

El diseño de experimentos en la docencia y la investigación

Carlos Rolz Asturias

Centro de Ingeniería Bioquímica, Instituto de Investigaciones, Universidad del Valle de Guatemala

carlosrolz@uvg.edu.gt

El diseño de experimentos, o DOE por sus siglas en inglés, es una estrategia experimental que permite obtener un conocimiento de un fenómeno con el objeto de conocerlo con más profundidad. El conocimiento se adquiere al realizar una serie programada de pruebas o experimentos.

La elaboración del diseño implica la definición de un número relativamente pequeño de experimentos que se ejecutan bajo condiciones controladas. Existen varios pasos a seguir durante la elaboración del mismo:

- Definir el objetivo del experimento
- Definir las variables que se controlarán durante los experimentos y los intervalos de su variación, o sea las variables de diseño (o independientes)
- Definir las variables respuesta (o dependientes) que se medirán durante el experimento y examinar su precisión
- Seleccionar entre los diseños disponibles el que más se adapta a las condiciones anteriores

Los objetivos experimentales siguientes son los que más se adaptan a las técnicas del DOE:

- Dentro de un grupo apreciable de variables de diseño identificar aquellas que ejercen una mayor influencia sobre la variable respuesta. Los diseños aplicables son los denominados factoriales fraccionados y los de Plackett-Burman
- Explorar los efectos que sobre la variable respuesta ejercen diferentes valores de las variables de diseño. Generalmente se aplican diseños factoriales
- Localización del valor máximo (o mínimo) de la variable respuesta en un espacio definido por el valor de las variables de diseño. Acá se usan diseños de optimización, siendo los más usados: el central compuesto (CCD por sus siglas en inglés) y los de Box Behnken
- Casos de optimización en donde las variables de diseño tienen restricciones. Existe una variedad de diseños, la mayoría exclusivos de diferente software y que generalmente se refieren como diseños D-óptimos.

- Composición óptima de mezclas conteniendo varios componentes. Se usan diseños de mezcla. En ellos también pueden existir situaciones en donde las variables de diseño tienen restricciones.

La selección de las variables de diseño y los valores que esas variables tendrán en las pruebas a realizar, conforman la información clave para resolver un problema exitosamente. Cuán difícil resulta la tarea dependerá, en una mayor parte, del conocimiento previo que se tenga del caso. La situación más complicada es cuando existe poca información previa y la experiencia del grupo investigador en el tema es limitada.

Los diferentes diseños experimentales existentes se encuentran descritos en detalle en los textos clásicos al respecto (ver Viñeta 1). También el software comercial incluye una amplia gama de diseños y recomiendan el más adecuado para cada caso en particular (ver Viñeta 2).

Los diseños factoriales

En estos diseños se combinan los niveles con las variables de diseño, de tal manera que se encuentran balanceados, es decir, cada nivel de una variable se estudia igual número de veces en combinación con cada nivel de las otras variables de diseño.

Suponer que se tienen tres variables de diseño y se han escogido dos niveles para cada variable, uno menor (-1) que el otro (1). La matriz de experimentos a realizar del diseño factorial será de 2^3 pruebas, o sean 8 pruebas. Las pruebas se realizan en duplicado de manera que pueda estimarse el error experimental y la varianza de la respuesta. En total, entonces, se realizarán 16 pruebas. Los Cuadros 1 y 2 ilustran la combinación de niveles y variables que definen las condiciones a las cuales se llevan a cabo los experimentos o pruebas. Notar que cada experimento se lleva a cabo bajo condiciones diferentes y también los signos en las tres columnas se alternan con un patrón definido.

La Gráfica 1 ilustra la distribución de las pruebas en los vértices de un cubo, y muestra los cambios de nivel de su valor mínimo al máximo de las variables de diseño A, B, y C. Notar que las ocho pruebas del factorial ocupan los ocho vértices del cubo.

Viñeta 1: Libros de Diseño de Experimentos en orden preferencial

Textos adecuados para cursos formales en DOE

Douglas C Montgomery *Design and Analysis of Experiments* Wiley 2005 QA 279.M66
 Lisa Custer, Daniel R. McCarville, Douglas C Montgomery *Design and Analysis of Experiments. Student solutions manual* Wiley 2006 QA 279.N66
 Mark J Anderson, Patrick J Whitcomb *DOE Simplified : Practical Tools for Effective Experimentation* Portland: Productivity 2000 QA 279.A5287
 Mark J Anderson, Patrick J Whitcomb *RSM Simplified : Optimizing Processes Using Response Surface Methods for Design of Experiments* Portland: Productivity 2005 QA 279.A528
 H Gutiérrez Pulido, Román de la Vara Salazar *Análisis y Diseño de Experimentos* México : McGraw-Hill, 2012 QA 279.G96

Textos para consulta en investigación y la práctica profesional

Paul D Berger, Robert E Maurer *Experimental Design: with Applications in Management, Engineering, and the Sciences* Duxbury/Thomson Learning 2002 QA 279.B467
 John A Cornell *Experiments with Mixtures : Designs, Models, and the Analysis of Mixture Data* Wiley 1990 QA 279.C66

Textos avanzados

Raymond H Myers, Douglas C Montgomery *Response Surface Methodology* Wiley 1995 QA 279.M67
 George EP Box, Norman R. Draper *Empirical Model-Building and Response Surfaces* Wiley 1987 QA 279.B67
 RH Myers *Response Surface Methodology : Process and Product Optimization using Designed Experiments* Wiley 1995 QA 279.M67

Viñeta 2. Software para Diseño de Experimentos

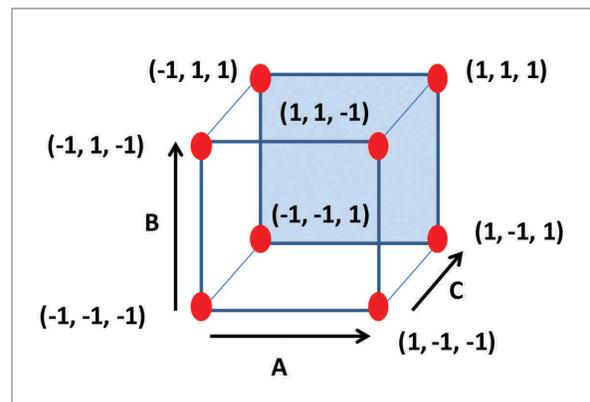
Minitab 17 www.minitab.com
 JMP 11 www.jmp.com
 Design Expert Version 9 www.statease.com
 The Unscrambler X www.camo.com

Cuadro 1. Definición de los niveles menor y mayor de las variables de diseño

| Variable | Nivel menor | Nivel mayor |
|----------|-------------|-------------|
| A | -1 | 1 |
| B | -1 | 1 |
| C | -1 | 1 |

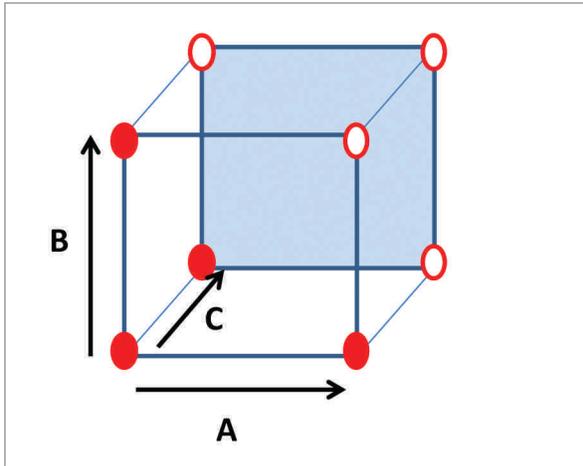
Cuadro 2. Matriz del diseño factorial de 8 pruebas con la variable respuesta evaluada en duplicado (R₁ y R₂)

| Prueba | A | B | C | R ₁ | R ₂ |
|--------|----|----|----|----------------|----------------|
| 1 | -1 | -1 | -1 | | |
| 2 | 1 | -1 | -1 | | |
| 3 | -1 | 1 | -1 | | |
| 4 | 1 | 1 | -1 | | |
| 5 | -1 | -1 | 1 | | |
| 6 | 1 | -1 | 1 | | |
| 7 | -1 | 1 | 1 | | |
| 8 | 1 | 1 | 1 | | |



Gráfica 1. Despliegue del diseño factorial de ocho pruebas.

La estrategia experimental definida por el factorial en donde las tres variables cambian simultáneamente permite encontrar no solo el efecto de cada variable en particular sino que también las posibles interacciones entre las variables. Tal estrategia contrasta con la técnica empleada con frecuencia en la experimentación, que consiste en cambiar una variable a la vez, manteniendo el resto constante. La Gráfica 2 muestra lo limitado de la exploración que se obtiene con esa estrategia.



Gráfica 2. Estrategia de una a la vez. En la primera prueba la variable A aumenta en dirección de la flecha. B y C permanecen en el origen. El segundo experimento cambia B y en la tercera prueba cambia C. Las otras dos, de nuevo, permanecen en el origen.

Aplicación de un diseño factorial en la investigación

Recientemente en uno de los proyectos en desarrollo en el Centro de Ingeniería Bioquímica del Instituto de Investigaciones decidimos emplear un diseño factorial para contestar algunas interrogantes relacionadas con la producción de etanol por levaduras cuando la concentración inicial de reaccionantes era alta. Se sabe que la levadura es el mejor microorganismo para producir etanol por su velocidad de producción y su tolerancia al mismo. Sin embargo, durante el proceso la levadura está sujeta a dos fuertes inhibiciones, una cuando los reaccionantes están inicialmente a una alta concentración; la otra, cuando la acumulación del etanol en el bio reactor llega a niveles no permisibles por la célula. Los mecanismos de las inhibiciones, o los de tolerancia, no se conocen con certeza, pero en algunos casos la adición de nutrientes ha mostrado ser un paliativo. A nosotros nos interesaba saber si el efecto de agregar esos nutrientes era igual para varias cepas de levaduras en nuestro laboratorio. Y nos preguntábamos, ¿qué es más importante, la adición de nutrientes o la cepa de levadura empleada?

El diseño empleado fue un factorial 2^4 (16 experimentos) con el punto central repetido seis veces. En total 22 experimentos

Las variables se encuentran definidas en el Cuadro 3 y debe notarse que una de ellas, levadura, no es numérica sino de categoría. La misma identifica las dos levaduras empleadas. Se

obtuvieron dos variables respuesta, el etanol producido y los azúcares consumidos.

La matriz de experimentos se muestra en el Cuadro 4. El orden de los experimentos ha sido diseñado al azar por el software *Design Expert 7.03*. El cuadro incluye el valor de dos variables de respuesta obtenidos de los resultados experimentales.

Notar lo siguiente:

- La cantidad de etanol experimental cambió en los experimentos 6.9 veces del valor mínimo encontrado de 1.17 % en la prueba 15. El consumo de azúcares varió 2 veces del valor mínimo encontrado de 43.38 % de la prueba 1. Lo anterior significa que las variables de diseño tuvieron un efecto sobre las dos variables de respuesta, cuando cambiaron en el intervalo designado.
- El punto central se repitió tres veces para la levadura A1 y tres veces para la levadura A2. En los tres replicados el etanol cambió 1.4 veces del valor mínimo para la levadura A1 y 1.86 veces para A2. Las cifras para el consumo de azúcares fueron 1.06 veces para la levadura A1 y 1.02 veces para la levadura A2. Lo anterior mostró que el error experimental era menor que los cambios obtenidos por el efecto de las variables de diseño.

El software estimó el valor de los efectos de cada variable sobre las variables respuesta. La Gráfica 3 ilustra cuales variables influyeron más sobre el etanol obtenido. La Gráfica 4 muestra el mismo resultado para el consumo de azúcares. En las gráficas las barras de color azul indican que el mayor efecto se logró en el nivel menor de la variable y las de color naranja el efecto en el nivel mayor de la variable. Por lo tanto:

La levadura, el Brix inicial y la interacción entre ambos fueron las variables significativas en la producción de etanol. En otras palabras la levadura A1 (CBS 381) fue superior a la A2 (CBS422). El cambio de un valor bajo de Brix inicial a un valor alto del Brix causó un efecto positivo sobre la producción de etanol. La interacción entre ambas variables indica que no puede decirse cuanto será el efecto del aumento en el Brix, sino antes se especifica cuál es la levadura.

El consumo de azúcares fue influenciado significativamente por la levadura, siendo la A1 (CBS 381) mejor que la A2 (CBS422).

Cuadro 3. Definición de los niveles mínimo y máximo y el punto central de las variables de diseño

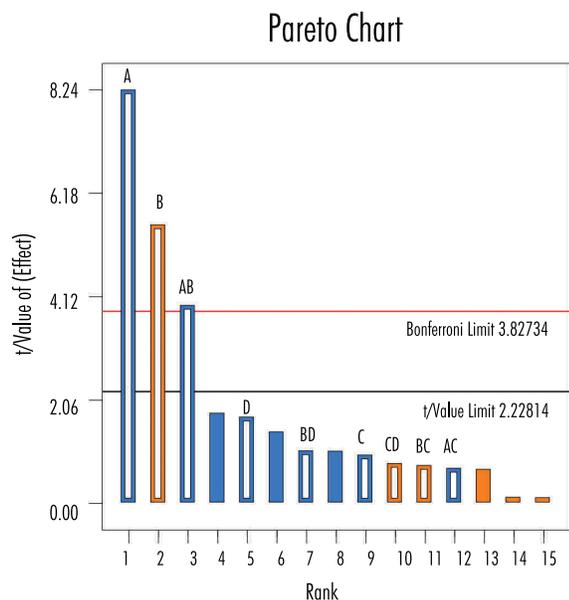
| Variable | Nivel menor | Punto central | Nivel mayor | Unidades |
|---------------------------|--------------|---------------|--------------|-------------|
| Levadura | CBS 381 (A1) | - | CBS 422 (A2) | Categoría |
| Reaccionante inicial | 10.0 | 12.5 | 15.0 | Grados Brix |
| Extracto de levadura (YE) | 5.0 | 7.5 | 10.0 | g/L |
| Fosfato de amonio (FA) | 1.0 | 2.0 | 3.0 | g/L |

Cuadro 4. Matriz de diseño de experimentos y variables de respuesta obtenidas

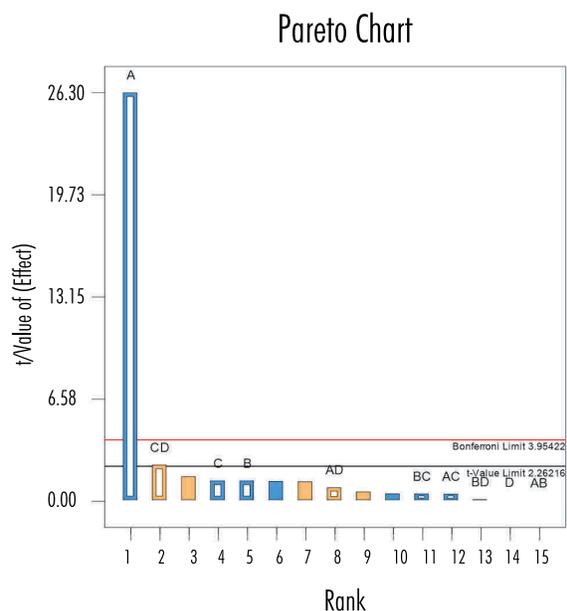
| Prueba | Levadura | Brix | YE | FA | Etanol % | Consumo Azúcares % |
|--------|----------|------|------|----|----------|--------------------|
| 1 | 422 | 12.5 | 7.50 | 2 | 1.40 | 43.38 |
| 2 | 381 | 12.5 | 7.50 | 2 | 5.32 | 83.45 |
| 3 | 422 | 10.0 | 10.0 | 1 | 1.82 | 46.43 |
| 4 | 422 | 10.0 | 5.0 | 3 | 1.80 | 48.26 |
| 5 | 422 | 12.5 | 7.50 | 2 | 2.60 | 49.55 |
| 6 | 422 | 15.0 | 5.0 | 1 | 3.21 | 50.51 |
| 7 | 381 | 12.5 | 7.50 | 2 | 4.72 | 84.53 |
| 8 | 422 | 10.0 | 10.0 | 3 | 1.45 | 53.41 |
| 9 | 381 | 10.0 | 10.0 | 3 | 3.52 | 88.42 |
| 10 | 422 | 15.0 | 5.0 | 3 | 2.61 | 50.51 |
| 11 | 381 | 15.0 | 10.0 | 3 | 7.09 | 82.56 |
| 12 | 381 | 10.0 | 5.0 | 1 | 4.88 | 87.88 |
| 13 | 422 | 12.5 | 7.50 | 2 | 1.82 | 46.68 |
| 14 | 422 | 15.0 | 10.0 | 1 | 2.54 | 45.32 |
| 15 | 422 | 15.0 | 10.0 | 3 | 1.17 | 51.21 |
| 16 | 422 | 10.0 | 5.0 | 1 | 1.96 | 56.80 |
| 17 | 381 | 12.5 | 7.50 | 2 | 3.80 | 84.76 |
| 18 | 381 | 10.0 | 10.0 | 1 | 8.07 | 85.50 |
| 19 | 381 | 15.0 | 5.0 | 3 | 6.00 | 85.25 |
| 20 | 381 | 15.0 | 5.0 | 3 | 3.08 | 85.99 |
| 21 | 381 | 15.0 | 5.0 | 1 | 7.37 | 88.07 |
| 22 | 381 | 10.0 | 10.0 | 1 | 2.24 | 86.06 |

Los efectos de los dos nutrientes (C y D) no fueron significativos y en ambos casos el efecto mayor fue en el nivel menor de la variable. Lo que indica que los mismos deberían emplearse en la menor concentración, es decir 5 y 1 g/L respectivamente.

La Gráfica 5 muestra la relación entre las dos variables respuesta para las levaduras. Se observa claramente la superioridad de la levadura CBS 381 para producir etanol del jugo de sorgo dulce.



Gráfica 3. Variables que influyeron sobre la producción de etanol.



Gráfica 4. Variables que influyeron sobre el consumo de azúcares.

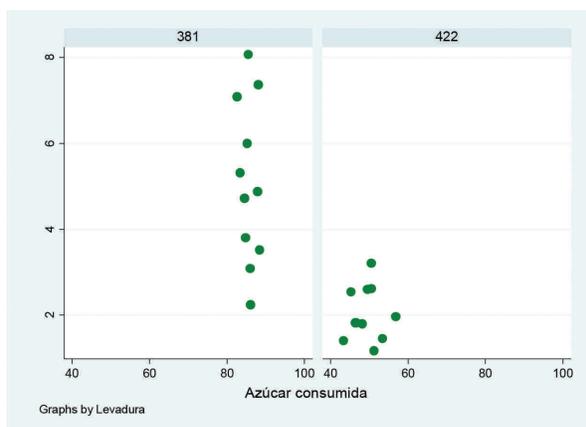


Gráfico 5. Relación entre etanol producido y consumo de azúcares para las dos levaduras.

El diseño de experimentos en la docencia

En el Departamento de Ingeniería Química se ha programado un Taller específico de Investigación ofrecido en dos ciclos del cuarto año de la carrera. El Taller es de asistencia voluntaria y en el mismo se ofrecen presentaciones en diseño de experimentos, análisis exploratorio de datos experimentales, técnicas de ajuste de datos y elaboración de modelos, análisis de variables múltiples y empleo de modelos probabilísticos.

La sección de diseño de experimentos incluye:

- Diseños de selección de variables (factoriales fraccionados y Plackett-Burman)
- Diseños de exploración de efectos (factoriales)
- Diseños optimización de superficie de la respuesta (Central compuesto o CCD, Box-Behnken)
- Diseños de mezcla
- Diseños D-óptimos

En las presentaciones semanales se discute la resolución de problemas seleccionados de artículos de investigación en revistas de corriente principal. Para esto el instructor emplea el software Design Expert. Los estudiantes tienen la oportunidad de resolver problemas escogidos en el equipo de cómputo de Operaciones Unitarias empleando el software The Unscrambler.

Uno de los ejemplos resueltos por los estudiantes se relacionó con el empleo de un diseño de mezcla. El problema consistía en encontrar la mezcla de tres jugos de frutas que gustara más a un grupo de expertos catadores. La empresa en cuestión procesaba melón para exportación como fruta fresca. Sin embargo, generalmente tenía un exceso de fruta, por lo tanto, una posibilidad era producir una bebida a base de melón, mezclada con jugo de naranja y jugo de piña. La mezcla debería contener por lo menos un 30 % de melón.

El intervalo de variación de los componentes se muestra en el Cuadro 5. Nótese que el valor mínimo del melón en la mezcla es de 30 %, lo cual restringe el valor máximo permisible para la naranja y la piña en 70 %.

El software The Unscrambler seleccionó para este problema un diseño de mezcla simplex con restricciones de acuerdo con la matriz de pruebas del Cuadro 6. Las pruebas están ilustradas en el diagrama en coordenadas triangulares mostrado en la Gráfica 6.

Los estudiantes realizaron las pruebas in silico ingresando los resultados de la evaluación organoléptica. El modelo fue significativo $P = 0.085$ y se obtuvo la respuesta en el diagrama mostrado en la Gráfica 7. La composición dentro de la zona que emplea el mayor porcentaje de jugo de melón tiene la composición siguiente: melón = 56 %, piña 20 % y naranja 24 %.

En la UVG se ofrece un curso, en el primer año, a todos los estudiantes de primer ingreso sobre Técnicas de Investigación, pero en el mismo no se incluye aspectos de diseño de experimentos. Todas las carreras llevan un curso introductorio de Estadística y solamente las carreras de Ingeniería y Ciencias y Humanidades, tales como Ingeniería Industrial e Ingeniería de Alimentos, Biología, Bioquímica y Nutrición, tienen en su pensum un segundo curso de Estadística Aplicada. En este curso sí se ven los diseños factoriales.

Cuadro 5. Definición de los niveles bajo y alto de las variables de diseño

| Ingrediente | Bajo % | Alto % |
|-------------|--------|--------|
| Melón | 30 | 100 |
| Piña | 0 | 70 |
| Naranja | 0 | 70 |

Cuadro 6. Diseño de mezcla simplex con restricciones

| Prueba | Melón % | Piña % | Naranja % |
|--------|---------|--------|-----------|
| 1 | 100 | 0 | 0 |
| 2 | 30 | 70 | 0 |
| 3 | 30 | 0 | 70 |
| 4 | 65 | 35 | 0 |
| 5 | 65 | 0 | 35 |
| 6 | 41.67 | 46.67 | 11.66 |
| 7 | 53.33 | 23.33 | 23.33 |
| 8 | 53.33 | 23.33 | 23.33 |
| 9 | 53.33 | 23.33 | 23.33 |
| 10 | 53.33 | 23.33 | 23.33 |
| 11 | 76.67 | 11.67 | 11.66 |
| 12 | 41.67 | 46.67 | 11.66 |
| 13 | 41.67 | 11.66 | 46.67 |

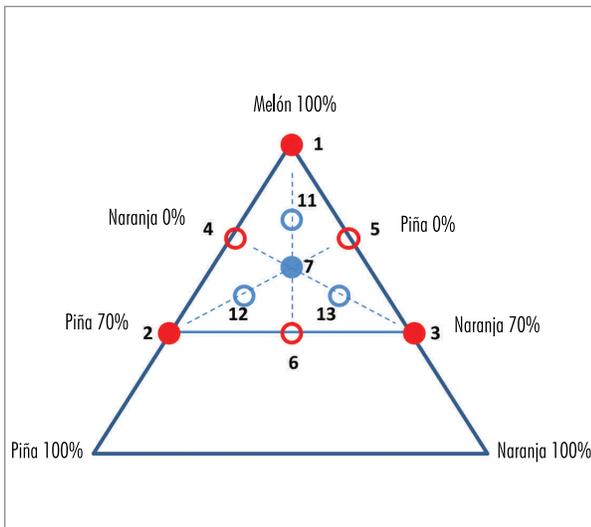


Gráfico 6. Diseño de mezcla simplex con restricciones. Los números indican la prueba.

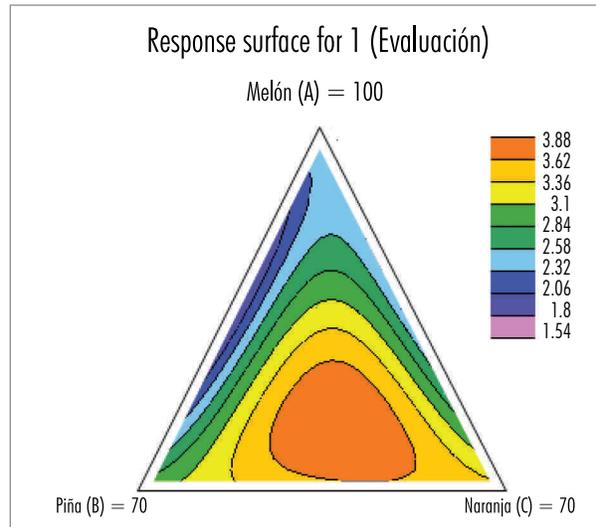


Gráfico 7. Superficie de la respuesta del diseño de mezcla elaborada a partir del modelo obtenido. La zona de mayor puntuación en color naranja limita las mezclas más apreciadas por los evaluadores.

Conclusiones

Los dos ejemplos ilustrativos muestran las características del uso del diseño de experimentos, en donde, con una serie de mínimas pruebas experimentales, se logra responder a las interrogantes iniciales, ofreciendo además información adicional complementaria que enriquece el conocimiento sobre el problema en cuestión.

La utilización del diseño de experimentos en la solución de problemas aplicados resulta ser una herramienta que los estudiantes de ciencias e ingeniería deberían aprender durante su formación profesional. Por lo tanto es pertinente recomendar que se incluya como un curso en los diferentes programas de estudio. De preferencia en el cuarto año previo a iniciar su trabajo de graduación.