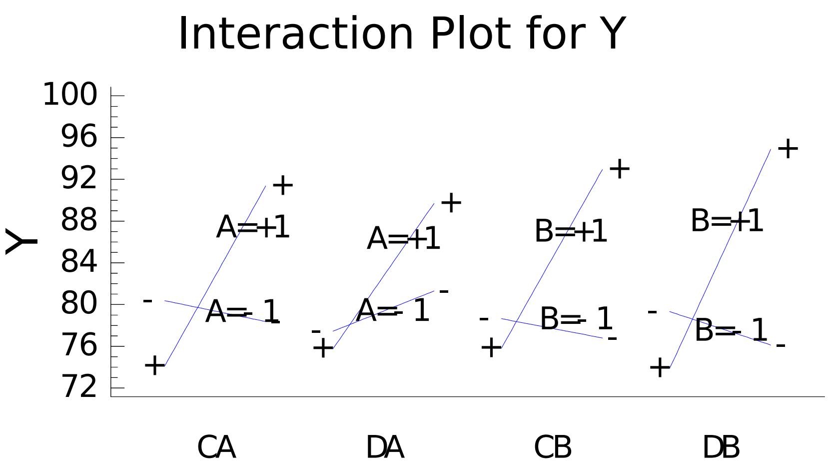
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
A:X1	45.2256	1	45.2256	11.47	0.0116
B:X2	177.556	1	177.556	45.05	0.0003
C:Z1	233.326	1	233.326	59.20	0.0001
D:Z2	315.951	1	315.951	80.16	0.0000
AC	371.526	1	371.526	94.26	0.0000
AD	101.506	1	101.506	25.75	0.0014
BC	360.051	1	360.051	91.35	0.0000
BD	584.431	1	584.431	148.28	0.0000
Total error	27.5894	7	3.94134		
Total (corr.)	2217.16	15			

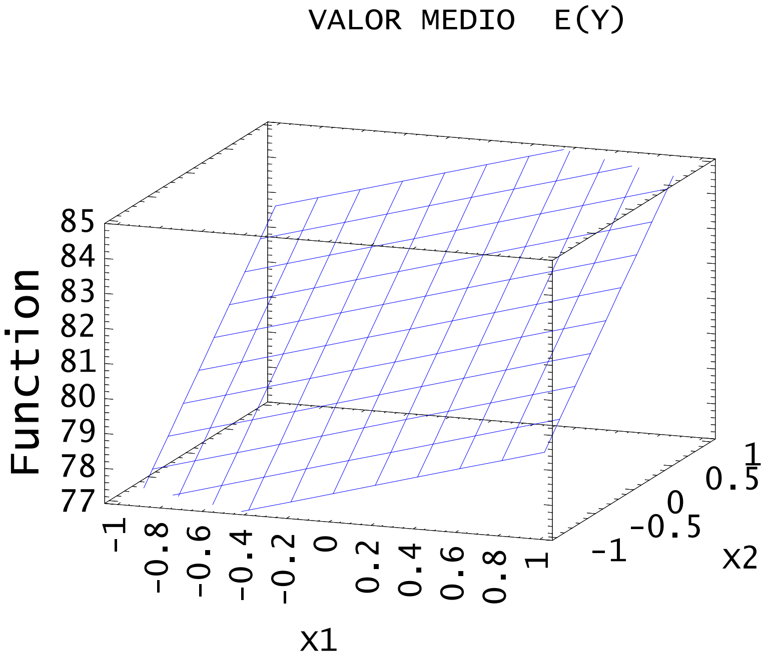
```
R-squared = 98.7556 percent
R-squared (adjusted for d.f.) = 97.3335 percent
Standard Error of Est. = 1.98528
Mean absolute error = 1.08437
Durbin-Watson statistic = 2.28289 (P=0.2409)
Lag 1 residual autocorrelation = -0.210867



Regression coeffs. for Y
----------------------------------------------------------------------
constant  = 81.0562
A:X1      = 1.68125
B:X2      = 3.33125
C:Z1      = 3.81875
D:Z2      = 4.44375
AC        = 4.81875
AD        = 2.51875
BC        = 4.74375
BD        = 6.04375
----------------------------------------------------------------------


The StatAdvisor
---------------
   This pane displays the regression equation which has been fitted to
the data.  The equation of the fitted model is

Y = 81.0562 + 1.68125*X1 + 3.33125*X2 + 3.81875*Z1 + 4.44375*Z2 +
4.81875 X1 Z1 + 2.51875 X1 Z2 + 4.74375 X2 Z1 + 6.04375 X2 Z2
```

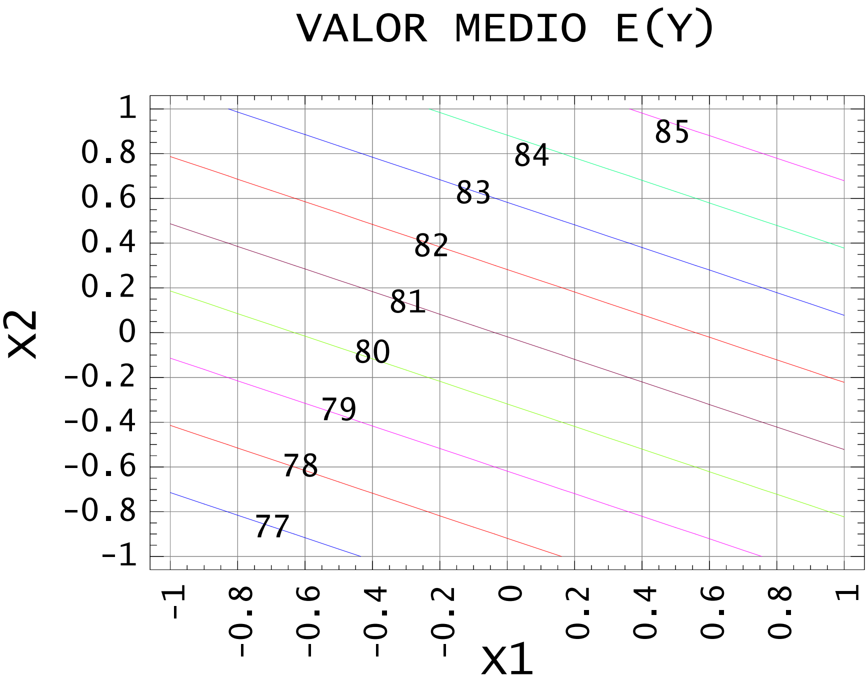
```
Analysis Summary      MODELO INICIAL
----------------
Ejemplo 3.1 PROCESOS ROBUSTOS.

Estimated effects for Y
----------------------------------------------------------------------
average = 81.0562 +/- 0.534719
A:X1    = 3.3625  +/- 1.06944
B:X2    = 6.6625  +/- 1.06944
C:Z1    = 7.6375  +/- 1.06944
D:Z2    = 8.8875  +/- 1.06944
AB      = -0.1875 +/- 1.06944
AC      = 9.6375  +/- 1.06944
AD      = 5.0375  +/- 1.06944
BC      = 9.4875  +/- 1.06944
BD      = 12.0875 +/- 1.06944
----------------------------------------------------------------------
Standard errors are based on total error with 6 d.f.
Analysis of Variance for Y
--------------------------------------------------------------------------------
Source                 Sum of Squares     Df    Mean Square    F-Ratio    P-Value
--------------------------------------------------------------------------------
A:X1                          45.2256      1        45.2256       9.89     0.0200
B:X2                          177.556      1        177.556      38.81     0.0008
C:Z1                          233.326      1        233.326      51.00     0.0004
D:Z2                          315.951      1        315.951      69.06     0.0002
AB                           0.140625      1       0.140625       0.03     0.8666
AC                            371.526      1        371.526      81.21     0.0001
AD                            101.506      1        101.506      22.19     0.0033
BC                            360.051      1        360.051      78.70     0.0001
BD                            584.431      1        584.431     127.75     0.0000
Total error                   27.4488      6        4.57479
--------------------------------------------------------------------------------
Total (corr.)                 2217.16     15

R-squared = 98.762 percent
R-squared (adjusted for d.f.) = 96.905 percent
Standard Error of Est. = 2.13888
Mean absolute error = 1.09531
Durbin-Watson statistic = 2.28434 (P=0.2642)
Lag 1 residual autocorrelation = -0.207273
Regression coeffs. for Y
----------------------------------------------------------------------
constant  = 81.0562
A:X1      = 1.68125
B:X2      = 3.33125
C:Z1      = 3.81875
D:Z2      = 4.44375
AB        = -0.09375
AC        = 4.81875
AD        = 2.51875
BC        = 4.74375
BD        = 6.04375
----------------------------------------------------------------------


The StatAdvisor
---------------
   This pane displays the regression equation which has been fitted to
the data.  The equation of the fitted model is

Y = 81.0562 + 1.68125*X1 + 3.33125*X2 + 3.81875*Z1 + 4.44375*Z2 -
0.09375*X1*X2 + 4.81875*X1*Z1 + 2.51875*X1*Z2 + 4.74375*X2*Z1 +
6.04375*X2*Z2
```

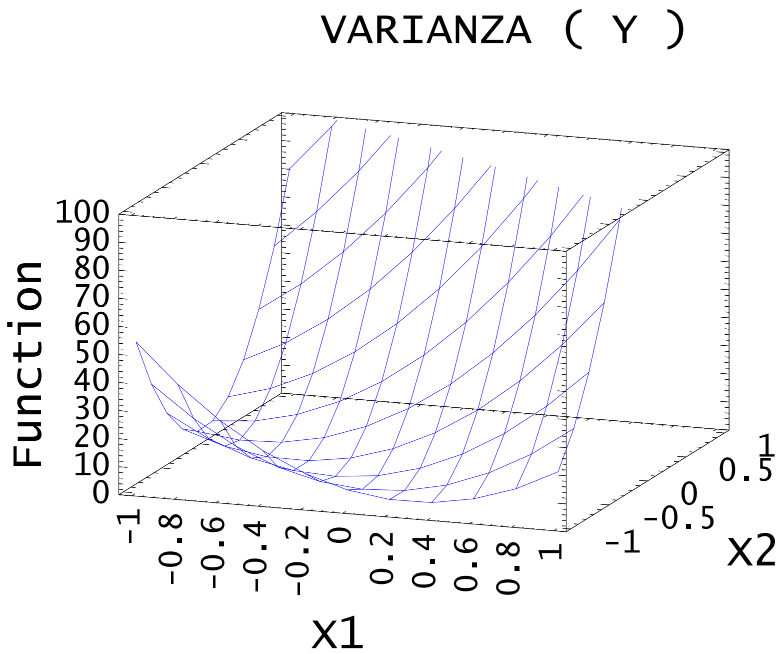

Standardized Pareto Char for Y
BD
AC
BC
D: Z2
C: Z1
B: X2
AD
A: X1
0
3
6
9
12
15
Standardized effect

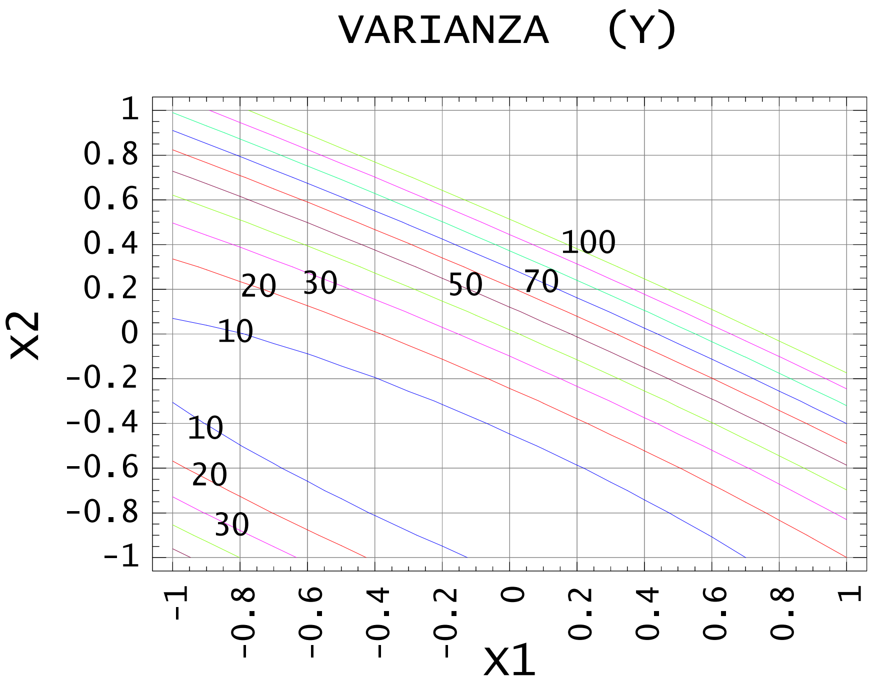

Main Effects Plot for Y
86
84
82
80
78
76
Y
X1
X2
Z1
Z2

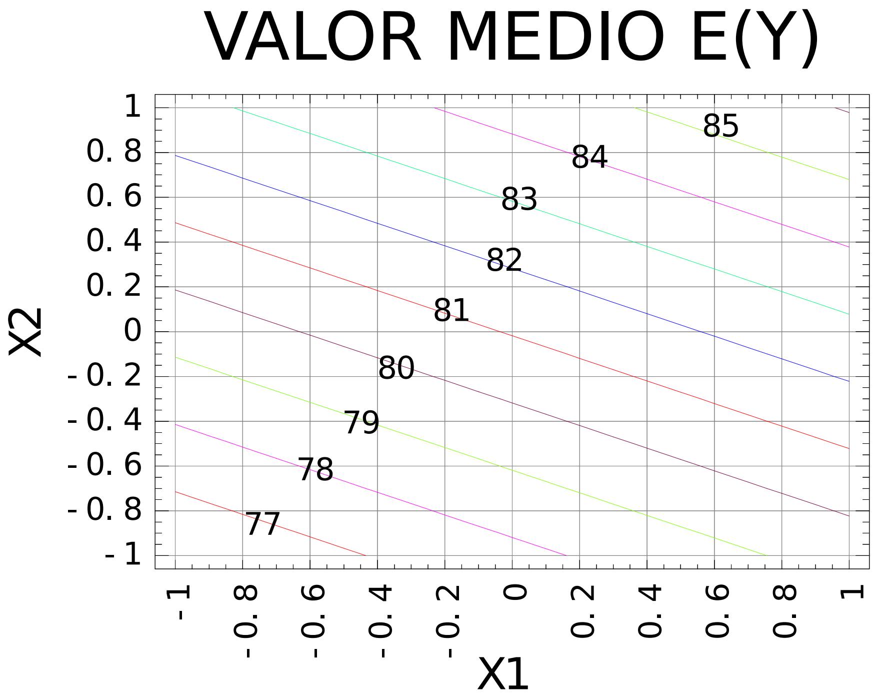

Interaction Plot for Y
100
96
92
88
84
80
76
72
Y
A=+1
A=-1
B=+1
B=-1
CA
DA
CB
DB

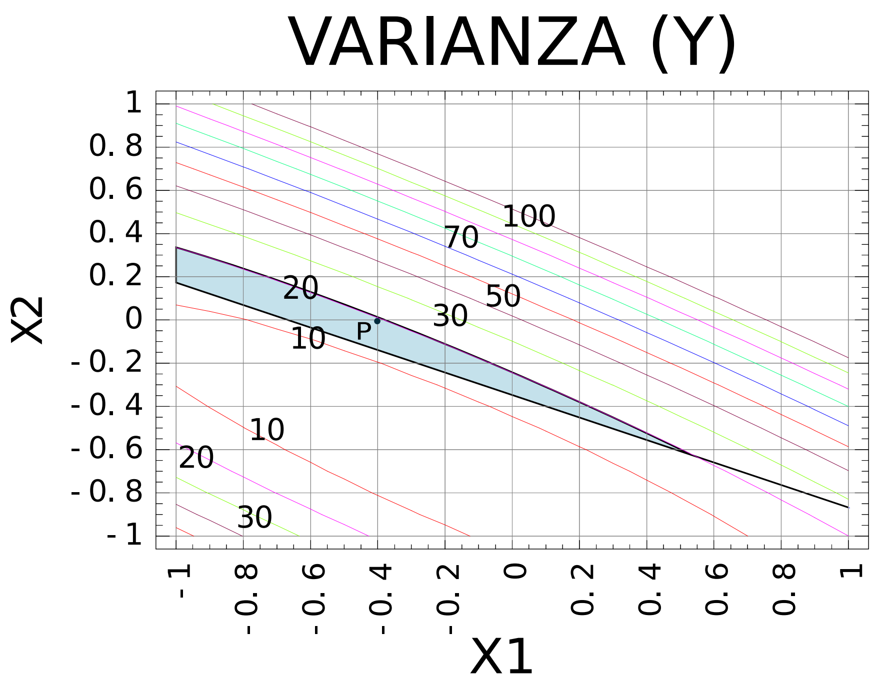

VALOR MEDIO E(Y)
Function
85
84
83
82
81
80
79
78
77
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
x1
1
0.5
0
-0.5
-1
x2


VALOR MEDIO E(Y)
x2
x1
77
78
79
80
81
82
83
84
85
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1


VARIANZA (Y)
Function
100
90
80
70
60
50
40
30
20
10
0
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
X1
-1
-0.5
0
0.5
1
X2


VARIANZA (Y)
X2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
X1
100
20
30
50
70
10
10
20
30


VALOR MEDIO E(Y)
X2
X1
1
0. 8
0. 6
0. 4
0. 2
0
- 0. 2
- 0. 4
- 0. 6
- 0. 8
- 1
85
84
83
82
81
80
79
78
77


VARIANZA (Y)
X2
X1
100
70
50
30
20
10
P
10
20
30

3.2. Ejemplo con una matriz de Diseño Central Compuesto

Objetivos y definición de variables

Se desea mejorar un proceso químico en cuanto al rendimiento y variabilidad del producto final obtenido. Para ello se realiza un estudio en el laboratorio para diseñar un Proceso Robusto y encontrar la región de variación de las dos variables de control: $x_1^{'}$ (temperatura) y $x_2^{'}$ (tiempo) que, con independencia de los valores de la variable de ruido $z_1^{'}$ (pureza de la materia prima), permita obtener un rendimiento "y" del producto final suficiente y con la mínima variabilidad posible.

Los rangos de variación de las variables independientes son:

$600 < x_1^{'} < 800$ºC con un valor medio de 700ºC

$200 < x_2^{'} < 300$ min con un valor medio de 250 min.

La variable de ruido con rango de variación $50 < z_i' < 60$ se supone que sigue una distribución normal $N(\mu_{z_1'}; \sigma_{z_1'}) = N(55; 5)$ con $\sigma_{z_1'} = 5$.

Las variables originales se codifican mediante las transformaciones:

$$x_1 = \frac{x_1^{'} - 700}{100}; x_2 = \frac{x_2^{'} - 250}{50} \text{ y } z_1 = \frac{z_1^{'} - 55}{5}$$

De esta forma se obtienen las siguientes equivalencias:

Variables originales			Variables codificadas		
$X_1^{'}$	$X_2^{'}$	$Z_1^{'}$	X_1	Z_1	Z_1
600	200	50	-1	-1	-1
700	250	55	0	0	0
800	300	60	+1	+1	+1

Modelo y matriz de Diseño

En este caso, el modelo a ajustar será un modelo de segundo orden en las variables de control con ecuación como la (10):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + (\nu_1 + \delta_{11} x_1 + \delta_{21} x_2) z_1 + \varepsilon$$

La matriz de diseño es la del **Diseño Central Compuesto** presentado en la pag 7 con 3 puntos centrales ($n_c = 3$), de caras centradas ($\alpha = 1$) y un total de $14 + n_c = 17$ puntos:

$$M = \begin{bmatrix} \mathbf{x_1} & \mathbf{x_2} & \mathbf{z_1} \\ -1 & -1 & -1 \\ +1 & -1 & -1 \\ -1 & +1 & -1 \\ +1 & +1 & -1 \\ -1 & -1 & +1 \\ +1 & -1 & +1 \\ -1 & +1 & +1 \\ +1 & +1 & +1 \\ -1 & 0 & 0 \\ +1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & +1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Los resultados junto con la matriz de diseño se recogen en la pag 26. En las pags 23 y 24 se presenta la discusión de resultados y en las 27 a 32 los cálculos y gráficos efectuados con el programa estadístico STATGRAPHICS.

Discusión de Resultados

- Tras ajustar un modelo inicial (pag 27) constituido por todos los términos de segundo orden e interacciones de dos variables que pueden formarse con las tres variables independientes, se excluyen los términos con coeficientes no significativos obteniéndose el modelo final de la pag 28 con la siguiente ecuación para la Superficie de Respuesta:

  $y = 39{,}61 + 3{,}13 x_1 + 4{,}29 x_2 + 5{,}13 x_1^2 + 3{,}63 x_2^2 + (4{,}52 + 6{,}61 x_1 + 5{,}14 x_2) z_1 + \varepsilon$

- La tabla ANOVA (pag 28) muestra que todos los coeficiente estimados son significativos con valores "p" muy próximos a cero. El modelo ajustado explica el 98.32 % de la variabilidad total con una varianza del error ε de $\sigma^2 = 2{,}45$. El modelo final carece del término de interacción $\beta_{12}x_1x_2$ ya que en el modelo inicial β_{12} no era significativo (pag 27). El diseño consta de un total de 17 experimentos para la estima de los 8 coeficientes del modelo. Los 9 grados de libertad restantes se aplican a la estimación del error experimental.

- Los gráficos de interacción del final de la pag 29 muestran la acusada interacción existente entre las variables de control x_1, x_2 con la variable de ruido z_1 lo que posibilita la aplicación del método.

- Aplicando los operadores E(y) y V(y) al modelo de superficie de respuesta tenemos:

  $E(y) = 39{,}75 + 3{,}13x_1 + 4{,}29x_2 + 5{,}13x_1^2 + 3{,}63x_2^2$

  $V(y) = (4{,}52 + 6{,}61x_1 + 5{,}14x_2)^2\sigma_{z_1}^2 + \sigma^2$

  Dado que $\sigma_{z_1}^2 = 1$ y $\sigma^2 = 2{,}45$ la expresión final para V(y) es:

  V(y) $= \sigma_y^2 = 22{,}88 + 59{,}75x_1 + 46{,}47x_2 + 67{,}95x_1x_2 + 43{,}69x_1^2 + 26{,}42x_2^2$

  Nótese que $\sigma_{z_1}^2 = 1$ porque $\sigma_{z_1'}^2 = 5^2$ y la transformación que liga a z_1 y z_1' es $z_1 = \dfrac{z_1' - 55}{5}$. Por tanto, $\sigma_{z_1}^2 = \dfrac{\sigma_{z_1'}^2}{5^2} = \dfrac{5^2}{5^2} = 1$

- La pag 30 incluye el gráfico tridimensional y las líneas de nivel del valor medio $E(y) = 39{,}75 + 3{,}13x_1 + 4{,}29x_2 + 5{,}13x_1^2 + 3{,}63x_2^2$

- La pag 31 muestra el gráfico tridimensional y el de líneas de nivel de la varianza

  $V(y) = \sigma_y^2 = 22{,}88 + 59{,}75x_1 + 46{,}47x_2 + 67{,}95x_1x_2 + 43{,}69x_1^2 + 26{,}42x_2^2$

- En la pag 32 se presentan los gráficos de líneas de nivel del valor medio E(y) y de la varianza V(y) cuyos valores deben ser, simultaneamente, optimizados. **El objetivo se sitúa en la obtención de un valor medio $\mu = E(y) \geq 46$ con la mínima variabilidad posible ($\sigma^2 \leq 20$).**

- Los puntos situados en la **región acotada** de los gráficos superior e inferior de la pag 32, cumplirán el objetivo al ser el valor medio $\mu \geq 46$ y la varianza $\sigma_y^2 \leq 20$. Por ejemplo, al punto P con $x_1 = -0{,}8y$ $x_2 = 0{,}9$ le corresponde un valor medio de $\mu = 47{,}33$ y una varianza $\sigma_y^2 = 17{,}35$.

En sus unidades originales, el punto P tendrá unos valores $x_1^{'} = 620^{\underline{o}}\text{C}$ y $x_2^{'} = 295$ min. para las variables de control temperatura y tiempo. Estos valores se deducen de las ecuaciones de transformación.

$$\begin{cases} x_1^{'} & = 700 + 100 \cdot x_1 = 620^{\underline{o}}\text{C} \\ x_2^{'} & 250 + 50 \cdot x_2 = 295 \text{ min.} \end{cases}$$

CONCLUSIÓN

Se ha conseguido acotar una región óptima de funcionamiento dependiente solamente de las variables de control temperatura y tiempo y robusta frente a la influencia de la variable de ruido dependiente de la materia prima.

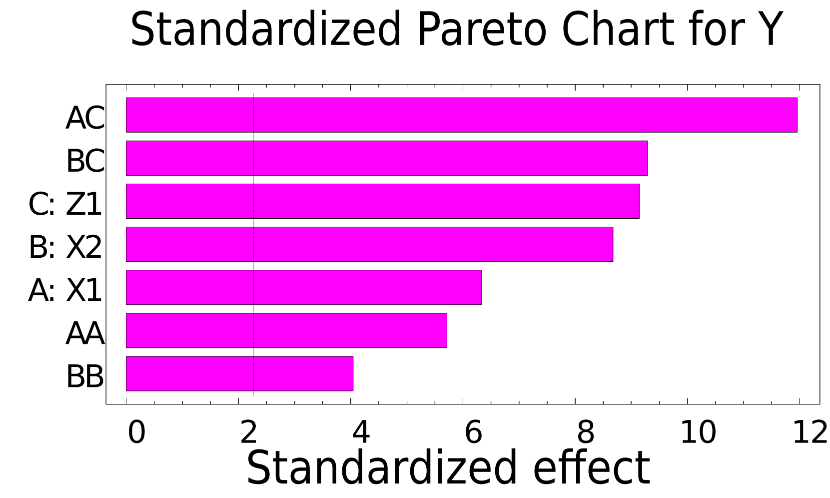
	BLOCK	X1	X2	Z1	Y
1	1	-1	-1	-1	48.9
2	1	1	-1	-1	41.7
3	1	-1	1	-1	47.5
4	1	1	1	-1	38.9
5	1	-1	-1	1	35.3
6	1	1	-1	1	52.6
7	1	-1	1	1	52.5
8	1	1	1	1	72.3
9	1	-1	0	0	39.1
10	1	1	0	0	49.1
11	1	0	-1	0	37.5
12	1	0	1	0	47.7
13	1	0	0	-1	35.2
14	1	0	0	1	44.7
15	1	0	0	0	38.3
16	1	0	0	0	40.3
17	1	0	0	0	41.8

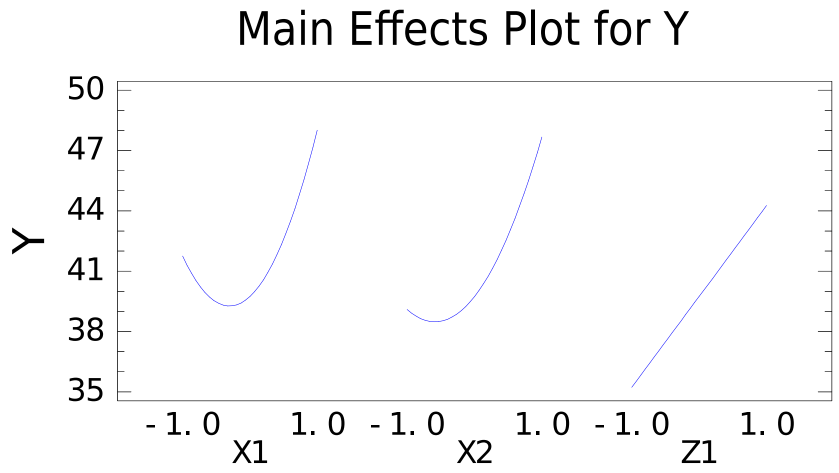
```
Analysis Summary
----------------
```

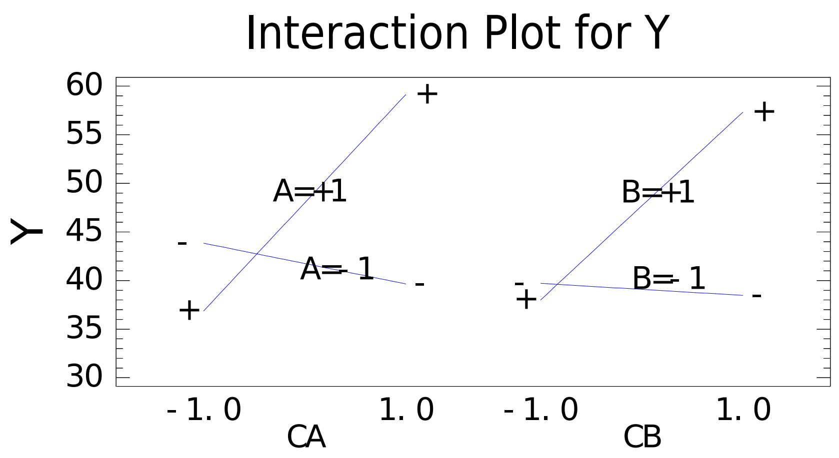
MODELO INICIAL

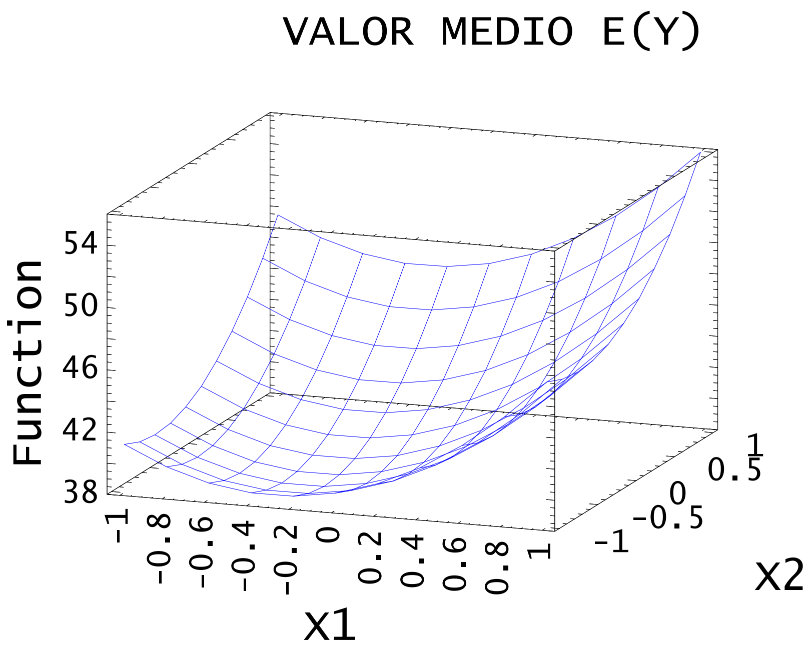
```
Estimated effects for Y
----------------------------------------------------------------------
average = 39.6085 +/- 0.730775
A:X1    = 6.26    +/- 1.08012
B:X2    = 8.58    +/- 1.08012
C:Z1    = 9.04    +/- 1.08012
AA      = 9.77042 +/- 2.08672
AB      = 0.275   +/- 1.20761
AC      = 13.225  +/- 1.20761
BB      = 6.77042 +/- 2.08672
BC      = 10.275  +/- 1.20761
CC      = 1.47042 +/- 2.08672
----------------------------------------------------------------------
Standard errors are based on total error with 7 d.f.
Analysis of Variance for Y
--------------------------------------------------------------------------------
Source                    Sum of Squares     Df    Mean Square    F-Ratio    P-Value
--------------------------------------------------------------------------------
A:X1                              97.969      1         97.969      33.59     0.0007
B:X2                             184.041      1        184.041      63.10     0.0001
C:Z1                             204.304      1        204.304      70.05     0.0001
AA                                63.941      1         63.941      21.92     0.0023
AB                               0.15125      1        0.15125       0.05     0.8264
AC                               349.801      1        349.801     119.93     0.0000
BB                               30.7032      1        30.7032      10.53     0.0142
BC                               211.151      1        211.151      72.40     0.0001
CC                               1.44823      1        1.44823       0.50     0.5038
Total error                      20.4165      7        2.91664
--------------------------------------------------------------------------------
Total (corr.)                    1311.31     16

R-squared = 98.443 percent
R-squared (adjusted for d.f.) = 96.4412 percent
Standard Error of Est. = 1.70782
Mean absolute error = 0.894051
Durbin-Watson statistic = 2.44178 (P=0.0765)
Lag 1 residual autocorrelation = -0.341318
Regression coeffs. for Y
----------------------------------------------------------------------
constant   = 39.6085
A:X1       = 3.13
B:X2       = 4.29
C:Z1       = 4.52
AA         = 4.88521
AB         = 0.1375
AC         = 6.6125
BB         = 3.38521
BC         = 5.1375
CC         = 0.735211
----------------------------------------------------------------------


The StatAdvisor
---------------
   This pane displays the regression equation which has been fitted to
the data.  The equation of the fitted model is

Y = 39.6085 + 3.13*X1 + 4.29*X2 + 4.52*Z1 + 4.88521*X1^2 +
0.1375*X1*X2 + 6.6125*X1*Z1 + 3.38521*X2^2 + 5.1375*X2*Z1 +
0.735211*Z1^2
```

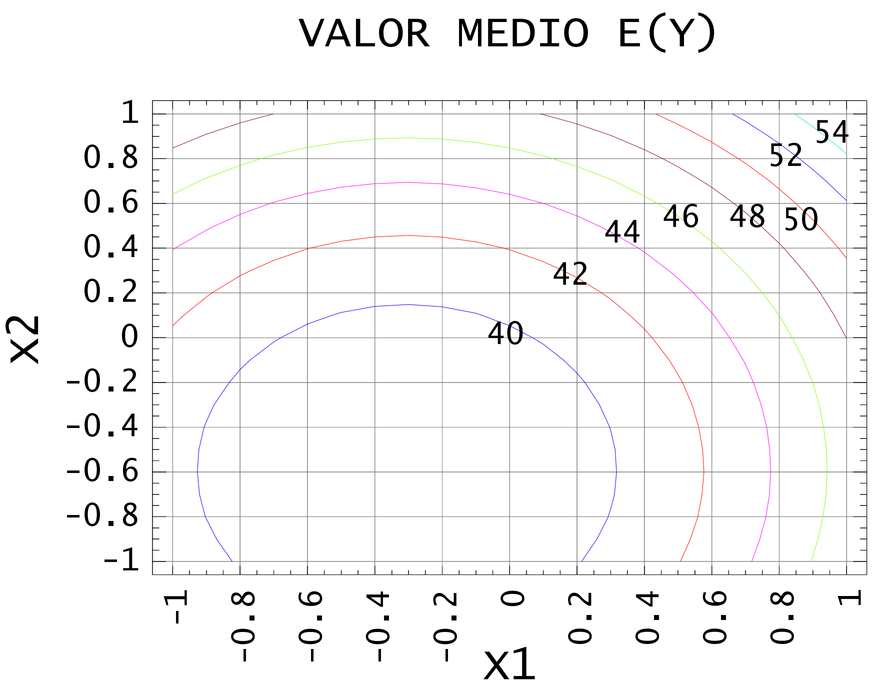
```
Analysis Summary     MODELO FINAL
----------------

Estimated effects for Y
----------------------------------------------------------------------
average = 39.7472 +/- 0.644512
A:X1    = 6.26    +/- 0.989185
B:X2    = 8.58    +/- 0.989185
C:Z1    = 9.04    +/- 0.989185
AA      = 10.2698 +/- 1.79746
AC      = 13.225  +/- 1.10594
BB      = 7.26981 +/- 1.79746
BC      = 10.275  +/- 1.10594
----------------------------------------------------------------------
Standard errors are based on total error with 9 d.f.
Analysis of Variance for Y
--------------------------------------------------------------------------------
Source                 Sum of Squares     Df    Mean Square    F-Ratio    P-Value
--------------------------------------------------------------------------------
A:X1                           97.969      1         97.969      40.05     0.0001
B:X2                          184.041      1        184.041      75.23     0.0000
C:Z1                          204.304      1        204.304      83.52     0.0000
AA                            79.8551      1        79.8551      32.64     0.0003
AC                            349.801      1        349.801     143.00     0.0000
BB                            40.0151      1        40.0151      16.36     0.0029
BC                            211.151      1        211.151      86.32     0.0000
Total error                    22.016      9        2.44622
--------------------------------------------------------------------------------
Total (corr.)                 1311.31     16

R-squared = 98.3211 percent
R-squared (adjusted for d.f.) = 97.0152 percent
Standard Error of Est. = 1.56404
Mean absolute error = 0.87758
Durbin-Watson statistic = 2.46788 (P=0.0938)
Lag 1 residual autocorrelation = -0.337102
Regression coeffs. for Y
----------------------------------------------------------------------
constant  = 39.7472
A:X1      = 3.13
B:X2      = 4.29
C:Z1      = 4.52
AA        = 5.13491
AC        = 6.6125
BB        = 3.63491
BC        = 5.1375
----------------------------------------------------------------------


The StatAdvisor
---------------
   This pane displays the regression equation which has been fitted to
the data.  The equation of the fitted model is

Y = 39.7472 + 3.13*X1 + 4.29*X2 + 4.52*Z1 + 5.13491*X1^2 +
6.6125*X1*Z1 + 3.63491*X2^2 + 5.1375*X2*Z1
```

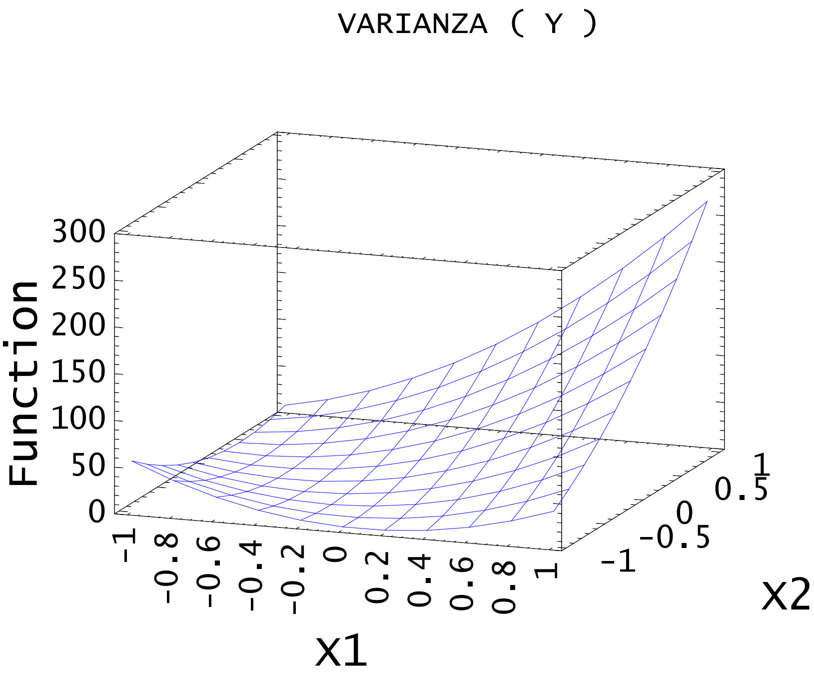

Standardized Pareto Chart for Y
AC
BC
C: Z1
B: X2
A: X1
AA
BB
0
2
4
6
8
10
12
Standardized effect

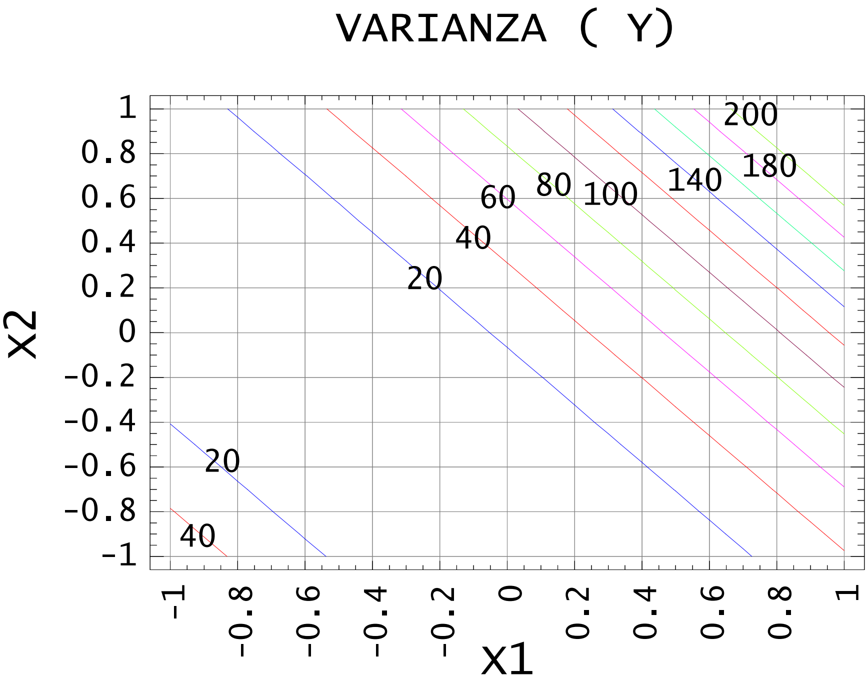

Main Effects Plot for Y
Y
50
47
44
41
38
35
-1.0
1.0
X1
X2
Z1

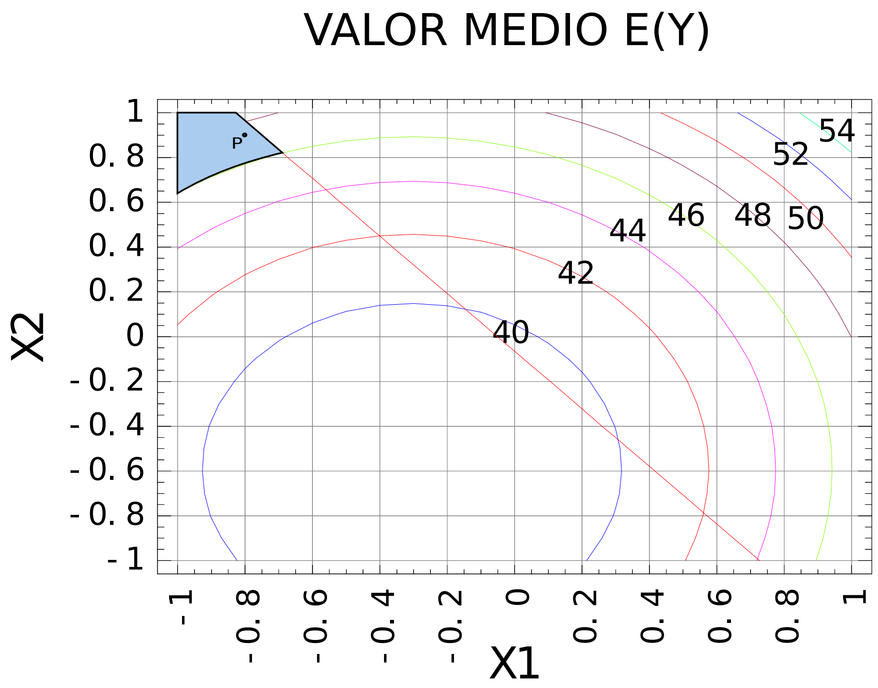

Interaction Plot for Y
Y
60
55
50
45
40
35
30
A=+1
A=-1
B=+1
B=-1
-1.0
1.0
CA
CB

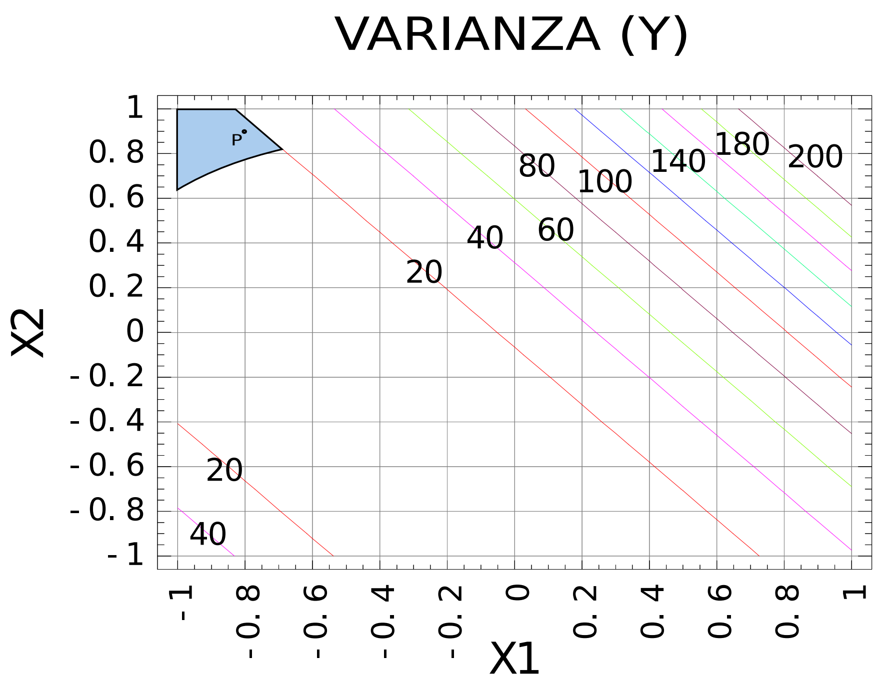

VALOR MEDIO E(Y)
Function
54
50
46
42
38
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
x1
-1 -0.5 0 0.5 1
x2


VALOR MEDIO E(Y)
40
42
44
46
48
50
52
54
x2
1 0.8 0.6 0.4 0.2 0 -0.2 -0.4 -0.6 -0.8 -1
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
x1


VARIANZA (Y)
Function
300
250
200
150
100
50
0
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
X1
-1
-0.5
0
0.5
1
X2


VARIANZA (Y)
200
180
140
100
80
60
40
20
20
40
X2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
X1


VALOR MEDIO E(Y)
X2
X1
P
40
42
44
46
48
50
52
54


VARIANZA (Y)
X2
X1
P
20
40
60
80
100
140
180
200

BIBLIOGRAFÍA

1. R.H. Myers y D.C.Montgomery. Response Surface Methodology, Second Edition(2002). John Wiley & Sons.

2. D.C. Montgomery. Design and Analysis of Experiments. 5th Edition (2001). John Wiley & Sons.

3. E. P. Box, W. G. Hunter and J.S. Hunter. Estadística para Investigadores (1989). Editorial Reverté.

4. A. Fernández de Trocóniz (1987). Modelos Lineales. Universidad del País Vasco.

5. Phillip J. Ross. Taguchi Techniques for Quality Engineering (1988). McGraw-Hill.

6. Programa Estadístico Statgraphics versión 5 de Statistical Graphics Corporation (2000).