

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
FACULTAD DE INGENIERÍA, ARQUITECTURA Y DISEÑO
MAESTRÍA Y DOCTORADO EN CIENCIAS E INGENIERÍA



TESIS

“OPTIMIZACIÓN ROBUSTA DE MÚLTIPLES VARIABLES DE RESPUESTA
MEDIANTE MÉTODOS BAYESIANOS”

Que para cubrir parcialmente los requisitos necesarios para obtener el grado de

DOCTOR EN CIENCIAS

presenta:

JORGE LIMÓN ROMERO

Director de tesis

DR. MANUEL ARNOLDO RODRÍGUEZ MEDINA

Ensenada, Baja California

Enero 2013.

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

Facultad de Ingeniería, Arquitectura y Diseño

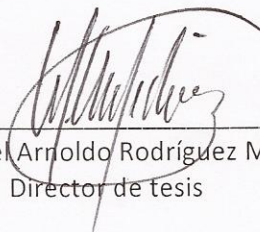
“OPTIMIZACIÓN ROBUSTA DE MÚLTIPLES RESPUESTAS MEDIANTE MÉTODOS BAYESIANOS”

TESIS

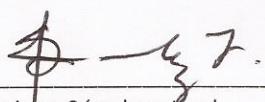
Que para obtener el grado de DOCTOR en CIENCIAS presenta:

JORGE LIMÓN ROMERO

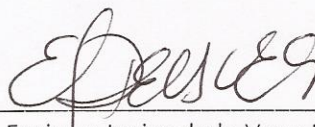
Aprobada por el siguiente comité:



Dr. Manuel Arnoldo Rodríguez Medina
Director de tesis



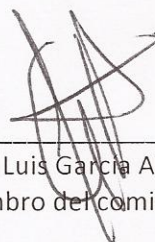
Dr. Jaime Sánchez Leal
Miembro del comité



Dr. Enrique Javier de la Vega Bustillos
Miembro del comité



Dr. Rafael García Martínez
Miembro del comité




Dr. Jorge Luis García Alcaraz
Miembro del comité

Ensenada, Baja California, México. Diciembre, 2012.

RESUMEN de la tesis de **JORGE LIMÓN ROMERO** presentada como requisito parcial para la obtención del grado de DOCTOR EN CIENCIAS del programa de Maestría y Doctorado en Ciencias e Ingeniería de la UABC. Ensenada Baja California, México, Diciembre de 2012.

“OPTIMIZACIÓN ROBUSTA DE MÚLTIPLES VARIABLES DE RESPUESTA MEDIANTE MÉTODOS BAYESIANOS”

Resumen aprobado por:



Dr. Manuel Arroldo Rodríguez Medina
Director de Tesis

En la actualidad se pueden encontrar muchos trabajos referentes a la optimización simultánea de múltiples variables de respuesta, dejando atrás la opción de optimizar una a una estas variables por separado, evitando con ello soluciones subóptimas o contradictorias. Para llevar a cabo esta optimización simultánea se han propuesto diversas variantes, las cuales conllevan también diversos niveles de complejidad técnica. Algunas de estas opciones se pueden encontrar incluso dentro de la paquetería del software estadístico comercial, lo que las ha beneficiado para que se conviertan en las más utilizadas en este campo, aunque no necesariamente sean las mejores alternativas.

En el presente trabajo se plantea la modificación de una de las opciones que ha surgido más recientemente para la optimización simultánea de múltiples respuestas. Esta se basa en los principios de la inferencia bayesiana, ya que hace uso de la función de distribución predictiva y la simulación Monte Carlo, para generar valores de las variables de respuesta en condiciones específicas de operación del proceso, los cuales son posteriormente contrastados contra sus respectivas especificaciones. Durante la simulación es posible incorporar un tratamiento especial para los factores de ruido con el objetivo de atender la robustez. Esta técnica utiliza la reducción de dimensionalidad del problema, al sustituir las variables de respuesta originales por una sola, la cual consiste en la probabilidad de concordancia de dichas variables, es decir, la probabilidad de que en una condición de operación del proceso se cumpla con sus requisitos o especificaciones. Esta nueva variable de respuesta debe modelarse y posteriormente optimizarse.

La modificación propuesta consiste en sustituir la probabilidad de concordancia conjunta, por la media geométrica de las probabilidades de concordancia individuales. Esta modificación permite trabajar con el caso en el que no todas las variables de respuesta son de igual importancia, mediante la asignación de ponderaciones de acuerdo con la relevancia de cada una de ellas y calculando así la media geométrica ponderada de dichas probabilidades, de tal forma que esta se convierte en la variable de respuesta unificadora, posteriormente se modela mediante regresión logística y por último se optimiza mediante el método del gradiente reducido generalizado.

En esta investigación se demuestra mediante la generación de distintos escenarios en dos casos de estudios tomados de la literatura, que las condiciones óptimas propuestas siempre se movieron hacia las regiones en donde las variables con los mayores pesos o ponderaciones tienen probabilidades de concordancia e índices *Cpks* más altos, a expensas del desempeño en las variables con menores ponderaciones. Este efecto fue comprobado de forma visual mediante gráficas comparativas del comportamiento de cada variable de respuesta en los distintos escenarios y posteriormente de manera estadística construyendo intervalos de confianza del 95% para las probabilidades de concordancia de cada variable de respuesta. Debido a esto, es posible concluir que la modificación planteada, sí ayuda a mejorar la flexibilidad de la herramienta tomada como referencia, además de conservar sus bondades originales.

Palabras Clave: *Múltiples respuestas, inferencia bayesiana, distribución predictiva, optimización simultánea, probabilidad de concordancia, media geométrica ponderada.*

A mis padres **Federico y Rosa Delia**
a mi esposa **Alma Judith**
y a mis hijos **Jorge Luis y Carolina**

AGRADECIMIENTOS

A **DIOS** por todas las bendiciones con que me ha favorecido, seguramente algunas de ellas, incluso, sin merecerlas y por hacerme sentir que siempre está a mi lado, en los momentos buenos para alegrarse conmigo y en los difíciles para darme la fortaleza, inteligencia y sabiduría que se requiere para sobreponerse a ellos.

Al **Dr. Manuel Arnoldo Rodríguez Medina** por el apoyo brindado, la asesoría y los atinados consejos que fueron una parte importante para la culminación de este proyecto.

Al **Dr. Enrique de la Vega Bustillos** por el apoyo brindado y sus sugerencias desde la planeación de mis estudios de doctorado.

Al **Dr. Rafael García Martínez** por el apoyo brindado para hacer posible este programa de doctorado y su asesoría durante la realización de este proyecto.

Al **Dr. Jaime Sánchez Leal** por el apoyo brindado, sus consejos y recomendaciones, buscando siempre enriquecer este trabajo de tesis.

Al **Dr. Jorge Luis García Alcaraz** por su solidaridad y apoyo desde los inicios de este programa doctoral.

En general a todos los doctores que conformaron mi **comité de tesis** que con su asesoría técnica y paciencia, fueron una parte fundamental para lograr este objetivo. A todos ellos mi más sincero agradecimiento.

A mis padres **Federico Limón y Rosa Delia Romero**, que siempre están pendientes de mí y cada vez que pueden, muy a su manera me hacen sentir su respaldo. Gracias por darme la vida y por demostrarme que confían en mí. Están y siempre estarán conmigo.

A mi esposa **Alma Judith Pedroza Gómez** que con su solidaridad, cariño y paciencia fue una parte elemental en el logro de este objetivo. Junto con mis niños Carito y Jorge Luis son el regalo más hermoso que dios me ha dado y mi mayor motivación para seguir adelante.

A mis hermanos (as): **Alma Guadalupe, Heriberto, Yolanda y Claudia** y sus respectivas familias, **Federico Araujo, Gabriela y Daniela; Rebeca, Carlos Enrique**

y Erick; Jorge y Mariel; César y Cesarito (Morita) que está por venir a alegrarnos con su nacimiento. A todos ellos por echarme la mano siempre que lo he requerido y por confiar siempre en mí. Su respaldo me da fortaleza y me motiva a seguir adelante.

Al Dr. **Oscar Roberto López Bonilla** por su apoyo para arrancar este programa de doctorado Y por estar siempre atento a nuestro avance.

Al Dr. **Juan Iván Nieto** que primeramente como coordinador de posgrado nos brindó su apoyo para iniciar este proyecto y posteriormente como director de la FIAD nos ayudó a llevarlo a buen término.

Al Dr. **Juan de Dios Sánchez**, coordinador de posgrado, por su disponibilidad para brindarnos la asesoría requerida para llevar a cabo los trámites correspondientes.

Al Dr. **Efrén García Guerrero** por su amistad y atinados consejos.

A mis amigos (as) de la Facultad: **Diego Tlapa, Yolanda Báez, Jesús Olguín, Claudia Camargo, Julián Aguilar, Miguel Cadena, Jesús Salinas, Dora Luz Flores, Claudia Gómez, Everardo Inzunza, Pablo Rousseau, Christian Navarro**, quienes al brindarme su apoyo hicieron más fácil la culminación de este objetivo.

A la **Universidad Autónoma de Baja California** y a la **Facultad de Ingeniería, Arquitectura y Diseño** por todas las facilidades brindadas para hacer posible la apertura de este programa de doctorado.

Ensenada B.C. México,
Enero 2013.

Jorge Limón Romero.

TABLA DE CONTENIDO

	Página
RESUMEN	iii
AGRADECIMIENTOS	vi
TABLA DE CONTENIDO	viii
LISTA DE FIGURAS	x
LISTA DE TABLAS	xii
1. INTRODUCCIÓN	1
1.1 Antecedentes	2
1.2 Planteamiento del Problema	2
1.2.1 Preguntas de Investigación.....	3
1.2.2 Hipótesis	3
1.2.3 Objetivos.....	4
1.2.3.1 Objetivo General.....	4
1.2.3.2 Objetivos Específicos	4
1.2.4 Justificación.....	4
2. REVISION DE LITERATURA	6
2.1 La Metodología de Superficie de Respuesta Dual	6
2.2 Optimización de Múltiples Variables de Respuesta	7
2.2.1 Sobreposición de Gráficas de Contorno	8
2.2.2 Problema de Optimización Restringida	12
2.2.3 Combinación de las Respuestas en una Sola	13
2.2.3.1 Función de Deseabilidad.....	13
2.2.3.2 Función de Pérdida	15
2.2.3.3 Índices de Capacidad de Procesos	16
2.2.3.3.1 Índice <i>Cpm</i>	16
2.2.3.3.2 Índices <i>Mcpm</i> y <i>Mcpk</i>	17
2.2.4 Enfoques que se Basan en la Proporción de Concordancia.....	18
2.2.4.1 Método de Chiao y Hamada	18
2.3 La Inferencia Bayesiana	19
2.3.1 El Teorema de Bayes.....	20
2.3.1.1 El Teorema de Bayes para Eventos	21
2.3.1.2 El Teorema de Bayes para Densidades	21

2.4	Las Cadenas de Markov con Simulación Monte Carlo (CMMC) y el Muestreo de Gibbs..	22
2.5	La Media Geométrica	26
2.6	El Enfoque Bayesiano para Optimizar Múltiples Variables de Respuesta	27
3.	MATERIALES Y MÉTODOS	29
3.1	Materiales	29
3.2	Métodos	30
4.	ANÁLISIS DE RESULTADOS.....	33
4.1	Caso de Estudio I: Análisis de Sistema HPLC	33
4.2	Caso de Estudio II: Porcentaje de Conversión y Actividad Térmica	55
4.3	Discusión de Resultados	77
4.3.1	Conclusiones Fase X: Análisis de Sistema HPLC	77
4.3.2	Conclusiones Fase X: Porcentaje de Conversión y Actividad Térmica	78
5.	CONCLUSIONES Y RECOMENDACIONES	80
5.1	Conclusiones de Acuerdo con las Hipótesis y los Objetivos	80
5.1.1	Conclusiones de Acuerdo con las Hipótesis.....	80
5.1.2	Conclusiones de Acuerdo con los Objetivos	81
5.1.3	Conclusiones Generales	81
5.2	Recomendaciones	82
5.2.1	Futuras Líneas de Investigación	83
6.	REFERENCIAS	84
7.	ANEXOS.....	87
7.1	ANEXO 1	87

LISTA DE FIGURAS

Figura 2.1 Superficie de respuesta para el rendimiento	9
Figura 2.2 Gráfica de contorno para el rendimiento.....	9
Figura 2.3 Superficie de respuesta para la viscosidad.....	10
Figura 2.4 Gráfica de contorno para viscosidad.....	10
Figura 2.5 Superficie de respuesta para el peso molecular	11
Figura 2.6 Gráfica de contorno para el peso molecular	11
Figura 2.7 Sobreposición de las gráficas de contorno del rendimiento, la viscosidad y el peso molecular.	12
Figura 2.8 Función de deseabilidad.....	14
Figura 2.9 Algoritmo del muestreo de Gibbs para muestrear de una distribución posterior de un vector de parámetros q -dimensional.	24
Figura 2.10 Comparación de tres histogramas de probabilidad de la distribución beta - binomial con $n = 16$, $\alpha = 2$ y $\beta = 4$	26
Figura 4.1 Comparación de los resultados obtenidos para Mg y Mgp	42
Figura 4.2 Diferencias en las probabilidades entre Mg y Mgp para cada corrida	43
Figura 4.3 Resultados simulados contra pronosticados para Mg	44
Figura 4.4 Resultados simulados contra pronosticados para Mgp	44
Figura 4.5 Superficie de respuesta para el modelo logit ajustado a Mg	45
Figura 4.6 Gráfica de Contornos para el modelo logit ajustado a Mg	45
Figura 4.7 Superficie de respuesta para el modelo logit ajustado a Mgp	46
Figura 4.8 Gráfica de contornos para el modelo logit ajustado a Mgp	46
Figura 4.9 Comparación de las probabilidades de concordancia en las condiciones óptimas propuestas para y_1 según Mg con ponderación 0.25 y Mgp con ponderación 0.1.....	53
Figura 4.10 Comparación de las probabilidades de concordancia en las condiciones óptimas propuestas para y_2 según Mg con ponderación 0.25 y Mgp con 0.4.	53
Figura 4.11 Comparación de las probabilidades de concordancia en las condiciones óptimas propuestas para y_3 según Mg con ponderación 0.25 y Mgp con 0.4.	54
Figura 4.12 Comparación de las probabilidades de concordancia en las condiciones óptimas propuestas para y_4 según Mg con ponderación 0.25 y Mgp con ponderación 0.1.....	54
Figura 4.13 Resultados simulados contra los pronosticados para Mgp_1	63
Figura 4.14 Resultados simulados contra los pronosticados para Mgp_2	63
Figura 4.15 Resultados simulados contra los pronosticados para Mgp_3	64
Figura 4.16 Resultados simulados contra los pronosticados para Mgp_4	64
Figura 4.17 Gráfica de superficie para Mgp_1	65
Figura 4.18 Gráfica de contornos para Mgp_1	65
Figura 4.19 Gráfica de superficie para Mgp_2	66
Figura 4.20 Gráfica de contornos para Mgp_2	66
Figura 4.21 Gráfica de superficie para Mgp_3	67

Figura 4.22 Gráfica de contornos para Mgp_3	67
Figura 4.23 Gráfica de superficie para Mgp_4	68
Figura 4.24 Gráfica de contornos para Mgp_4	68
Figura 4.25 Comportamiento de la probabilidad de concordancia para y_1 en las condiciones óptimas propuestas para los cuatro escenarios Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4 con ponderaciones 0.5, 0.6, 0.8 y 0.2 respectivamente.	76
Figura 4.26 Comportamiento de la probabilidad de concordancia para y_2 en las condiciones óptimas propuestas para los cuatro escenarios Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4 con ponderaciones 0.5, 0.4, 0.2 y 0.8 respectivamente.	76

LISTA DE TABLAS

Tabla 4.1 Datos Originales del Experimento de la Cromatografía	34
Tabla 4.2 Datos Transformados Ortogonalmente del Experimento de la Cromatografía	34
Tabla 4.3 Combinaciones de los Factores de Control Utilizadas para Formar la Malla en la Simulación: Caso de Estudio I.....	37
Tabla 4.4 Pesos Dados a las Variables de Respuesta para Obtener Mg y Mgp	38
Tabla 4.5 Resultados de la Simulación del Arreglo 11^2 : Caso de Estudio I.....	39
Tabla 4.6 Resultados Óptimos Encontrados para Mg y Mgp	47
Tabla 4.7 Resultados Obtenidos por Simulación y por Sustitución en los Modelos Logísticos en las Condiciones Óptimas.	48
Tabla 4.8 Resumen de las 30 Iteraciones en la Condición Óptima para Mg ($x_1 = 0.514$, $x_2 = -1$)	49
Tabla 4.9 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp ($x_1 = 0.552$, $x_2 = -1$).....	51
Tabla 4.10 Estadísticos Resumen para las Condiciones Óptimas Propuestas según Mg y Mgp	52
Tabla 4.11 Datos Tomados para Ilustrar la Metodología: Caso de Estudio II	56
Tabla 4.12 Combinaciones de los Factores de Control Utilizadas para Formar la Malla en la Simulación: Caso de Estudio II.....	57
Tabla 4.13 Ponderaciones Consideradas para las Variables de Respuesta.....	58
Tabla 4.14 Resultados de la Simulación del Arreglo 11^2 : Caso de Estudio II.....	59
Tabla 4.15 Resultados de la Optimización: Caso de Estudio II.....	70
Tabla 4.16 Evaluación por Simulación y según los Modelos Logísticos en las Condiciones Óptimas....	70
Tabla 4.17 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp_1 ($x_2 = 1.663$, $x_3 = -0.767$)	71
Tabla 4.18 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp_2 ($x_2 = 1.682$, $x_3 = -0.731$)	72
Tabla 4.19 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp_3 ($x_2 = 1.682$, $x_3 = -0.597$)	73
Tabla 4.20 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp_4 ($x_2 = 1.4817$, $x_3 = -0.848$)	74
Tabla 4.21 Estadísticos Resumen para las Condiciones Óptimas Propuestas según Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4	75
Tabla 4.22 Intervalos de Confianza para las Probabilidades de Concordancia de cada Variable de Respuesta en las Condiciones Óptimas para el Caso de Estudio I.....	78
Tabla 4.23 Intervalos de Confianza para las Probabilidades de Concordancia de cada Variable de Respuesta en las Condiciones Óptimas para el Caso de Estudio II.....	79

1. INTRODUCCIÓN

Uno de los principales aspectos que determinan la permanencia en el mercado de cualquier negocio en la actualidad, sea este dedicado a la manufactura o a la prestación de servicios, sin duda alguna, es su nivel de calidad, ya que este aspecto está relacionado directamente con la satisfacción de sus clientes. Los productos o servicios de calidad ayudan a lograr clientes cautivos y por tanto aumentan la posibilidad de seguir haciendo negocios con ellos. Es por esto que en estos días la mejora de los niveles de la calidad de sus productos y servicios es un factor preponderante para las organizaciones.

Según detallan Gutiérrez y de la Vara (2004), uno de los elementos fundamentales del control de la calidad ha sido el enfocar mayores esfuerzos hacia la prevención en detrimento de la detección (inspección). Es decir, se movió de la simple inspección de los productos una vez ya elaborados, para separar los productos malos de los buenos, a una segunda etapa la cual consistió en el monitoreo de la calidad en el momento en que se elaboraban mediante el control estadístico de procesos y de esta a una tercera etapa, en la que se busca la calidad en la etapa de diseño de los productos y los procesos.

En cada una de las fases antes mencionadas ha sido preponderante el uso de herramientas estadísticas y es por esto que como mencionan Tuck, et al. (1993), la alta administración se ha convencido de que la inversión en experimentos planeados y el análisis estadístico es una parte esencial en el mejoramiento continuo de la calidad y es vital para la competitividad de sus compañías.

Dentro de estas técnicas estadísticas una de las más importantes para optimizar la calidad es la metodología de superficie de respuesta, la cual, según Montgomery (2005), 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. Para construir esta superficie de respuesta, se requiere de modelos polinomiales, los cuales se ajustan empleando técnicas como los mínimos cuadrados para obtener sus parámetros.

Cuando se trabaja con la optimización de productos, es probable que estos cuenten con varias características de calidad que se deban considerar en el análisis. Una alternativa es optimizar de forma independiente cada una de ellas, lo cual según diversos autores no es recomendable ya que se pueden proponer tantas combinaciones distintas de las variables de control, como variables de respuesta se consideren en el estudio, lo que generaría un conflicto al decidir entre ellas. La mejor forma de abordar esta situación es considerando todas las respuestas de forma simultánea, para lo cual se han propuesto diversas técnicas, algunas más completas respecto a las distintas consideraciones que se deben tener presentes durante la optimización. También se pueden encontrar alternativas con distintos niveles de complejidad técnica, con algunas de ellas incluso, ya contempladas dentro de la paquetería del software estadístico comercial, lo que contribuye a que sean las más conocidas y por tanto utilizadas, aunque no necesariamente sean las mejores opciones.

1.1 Antecedentes

Como ya se mencionaba anteriormente son diversas las opciones que han surgido para trabajar con el problema de la optimización simultánea de múltiples respuestas, siendo las más comunes las basadas en la función de pérdida, las que utilizan índices de capacidad de proceso, las que emplean una función de deseabilidad, la sobreposición de gráficas de superficie de respuesta y recientemente han surgido las técnicas que optimizan la probabilidad de concordancia, teniendo cada una de estas sus ventajas o desventajas con respecto a las otras. Algunas de estas técnicas surgieron para contrarrestar una debilidad que presentaban las otras opciones o incluso para complementar sus versiones anteriores evolucionando sobre la misma idea incorporando alguna modificación que las hiciera más completas.

Los aspectos que deben considerar las diferentes propuestas para llevar a cabo la optimización simultánea de las variables de respuesta son la correlación entre las variables de respuesta, debe también contemplar los factores de ruido que afectan las características de calidad analizadas y proponer condiciones de operación de los procesos que sean robustas a ellos, asimismo debe incorporar la incertidumbre en los parámetros de los modelos, todo esto sin perder de vista la simplicidad de la herramienta, por lo que la herramienta a utilizar, ha de seleccionarse entonces considerando algunos de estos elementos.

Han sido muchos los esfuerzos realizados por diversos investigadores para lograr diseñar una herramienta que comprenda la mayor cantidad posible de los aspectos antes mencionados y se puede encontrar mucha información al respecto en las diferentes revistas de investigación. Posteriormente en éste trabajo, se podrá encontrar información referente a las opciones más interesantes que han surgido hasta estos días. En este estudio se plantea modificar una técnica existente para la optimización simultánea de múltiples respuestas que ya considera los aspectos antes mencionados. Con dicha modificación se pretende hacerla más flexible, incorporándole la posibilidad de que permita introducir diferentes pesos a las variables de respuesta, de acuerdo a su grado de importancia en la calidad o el costo del producto final.

1.2 Planteamiento del Problema

Cuando se trabaja con la optimización de respuestas múltiples la mayoría de las técnicas existentes no toman en cuenta la incertidumbre en los parámetros estimados para los modelos utilizados en la optimización, la correlación entre las variables de respuesta y la robustez, de manera conjunta. En este trabajo se utilizará el enfoque bayesiano propuesto por Peterson (2004), el cual considera la correlación entre las respuestas de interés y la incertidumbre en la estimación de los parámetros del modelo de una manera formal. Este sugiere un cambio de las variables a optimizar, ya que sustituye las variables de respuesta originales por la probabilidad de concordancia simultánea, es decir, la probabilidad de que cada una estas respuestas cumplan con sus respectivas especificaciones al mismo tiempo, en cada condición de operación del proceso. Además se extiende el enfoque

anteriormente mencionado, al diseño robusto de parámetros al incorporar variables de ruido al análisis, como lo proponen Miró-Quesada et al. (2004), partiendo del trabajo elaborado por Peterson (2004).

La aportación que se hace con el presente trabajo tiene un impacto directo en la optimización de múltiples respuestas, ya que plantea una modificación de las propuestas mencionadas en el párrafo anterior. Esta modificación consiste en la posibilidad de incorporar una ponderación para cada una de las variables de respuesta, de acuerdo con la importancia que tenga cada una en la funcionalidad o costos de los productos, para atender una situación que pudiera presentarse en la práctica, en la cual no todas las variables consideradas son de igual importancia. La anterior modificación se lleva a cabo reemplazando como variable a optimizar la probabilidad de concordancia conjunta, por la media geométrica ponderada de las probabilidades de concordancia individuales para cada variable de respuesta. Al igual que en las técnicas tomadas como base se logra reducir la dimensión del problema de optimización, pasando de un problema p -dimensional a uno unidimensional.

1.2.1 Preguntas de Investigación

- ¿Cómo se comporta la metodología propuesta por Peterson (2004) y complementada por Miró-Quesada et al., (2004) si las variables de respuesta originales se sustituyen por la media geométrica ponderada de las probabilidades de concordancia individuales y no por la probabilidad de concordancia simultánea de todas las variables de respuesta?
- ¿Qué tipo de modelo se podrá utilizar para representar de manera adecuada la relación entre la media geométrica ponderada de las probabilidades de concordancia individuales y los factores de control?
- ¿Con qué técnica se podrá optimizar el modelo que relacione la media geométrica ponderada de las probabilidades de concordancia individuales y los factores de control?
- ¿Cuál será el efecto en las conclusiones respecto a las condiciones óptimas de operación del proceso si se ajusta el modelo a optimizar con respecto a la media geométrica ponderada para las probabilidades de concordancia individuales utilizando diferentes ponderaciones para las variables de respuesta?

1.2.2 Hipótesis

- La metodología propuesta por Peterson (2004) y complementada por Miró-Quesada et al. (2004) funciona correctamente si las variables de respuesta originales se sustituyen por la media geométrica ponderada de las probabilidades de concordancia individuales.
- La media geométrica ponderada se puede modelar utilizando regresión logística binaria, ya que será calculada a partir de probabilidades individuales, las cuales son el resultado de una

variable de respuesta con únicamente dos resultados posibles, evaluando si cada valor generado mediante simulación cumple o no con su respectiva especificación.

- Se puede utilizar el *MGRG* para optimizar el polinomio resultante de la regresión logística, lo que es deseable, ya que este permite la optimización restringida.
- La condición óptima resultante para operar el proceso será diferente si se consideran diferentes ponderaciones y tenderá hacia donde las variables de respuesta más importantes se comportan de la mejor manera, si se utiliza como variable de respuesta a optimizar la media geométrica ponderada de las probabilidades de concordancia individuales, dando ponderaciones más altas a las variables de respuesta más importantes.

1.2.3 Objetivos

1.2.3.1 Objetivo General

Adaptar la metodología propuesta por Peterson (2004) y complementada por Miró-Quesada et al. (2004), para que permita optimizar múltiples variables de respuesta simultáneamente en escenarios en donde no todas estas variables son igual de importantes.

1.2.3.2 Objetivos Específicos

- Mostrar mediante casos de estudio propuestos en la literatura que la metodología de referencia funciona adecuadamente, si se sustituye la probabilidad de concordancia conjunta, por la media geométrica ponderada de las probabilidades de concordancia individuales.
- Modelar mediante regresión logística la relación entre los factores de control y la media geométrica ponderada correspondiente.
- Mostrar que los resultados óptimos alcanzados en los distintos escenarios propuestos tienden a favorecer a las variables de respuesta con los mayores pesos o ponderaciones.

1.2.4 Justificación

Un problema común en el diseño de productos o procesos es la selección de niveles de los parámetros para optimizar múltiples respuestas simultáneamente, a lo que se le conoce como problema de multirespuestas (Myers y Montgomery, 1995). Ames et al. (1997), mencionan que la mayoría de los productos manufacturados tienen múltiples criterios de calidad, a su vez, Chiao y Hamada (2001), mencionan que para la mayoría de los productos la calidad es multidimensional, por lo que es común observar múltiples respuestas en las unidades experimentales, también mencionan que optimizar estas respuestas por separado puede llevar a resultados no satisfactorios. Es por esto que este trabajo trata de la optimización de múltiples características de calidad simultáneamente.

Además es un hecho que en muchas ocasiones existe diferencia en la relevancia de las diferentes variables de respuesta a la hora de optimizarlas simultáneamente ya que todas ellas pudieran afectar de diferentes formas la funcionalidad o los costos de los productos, por lo cual un buen método de optimización debería de considerar estas diferencias y asegurar que la condición óptima de operación para el proceso sea tal que se tenga un buen desempeño en estas características y quizá no tan bueno en las otras respuestas menos críticas. Este aspecto se aborda en este trabajo dando ponderaciones distintas a las respuestas dependiendo de su importancia, mediante la media geométrica ponderada.

El problema antes descrito se abordará utilizando métodos bayesianos considerando sus siguientes bondades descritas por Retzer (2006):

- Estos formalizan el proceso de aprendizaje modificando las creencias conforme se va recabando más información. Esta idea es atractiva pero es difícil de implementar con la estadística tradicional. La regla de Bayes provee el mecanismo para actualizar el conocimiento previo con los nuevos datos.
- Permite la estimación de modelos complejos, los cuales pueden no ser posibles usando las técnicas tradicionales.
- El resultado final de cualquier análisis bayesiano es una distribución del (los) parámetro(s) de interés (distribución a posteriori) y conociendo esta distribución de los parámetros nos da el acceso a toda la información que pudiera interesar acerca del parámetro.

El enfoque bayesiano no se había usado tan frecuentemente en años anteriores porque puede ser matemáticamente complejo, ya que estimar parámetros usando las distribuciones de probabilidad requiere frecuentemente de la integración y desafortunadamente muchas de las integraciones necesarias para el análisis bayesiano son demasiado complejas para resolverlas usando cálculo simbólico. Pero esta problemática ha desaparecido recientemente gracias a la simulación bayesiana, la cual es capaz de resolver mediante la computadora los más complejos problemas de integración. La simulación bayesiana es basada en modernas técnicas de muestreo conocidas como Cadenas de Markov Monte Carlo (CMMC). La combinación de estos dos fenómenos, innovación matemática (CMMC) y el creciente poder de la computadora, han permitido que el análisis bayesiano emerja como una solución práctica para ser utilizada en nuevas investigaciones.

2. REVISION DE LITERATURA

2.1 La Metodología de Superficie de Respuesta Dual

La metodología de superficie de respuesta (MSR) es una colección de técnicas estadísticas y matemáticas útiles para desarrollar, mejorar y optimizar procesos. Esta metodología también tiene importantes aplicaciones en el diseño, desarrollo y formulación de nuevos productos, así como en la mejora de los productos existentes (Myers y Montgomery, 1995).

La MSR es una herramienta importante en la ingeniería de calidad moderna. La idea básica es ajustar un modelo para la variable de respuesta y después explorar varios niveles de interés para las variables explicatorias. El principal énfasis ha sido en maximizar (o minimizar) el valor de la media (localización) de la respuesta Y . Taguchi enfatizó en la necesidad de desarrollar una metodología estadística que pudiera simultáneamente optimizar la media y la varianza de la característica que se está investigando. Esta necesidad se presenta en muchos problemas industriales para los cuales se requiere lograr de manera simultánea un valor objetivo y mantener una varianza pequeña (Ding, et al., 2004).

Algunos autores como Vining y Myers (1990), Copeland y Nelson (1996), Lin y Tu (1995), entre otros tratan de alguna forma el problema del desarrollo de análisis estadístico para múltiples respuestas tales como la media y la varianza de las características de calidad de interés. Así, el objetivo deberá ser determinar las condiciones de ajuste óptimo de las variables independientes (x 's), de manera que la respuesta (y) sea lo más cercana a un valor objetivo establecido y reduciendo al mismo tiempo su variación.

Vining y Myers (1990) proponen ajustar modelos polinomiales de segundo orden de manera separada, para optimizar la media muestral (w_μ) y la desviación estándar de la muestra (w_σ) de la siguiente manera:

$$w_\mu = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j}^k \beta_{ij} x_i x_j + \varepsilon_\mu \quad (2.1)$$

$$w_\sigma = \gamma_0 + \sum_{i=1}^k \gamma_i x_i + \sum_{i=1}^k \gamma_{ii} x_i^2 + \sum_{i < j}^k \gamma_{ij} x_i x_j + \varepsilon_\sigma \quad (2.2)$$

y establecen como procedimiento de optimización

$$\begin{array}{ll} \text{Minimizar} & \hat{w}_\sigma \\ \text{Sujeto a} & \hat{w}_\mu = T, \text{ donde } T \text{ es el valor objetivo} \end{array} \quad (2.3)$$

Por otra parte Lin y Tu (1995) proponen usar el criterio del cuadrado medio del error (MSE), de la siguiente manera:

$$\text{Minimizar } MSE = (\hat{w}_\mu - T)^2 + \hat{w}_\sigma^2 \quad (2.4)$$

A su vez, Copeland y Nelson (1996), especifican una desviación de w_μ de T , y establecen el modelo:

$$\begin{aligned} \text{Minimizar } & \hat{w}_\sigma \\ \text{Sujeto a } & (\hat{w}_\mu - T)^2 \leq \Delta^2 \end{aligned} \quad (2.5)$$

Mientras que Ding et. al, (2004) consideran que en el problema de respuesta dual, es deseable que la media sea cercana al valor objetivo y la varianza sea pequeña, y utilizan un modelo ponderando la media y la varianza en el criterio del cuadrado medio del error (WMSE), esto es:

$$\begin{aligned} \text{Minimizar } & WMSE = \lambda(\hat{w}_\mu(x) - T)^2 + (1 - \lambda)\hat{w}_\sigma(x) \\ \text{Sujeto a } & X \in \Omega \\ \text{donde } & \lambda \in [0,1] \end{aligned} \quad (2.6)$$

2.2 Optimización de Múltiples Variables de Respuesta

En el mercado tan altamente competitivo de hoy en día, las compañías son inducidas a mejorar constantemente la calidad de sus productos. Estos productos pueden involucrar el estudio simultáneo de varias respuestas teniendo cada una de ellas su importancia relativa para los clientes. Por lo tanto el reto es cómo determinar el ajuste adecuado para las variables de control, de tal forma que se obtengan los niveles deseados de calidad (Najafi, et al., 2011). En la optimización de estas múltiples respuestas, los investigadores buscan optimizar p características de calidad o variables de respuesta y como menciona Costa et al. (2011) esas respuestas se deben optimizar simultáneamente ya que analizarlas por separado puede resultar en soluciones incompatibles. Al respecto Pal y Gauri (2010), mencionan que frecuentemente las condiciones del proceso requeridas para dos o más variables de respuesta son contradictorias, por lo que generalmente es difícil encontrar un nivel de operación de las variables de control, que pueda resultar en valores cercanos a sus respectivos valores objetivo para todas las variables de respuesta, por lo que según mencionan, el objetivo de la optimización de múltiples respuestas es encontrar los valores para las variables de entrada con los cuales se logre un compromiso óptimo entre las respuestas.

En los últimos años se ha publicado una gran cantidad de artículos referentes a la optimización de múltiples respuestas y es un tópico común en la bibliografía referente a optimización de procesos y productos. Como resultado de ello, han surgido una amplia variedad de opciones en cuanto a metodologías y herramientas con este propósito, algunas similares entre ellas pero muy distintas a otras. Debido a esto, algunos autores tratan de agrupar estas propuestas de acuerdo a una

característica común, por ejemplo, Pal y Gauri (2010), clasifican estas propuestas en las siguientes tres categorías: (1) Técnicas matemáticamente rigurosas, (2) Índices de desempeño del proceso simplificados y (3) Técnicas basadas en inteligencia artificial. En su trabajo reportan algunas características propias de cada categoría.

Por otra parte Taleghani (2009), categoriza los métodos existentes para la optimización de múltiples respuestas en las siguientes cuatro categorías básicas: (1) sobreposición de gráficas de contorno, (2) problema de optimización restringida, (3) combinación de las respuestas en una sola y (4) enfoques que se basan en la proporción de concordancia. En su trabajo pueden encontrarse detalles de cada una de estas categorías.

A continuación se presentan algunos de los métodos de optimización múltiple más populares, agrupándolos de acuerdo con las categorías propuestas por Taleghani (2009).

2.2.1 Sobreposición de Gráficas de Contorno

Menciona Peterson (2004), que una forma simple de trabajar con múltiples superficies de respuestas cuando solo existe un pequeño número de factores es hacer una Sobreposición de las gráficas de contornos de las superficies de respuestas, lo cual puede indicar una región en el espacio de los factores en la que cada respuesta media satisface los requerimientos del experimentador o se puede obtener una solución compromiso.

En Montgomery (2005), se muestra esta técnica para la optimización de tres variables de respuesta: Rendimiento, Viscosidad y Peso molecular en un proceso químico. Después de recabar información de acuerdo a un diseño central compuesto, se ajustaron los siguientes modelos considerando los dos factores de control (X_1 : Tiempo y X_2 : Temperatura).

$$y_{\text{rendimiento}} = 79.94 + 0.99x_1 + 0.52x_2 - 1.38x_1^2 - 1.00x_2^2 + 0.25x_1x_2$$

$$y_{\text{viscosidad}} = 70.00 - 0.16x_1 - 0.95x_2 - 0.69x_1^2 - 6.69x_2^2 - 1.25x_1x_2$$

$$y_{\text{pesomolecular}} = 3386.2 + 205.1x_1 + 17.4x_2$$

Con estos modelos se procede a construir las respectivas superficies de respuesta y gráficas de contorno, para el rendimiento, la viscosidad y el peso molecular, las cuales se muestran en las figuras 2.1 y 2.2, 2.3 y 2.4, y 2.5 y 2.6 respectivamente. Los requerimientos para las variables de respuesta son:

$$y_1 \geq 78.5, \quad 62 \leq y_2 \leq 68, \quad y_3 \leq 3400.$$

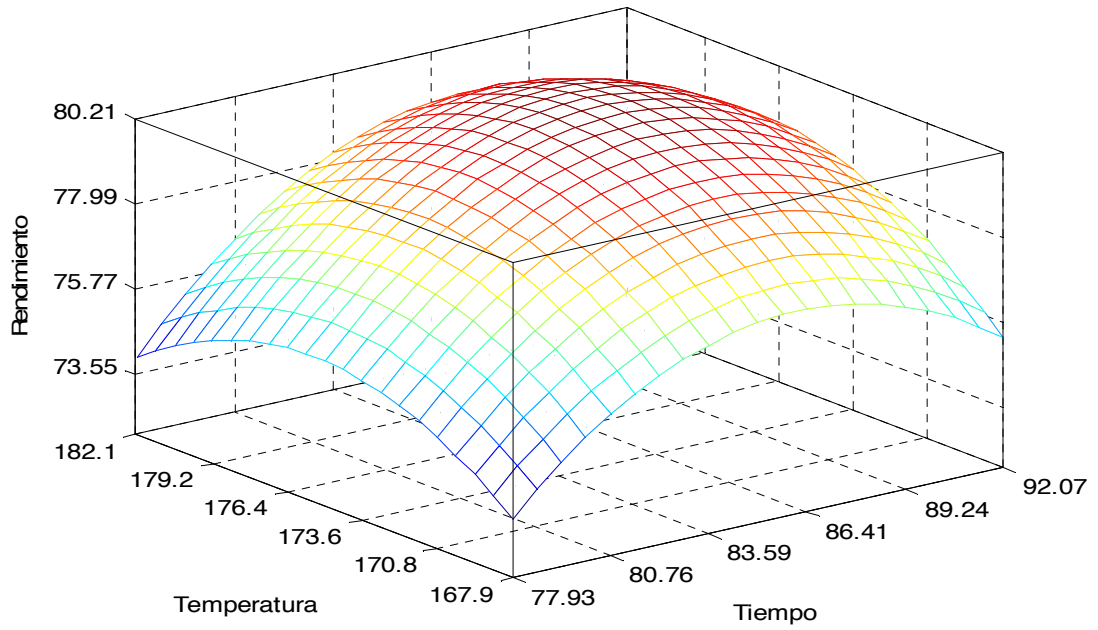


Figura 2.1 Superficie de Respuesta para el Rendimiento

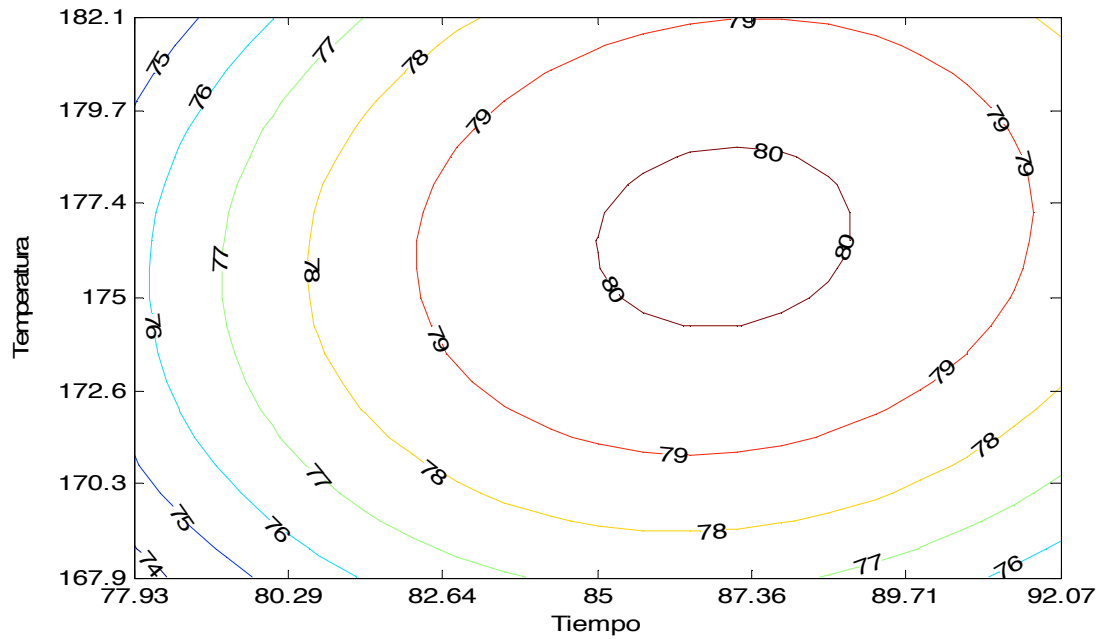


Figura 2.2 Gráfica de Contorno para el Rendimiento

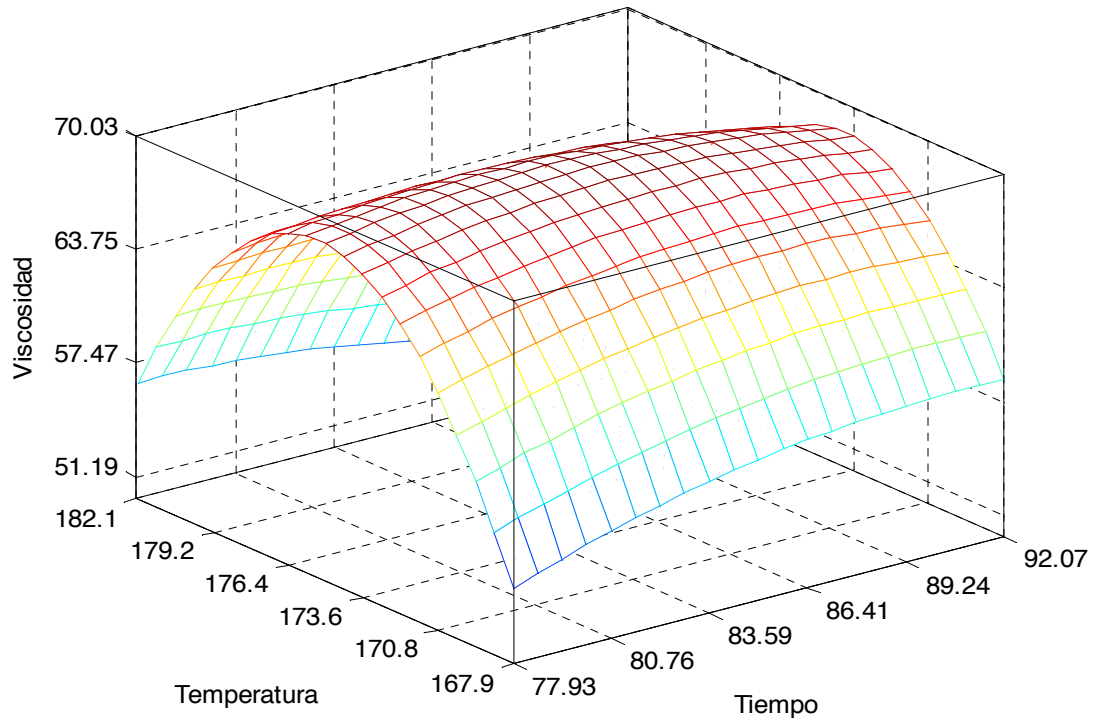


Figura 2.3 Superficie de Respuesta para la Viscosidad

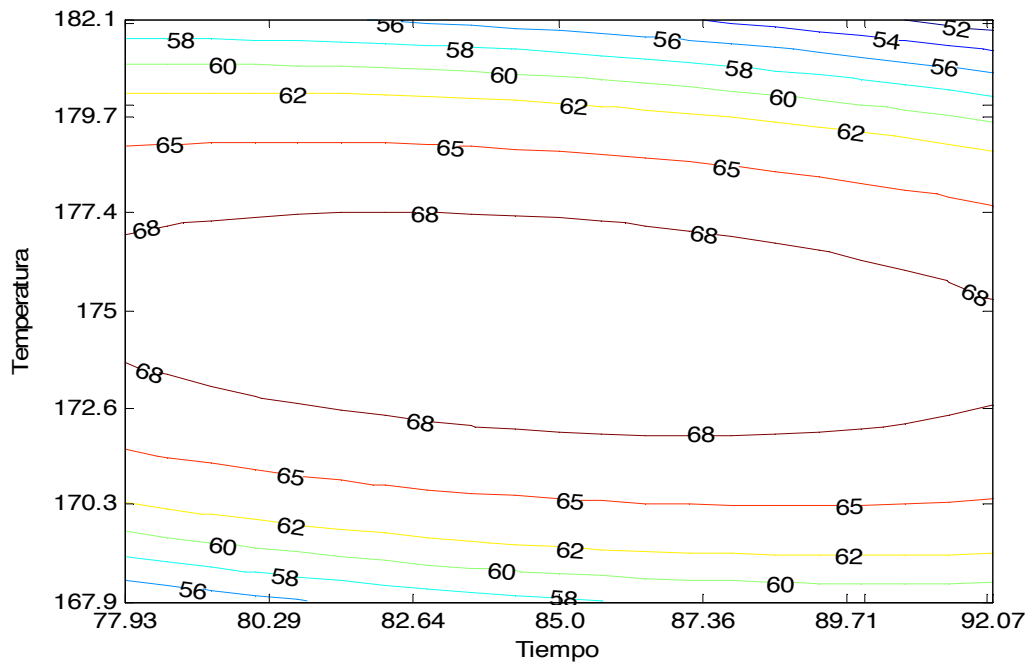


Figura 2.4 Gráfica de Contorno para Viscosidad

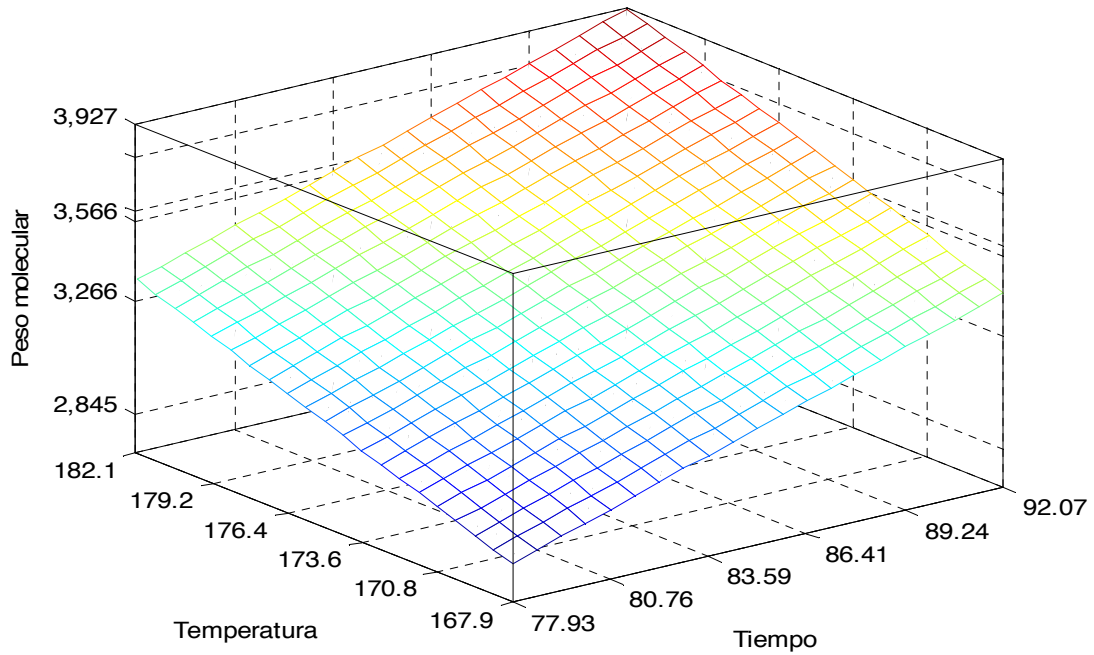


Figura 2.5 Superficie de Respuesta para el Peso Molecular

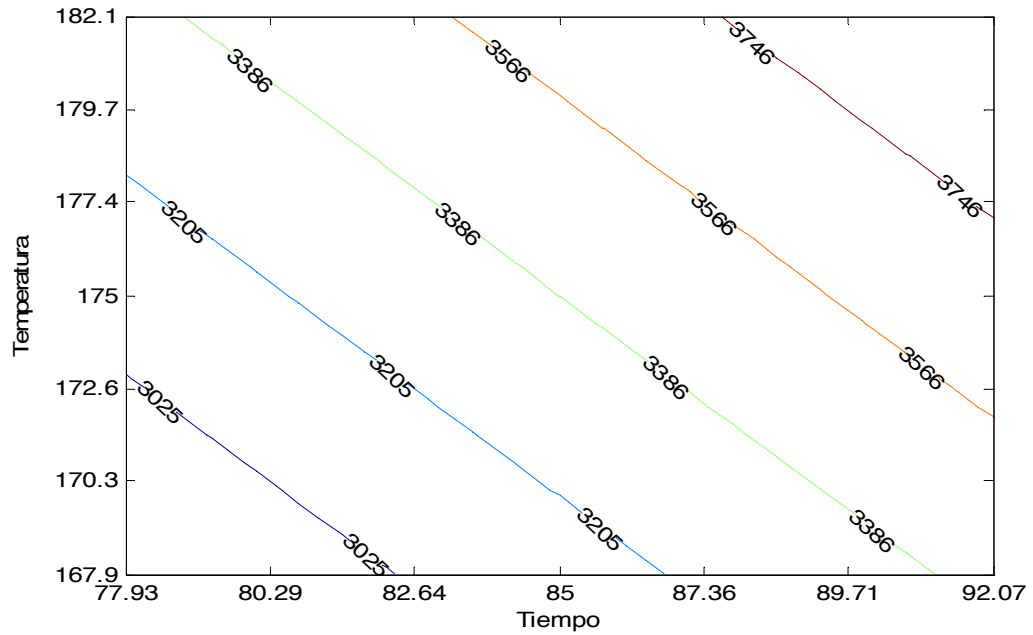


Figura 2.6 Gráfica de Contorno para el Peso Molecular

Una vez que se tienen las tres gráficas de contorno, estas se superponen en una sola para buscar la zona en la que se cumple con las especificaciones simultáneamente. Esta Sobreposición se muestra en la Figura 2.7 y las regiones en las que se cumple con los requisitos simultáneamente son las que aparecen de color verde y como puede apreciarse, son varias las combinaciones de tiempo y

temperatura en las que la respuesta media para las tres variables de respuesta cumplen con sus respectivas especificaciones.

Como puede verse esta técnica es relativamente sencilla y directa, pero también tiene algunos inconvenientes, por ejemplo, Montgomery (2005) menciona que cuando hay más de tres variables de diseño, se hace muy complicada la Sobreposición de las gráficas de contorno, ya que esta es bidimensional, y $k-2$ de las variables de control deben de mantenerse constantes para construir la gráfica. Además también menciona que con frecuencia se necesita una gran cantidad de ensayo y error para determinar cuáles son los factores que deben mantenerse constantes y que niveles seleccionar para obtener la mejor vista de la superficie, por lo que existe interés práctico en métodos de optimización más formales para las respuestas múltiples.

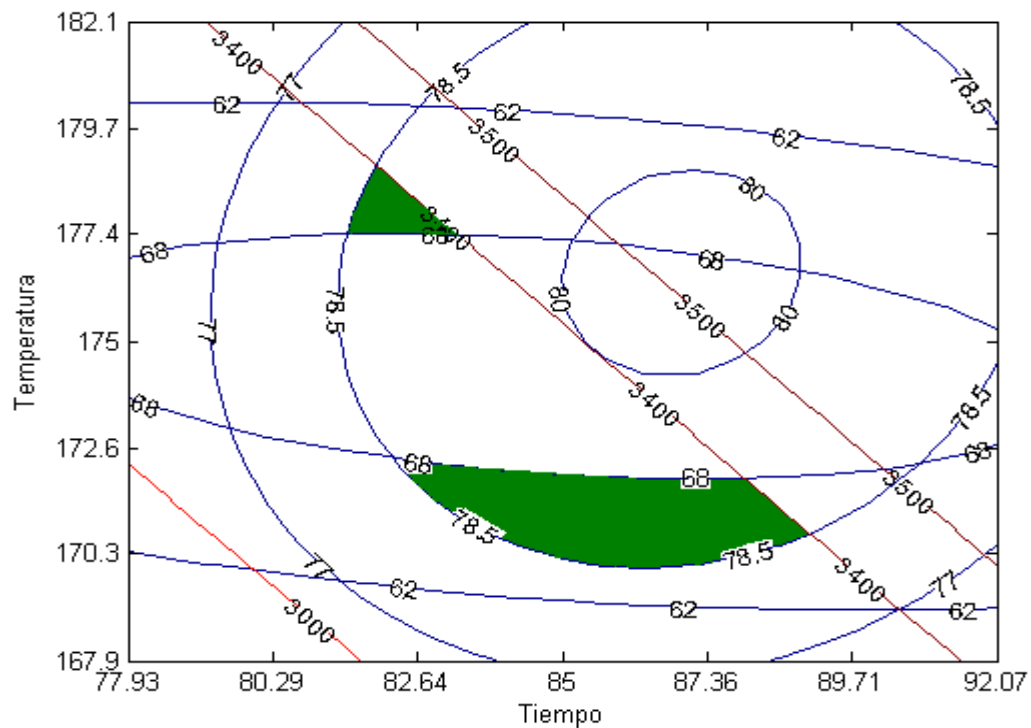


Figura 2.7 Sobreposición de las Gráficas de Contorno del Rendimiento, la Viscosidad y el Peso Molecular.

Adicionalmente en la sobreposición de gráficas de contorno, tampoco es posible considerar algunos aspectos clave en la optimización simultánea, como lo es la correlación entre las variables de respuesta, la incertidumbre en los parámetros de los modelos y la robustez.

2.2.2 Problema de Optimización Restringida

En este enfoque se selecciona la variable de respuesta más importante y se utiliza como función objetivo, mientras que las respuestas restantes se consideran como restricciones.

Esta propuesta puede resumirse de la siguiente manera:

Optimizar Respuesta principal

S.a. Requerimientos para las otras respuestas

$$x \in \Omega$$

De lo anterior puede decirse que es posible que escoger una de las respuestas como función objetivo no sea una tarea sencilla. Además, este enfoque no utiliza la idea básica de considerar todas las respuestas simultáneamente.

2.2.3 Combinación de las Respuestas en una Sola

Un enfoque popular para la optimización de múltiples respuestas ha sido el uso de la estrategia de la reducción de su dimensión. Esta estrategia convierte un problema de múltiples respuestas, en uno con una sola medición, el cual se resuelve como un problema de optimización de un objetivo simple (Ko et al., 2005). Entre estas propuestas surgen como las más utilizadas, las que emplean la función de deseabilidad, la función de pérdida o las que trabajan con índices de capacidad del proceso multivariable. Posteriormente, esta única variable de respuesta que de alguna manera resume la información referente a todas las variables originales, se modela y optimiza generalmente con la ayuda de técnicas de métodos numéricos. A continuación se muestra de manera general algunas propuestas de esta categoría.

2.2.3.1 Función de Deseabilidad

Derringer y Suich (1980) proponen utilizar funciones de deseabilidad las cuales convierten el problema de múltiples respuestas, en un problema de una sola respuesta; esto es, la respuesta que se analiza es la deseabilidad total $D = (d_1(Y_1) \cdot \dots \cdot d_m(Y_m))^{1/m}$, donde (Y_1, \dots, Y_m) son las m respuestas y d_1, \dots, d_m son las deseabilidades individuales. Para convertir las respuesta y_i a d_i , estos autores, proponen la siguiente transformación:

$$d_i = \begin{cases} \left[\frac{\hat{y}_i(x) - EI_i}{T_i - EI_i} \right]^s & EI_i \leq \hat{y}_i(x) \leq T_i \\ \left[\frac{\hat{y}_i(x) - ES_i}{T_i - ES_i} \right]^t & T_i \leq \hat{y}_i(x) \leq ES_i \\ 0 & \hat{y}_i(x) < EI_i \quad \text{o} \quad \hat{y}_i(x) > ES_i \end{cases} \quad (2.7)$$

Donde s y t son exponentes que sirven para elegir la forma deseada de la transformación y con ello reflejar los deseos del experimentador: si se toman valores grandes (por ejemplo $s, t \geq 10$) significa

que la deseabilidad d_i sólo toma valores grandes cuando \hat{y}_i cae cerca de su valor objetivo; si se toman valores pequeños para s y t ($s, t \leq 0.1$) significa que cualquier valor de \hat{y}_i dentro del intervalo $[E_i, ES_i]$ es igualmente deseable. Lo anterior puede apreciarse en la figura 2.8.

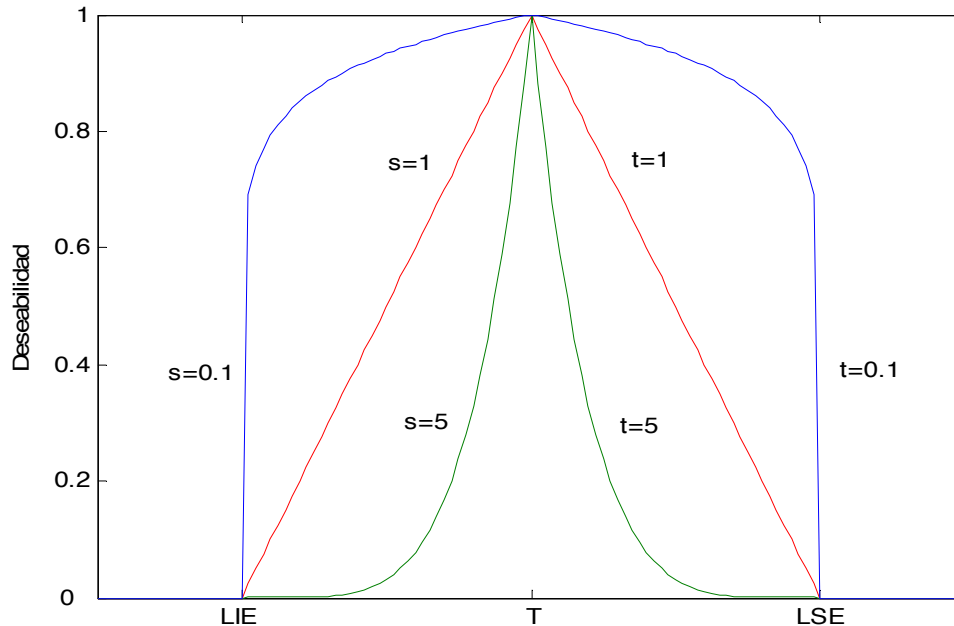


Figura 2.8 Función de Deseabilidad

Según Gutiérrez y De la Vara (2012), el valor por omisión de estos exponentes es uno, lo que sugiere un incremento lineal de la deseabilidad hacia el valor objetivo (véase figura 2.8). Además si la respuesta tiene especificaciones de un solo lado, lo que se hace es tomar el valor objetivo (T) igual al valor a partir del cual se considera que no hay ganancia adicional a la calidad de la respuesta, es decir, en la función 3.7, desaparece una de las restricciones y la figura 2.8, se reduce a uno de los lados del valor objetivo (T).

El optimizar múltiples variables de respuesta mediante la función de deseabilidad según Chiao y Hamada (2001) tiene dos desventajas: primero, D puede ser difícil de modelar ya que es una función compleja de las m respuestas; segundo, es difícil decir que es lo que significa la diferencia entre valores esperados de D excepto que el mayor es el mejor. Otra desventaja de este enfoque según Ko et al. (2005) es que no considera la estructura de varianzas-covarianzas de las respuestas y según mencionan ignorar esta información puede llevar a una solución no real, si en realidad las respuestas tienen varianzas significativamente diferentes o si están altamente correlacionadas. Se han propuesto otros enfoques basados en la deseabilidad para corregir algunos de los detalles mencionados anteriormente de la propuesta original, véase por ejemplo, Del Castillo et al. (1996), Ortiz et al. (2004), Ch'ng, Quah y Low (2005) y Lee y Kim (2007).

2.2.3.2 Función de Pérdida

Ames et al. (1997) proponen las funciones de pérdida para optimizar múltiples características de calidad. Mencionan que las funciones de pérdida proveen un marco conceptual que permite combinar criterios diferentes en una función objetivo única. Una función de pérdida describe la pérdida económica de la sociedad, provocada por las desviaciones de las características con respecto a su valor objetivo, debida a variaciones aleatorias y errores sistemáticos. La pérdida total a la sociedad debida a las desviaciones en todas las características del producto se puede tomar como la suma de las pérdidas debidas a las características individuales.

Según Ko et al. (2005), la mayor ventaja de la función de pérdida es su habilidad para incorporar la estructura de varianzas-covarianzas de las respuestas, así como la economía de los procesos. En su trabajo Ames et al. (1997) obtienen sus funciones de pérdida de la siguiente manera:

Sea V_r la medición de una variable de respuesta para una cierta característica de calidad r y T_r el valor objetivo o ideal para V_r . Cada resultado V_r puede variar de su valor objetivo T_r , debido a la variación aleatoria o al error sistemático. La función de pérdida de la calidad global (PCG) describe la pérdida de la sociedad que resulta de estas desviaciones. Bajo esta lógica usan una aproximación cuadrática para esta función de pérdida, de la siguiente manera:

$$PCG = \sum_{r=1}^{R_g} W_r \{V_r - T_r\}^2 \quad (2.8)$$

Donde los factores de ponderación (W_r) escalan la importancia relativa para los diferentes criterios y su suma se toma sobre todas las respuestas (R_g) que contribuyen a la calidad del producto.

Se puede tratar con la reducción de la función de pérdida mediante la elección apropiada de las variables controlables de entrada de un proceso (X_1, X_2, \dots) utilizando un modelo que relacione estas variables con las V_r respuestas de la siguiente manera:

$$V_r = Y_r(X_1, X_2, \dots) + e_r \quad (2.9)$$

Donde e_r es el término del error que incorpora tanto el componente aleatorio como el sistemático. Las funciones Y_r utilizadas son modelos polinomiales ajustados mediante el uso de experimentos diseñados. En particular estos modelos pueden ser modelos cuadráticos de superficie de respuesta (MCSR). Una vez construidos estos modelos la función de pérdida (2.8) se puede reemplazar por $PCGP$ (pérdida de calidad global polinomial):

$$PCGP = \sum_{r=1}^R W_r \{Y_r(X_1, X_2, \dots) + e_r - T_r\}^2 \quad (2.10)$$

Donde R es el número de características de calidad para las que se puede construir un modelo de superficie de respuestas Y_r (el cual será un polinomio de grado 2 en las variables de control X_i) y puede ser menor a R_g . Finalmente si los errores e_r , son pequeños comparados con los errores con respecto a los valores objetivo $(Y_r - T_r)$, es posible desprestigiar e_r y entonces PCGP se puede reescribir como:

$$PCP = \sum_{r=1}^R W_r \{Y_r(X_1, X_2, \dots) - T_r\}^2 \quad (2.11)$$

Estas funciones de pérdida se deben minimizar para obtener las mejores condiciones de operación del proceso ya que esto minimizaría la pérdida del cliente o de la sociedad. Según Ames et al. (1997), al minimizar la función de pérdida además de la calidad más alta, también se produce la mayor estabilidad o robustez en el proceso contra las variaciones en la calidad debida a las variables de entrada. Para opciones de la función de pérdida más completas, es decir, que consideren más aspectos relevantes para la optimización simultanea, refiérase a Ko et al. (2005) y Romano et al. (2004).

2.2.3.3 Índices de Capacidad de Procesos

2.2.3.3.1 Índice C_{pm}

Ch'ng et al. (2005) proponen modificar el índice C_{pm} de tal forma que se obtiene la siguiente expresión:

$$C_{pm} = \frac{\min(LSE - T, T - LIE)}{3\sqrt{(\hat{\omega}_\mu - T)^2 + \hat{\omega}_\sigma^2}} \quad (2.12)$$

Entre más grande sea el valor de este índice mejor será la configuración del proceso. Si lo que se desea es optimizar varias características de calidad al mismo tiempo Ch'ng et al. (2005) sugieren una modificación extra resultando la siguiente expresión:

$$Total \quad C_{pm} = \sum_{i=1}^p e_i \frac{\min(LSE_i - T_i, T_i - LIE_i)}{3\sqrt{(\hat{\omega}_{i\mu} - T_i)^2 + \hat{\omega}_{i\sigma}^2}} \quad (2.13)$$

Cuando están disponibles el límite superior de especificación (LSE), el valor objetivo y el límite inferior de especificación (LIE) la ecuación anterior se puede utilizar para el caso "nominal es lo mejor". Pero además esta ecuación también permite la optimización de características de calidad unilaterales. Por

ejemplo cuando solo se conocen el valor objetivo y el LIE se procede haciendo el $LSE = \infty$. Así las ecuaciones 2.12 y 2.13 quedan de la siguiente manera respectivamente:

$$Cpm = \frac{T - LIE}{3\sqrt{(\hat{\omega}_\mu - T)^2 + \hat{\omega}_\sigma^2}} \quad (2.14)$$

$$Total \quad Cpm = \sum_{i=1}^p e_i \frac{T_i - LIE_i}{3\sqrt{(\hat{\omega}_{i\mu} - T_i)^2 + \hat{\omega}_{i\sigma}^2}} \quad (2.15)$$

Además cuando solo se dispone del LSE y el valor objetivo se puede hacer $LIE = -\infty$ resultando:

$$Cpm = \frac{LSE - T}{3\sqrt{(\hat{\omega}_\mu - T)^2 + \hat{\omega}_\sigma^2}} \quad (2.16)$$

$$Total \quad Cpm = \sum_{i=1}^p e_i \frac{LSE_i - T_i}{3\sqrt{(\hat{\omega}_{i\mu} - T_i)^2 + \hat{\omega}_{i\sigma}^2}} \quad (2.17)$$

Para optimizar estas expresiones en base a el ajuste del proceso se pueden utilizar métodos numéricos como el *MGRG*, el cual es utilizado por Ch'ng et al. (2005) en su trabajo. Otra alternativa es hacer programas que busquen en todas las combinaciones posibles de las variables la que sería la óptima con la desventaja de que el tiempo de procesamiento sería más lento, pero con la ventaja de que no arrojaría óptimos locales.

2.2.3.3.2 Índices Mcpm y Mcpk

Plante (2001) propone un índice de capacidad del proceso llamado *MCpk*, el cual es igual a la media geométrica de las medidas de comportamiento *Cpk's*, para que sea utilizado como un criterio para evaluar el diseño de productos o procesos con múltiples respuestas. Expresa el índice *MCpk* como una función de los factores de diseño $X=(X_1, X_2, \dots, X_n)$ de la siguiente manera:

$$MCpk = \left\{ \prod_{i=1}^m \text{Min} \left[\frac{(\mu_i - LS_i)}{3\sigma_i}, \frac{(US_i - \mu_i)}{3\sigma_i} \right] \right\}^{(1/m)} \quad (2.18)$$

Donde:

σ_i = la desviación estándar del *i-ésimo* término del error, $\sigma_{\epsilon i}$.

Notar que cuando $m=1$, este resultado coincide con el índice Cpk .

Según Plante (2001), la filosofía fundamental detrás del índice $MCpk$ tiene profundas implicaciones en la calidad. Esto es, si una o más de las medidas de desempeño es inadecuada (índice cercano a cero), entonces el producto completo se vuelve inaceptable, sin importar los índices de capacidad de las restantes medidas de desempeño. De hecho el índice $MCpk$ es fuertemente influenciado por las características de calidad que tienen una capacidad baja. Además este criterio es mas universalmente aplicable, puesto que no requiere de un valor objetivo para cada medida de desempeño.

Similar al índice $MCpk$, en el índice $MCpm$ también se utiliza la media geométrica de las medidas de desempeño Cpm individuales. El criterio de capacidad $MCpm$ se expresa de la siguiente manera:

$$MCPM = \left\{ \prod_{i=1}^m \left[\frac{(US_i - LS_i)}{6\sqrt{[\sigma_i^2 + (\mu_i - T_i)^2]}} \right] \right\}^{(1/m)} \quad (2.19)$$

Donde: T_i es igual al valor objetivo de la i -ésima característica de calidad.

Sin embargo, el índice $MCpm$, como el índice Cpm , requiere de valores objetivo para cada característica de calidad, por lo que no es tan flexible con el índice $MCpk$. Posteriormente se utilizan formulaciones de programación no lineal para maximizar tanto el índice $MCpm$ como el índice $MCpk$.

Plante (2001), reconoce que en algunas ocasiones no todas las características de calidad son igual de importantes, por lo que según dice, puede ser más importante lograr mayores índices de capacidad en una o más características a expensas de algunas otras. Es por esto que también propone un enfoque que considera la importancia relativa de las características de calidad.

2.2.4 Enfoques que se Basan en la Proporción de Concordancia

2.2.4.1 Método de Chiao y Hamada

Según mencionan Chiao y Hamada (2001), cuando se analizan datos de múltiples respuestas, analizarlas por separado no ofrece resultados satisfactorios, especialmente cuando estas variables están correlacionadas. El realizar análisis separados puede llevar a conclusiones contradictorias respecto a los niveles de los factores importantes, ya que ciertos niveles de las variables pueden mejorar la calidad de alguna respuesta, pero empeorar la de otras. En el método que proponen se considera la matriz de varianzas-covarianzas dependiente de los factores experimentales y el criterio de decisión es la probabilidad de que todas las respuestas satisfagan simultáneamente sus respectivas especificaciones, por lo que la configuración óptima del proceso será aquella que maximice esta probabilidad de que las variables de respuesta estén en la región S , lo cual se simboliza de la siguiente manera:

$$P(Y \in S) \quad (2.20)$$

Debido a que según Chiao y Hamada (2001) la proporción de concordancia se determina por las medias, varianzas y correlaciones (ρ) de la distribución multivariada de las respuestas, proponen construir modelos para estos parámetros de la siguiente manera:

$$\mu_k = x\alpha_k, \quad k = 1, \dots, m, \quad (2.21)$$

$$\log(\sigma_k^2) = x\beta_k, \quad k = 1, \dots, m, \quad (2.22)$$

$$\tanh^{-1}(\rho_{kl}) = x\gamma_{kl}, \quad 1 \leq k < l \leq m. \quad (2.23)$$

En las ecuaciones anteriores \mathbf{x} denota un vector fila de covariables asociadas con un conjunto de factores los cuales pueden incluir una constante, efectos principales, interacciones dobles, etcétera. Con los modelos anteriores se obtienen diferentes valores de los parámetros para diferentes niveles de los factores de interés (X 's), los cuales se sustituyen en la función de densidad de probabilidad conjunta Y , la cual se considera una normal multivariada m dimensional con media $\mu = (\mu_1, \mu_2, \dots, \mu_m)^T$ y matriz de varianza-covarianza Σ , $f(Y; \mu, \Sigma)$ dada por:

$$f(Y; \mu, \Sigma) = (2\pi)^{-(m/2)} |\Sigma|^{-1/2} \exp\left[-\frac{1}{2}(Y - \mu)' \Sigma^{-1} (Y - \mu)\right] \quad (2.24)$$

El enfoque bayesiano propuesto por Peterson (2004), también considera la probabilidad de que el proceso cumpla con los requisitos de calidad de las características del producto a optimizar consideradas en el análisis.

2.3 La Inferencia Bayesiana

De acuerdo con Jackman (2004) el teorema de Bayes lleva ese nombre por el reverendo Thomas Bayes quien murió en 1761. Este teorema apareció por primera vez en un ensayo atribuido a Bayes y comunicado a la Real Sociedad después de su muerte por Richard Price en 1763, el cual se ha vuelto a publicar en muchas ocasiones. Bayes solamente anunció el resultado para una previa uniforme. En 1774 Laplace, aparentemente sin conocimiento del trabajo de Bayes, anunció el teorema en su forma más general.

Berger (2000) hace un recuento de la proliferación en los últimos años de las actividades referentes a la difusión de la inferencia bayesiana. Menciona que de 1974 a 1994 la tendencia era lineal con aproximadamente el doble de artículos cada 10 años. Sin embargo en los últimos cinco años (con referencia a 1999, que fue cuando se publicó este artículo) ha habido un aumento muy dramático tanto en la cantidad como en la tasa de incremento de los artículos bayesianos. Este mismo fenómeno también es visible en el número de libros escritos sobre este análisis. Menciona que

durante los primeros 200 años de (1769 - 1969), hubo quizá unos 15 libros al respecto. En los años de 1970 a 1989 el número de libros fue de 30. De 1990 a 1999, aproximadamente 60 libros han sido escritos, sin contar las muchas docenas de conferencias, proceedings y colección de artículos sobre el tema.

Por análisis bayesiano de datos se entiende, métodos prácticos para hacer inferencia de datos usando modelos de probabilidad para cantidades observadas y cantidades de las cuales deseamos aprender. La característica esencial de los métodos bayesianos es su uso explícito de la probabilidad para cuantificar la incertidumbre en inferencias basadas en el análisis estadístico de datos (Gelman, et. al, 1995).

Withers (2002) menciona que en la última década, la emergencia de los métodos bayesianos en las ciencias y la estadística ha ocurrido como una fiebre y refiere que están teniendo un impacto fundamental en prácticamente todas las metodologías estadísticas.

La definición bayesiana de probabilidad ajusta el término probabilidad al lenguaje cotidiano, y así la teoría de probabilidad bayesiana parece estar relacionada de una manera mucho más cercana al razonamiento intuitivo de las personas en la presencia de incertidumbre (Lecoutre, 2006). El análisis bayesiano a diferencia del enfoque frecuentista es de naturaleza secuencial (Wooley, 2004). A este respecto Retzer (2006) comenta que la estimación bayesiana consiste en la actualización de nuestras creencias actuales con nueva información procedente de los datos, lo cual describe el proceso fundamental de aprendizaje utilizado tanto en la ciencia como en la vida diaria.

En algunas ocasiones resulta adecuado considerar a un parámetro θ desconocido como el valor de una variable aleatoria de una distribución de probabilidad dada. Esto se presenta cuando antes de observar los resultados de los datos X_1, \dots, X_n , se tiene alguna información acerca del valor de θ y dicha información se puede expresar en términos de distribución de probabilidad, acertadamente llamada distribución a priori de θ (Ross, 2002). Esta distribución previa la cual se supone que representa lo que se conoce acerca de los parámetros desconocidos antes de tener datos disponibles, asume un papel importante en el análisis bayesiano. Dicha distribución se puede utilizar para representar el conocimiento previo o la ignorancia relativa (Box y Tiao, 1992).

Según comenta Retzer (2006) el análisis bayesiano es una reformulación de la teoría estadística básica, no una rama de estadísticos. Esto significa que cualquier cosa que se haga con la estadística tradicional también puede ser hecha con los métodos bayesianos. En realidad, la flexibilidad fundamental de este enfoque permite hacer muchas cosas que no son posibles con el análisis estadístico tradicional.

2.3.1 El Teorema de Bayes

A continuación se presenta el teorema de Bayes para eventos y posteriormente para densidades según Colosimo y Del Castillo (2007).

2.3.1.1 El Teorema de Bayes para Eventos

Para los eventos A y B en algún espacio muestral S , partiendo de la definición de la probabilidad condicional tenemos,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2.25)$$

y

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad (2.26)$$

En donde estas expresiones serán verdaderas si y solo si

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

Por lo cual,

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)} \quad (2.27)$$

Esta fórmula muestra la esencia del Teorema de Bayes: si el evento B representa alguna información adicional disponible, entonces $P(A|B)$ es la probabilidad después de incorporar esta información, es decir, la *posterior* de A y $P(A)$ es la probabilidad antes de que esté disponible la información, es decir, la *previa* de A .

2.3.1.2 El Teorema de Bayes para Densidades

Si y representa datos y θ algún parámetro o vector de parámetros, entonces de la definición de densidad condicional, se tiene que,

$$p(y|\theta) = \frac{p(\theta, y)}{p(\theta)} \quad (2.28)$$

y

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} \quad (2.29)$$

Lo cual implica que la densidad conjunta es

$$p(\theta, y) = p(\theta)p(y | \theta) = p(y)p(\theta | y),$$

lo cual implica

$$p(\theta | y) = \frac{p(\theta)p(y | \theta)}{p(y)} \tag{2.30}$$

En donde el denominador tanto en el caso discreto como en el continuo según Box y Tiao (1992), sería:

$$c^{-1} = \begin{cases} p(y) = \int p(y | \theta)p(\theta)d(\theta) & \text{caso continuo} \\ p(B) = \sum_{\text{todo } A} p(B | A)p(A) & \text{caso discreto} \end{cases}$$

Por lo que c es meramente una constante de normalización necesaria para asegurar que la distribuciones posteriores $P(\theta | y)$ o $P(A | B)$ integran o suman uno respectivamente.

Según Retzer (2006) una descripción intuitiva de los pasos básicos necesarios en el análisis bayesiano se muestra enseguida:

- Se empieza el análisis con el conocimiento previo acerca de los parámetros antes de recolectar datos.
- Luego, se recolectan los datos y se representan en términos de lo que se conoce como función de verosimilitud.
- Finalmente se actualiza el conocimiento previo usando la información contenida en los datos recabados. Esta actualización de la información se refleja en la distribución *a posteriori* de los parámetros de interés. La actualización implica multiplicar *a priori* veces la verosimilitud.

Finalmente, respecto a la importancia de la estadística bayesiana, Berger (2000) redacta que es importante reconocer que el genio no se puede regresar a la botella. La “máquina” bayesiana junto con las cadenas de Markov y la simulación Monte Carlo (CMMC), es posiblemente el mecanismo más poderoso jamás creado para el procesamiento de datos y conocimiento.

2.4 Las Cadenas de Markov con Simulación Monte Carlo (CMMC) y el Muestreo de Gibbs

Según Jackman (2004), los métodos basados en CMMC probablemente son el desarrollo más excitante en la estadística en los últimos años. Según Geyer (1992), la idea básica de las CMMC es

muy simple. Si no es posible encontrar una manera de simular datos independientes de algún proceso estocástico complicado, casi tiene la misma utilidad poder simular datos dependientes x_1, x_2, \dots que formen una cadena de Markov irreducible con la distribución $f(x)$ de interés como la distribución estacionaria. El algoritmo de Metropolis – Hastings proporciona estas cadenas. Según Chib y Greenberg (1995) este algoritmo es extremadamente versátil y presenta el muestreo de Gibbs como un caso especial.

Mencionan Chib y Greenberg (1995) que el muestreo de Gibbs es uno de los métodos más conocidos de las CMMC para simular distribuciones multivariadas complejas sin una forma estándar, cuyo impacto en la estadística bayesiana ha sido inmenso. Según Casella y George (1992), el muestreo de Gibbs es una técnica para generar variables aleatorias de una distribución (marginal) indirectamente, sin tener que calcular su densidad, que se basa únicamente en las propiedades elementales de las cadenas de Markov. Para la inferencia bayesiana el muestreo de Gibbs se utiliza principalmente para generar distribuciones posteriores y su utilidad se incrementa grandemente conforme se incrementan las dimensiones del problema, debido a que evita el cómputo de integrales que pueden ser prohibitivas.

Mencionan Hamada et al. (2010), que generalmente un mejor método de CMMC para muestrear de una distribución posterior se puede obtener reemplazando las densidades propuestas genéricas del algoritmo de Metropolis-Hastings por la distribución condicional del elemento que compone el vector de parámetros que se está muestreando. Los algoritmos CMMC que siguen esta estrategia se nombran muestreadores de Gibbs. Para describir un muestreador de Gibbs, suponga que el vector de parámetros θ se puede particionar en q componentes $\theta = (\theta_1, \theta_2, \dots, \theta_q)$, y que las distribuciones posteriores condicionales (distribuciones condicionales completas) se pueden escribir como:

$$\begin{aligned} & p_1(\theta_1 | \theta_2, \dots, \theta_q, \text{datos}) \\ & p_2(\theta_2 | \theta_1, \theta_3, \dots, \theta_q, \text{datos}) \\ & \vdots \\ & p_q(\theta_q | \theta_1, \dots, \theta_{q-1}, \text{datos}). \end{aligned}$$

En el caso de simple de $q = 2$, la densidad de probabilidad p_1 representa la densidad posterior del componente θ_1 , condicionado con el valor del componente θ_2 . Asimismo, p_2 es la densidad posterior condicional de θ_2 dado θ_1 .

En muchos modelos, es difícil simular directamente del vector completo de parámetros θ . Sin embargo frecuentemente es posible generar valores simulados de cada una de las densidades condicionales completas p_1, p_2, \dots, p_q . Cuando este es el caso, el muestreo de Gibbs se puede definir de acuerdo con el algoritmo que se muestra en la figura 2.9.

Casella y George (1992), reportan el siguiente ejemplo para ilustrar el funcionamiento del muestreo de Gibbs. Suponer que interesa calcular algunas características de la distribución marginal $f(x)$ de X , para la siguiente distribución conjunta de X y Y :

$$f(x, y) \propto \binom{n}{x} y^{x+\alpha-1} (1-y)^{n-x+\beta-1}, \quad (2.31)$$

$$x = 0, 1, \dots, n \quad 0 \leq y \leq 1$$

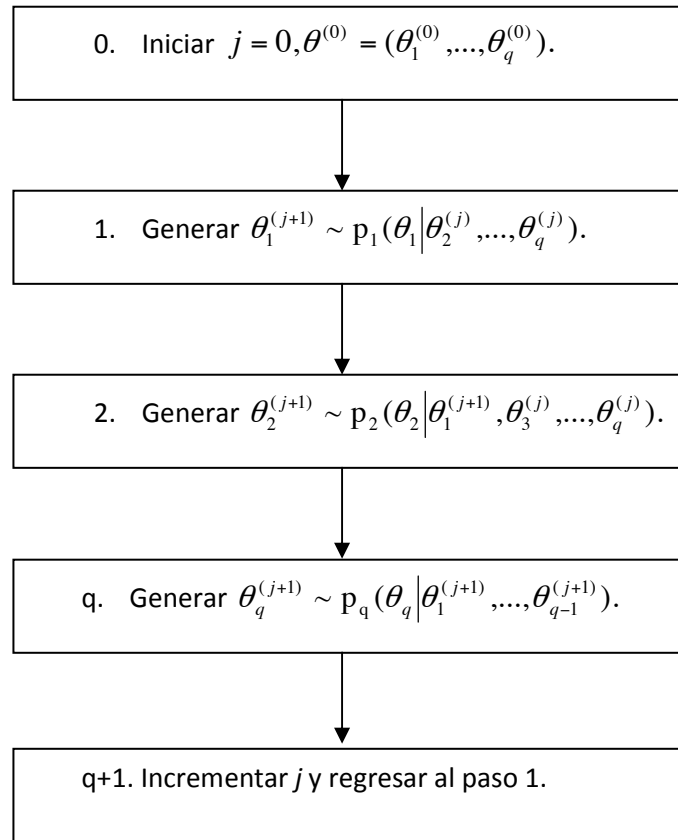


Figura 2.9 Algoritmo del Muestreo de Gibbs para Muestrear de una Distribución Posterior de un Vector de Parámetros q -Dimensional.

El muestreo de Gibbs permite generar muestras de esta marginal de la siguiente manera. De la ecuación 2.31 se sigue que:

$$f(x | y) \text{ es binomial } (n, y) \quad (2.32)$$

$$f(y | x) \text{ es Beta } (x + \alpha, n - x + \beta)$$

Para obtener información de esta variable X , es necesario generar una secuencia de Gibbs de variables aleatorias de la siguiente manera:

$$Y'_0, X'_0, Y'_1, X'_1, Y'_2, X'_2, \dots, Y'_K, X'_K \quad (2.33)$$

Donde el valor inicial $Y'_0 = y'_0$ se especifica de antemano y el resto de la secuencia es obtenido iterativamente, generando de forma alternada valores según las ecuaciones en 2.34.

$$\begin{aligned} X'_j &\sim f(x | Y'_j = y'_j) \\ Y'_{j+1} &\sim f(y | X'_j = x'_j) \end{aligned} \quad (2.34)$$

Resulta que bajo condiciones razonablemente generales, la distribución de X'_K converge hacia $f(x)$ (la distribución marginal verdadera de X) conforme $k \rightarrow \infty$. Así, para k lo suficientemente grande, las observaciones finales en 2.33, es decir, $X'_K = x'_k$ son efectivamente una muestra de $f(x)$. Una forma de obtener las muestras es tomando solo los últimos valores generados en la secuencia de tamaño k como resultados para la variable X y repetir este proceso m veces.

Si ahora se aplica el esquema iterativo 2.34 a las distribuciones en 2.32, es posible generar una muestra X_1, X_2, \dots, X_m de $f(x)$ y utilizar esta muestra para calcular cualquier característica deseada. En este caso en específico el muestreo de Gibbs no es necesario, ya que se puede notar que $f(x)$ se puede obtener analíticamente utilizando la ecuación 2.35, la cual es la distribución beta – binomial y con ella se pueden obtener las características de $f(x)$ directamente. Sin embargo este caso es útil para ilustrar como el muestreo de Gibbs funciona adecuadamente, ya que las muestras obtenidas mediante el algoritmo iterativo llevan prácticamente a los resultados obtenidos con la distribución exacta cuando el número de iteraciones es lo suficientemente grande.

$$\begin{aligned} f(x) &= \binom{n}{x} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(x + \alpha)\Gamma(n - x + \beta)}{\Gamma(\alpha + \beta + n)}, \\ &x = 0, 1, \dots, n \end{aligned} \quad (2.35)$$

En la figura 2.10 se muestran los resultados obtenidos mediante el muestreo de Gibbs partiendo de dos valores iniciales distintos y mediante la distribución beta – binomial utilizando los valores $n = 16$, $\alpha = 2$ y $\beta = 4$. En esta figura se puede apreciar que los resultados alcanzados mediante el muestreo son prácticamente los mismos independientemente del valor inicial, y a su vez, casi idénticos a los obtenidos con la distribución exacta, con lo que queda demostrado que los valores generados X_1, X_2, \dots, X_m mediante el muestreo de Gibbs se pueden considerar como una muestra de $f(x)$. Este hecho es de suma importancia en problemas de mayor complejidad. En esta figura el

histograma en azul se construyó partiendo del valor inicial $y_0 = 0.35$, el histograma en verde con $y_0 = 0.65$ utilizando en ambos casos 50000 simulaciones y $k = 15$. El histograma de color rojo se obtuvo con las probabilidades exactas de la distribución beta-binomial.

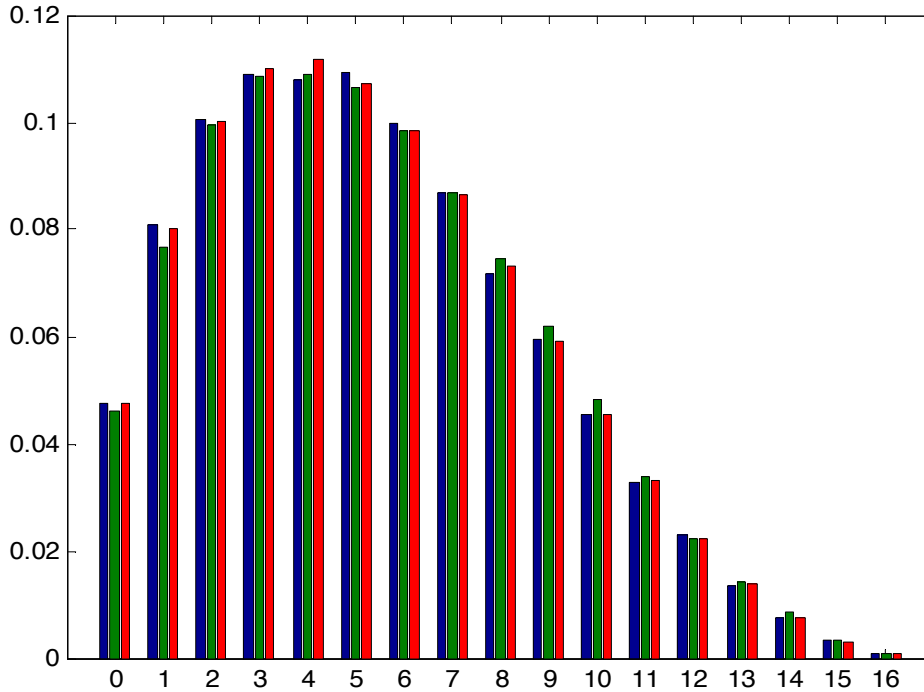


Figura 2.10 Comparación de Tres Histogramas de Probabilidad de la Distribución Beta - Binomial con $n = 16$, $\alpha = 2$ y $\beta = 4$.

2.5 La Media Geométrica

Según Spiegel y Stephens (2002) la media geométrica (Mg) de un conjunto de N números positivos $x_1, x_2, x_3, \dots, x_N$ es la raíz N -ésima del producto de esos números:

$$G = \sqrt[N]{X_1 X_2 X_3 \dots X_N} \quad (2.36)$$

Los mismos autores utilizan también la fórmula:

$$G = \sqrt[N]{X_1^{f_1} X_2^{f_2} \dots X_k^{f_k}} \quad (2.37)$$

donde $N = \sum f$. A esto se le denomina media geométrica ponderada (Mgp). Según los Spiegel y Stephens (2002), la Mg es más adecuada que la media aritmética para promediar proporciones.

2.6 El Enfoque Bayesiano para Optimizar Múltiples Variables de Respuesta

Esta propuesta es un híbrido entre las categorías 3 y 4 mencionadas previamente, ya que al inicio mediante simulación calcula probabilidades de concordancia para las variables de respuesta analizadas, de acuerdo con sus especificaciones y posteriormente esta probabilidad de concordancia general reemplaza a las variables de respuesta originales, convirtiendo un problema de optimización de múltiples respuestas en un problema de una sola variable de respuesta.

Miró-Quesada et al. (2004) mencionan que el problema de la optimización de respuestas múltiples consiste en escoger los valores de los k factores controlables x_i , tal que y tenga ciertas propiedades deseadas. Frecuentemente se presenta el caso en que estas propiedades son que el vector de respuestas y satisfaga las especificaciones. Digamos que A representa la región del espacio de respuestas definido por estas especificaciones. La región A puede tener una forma arbitraria. Entonces la optimización de respuestas múltiples se puede resumir en un simple objetivo, el cual consiste en maximizar la probabilidad de tener el vector de respuestas dentro de la región de especificación A , por ejemplo, maximizar la siguiente probabilidad de concordancia:

$$\begin{aligned} \max p(x) &= p(y \in A | x, \text{datos}) \\ \text{sujeto a:} & \\ x &\in R \end{aligned} \tag{2.38}$$

Donde $y = B'x + \varepsilon$ y R es la región en donde el modelo es válido, la cual normalmente se toma como la región experimental definida por la matriz de diseño X .

Según Peterson (2004) partiendo del modelo típico de regresión multivariada se tiene que:

$$Y = Bx + e \tag{2.39}$$

Donde B es una matriz $p \times q$ de los coeficientes de regresión y x es el vector $q \times 1$ de valores de las variables controlables, en los que se evalúa la ecuación 2.39. Además el vector e tiene una distribución normal multivariada con vector de medias igual a $\mathbf{0}$ y matriz de varianzas-covarianzas Σ . Para considerar la incertidumbre en los parámetros del modelo B y Σ se puede usar la densidad posterior predictiva $f(y | x, \text{datos})$. Utilizando la distribución previa conjunta no informativa clásica para B y Σ y el modelo de la ecuación 2.39, la densidad predictiva bayesiana para y se puede obtener en forma cerrada. Esta densidad predictiva para el vector de valores x en los que se evaluará la función, está dada por una distribución t multivariada con ν grados de libertad (gl), de la siguiente manera:

$$f(y|x, \text{datos}) = \frac{\Gamma(\frac{\nu+p}{2})}{(\pi\nu)^{p/2} \Gamma(\frac{\nu}{2})} \sqrt{|H|} \left\{ 1 + \frac{1}{\nu} (y - \hat{B}'x)' H (y - \hat{B}'x) \right\}^{-\frac{\nu+p}{2}} \quad (2.40)$$

$$H = \left(\frac{\nu}{n-q} \right) \frac{\sum^{-1}}{1 + x'(X'X)^{-1}x}$$

$$p(x) = \frac{\# \text{ de valores en } A}{\# \text{ generados}} \quad (2.41)$$

Aquí p es la cantidad de respuestas a optimizar, \mathbf{X} es la matriz $q \times n$ formada por los $n(x_i)$ vectores de covariables, $\hat{\mathbf{B}}$ es la estimación por mínimos cuadrados de \mathbf{B} , $\Gamma(\cdot)$ se refiere a la función gamma, $\hat{\Sigma}$ es el estimador usual de Σ calculado con los residuales del ajuste de la regresión multivariada, $\nu = n - p - q + 1$ y n es el tamaño de muestra.

Según Peterson (2004), debido a que la ecuación 2.40, sigue una distribución t multivariada, es fácil simular los valores Y , de esta densidad predictiva. Jhonson (1987) en su libro *Multivariate statistical simulation* menciona que se puede simular una variable aleatoria t - *variada* Y , mediante la simulación de una variable aleatoria normal multivariada y una variable aleatoria chi-cuadrada independiente (Peterson, 2004). Si W es una variable aleatoria normal con vector de medias igual a cero y una matriz de varianzas-covarianzas igual a H^{-1} y además U es una variable aleatoria chi-cuadrada que es independiente de W , entonces:

$$Y_j = (\sqrt{\nu}W_j / \sqrt{U}) + \hat{\mu}_j, \quad \text{para } j=1, \dots, p \quad (2.42)$$

Donde Y_j es el j -ésimo elemento de Y , W_j es el j -ésimo elemento de W y $\hat{\mu}_j$ es el j -ésimo elemento de $\hat{\mu} = \hat{B}z(x)$. De lo anterior menciona Peterson (2004), se sigue que Y sigue una distribución t multivariada con ν grados de libertad. Es en el instante de la simulación mediante la ecuación 2.42 que se considera la robustez, ya que como proponen Miró-Quesada et al. (2004) una vez que se genera el vector de respuestas en una combinación específica de los factores de control, mientras estos permanecen constantes, los factores de ruido se generan de una distribución de probabilidad que se supone siguen, por lo que de existir una combinación de los factores de control que a pesar del comportamiento de los factores de ruido resulte en una probabilidad de concordancia alta, se puede decir que esta condición de operación del proceso es robusta a los factores de ruido.

3. MATERIALES Y MÉTODOS

En este capítulo se detallan los equipos y el software utilizado en el desarrollo del presente trabajo, así también, se da una descripción breve de las etapas o fases que conforman la metodología empleada para resolver los casos de estudio analizados. En el siguiente capítulo se ilustrará con dos ejemplos cada una de las fases de esta metodología.

3.1 Materiales

El equipo y el software que se utilizaron en el desarrollo de este trabajo fueron los siguientes:

Equipo: Computadora Lap top Sony Vaio con procesador Intel Core 2 Duo, 2.40GHz y 2.0 GB de RAM para el modelado, simulación, optimización y finalmente para la redacción de la memoria o tesis. Respecto al software se requirieron los siguientes:

Minitab 15 en español: El cual es un software ampliamente utilizado para realizar análisis estadístico de un conjunto de datos por su facilidad para usarse, prueba de su popularidad es que existe una gran cantidad de artículos científicos que reportan haber basado sus hallazgos en los resultados arrojados por este software ya sea total o parcialmente. De la misma manera es posible encontrar bastantes libros escritos basándose en su uso. Este paquete computacional fue requerido para:

- 1.- El ajuste de los modelos de regresión para las variables de respuesta utilizando el método de mínimos cuadrados ordinarios en la fase III de la metodología propuesta.
- 2.- El ajuste de los modelos logísticos requeridos en la fase VI, para modelar la relación entre las variables de control y la media geométrica de las probabilidades encontradas en la fase V.

Matlab 6.0: Para poder utilizar este paquete computacional primeramente fue necesario tomar un curso sobre su lenguaje y lógica de programación, ya que en este caso es necesario programar para poder realizar algunas tareas específicas de esta metodología de optimización, que no se encuentran previamente cargadas en ningún software comercial. Se requirió matlab para:

- 1.- Elaborar el programa que se utilizó para la simulación de la distribución predictiva, con el cual se obtuvo el cálculo de las probabilidades de concordancia individuales y sus respectivas medias geométricas ponderadas, siendo estas últimas utilizadas como las variables de respuesta a optimizar. Este programa se muestra en el anexo 1, en el cual se encuentra la información referente al caso de estudio I y se indican las modificaciones necesarias para hacer la corrida de validación. Para utilizar este programa para el caso de estudio II, solo es necesario cargar su propia información.
- 2.- Realizar las corridas de validación en la condición óptima encontrada según el paquete Solver de Excel. Esto se realiza en la fase IX de la metodología.

Excel XP: En cuanto al paquete Excel, se requirieron tanto su hoja de cálculo tradicional, como la paquetería de Solver. Respecto a la hoja de cálculo esta fue requerida para:

1.- Pronosticar la *Mgp* en las distintas combinaciones de los factores de control consideradas en la malla, utilizando los modelos logísticos dentro de la fase VII.

2.- La optimización en la fase VIII utilizando el paquete Solver, el cual previamente debe ser activado, ya que no viene en los paquetes que se abren en automático, pero sí está cargado dentro del Excel solo que hay que activarlo entrando a la opción *Add-Ins* dentro de la ventana de *Excel options* que aparece cuando se acciona el logotipo de Windows que se encuentra en la pantalla principal de Excel.

Word XP: Este último paquete se utilizó para la redacción del documento de tesis.

3.2 Métodos

La metodología empleada para llevar a cabo la optimización en ambos casos de estudio, constó de las siguientes etapas:

Etapas I: Identificar las variables de control y de ruido, así como las variables de respuesta que se considerarán en el estudio. Este punto no es único para este trabajo, ya que este es un aspecto general a considerar en todo experimento diseñado.

Etapas II: Planear y realizar el experimento de tal forma que la información obtenida permita ajustar modelos para las diferentes variables de respuesta a optimizar. Se recomienda que el experimento diseñado sea tal que permita ajustar modelos de segundo orden en cada una de las variables de respuesta.

Etapas III: Ajustar modelos de regresión multivariada utilizando mínimos cuadrados ordinarios y verificar que tengan un ajuste adecuado ($R^2 > 70\%$). Si alguna variable de respuesta no cumple con este requisito se saca del estudio. Es importante mencionar en este punto que en cada modelo se deben emplear el mismo grupo de covariables.

Etapas IV: Determinar la cantidad de combinaciones de los factores de control a considerarse en la simulación y generar en cada una de estas combinaciones mediante la ecuación 2.42, los vectores de respuestas necesarios para el cálculo de las probabilidades de concordancia. Con estas combinaciones se deberá formar una malla lo suficientemente fina que permita ajustar un modelo logístico, utilizando como variables exploratorias las variables controlables y la *Mgp* como variable de respuesta. Es también en esta etapa en la que se atiende la robustez, debido a que para cada una de las combinaciones de las variables controlables, mientras estas permanecen fijas en la simulación de los vectores de respuestas, la o las variables de ruido toman valores distintos en cada iteración para formar el vector de covariables (x_c, x_r). Es decir, durante la simulación, mientras los valores de X_c permanecen constantes, los valores X_r se generan de una distribución de probabilidad específica bajo

la cual se supone que están distribuidas dichas variables de ruido. Por esta razón se puede decir que si de entre las combinaciones de las variables de control existe alguna que de una probabilidad de concordancia alta, también se está asegurando de que esta combinación será robusta a cambios en las variables de ruido.

Etapa V: Calcular las probabilidades de concordancia individuales y la *Mgp* con las ponderaciones que se consideren adecuadas, utilizando los vectores de respuestas generados en la etapa anterior. Para calcular estas probabilidades se utilizan tantos contadores en el programa de matlab como variables de respuesta se estén analizando, es decir, un contador por cada variable de respuesta. Cuando se lleva a cabo la simulación en cada una de las combinaciones de los factores de control, en cada vector que se genera, cada una de las respuestas es comparada contra su especificación y su respectivo contador se incrementa en uno si cumple con ella o permanece igual en caso contrario. De esta forma cuando se han completado las iteraciones cada contador se divide entre el número de iteraciones efectuadas, con lo que se obtienen las estimaciones de las probabilidades de que cada una de las respuestas cumpla con sus especificaciones para dichas condiciones de operación del proceso. Partiendo de estas estimaciones se obtienen las *Mgp* utilizando las ponderaciones para las variables de respuesta que se juzguen convenientes, de acuerdo a la importancia de cada una de ellas, dando ponderaciones más altas a las variables más importantes.

Para la simulación y cálculos realizados en las etapas IV y V se escribió un programa en matlab tomando como referencia el propuesto por Del Castillo en su página web (Del Castillo, 2006).

Etapa VI: Ajustar el modelo logístico para la *Mgp* utilizando como predictores los factores de control con las combinaciones de la malla construida en la etapa IV, aumentando el orden en cada uno de ellos y agregando interacciones según sea requerido para formar el vector de regresores completo. Como ya se mencionó, en los casos de estudio mostrados como ilustración, se utilizó minitab 15 para ajustar estos modelos.

Etapa VII: Utilizando el modelo ajustado en la etapa anterior, pronosticar la *Mgp* en cada una de las combinaciones de los factores de control consideradas en la simulación y construir un gráfico de líneas para comparar los resultados obtenidos con la simulación y el modelo logístico. Si el modelo se ajusta de manera adecuada a los datos simulados, el comportamiento de ambas gráficas debiera ser muy parecido, es decir, las dos series de resultados debieran prácticamente estar sobrepuestas. Previamente se verifica la significancia de los términos incluidos en el modelo según la salida de minitab.

Etapa VIII: Toda vez que se ha verificado que el modelo logístico representa de forma adecuada la información contenida en la simulación, entonces es posible optimizar (maximizar) el modelo logístico resultante utilizando el paquete Solver de Excel para interpolar y buscar el resultado óptimo entre cualquier combinación factible de los factores de control. Es necesario recordar en este punto, que en la malla creada con anterioridad solo se consideró un número limitado de combinaciones de los factores de control por lo que posteriormente es necesario interpolar entre estas combinaciones

en la búsqueda de un mejor resultado, lo cual se logra mediante la optimización del modelo logístico utilizando métodos numéricos. Para llevar a cabo esta optimización se empleó el *MGRG* el cual permite minimizar o maximizar un polinomio sujeto a ciertas restricciones, las cuales deben consistir en acotar los valores de las variables de control a la región en la que se consideraron en el experimento diseñado inicial, esto para evitar la extrapolación. Según Chapra y Canale (2011) este es uno de los métodos directos más populares. Para proceder con la búsqueda del punto óptimo, el algoritmo se inició dando diferentes valores a las variables de control, para ver en cuantos resultados distintos convergía. En específico se consideraron 15 escenarios distintos, es decir, 15 combinaciones distintas de las variables de control.

Etapa IX: En las condiciones propuestas en el paso anterior para cada uno de los escenarios considerados, calcular la *Mgp* de las probabilidades según el modelo logístico y con una simulación lo suficientemente grande y verificar que ambos resultados sean cercanos para comprobar su consistencia. Posteriormente, seleccionar como la mejor condición aquella que tenga la *Mgp* más alta. En esta condición óptima realizar varias simulaciones para poder apreciar de mejor manera el efecto de las ponderaciones en los resultados obtenidos para cada variable de respuesta, es decir, el efecto que estas tuvieron en sus respectivas probabilidades de concordancia.

Etapa X: Concluir de acuerdo a los resultados alcanzados en la etapa anterior.

Como se mencionó al inicio de este apartado en el siguiente capítulo se mostrarán dos casos de estudio que se abordan siguiendo en cada uno de ellos, las etapas mencionadas anteriormente, por lo que quedará más claro el trabajo realizado en cada una de ellas.

4. ANÁLISIS DE RESULTADOS

En este capítulo se muestra el desarrollo de la metodología antes mencionada, utilizando dos casos de estudio tomados de la literatura. En el primer caso se utilizan los datos de Miró-Quesada, et al. (2004), mientras que en el segundo se utilizan datos reportados por Myers y Montgomery (1995). A continuación se presenta el primer ejemplo, para mostrar nuestra propuesta.

4.1 Caso de Estudio I: Análisis de Sistema HPLC

Fases I y II: En este ejemplo se aborda la optimización de un sistema de cromatografía líquida de alto rendimiento (HPLC), para detectar mezclas de impurezas. El análisis del desempeño se basó en cuatro variables de respuesta cuantitativas, la resolución crítica (RC), tiempo total de la corrida (TC), razón señal a ruido del último pico (S/N) y el factor de asimetría del pico mayor (As). Se utilizaron tres factores controlables que afectaban el análisis HPLC, %IPA (x_r), temperatura (x_1) y pH (x_2), pero al igual que en Miró-Quesada et al. (2004), el factor controlable %IPA se consideró como variable de ruido (x_r) y además se asumió que esta normalmente distribuido con media igual a cero y una desviación estándar de 0.1. El diseño utilizado es un Box-Behnken con tres puntos centrales. Los datos originales se presentan en la tabla 4.1, mientras que los datos transformados ortogonalmente se presentan en la tabla 4.2. Para construir este diseño se requirió de 3 diseños 2^2 , cada uno de ellos considerando una combinación entre dos de los tres factores analizados, mientras el tercero permanece fijo en su nivel medio y por último se le agregaron tres puntos centrales, es decir, tres corridas con todos los factores en su nivel medio. Respecto a la transformación ortogonal esta se logra codificando los niveles bajo, medio y alto de las variables naturales, utilizando la notación -1, 0 y 1 respectivamente, con lo cual puede comprobarse que los vectores columna son ortogonales entre sí, ya que el producto punto entre cualquier par es igual a cero.

Fase III: Se ajustaron modelos de superficie de respuesta de segundo orden completos para las cuatro respuestas usando las versiones codificadas de los factores, al igual que en el trabajo de referencia. Los estadísticos R^2 resultantes fueron mayores de 99.6% para los cuatro modelos, lo que indica que estos modelos se ajustan muy bien a los datos. Los modelos ajustados para las cuatro variables de respuesta y_{RC} , y_{TC} , $y_{S/N}$ y y_{AS} , así como sus respectivos valores R^2 se muestran en las ecuaciones 4.1, 4.2, 4.3 y 4.4 respectivamente.

Tabla 4.1 Datos Originales del Experimento de la Cromatografía

Factores			Respuestas			
%IPA	Temp	pH	RC	TC	S/N	As
65	30	0.175	2.14	22	172	0.76
65	50	0.175	1.73	12	311	0.88
65	40	0.05	1.93	16	251	0.8
65	40	0.3	1.95	16	241	0.8
70	40	0.175	2.17	14	278	0.79
70	50	0.05	1.97	11	371	0.86
70	30	0.3	2.38	19	194	0.74
70	50	0.3	1.98	11	360	0.86
70	30	0.05	2.37	18	204	0.74
70	40	0.175	2.20	14	280	0.78
75	40	0.3	2.42	13	314	0.78
75	30	0.175	2.61	17	223	0.73
75	50	0.175	2.14	10	410	0.85
75	40	0.05	2.42	12	324	0.78
70	40	0.175	2.20	14	281	0.79

Tabla 4.2 Datos Transformados Ortogonalmente del Experimento de la Cromatografía

Factores			Respuestas			
%IPA	Temp	pH	RC	TC	S/N	As
-1	-1	0	2.14	22	172	0.76
-1	1	0	1.73	12	311	0.88
-1	0	-1	1.93	16	251	0.8
-1	0	1	1.95	16	241	0.8
0	0	0	2.17	14	278	0.79
0	1	-1	1.97	11	371	0.86
0	-1	1	2.38	19	194	0.74
0	1	1	1.98	11	360	0.86
0	-1	-1	2.37	18	204	0.74
0	0	0	2.20	14	280	0.78
1	0	1	2.42	13	314	0.78
1	-1	0	2.61	17	223	0.73
1	1	0	2.14	10	410	0.85
1	0	-1	2.42	12	324	0.78
0	0	0	2.20	14	281	0.79

$$y_{RC} = 2.190 + 0.230x_r - 0.210x_1 + 0.005x_2 - 0.015x_r x_1 - 0.005x_r x_2 - 0.000x_1 x_2 - 0.015x_r^2 - 0.020x_1^2 + 0.0057x_2^2 \quad \text{con } R^2 = 99.7\% \quad (4.1)$$

$$y_{TC} = 14.000 - 1.750x_r - 4.000x_1 + 0.250x_2 + 0.750x_r x_1 + 0.250x_r x_2 - 0.250x_1 x_2 + 0.375x_r^2 + 0.875x_1^2 - 0.125x_2^2 \quad \text{con } R^2 = 99.7\% \quad (4.2)$$

$$y_{S/N} = 279.667 + 37.000x_r + 82.375x_1 - 5.125x_2 + 12.000x_r x_1 - 0.000x_r x_2 - 0.250x_1 x_2 - 0.2083x_r^2 - 0.4583x_1^2 + 3.0417x_2^2 \quad \text{con } R^2 = 100\% \quad (4.3)$$

$$y_{AS} = 0.78667 - 0.0125x_r + 0.0600x_1 + 0.0000x_2 - 0.0000x_r x_1 - 0.0000x_r x_2 - 0.0000x_1 x_2 + 0.0042x_r^2 + 0.0142x_1^2 - 0.0008x_2^2 \quad \text{con } R^2 = 99.6\% \quad (4.4)$$

Fase IV: En esta etapa se debe determinar la cantidad de combinaciones de los factores de control a simularse, en cada una de las cuales se ha de calcular las probabilidades de concordancia individuales y posteriormente su respectiva *Mgp*. Utilizando las ponderaciones que se juzgue conveniente para cada respuesta. En este caso se decidió dividir en once niveles cada uno de los factores de control x_1 y x_2 utilizando para ello los valores:

{-1.0, -0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.8, 1.0}

Lo que da lugar a 121 combinaciones diferentes, en cada una de las cuales se llevó a cabo la simulación de 5000 vectores de respuestas mediante la ecuación 2.42. Estas combinaciones se muestran la tabla 4.3.

Fase V: Con los vectores simulados se calcularon las probabilidades de concordancia individuales y posteriormente la *Mgp* de acuerdo con los pesos indicados en la tabla 4.4, con la cual se sustituyen las variables de respuesta originales, reduciendo con ello la dimensión del problema. En la tabla 4.5 se muestran los resultados de la simulación del diseño 11², es decir, de las 121 combinaciones resultantes entre las variables x_1 y x_2 , utilizando el programa hecho en matlab. En estos datos se puede apreciar que los valores más altos tanto para la media geométrica normal como para la media geométrica ponderada de las probabilidades de concordancia individuales se encuentran en la zona que comprende a los valores 0.4 y 0.6 de la variable x_1 , por lo que habría de esperarse que la región óptima resultante se ubique entre estos valores. Los resultados de la columna *Mg* se calcularon utilizando la siguiente fórmula:

$$Mg = \sqrt[4]{p_1 * p_2 * p_3 * p_4} = p_1^{0.25} * p_2^{0.25} * p_3^{0.25} * p_4^{0.25} \quad (4.5)$$

Mientras que los valores de la columna *Mgp* se calcularon utilizando la ecuación:

$$Mgp = p_1^{0.1} * p_2^{0.4} * p_3^{0.4} * p_4^{0.1} \quad (4.6)$$

En las ecuaciones anteriores los exponentes se refieren a los pesos dados a cada una de las variables de respuesta, los cuales se indican en la tabla 4.4.

Como puede observarse por los pesos de la columna *Mgp*, se le está dando mayor importancia a las variables y_2 y y_3 , ya que para ilustrar el enfoque propuesto en este trabajo, hipotéticamente se consideraron como las dos características de calidad más importantes para este experimento.

Por otra parte en la columna D_i de la tabla 4.5, se muestran las diferencias que se observaron entre *Mgp* y *Mg* ($D_i = Mgp - Mg$). Si se considera la región 0.4 y 0.6 de la variable x_1 , por ejemplo, en las corridas para $x_1 = 0.4$ la *Mgp* se reduce en comparación con la *Mg* normal hasta en casi 0.025, mientras que en las corridas para $x_1 = 0.6$ se tiene el efecto contrario, ya que las probabilidades para la *Mgp* aumentan hasta en un poco más de 0.018 en relación con la probabilidad obtenida con la *Mg* normal. Esto es debido a que las condiciones del proceso con $x_1 = 0.4$, tienen un comportamiento no tan bueno en las variables y_2 y y_3 que son las que tienen mayor peso, por lo que son penalizadas, mientras que las condiciones del proceso con $x_1 = 0.6$ se comportan de una forma más adecuada en estas mismas variables, lo que es considerado por *Mgp* aumentando su valor. Para realizar los cálculos de las columnas P_1 , P_2 , P_3 y P_4 se consideraron las siguientes especificaciones para cada característica de calidad:

$$y_{RC} \geq 1.8, y_{TC} \leq 13.5, y_{S/N} \geq 302, 0.75 \leq y_{AS} \leq 0.84$$

Tabla 4.3 Combinaciones de los Factores de Control Utilizadas para Formar la Malla en la Simulación: Caso de Estudio I

Corrida	1	2	3	4	5	6	7	8	9	10	11	12	13
X_1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-0.8	-0.8
X_2	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1	-1	-0.8
Corrida	27	28	29	30	31	32	33	34	35	36	37	38	39
X_1	-0.6	-0.6	-0.6	-0.6	-0.6	-0.6	-0.6	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4
X_2	-0.2	0	0.2	0.4	0.6	0.8	1	-1	-0.8	-0.6	-0.4	-0.2	0
Corrida	40	41	42	43	44	45	46	47	48	49	50	51	52
X_1	-0.4	-0.4	-0.4	-0.4	-0.4	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
X_2	0.2	0.4	0.6	0.8	1	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4
Corrida	53	54	55	56	57	58	59	60	61	62	63	64	65
X_1	-0.2	-0.2	-0.2	0	0	0	0	0	0	0	0	0	0
X_2	0.6	0.8	1	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
Corrida	66	67	68	69	70	71	72	73	74	75	76	77	78
X_1	0	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.4
X_2	1	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1	-1
Corrida	79	80	81	82	83	84	85	86	87	88	89	90	91
X_1	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.6	0.6	0.6
X_2	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1	-1	-0.8	-0.6
Corrida	92	93	94	95	96	97	98	99	100	101	102	103	104
X_1	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.8	0.8	0.8	0.8	0.8
X_2	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1	-1	-0.8	-0.6	-0.4	-0.2
Corrida	105	106	107	108	109	110	111	112	113	114	115	116	117
X_1	0.8	0.8	0.8	0.8	0.8	0.8	1	1	1	1	1	1	1
X_2	0	0.2	0.4	0.6	0.8	1	-1	-0.8	-0.6	-0.4	-0.2	0	0.2
Corrida	118	119	120	121									
X_1	1	1	1	1									
X_2	0.4	0.6	0.8	1									

Tabla 4.4 Pesos Dados a las Variables de Respuesta para Obtener Mg y Mgp

Respuesta	Ponderaciones	
	Mg	Mgp
y_1	0.25	0.1
y_2	0.25	0.4
y_3	0.25	0.4
y_4	0.25	0.1

Donde algunas de ellas se modificaron ligeramente respecto al trabajo de referencia. Un aspecto importante que puede observarse en estas especificaciones es que la metodología permite trabajar tanto con especificaciones bilaterales como unilaterales del tipo lo mayor o lo menor es lo mejor, sin la necesidad de inventarle un valor objetivo, o un límite, por lo que puede decirse que es flexible en este aspecto.

Continuando con la revisión de la información presentada en la tabla 4.5, en la figura 4.1 se presenta una comparación gráfica entre los resultados obtenidos tanto para la Mg como para la Mgp para cada condición de operación del proceso. En esta puede verse que las diferentes ponderaciones pueden cambiar drásticamente los resultados, habiendo condiciones que aumentaron en casi 0.25 su valor resultante, mientras que en otras hubo una disminución de hasta casi un 0.17.

En la figura 4.2 se muestra la gráfica de las diferencias (columna D_i) entre Mg y Mgp , en la que puede apreciarse que las mayores penalizaciones se ubican en $x_1 = 0$ y 0.2, mientras que los mayores aumentos se localizan en $x_1 = 0.8$ y 1.

Tabla 4.5 Resultados de la Simulación del Arreglo 11²: Caso de Estudio I

X_1	X_2	P_1	P_2	P_3	P_4	Mg	Mgp	$D_i = Mgp - Mg$
-1.0	-1.0	0.9980	0.0060	0.0006	0.4968	0.0366	0.0062	-0.0304
-1.0	-0.8	0.9972	0.0068	0.0010	0.5094	0.0431	0.0080	-0.0351
-1.0	-0.6	0.9978	0.0052	0.0006	0.5292	0.0358	0.0059	-0.0299
-1.0	-0.4	0.9982	0.0058	0.0006	0.5250	0.0367	0.0061	-0.0306
-1.0	-0.2	0.9972	0.0052	0.0002	0.5194	0.0271	0.0038	-0.0233
-1.0	0.0	0.9976	0.0056	0.0006	0.5360	0.0366	0.0061	-0.0305
-1.0	0.2	0.9972	0.0064	0.0006	0.5330	0.0378	0.0064	-0.0314
-1.0	0.4	0.9986	0.0050	0.0002	0.5372	0.0271	0.0037	-0.0234
-1.0	0.6	0.9990	0.0036	0.0000	0.5202	0.0000	0.0000	0.0000
-1.0	0.8	0.9982	0.0072	0.0006	0.5074	0.0385	0.0067	-0.0318
-1.0	1.0	0.9980	0.0066	0.0004	0.5056	0.0340	0.0055	-0.0285
-0.8	-1.0	0.9974	0.0124	0.0000	0.7272	0.0000	0.0000	0.0000
-0.8	-0.8	0.9976	0.0108	0.0008	0.7340	0.0502	0.0091	-0.0411
-0.8	-0.6	0.9980	0.0108	0.0002	0.7676	0.0359	0.0053	-0.0306
-0.8	-0.4	0.9974	0.0096	0.0000	0.7618	0.0000	0.0000	0.0000
-0.8	-0.2	0.9972	0.0102	0.0008	0.7648	0.0499	0.0090	-0.0409
-0.8	0.0	0.9978	0.0080	0.0002	0.7670	0.0333	0.0047	-0.0286
-0.8	0.2	0.9984	0.0072	0.0014	0.7620	0.0526	0.0098	-0.0428
-0.8	0.4	0.9978	0.0092	0.0000	0.7546	0.0000	0.0000	0.0000
-0.8	0.6	0.9982	0.0074	0.0000	0.7364	0.0000	0.0000	0.0000
-0.8	0.8	0.9982	0.0054	0.0016	0.7364	0.0502	0.0091	-0.0411
-0.8	1.0	0.9978	0.0104	0.0004	0.7260	0.0417	0.0068	-0.0349
-0.6	-1.0	0.9972	0.0250	0.0006	0.8660	0.0600	0.0116	-0.0484
-0.6	-0.8	0.9970	0.0222	0.0010	0.8822	0.0665	0.0136	-0.0529
-0.6	-0.6	0.9962	0.0200	0.0012	0.8788	0.0677	0.0140	-0.0537
-0.6	-0.4	0.9964	0.0190	0.0010	0.8806	0.0639	0.0128	-0.0511
-0.6	-0.2	0.9970	0.0170	0.0006	0.8940	0.0549	0.0100	-0.0449
-0.6	0.0	0.9972	0.0160	0.0010	0.8920	0.0614	0.0119	-0.0495
-0.6	0.2	0.9972	0.0122	0.0010	0.8878	0.0573	0.0107	-0.0466
-0.6	0.4	0.9978	0.0156	0.0006	0.8912	0.0537	0.0096	-0.0441
-0.6	0.6	0.9974	0.0160	0.0012	0.8880	0.0642	0.0128	-0.0514
-0.6	0.8	0.9974	0.0138	0.0010	0.8818	0.0590	0.0112	-0.0478
-0.6	1.0	0.9954	0.0160	0.0012	0.8684	0.0638	0.0128	-0.0510

Tabla 4.5 Continuación

X_1	X_2	P_1	P_2	P_3	P_4	Mg	Mgp	$D_i = Mgp - Mg$
-0.4	-1.0	0.9956	0.0526	0.0020	0.9344	0.0995	0.0254	-0.0741
-0.4	-0.8	0.9978	0.0358	0.0026	0.9380	0.0966	0.0242	-0.0724
-0.4	-0.6	0.9972	0.0362	0.0006	0.9432	0.0672	0.0136	-0.0536
-0.4	-0.4	0.9962	0.0368	0.0012	0.9396	0.0802	0.0180	-0.0622
-0.4	-0.2	0.9970	0.0340	0.0012	0.9360	0.0786	0.0174	-0.0612
-0.4	0.0	0.9972	0.0276	0.0014	0.9440	0.0777	0.0171	-0.0606
-0.4	0.2	0.9950	0.0318	0.0008	0.9356	0.0698	0.0144	-0.0554
-0.4	0.4	0.9974	0.0270	0.0012	0.9414	0.0743	0.0159	-0.0584
-0.4	0.6	0.9972	0.0246	0.0014	0.9454	0.0755	0.0163	-0.0592
-0.4	0.8	0.9968	0.0250	0.0006	0.9368	0.0612	0.0117	-0.0495
-0.4	1.0	0.9962	0.0276	0.0008	0.9368	0.0674	0.0136	-0.0538
-0.2	-1.0	0.9960	0.1322	0.0036	0.9670	0.1463	0.0467	-0.0996
-0.2	-0.8	0.9962	0.1130	0.0048	0.9580	0.1508	0.0492	-0.1016
-0.2	-0.6	0.9952	0.1058	0.0034	0.9626	0.1362	0.0417	-0.0945
-0.2	-0.4	0.9964	0.0846	0.0036	0.9636	0.1308	0.0391	-0.0917
-0.2	-0.2	0.9962	0.0814	0.0024	0.9584	0.1169	0.0327	-0.0842
-0.2	0.0	0.9954	0.0806	0.0022	0.9598	0.1141	0.0314	-0.0827
-0.2	0.2	0.9964	0.0662	0.0032	0.9636	0.1194	0.0338	-0.0856
-0.2	0.4	0.9968	0.0698	0.0028	0.9612	0.1170	0.0327	-0.0843
-0.2	0.6	0.9972	0.0672	0.0016	0.9622	0.1008	0.0258	-0.0750
-0.2	0.8	0.9974	0.0616	0.0018	0.9674	0.1017	0.0261	-0.0756
-0.2	1.0	0.9962	0.0654	0.0024	0.9540	0.1105	0.0299	-0.0806
0.0	-1.0	0.9932	0.4340	0.0266	0.9682	0.3246	0.1672	-0.1574
0.0	-0.8	0.9954	0.3784	0.0198	0.9724	0.2918	0.1407	-0.1511
0.0	-0.6	0.9966	0.3380	0.0158	0.9706	0.2681	0.1229	-0.1452
0.0	-0.4	0.9958	0.2982	0.0122	0.9666	0.2433	0.1054	-0.1379
0.0	-0.2	0.9940	0.2692	0.0152	0.9672	0.2504	0.1104	-0.1400
0.0	0.0	0.9934	0.2504	0.0116	0.9624	0.2296	0.0962	-0.1334
0.0	0.2	0.9954	0.2214	0.0082	0.9698	0.2046	0.0798	-0.1248
0.0	0.4	0.9946	0.2188	0.0084	0.9680	0.2051	0.0802	-0.1249
0.0	0.6	0.9966	0.2026	0.0076	0.9732	0.1966	0.0748	-0.1218
0.0	0.8	0.9956	0.2026	0.0082	0.9682	0.2000	0.0770	-0.1230
0.0	1.0	0.9930	0.1976	0.0090	0.9664	0.2033	0.0791	-0.1242

Tabla 4.5 Continuación

X_1	X_2	P_1	P_2	P_3	P_4	Mg	Mgp	$D_i = Mgp - Mg$
0.2	-1.0	0.9950	0.7660	0.6570	0.9650	0.8338	0.7568	-0.0770
0.2	-0.8	0.9954	0.7500	0.5038	0.9694	0.7771	0.6751	-0.1020
0.2	-0.6	0.9932	0.7212	0.3874	0.9658	0.7195	0.5980	-0.1215
0.2	-0.4	0.9926	0.6994	0.2852	0.9710	0.6622	0.5228	-0.1394
0.2	-0.2	0.9934	0.6528	0.2184	0.9638	0.6078	0.4568	-0.1510
0.2	0.0	0.9954	0.6396	0.1644	0.9706	0.5646	0.4048	-0.1598
0.2	0.2	0.9938	0.6306	0.1338	0.9644	0.5333	0.3704	-0.1629
0.2	0.4	0.9954	0.6132	0.1126	0.9692	0.5080	0.3421	-0.1659
0.2	0.6	0.9942	0.6120	0.1008	0.9712	0.4940	0.3270	-0.1670
0.2	0.8	0.9934	0.6002	0.0862	0.9656	0.4720	0.3046	-0.1674
0.2	1.0	0.9952	0.6082	0.0894	0.9724	0.4789	0.3110	-0.1679
0.4	-1.0	0.9942	0.9136	0.9878	0.9456	0.9597	0.9539	-0.0058
0.4	-0.8	0.9906	0.9076	0.9814	0.9510	0.9571	0.9491	-0.0080
0.4	-0.6	0.9896	0.8986	0.9716	0.9548	0.9530	0.9418	-0.0112
0.4	-0.4	0.9920	0.8982	0.9700	0.9514	0.9523	0.9409	-0.0114
0.4	-0.2	0.9922	0.8894	0.9524	0.9496	0.9452	0.9302	-0.0150
0.4	0.0	0.9918	0.8740	0.9422	0.9408	0.9363	0.9189	-0.0174
0.4	0.2	0.9942	0.8776	0.9300	0.9514	0.9374	0.9168	-0.0206
0.4	0.4	0.9922	0.8694	0.9260	0.9488	0.9330	0.9114	-0.0216
0.4	0.6	0.9914	0.8728	0.9082	0.9524	0.9301	0.9060	-0.0241
0.4	0.8	0.9904	0.8596	0.9092	0.9456	0.9250	0.9002	-0.0248
0.4	1.0	0.9914	0.8748	0.9050	0.9428	0.9275	0.9047	-0.0228
0.6	-1.0	0.9876	0.9472	0.9956	0.8388	0.9401	0.9586	0.0185
0.6	-0.8	0.9904	0.9538	0.9966	0.8478	0.9452	0.9629	0.0177
0.6	-0.6	0.9886	0.9468	0.9966	0.8570	0.9456	0.9610	0.0154
0.6	-0.4	0.9892	0.9450	0.9958	0.8476	0.9425	0.9589	0.0164
0.6	-0.2	0.9882	0.9468	0.9936	0.8372	0.9393	0.9575	0.0182
0.6	0.0	0.9906	0.9426	0.9944	0.8398	0.9397	0.9567	0.0170
0.6	0.2	0.9904	0.9468	0.9938	0.8450	0.9420	0.9587	0.0167
0.6	0.4	0.9902	0.9458	0.9932	0.8446	0.9415	0.9580	0.0165
0.6	0.6	0.9894	0.9454	0.9940	0.8420	0.9406	0.9578	0.0172
0.6	0.8	0.9884	0.9414	0.9934	0.8534	0.9424	0.9571	0.0147
0.6	1.0	0.9894	0.9454	0.9896	0.8442	0.9402	0.9563	0.0161

Tabla 4.5 Continuación

X_1	X_2	P_1	P_2	P_3	P_4	Mg	Mgp	$D_i = Mgp - Mg$
0.8	-1.0	0.9804	0.9670	0.9984	0.3796	0.7742	0.8932	0.1190
0.8	-0.8	0.9856	0.9672	0.9972	0.3944	0.7825	0.8968	0.1143
0.8	-0.6	0.9892	0.9678	0.9992	0.3598	0.7659	0.8898	0.1239
0.8	-0.4	0.9838	0.9648	0.9976	0.3566	0.7623	0.8869	0.1246
0.8	-0.2	0.9842	0.9652	0.9990	0.3652	0.7673	0.8897	0.1224
0.8	0.0	0.9804	0.9652	0.9988	0.3604	0.7640	0.8881	0.1241
0.8	0.2	0.9802	0.9694	0.9970	0.3652	0.7669	0.8902	0.1233
0.8	0.4	0.9848	0.9656	0.9962	0.3490	0.7583	0.8849	0.1266
0.8	0.6	0.9838	0.9662	0.9964	0.3684	0.7686	0.8899	0.1213
0.8	0.8	0.9848	0.9670	0.9978	0.3648	0.7673	0.8899	0.1226
0.8	1.0	0.9852	0.9628	0.9978	0.3906	0.7798	0.8945	0.1147
1.0	-1.0	0.9596	0.9668	0.9980	0.0900	0.5373	0.7716	0.2343
1.0	-0.8	0.9666	0.9764	0.9988	0.0838	0.5302	0.7700	0.2398
1.0	-0.6	0.9662	0.9746	0.9980	0.0766	0.5180	0.7623	0.2443
1.0	-0.4	0.9712	0.9736	0.9992	0.0700	0.5071	0.7559	0.2488
1.0	-0.2	0.9724	0.9786	0.9990	0.0708	0.5093	0.7583	0.2490
1.0	0.0	0.9712	0.9750	0.9984	0.0724	0.5115	0.7586	0.2471
1.0	0.2	0.9730	0.9748	0.9982	0.0736	0.5138	0.7599	0.2461
1.0	0.4	0.9766	0.9786	0.9982	0.0692	0.5069	0.7567	0.2498
1.0	0.6	0.9694	0.9730	0.9988	0.0752	0.5159	0.7609	0.2450
1.0	0.8	0.9702	0.9746	0.9988	0.0858	0.5335	0.7715	0.2380
1.0	1.0	0.9652	0.9702	0.9984	0.0962	0.5476	0.7785	0.2309

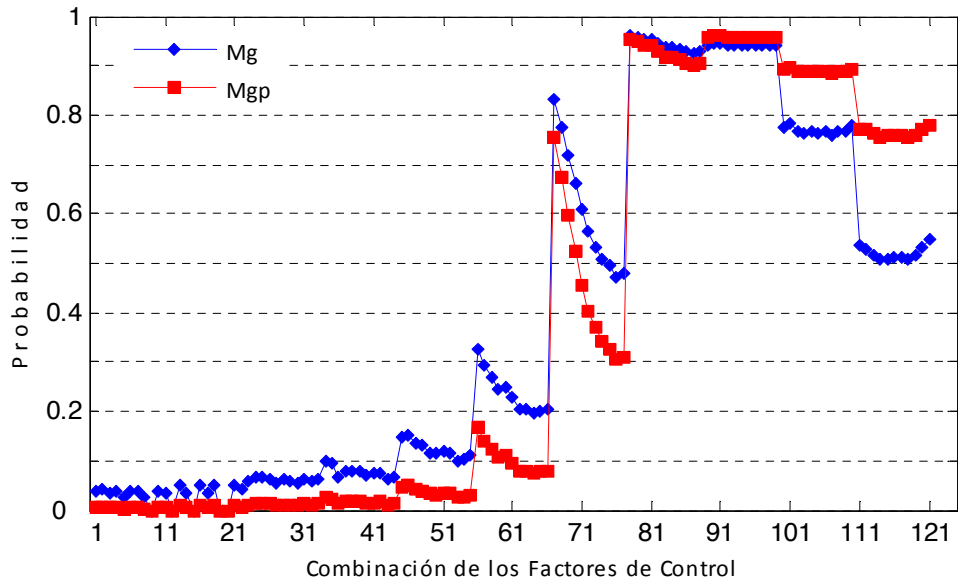


Figura 4.1 Comparación de los Resultados Obtenidos para Mg y Mgp .

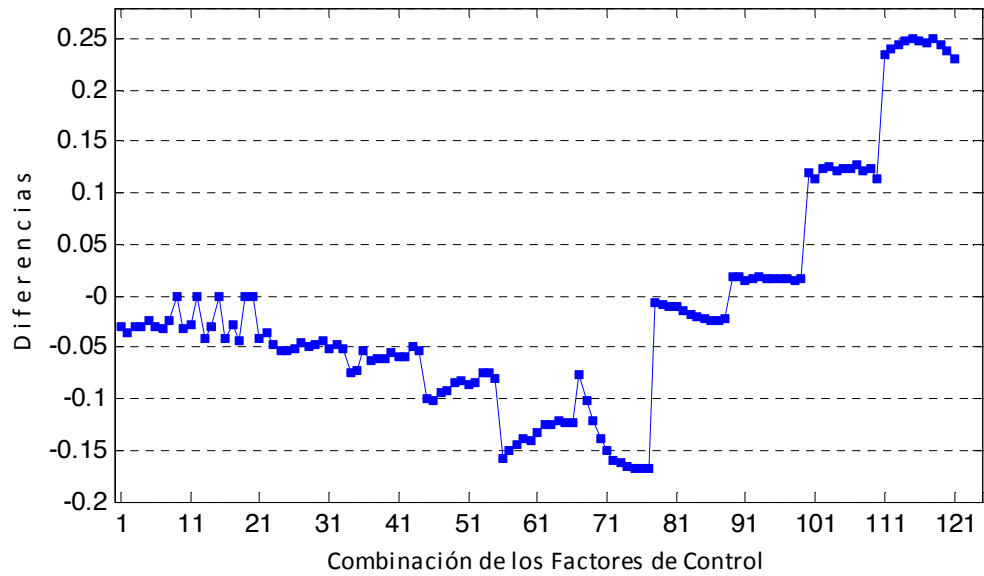


Figura 4.2 Diferencias en las Probabilidades entre Mg y Mgp para Cada Corrida

Fase VI: Una vez obtenidas Mg y Mgp , se ajustaron modelos para cada una de ellas, utilizando regresión logística y el software minitab 15, de lo cual resultaron los modelos de las ecuaciones 4.7 y 4.8 respectivamente.

Respecto a la columna Mg , se obtuvo:

$$\begin{aligned}
 \log it(\hat{p}(x_c)) = & -1.282 + 7.076x_1 - 0.509x_2 - 0.919x_1x_2 + 12.919x_1^2 - 0.750x_2^2 \\
 & + 1.631x_1^2x_2 + 1.565x_1x_2^2 - 0.199x_1^2x_2^2 - 8.069x_1^3 - 0.287x_2^3 + 2.597x_1^3x_2 + 0.245x_1x_2^3 \\
 & - 32.254x_1^4 + 2.553x_2^4 - 1.489x_1^4x_2 - 3.878x_1x_2^4 + 2.607x_1^5 + 0.341x_2^5 - 1.561x_1^5x_2 \\
 & - 0.352x_1x_2^5 + 19.010x_1^6 - 1.637x_2^6 + 0.291x_1^6x_2 + 2.487x_1x_2^6
 \end{aligned} \tag{4.7}$$

Mientras que respecto a la columna Mgp , se obtuvo:

$$\begin{aligned}
 \log it(\hat{p}(x_c)) = & -2.449 + 8.724x_1 - 0.688x_2 - 1.398x_1x_2 + 15.375x_1^2 - 0.072x_2^2 \\
 & + 3.854x_1^2x_2 + 0.691x_1x_2^2 - 0.385x_1^2x_2^2 - 7.087x_1^3 - 0.258x_2^3 + 4.710x_1^3x_2 + 0.198x_1x_2^3 \\
 & - 35.989x_1^4 + 1.076x_2^4 - 7.495x_1^4x_2 - 1.527x_1x_2^4 + 1.466x_1^5 + 0.269x_2^5 - 3.250x_1^5x_2 \\
 & - 0.239x_1x_2^5 + 21.102x_1^6 - 0.687x_2^6 + 4.294x_1^6x_2 + 0.974x_1x_2^6
 \end{aligned} \tag{4.8}$$

Fase VII: Después de ajustados los modelos, se procede con ellos a pronosticar las probabilidades de éxito para cada una de las combinaciones consideradas en la malla simulada. En las figuras 4.3 y 4.4 se muestra la comparación gráfica de las probabilidades de éxito obtenidas al hacer un pronóstico utilizando los modelos ajustados para las columnas Mg y Mgp contra las probabilidades de éxito obtenidas según las respectivas simulaciones. En ambas gráficas se puede apreciar que los dos

modelos ajustan bastante bien, ya que estos pronostican las probabilidades de éxito simuladas con bastante precisión. Por lo anterior se puede proceder con ellos a hacer la optimización utilizando métodos numéricos, en específico el *MGRG*, teniendo como función objetivo la maximización de los modelos logit. Previamente se construyen gráficas de superficie y de contorno para ubicar de forma aproximada la región óptima, la cual debe ser congruente con los resultados alcanzados en la optimización. Las gráficas de superficie y de contorno para *Mg* se muestran en las figuras 4.5 y 4.6, mientras que para *Mgp* se muestran en las figuras 4.7 y 4.8. En ambos casos se puede ver que la región óptima se encuentra en x_1 entre 0.4 y 0.6, mientras la variable x_2 es menos relevante.

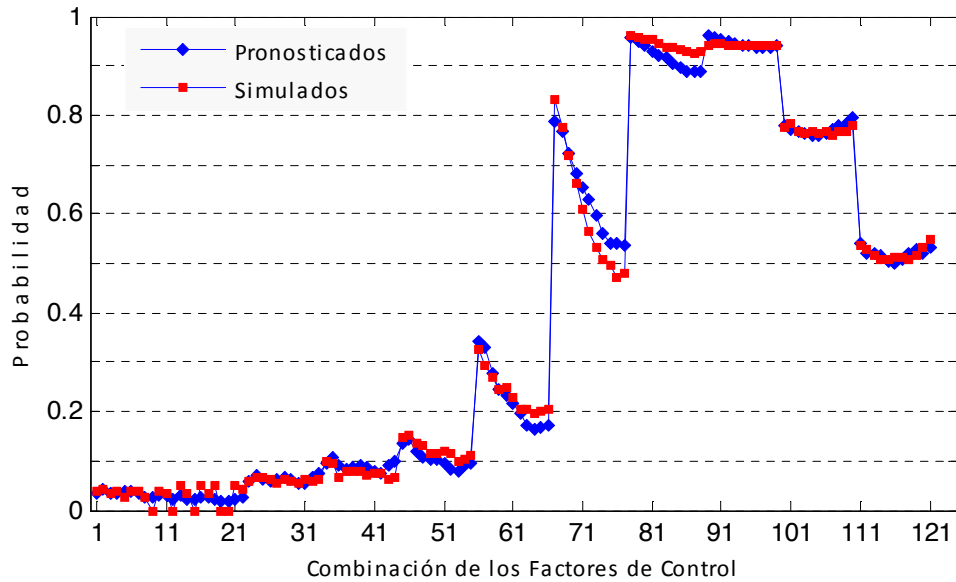


Figura 4.3 Resultados Simulados contra Pronosticados para *Mg*.

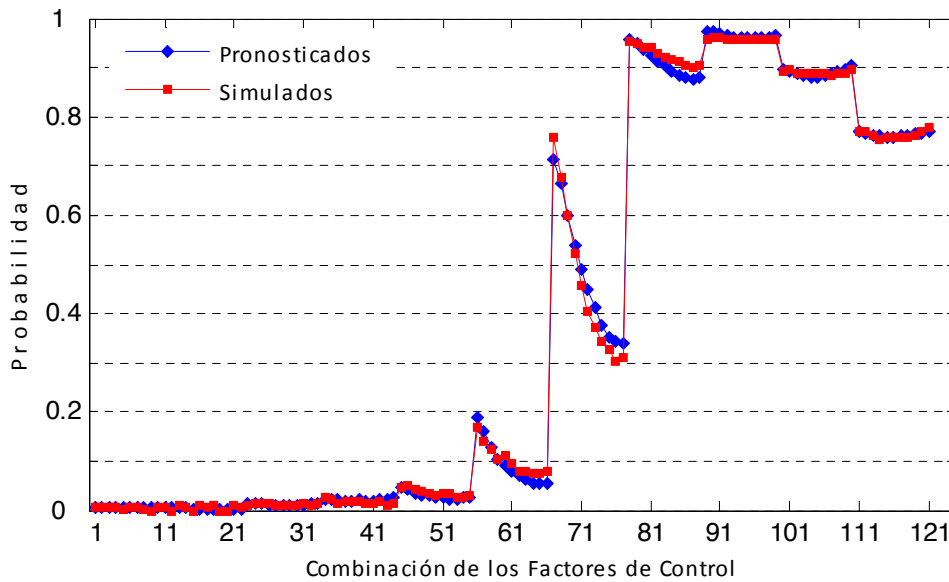


Figura 4.4 Resultados Simulados contra Pronosticados para *Mgp*.

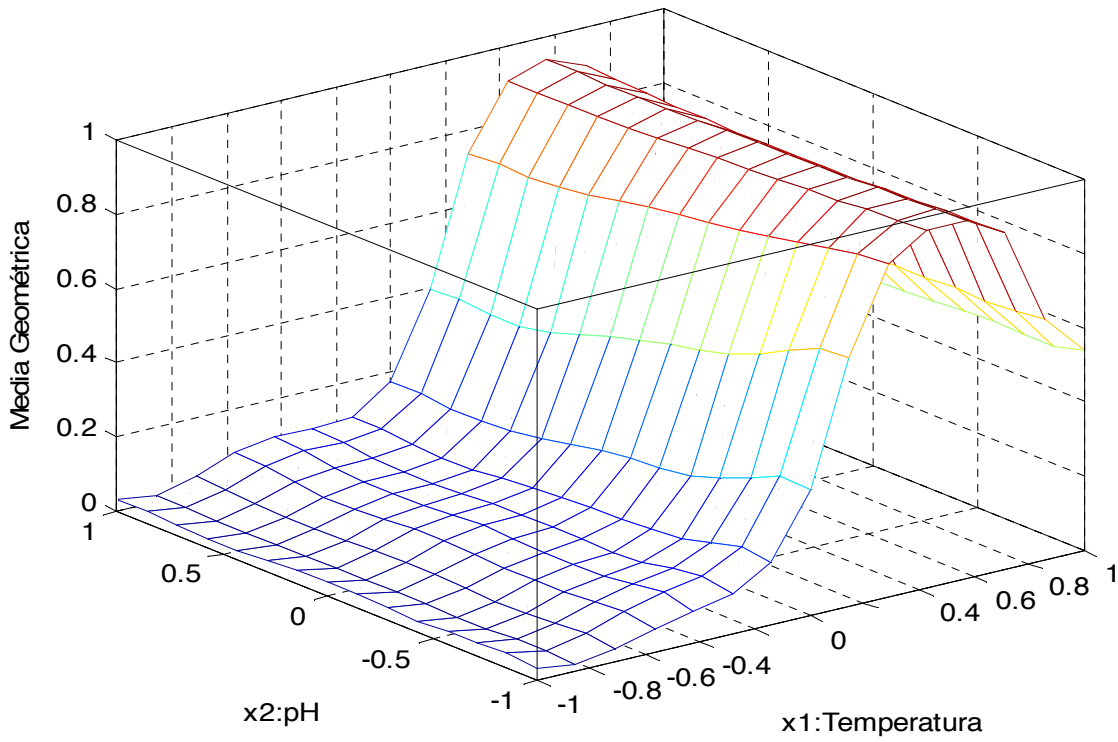


Figura 4.5 Superficie de Respuesta para el Modelo Logit Ajustado a Mg

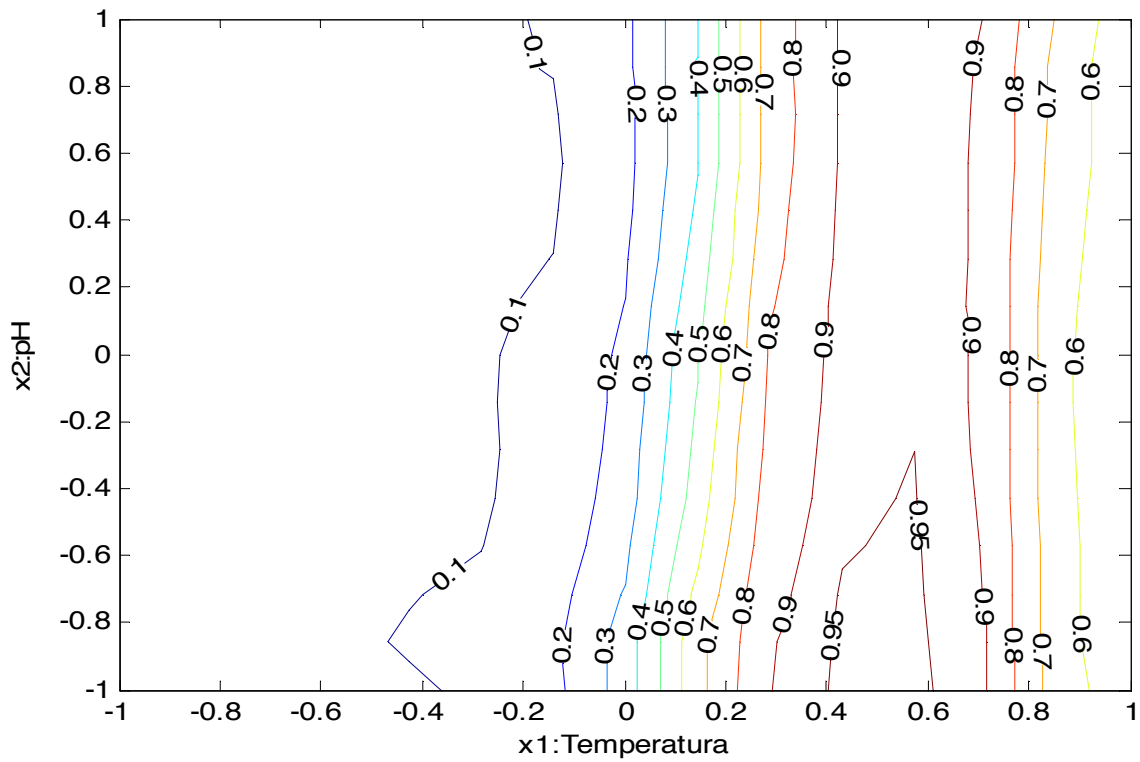


Figura 4.6 Gráfica de Contornos para el Modelo Logit Ajustado a Mg

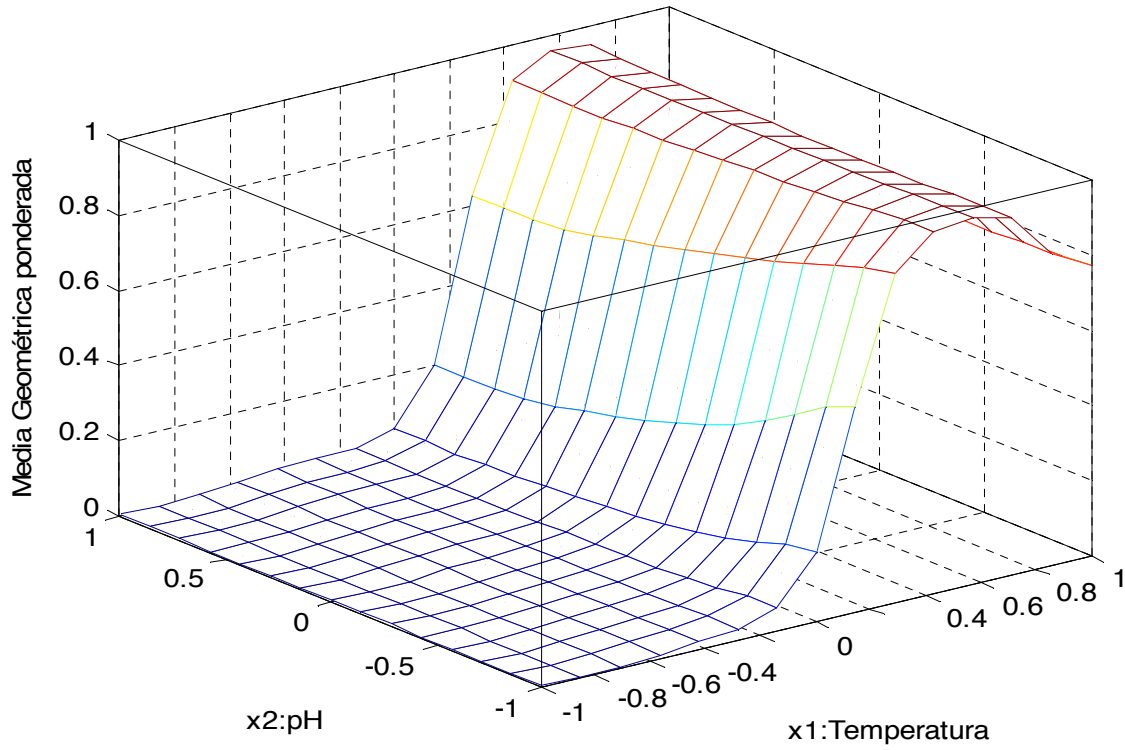


Figura 4.7 Superficie de Respuesta para el Modelo Logit Ajustado a *Mgp*

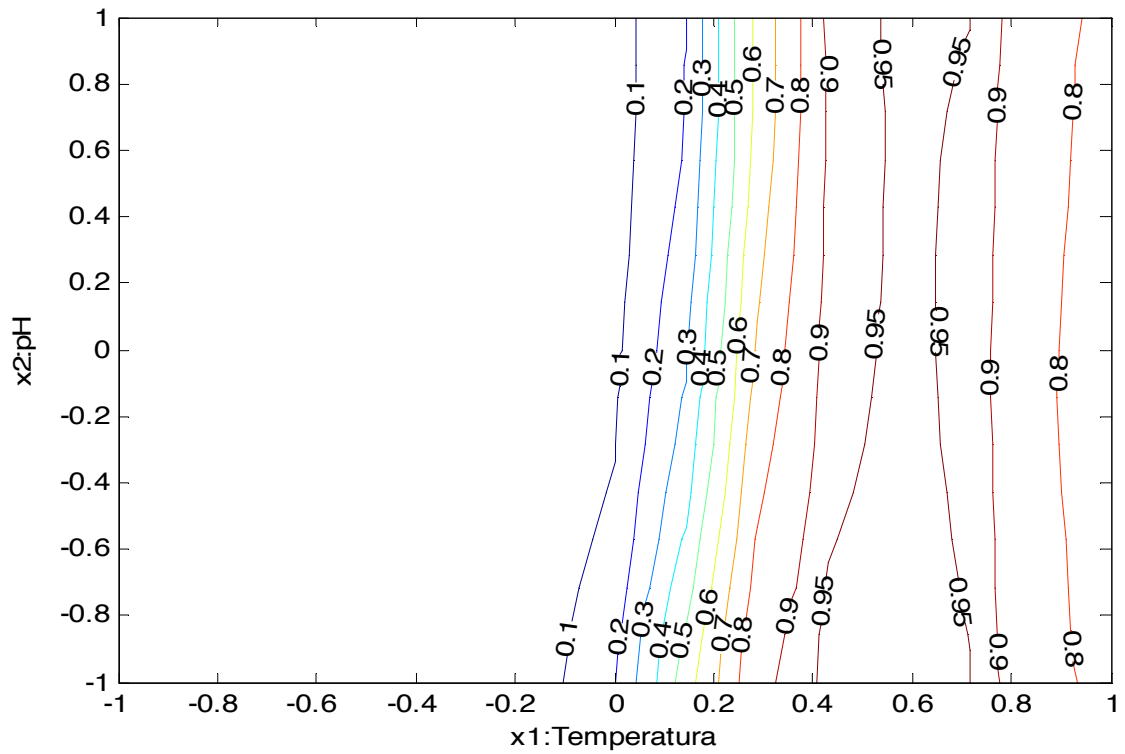


Figura 4.8 Gráfica de Contornos para el Modelo Logit Ajustado a *Mgp*

Fase VIII: En esta etapa se procedió a la optimización de los modelos logísticos mediante el *MGRG* del Excel solver. Como se comentó previamente este algoritmo de optimización permite incorporar restricciones en las zonas de búsqueda, lo que proporciona una ventaja, ya que de esta manera podremos asegurarnos de que la solución propuesta estará dentro de la región en la que se diseñó el experimento. Para esto se optimizaron ambos modelos partiendo de 15 valores iniciales distintos, tomados del rango permisible de valores para ambas variables de interés. Los resultados encontrados se muestran en la tabla 4.6. En esta tabla puede verse que independientemente de los valores iniciales el método únicamente converge a dos resultados distintos para ambos modelos.

Tabla 4.6 Resultados Óptimos Encontrados para *Mg* y *Mgp*

Valores iniciales		Resultados <i>Mg</i>		Resultados <i>Mgp</i>	
x_1	x_2	x_1	x_2	x_1	x_2
0.40	-0.40	0.513969	-1	0.551957	-1
0.20	-1.00	0.513969	-1	0.551957	-1
0.70	0.70	0.565974	1	0.608722	1
-0.30	-0.70	0.513968	-1	0.551957	-1
0.90	-0.90	0.513969	-1	0.551957	-1
0.80	0.25	0.513969	-1	0.551957	-1
0.00	0.00	0.513969	-1	0.551957	-1
0.50	-0.50	0.513969	-1	0.551957	-1
0.30	0.50	0.513968	-1	0.608722	1
0.25	0.25	0.513969	-1	0.551957	-1
0.47	0.47	0.513968	-1	0.608722	1
0.94	0.80	0.565976	1	0.608722	1
0.10	-0.10	0.513969	-1	0.551957	-1
-0.10	0.10	0.513969	-1	0.551957	-1
-0.50	-0.30	0.513969	-1	0.551957	-1

Fase IX: Para los resultados óptimos encontrados se procede a estimar su *Mg* y *Mgp* tanto por simulación con 100000 iteraciones, como por medio de sus respectivos modelos logísticos. Los resultados se muestran en la tabla 4.7, en la cual puede verse que de los dos resultados distintos propuestos por el *MGRG* para *Mg*, el resultado (0.514, -1), ofrece mejores rendimientos tanto por la sustitución de valores en el modelo logístico, como por la simulación en matlab, por lo que sería la condición de operación para el proceso que debiera de seleccionarse para la producción futura en el caso de *Mg* y en el caso de *Mgp* puede verse que el resultado (0.552, -1), ofrece mejores rendimientos también en ambos casos.

Tabla 4.7 Resultados Obtenidos por Simulación y por Sustitución en los Modelos Logísticos en las Condiciones Óptimas.

	$(x_1, x_2)^*$	Resultado	
		modelo	simulación
<i>Mg</i>	(0.514, -1)	0.9684	0.9569
<i>Mg</i>	(0.566, 1)	0.9420	0.9459
<i>Mgp</i>	(0.552, -1)	0.9761	0.9632
<i>Mgp</i>	(0.6087, 1)	0.9653	0.9546

Una vez que se seleccionaron y evaluaron las combinaciones óptimas, se hicieron 30 simulaciones de 100000 valores de cada variable de respuesta en cada una de estas condiciones, para observar el comportamiento que tenían en los distintos escenarios (en este caso *Mg* y *Mgp*) y apreciar el efecto que tuvieron las ponderaciones en los resultados, es decir, como se comportaron las variables en relación a sus probabilidades de concordancia en *Mg* con respecto a *Mgp*.

En las tablas 4.8 y 4.9 se encuentra el resultado de las 30 simulaciones para *Mg* y *Mgp* respectivamente, en estas se muestran las probabilidades de concordancia, los promedios y las desviaciones estándares para cada respuesta, así como las medias geométricas ponderadas obtenidas en cada una de las simulaciones. Posteriormente en la tabla 4.10 se encuentran los promedios de cada uno de los estadísticos de las tablas anteriores y con esta información se procedió a calcular los índices *Cpk*s y sus respectivos *Mcpk*s en cada una de las condiciones óptimas propuestas y como puede observarse las variables que aumentaron su ponderación en *Mgp* (y_2 y y_3) aumentaron su índice *Cpk*, mientras que en aquellas en las que se redujo la ponderación (y_1 y y_4) también se redujo su *Cpk*. Además utilizando los *Cpk*'s individuales también se calculó $Mcpk_1$ y $Mcpk_2$ utilizando las ponderaciones (0.25, 0.25, 0.25 y 0.25) y (0.1, 0.4, 0.4 y 0.1) respectivamente, los valores más altos para cada *Mcpk* coincidieron con las condiciones óptimas para cada escenario.

Las comparaciones del comportamiento de las probabilidades de concordancia para y_1 , y_2 , y_3 y y_4 en ambos escenarios se muestran de forma gráfica en las figuras 4.9, 4.10, 4.11 y 4.12 respectivamente. En estas puede verse que en el caso de y_1 y de y_4 sus probabilidades se redujeron para *Mgp*, lo cual puede ser explicado ya que su ponderación también se redujo de 0.25 a 0.1. En el caso de y_2 y y_3 el efecto fue contrario ya que sus probabilidades aumentaron para *Mgp*, respondiendo también al cambio en la ponderación ya que esta aumentó de 0.25 a 0.4 en ambos casos. Por lo anterior se puede decir que en todas las respuestas la metodología reaccionó de acuerdo a las ponderaciones, moviéndose hacia zonas con mejores resultados para las variables más importantes.

Tabla 4.8 Resumen de las 30 Iteraciones en la Condición Óptima para Mg ($x_1 = 0.514$, $x_2 = -1$)

Simulación	1	2	3	4	5	6	7	8	9	10
Y ₁	2.078	2.076	2.076	2.077	2.077	2.077	2.076	2.077	2.076	2.077
Y ₂	11.938	11.924	11.923	11.939	11.927	11.940	11.926	11.939	11.927	11.944
Y ₃	330.208	330.152	330.101	330.137	330.146	330.172	330.147	330.206	330.179	330.203
Y ₄	0.821	0.821	0.821	0.821	0.820	0.820	0.820	0.820	0.820	0.820
S ₁	0.134	0.153	0.161	0.125	0.139	0.138	0.122	0.150	0.125	0.146
S ₂	1.867	2.327	2.431	1.791	1.831	1.968	1.956	2.396	1.766	3.028
S ₂	11.559	10.720	10.548	11.196	10.159	11.040	10.531	11.468	10.093	13.453
S ₂	0.034	0.041	0.030	0.028	0.029	0.031	0.028	0.038	0.028	0.034
P ₁	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990
P ₂	0.939	0.939	0.939	0.938	0.939	0.938	0.939	0.939	0.939	0.938
P ₃	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994
P ₄	0.908	0.908	0.910	0.908	0.909	0.908	0.908	0.909	0.909	0.907
<i>Mg</i> (0.25, 0.25, 0.25, 0.25)	0.957	0.957	0.958	0.957	0.957	0.957	0.957	0.957	0.957	0.957
Simulación	11	12	13	14	15	16	17	18	19	20
Y ₁	2.077	2.076	2.077	2.077	2.077	2.076	2.076	2.077	2.077	2.077
Y ₂	11.924	11.935	11.924	11.930	11.920	11.942	11.924	11.943	11.922	11.936
Y ₃	330.145	330.142	330.137	330.150	330.145	330.232	330.141	330.196	330.132	330.166
Y ₄	0.821	0.820	0.821	0.821	0.821	0.820	0.821	0.821	0.820	0.821
S ₁	0.141	0.197	0.134	0.136	0.126	0.140	0.125	0.131	0.209	0.207
S ₂	2.015	3.121	2.080	2.077	2.047	2.107	1.938	1.988	3.861	3.332
S ₂	11.513	12.347	10.478	10.635	11.423	11.417	10.181	11.051	18.971	12.139
S ₂	0.034	0.031	0.030	0.029	0.029	0.036	0.029	0.030	0.106	0.061
P ₁	0.989	0.990	0.990	0.990	0.990	0.990	0.990	0.989	0.990	0.990
P ₂	0.938	0.940	0.940	0.938	0.940	0.938	0.939	0.939	0.939	0.939
P ₃	0.993	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994
P ₄	0.907	0.909	0.907	0.908	0.907	0.906	0.908	0.907	0.908	0.908
<i>Mg</i> (0.25, 0.25, 0.25, 0.25)	0.956	0.957	0.957	0.957	0.957	0.956	0.957	0.957	0.957	0.957

Tabla 4.8 Continuación

Simulación	21	22	23	24	25	26	27	28	29	30	Promedios
Y ₁	2.078	2.076	2.077	2.076	2.076	2.077	2.076	2.076	2.076	2.077	2.077
Y ₂	11.942	11.936	11.935	11.919	11.943	11.931	11.945	11.936	11.924	11.932	11.932
Y ₃	330.221	330.195	330.191	330.077	330.175	330.156	330.179	330.145	330.171	330.164	330.164
Y ₄	0.821	0.820	0.821	0.821	0.820	0.821	0.821	0.821	0.821	0.821	0.821
S ₁	0.133	0.125	0.128	0.137	0.128	0.124	0.142	0.146	0.144	0.123	0.142
S ₂	2.113	1.920	1.869	2.174	2.035	1.831	2.052	2.011	2.137	1.915	2.199
S ₂	10.640	10.566	10.862	11.754	10.683	10.356	10.802	10.501	12.820	11.426	11.378
S ₂	0.030	0.035	0.032	0.035	0.029	0.028	0.030	0.029	0.036	0.028	0.035
P ₁	0.990	0.989	0.990	0.990	0.989	0.990	0.990	0.989	0.990	0.990	0.990
P ₂	0.938	0.938	0.938	0.939	0.936	0.938	0.938	0.938	0.940	0.939	0.939
P ₃	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.993	0.994	0.994	0.994
P ₄	0.907	0.908	0.908	0.908	0.908	0.909	0.907	0.908	0.908	0.907	0.908
Mg(0.25, 0.25, 0.25, 0.25)	0.957	0.957	0.957	0.957	0.956	0.957	0.957	0.957	0.957	0.957	0.957

Tabla 4.9 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp ($x_1 = 0.552$, $x_2 = -1$)

Simulación	1	2	3	4	5	6	7	8	9	10
Y_1	2.069	2.067	2.068	2.068	2.068	2.068	2.068	2.068	2.068	2.068
Y_2	11.833	11.806	11.838	11.826	11.818	11.829	11.827	11.835	11.830	11.822
Y_3	333.291	333.210	333.330	333.276	333.263	333.317	333.271	333.377	333.285	333.281
Y_4	0.824	0.823	0.823	0.823	0.823	0.823	0.823	0.823	0.823	0.823
S_1	0.138	0.158	0.145	0.120	0.227	0.180	0.136	0.176	0.120	0.146
S_2	1.935	2.749	2.944	2.079	4.042	2.609	2.526	2.595	1.828	1.903
S_2	10.672	18.701	16.028	10.782	24.920	12.313	13.180	13.814	9.862	11.190
S_2	0.029	0.032	0.035	0.030	0.071	0.048	0.034	0.049	0.030	0.030
P_1	0.989	0.988	0.989	0.989	0.989	0.989	0.989	0.990	0.989	0.989
P_2	0.945	0.945	0.944	0.945	0.945	0.945	0.945	0.945	0.944	0.945
P_3	0.995	0.994	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995
P_4	0.885	0.885	0.887	0.885	0.887	0.887	0.885	0.888	0.884	0.884
$Mgp(0.1, 0.4, 0.4, 0.1)$	0.963	0.962	0.962	0.962	0.963	0.963	0.963	0.963	0.962	0.962
Simulación	11	12	13	14	15	16	17	18	19	20
Y_1	2.068	2.068	2.069	2.068	2.068	2.068	2.069	2.069	2.068	2.068
Y_2	11.822	11.830	11.830	11.828	11.829	11.833	11.845	11.829	11.828	11.820
Y_3	333.299	333.346	333.344	333.251	333.378	333.339	333.421	333.317	333.364	333.336
Y_4	0.823	0.823	0.823	0.823	0.823	0.823	0.823	0.823	0.823	0.823
S_1	0.133	0.121	0.139	0.132	0.163	0.175	0.130	0.160	0.130	0.151
S_2	1.959	1.922	2.089	1.941	2.026	2.216	2.803	2.047	1.888	2.056
S_2	11.820	10.378	12.578	11.499	11.283	11.406	13.110	11.477	10.563	11.358
S_2	0.030	0.030	0.032	0.028	0.032	0.034	0.038	0.036	0.033	0.033
P_1	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.989
P_2	0.945	0.946	0.944	0.943	0.944	0.944	0.944	0.945	0.946	0.946
P_3	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995
P_4	0.886	0.886	0.885	0.886	0.885	0.887	0.886	0.885	0.888	0.888
$Mgp(0.1, 0.4, 0.4, 0.1)$	0.963	0.963	0.963	0.962	0.963	0.963	0.963	0.963	0.964	0.964

Tabla 4.9 Continuación

Simulación	21	22	23	24	25	26	27	28	29	30	Promedios
Y ₁	2.068	2.068	2.068	2.067	2.067	2.069	2.067	2.067	2.067	2.068	2.068
Y ₂	11.830	11.841	11.835	11.821	11.827	11.831	11.815	11.839	11.817	11.837	11.828
Y ₃	333.277	333.348	333.295	333.296	333.249	333.372	333.226	333.300	333.307	333.411	333.313
Y ₄	0.824	0.823	0.823	0.823	0.823	0.823	0.823	0.823	0.823	0.823	0.823
S ₁	0.119	0.126	0.126	0.120	0.183	0.137	0.157	0.137	0.140	0.129	0.145
S ₂	1.881	2.083	2.085	1.783	1.932	2.428	2.218	2.172	2.074	1.982	2.227
S ₂	11.274	11.198	11.075	10.629	11.063	13.647	12.304	12.988	10.306	10.547	12.399
S ₂	0.043	0.030	0.032	0.029	0.031	0.029	0.034	0.035	0.032	0.030	0.035
P ₁	0.989	0.989	0.989	0.989	0.990	0.989	0.989	0.988	0.989	0.989	0.989
P ₂	0.945	0.944	0.944	0.945	0.945	0.945	0.945	0.943	0.944	0.943	0.945
P ₃	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995	0.995
P ₄	0.886	0.886	0.885	0.885	0.888	0.886	0.885	0.885	0.887	0.886	0.886
<i>Mgp</i> (0.1, 0.4, 0.4, 0.1)	0.963	0.962	0.962	0.963	0.963	0.963	0.963	0.962	0.963	0.962	0.963

Tabla 4.10 Estadísticos Resumen para las Condiciones Óptimas Propuestas según *Mg* y *Mgp*

X ₁	X ₂	P ₁	P ₂	P ₃	P ₄	(Y ₁ , S ₁)	(Y ₂ , S ₂)	(Y ₃ , S ₃)	(Y ₄ , S ₄)	Cpk _{y1}	Cpk _{y2}	Cpk _{y3}	Cpk _{y4}	Mcck ₁	Mcck ₂
0.514	-1	0.990	0.939	0.994	0.908	(2.077, 0.142)	(11.932, 2.199)	(330.164, 11.378)	(0.821, 0.035)	0.648	0.238	0.825	0.186	0.392	0.422
0.552	-1	0.989	0.945	0.995	0.886	(2.068, 0.145)	(11.828, 2.227)	(333.313, 12.399)	(0.823, 0.035)	0.616	0.250	0.842	0.161	0.380	0.426

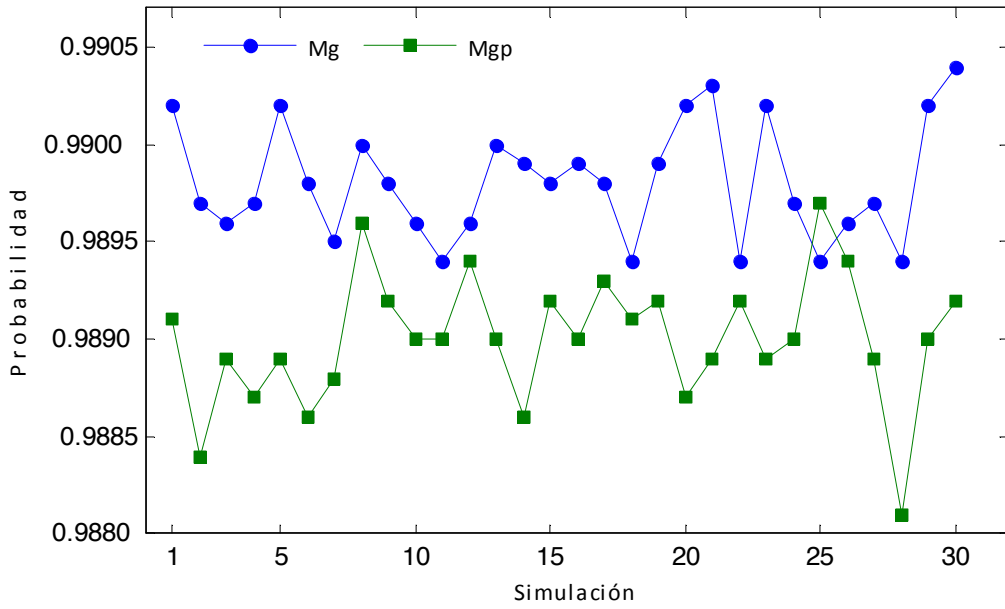


Figura 4.9 Comparación de las Probabilidades de Concordancia en las Condiciones Óptimas Propuestas para y_1 según Mg con Ponderación 0.25 y Mgp con Ponderación 0.1.

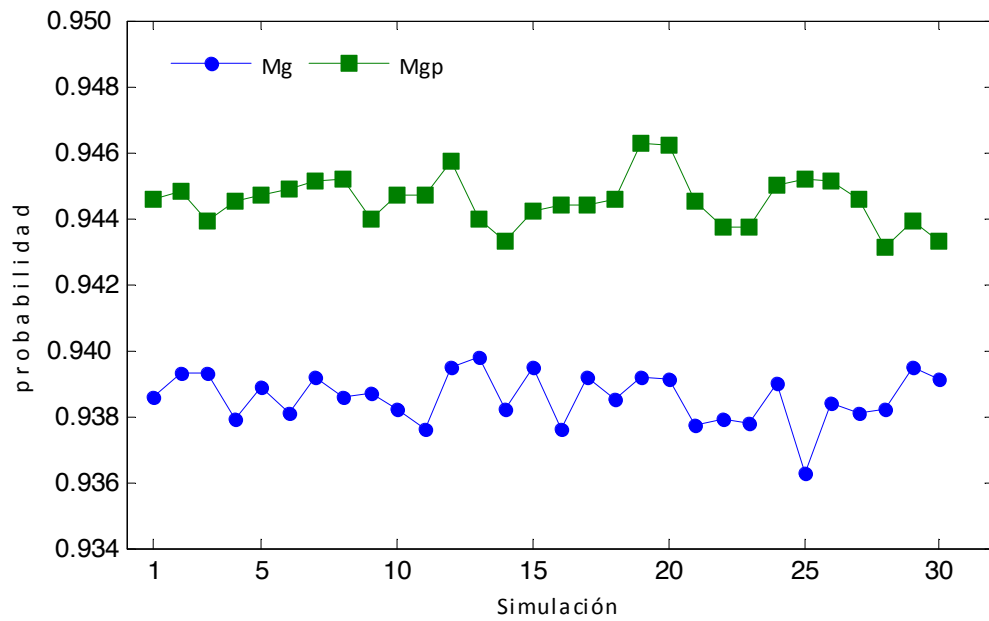


Figura 4.10 Comparación de las Probabilidades de Concordancia en las Condiciones Óptimas Propuestas para y_2 según Mg con Ponderación 0.25 y Mgp con 0.4.

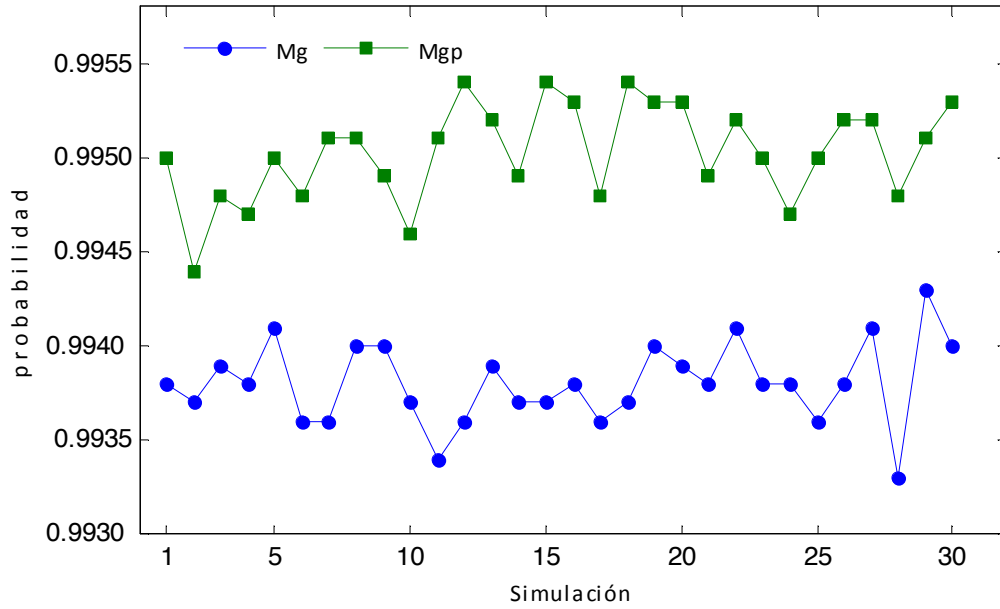


Figura 4.11 Comparación de las Probabilidades de Concordancia en las Condiciones Óptimas Propuestas para y_3 según Mg con Ponderación 0.25 y Mgp con 0.4.

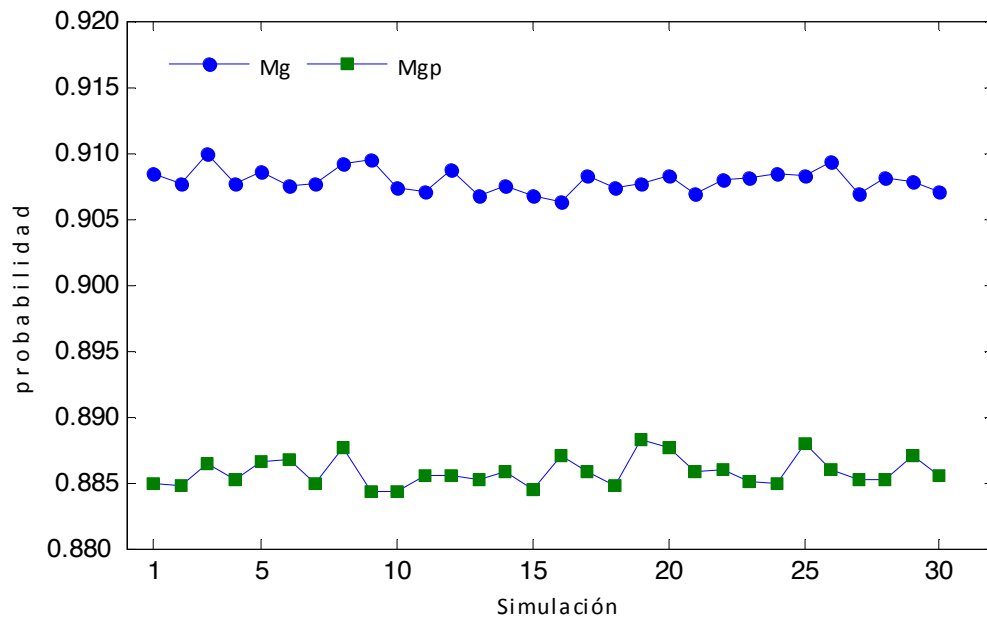


Figura 4.12 Comparación de las Probabilidades de Concordancia en las Condiciones Óptimas Propuestas para y_4 según Mg con Ponderación 0.25 y Mgp con Ponderación 0.1.

Fase X: Esta fase se explica en el apartado 4.3 del presente capítulo.

4.2 Caso de Estudio II: Porcentaje de Conversión y Actividad Térmica

Fases I y II: Para ilustrar nuevamente el funcionamiento de la metodología, se toman los datos del problema propuesto por Myers y Montgomery (1995), en el cual se consideran tres variables de diseño, tiempo de reacción (x_1), temperatura (x_2) y porcentaje de catalizador (x_3) y dos variables de respuesta, porcentaje de conversión (y_1) y actividad térmica (y_2) y se utiliza un diseño central compuesto con seis corridas centrales. En este caso se consideró a x_1 como variable de ruido (x_r), por lo que la optimización se realizará con respecto a las variables de control (x_c) x_2 y x_3 . Los datos se muestran en la tabla 4.9.

Fase III: Con los datos de este experimento, el modelo de segundo orden ajustado para la variable conversión es el siguiente:

$$y_1 = 81.1 + 1.03x_1 + 4.04x_2 + 6.20x_3 - 1.83x_1^2 + 2.94x_2^2 - 5.19x_3^2 + 2.12x_1x_2 + 11.4x_1x_3 - 3.88x_2x_3 \quad \text{con } R^2 = 92.0\% \quad (4.9)$$

Mientras que el modelo de segundo orden ajustado para la variable actividad es el siguiente:

$$y_2 = 59.8 + 3.58x_1 + 0.255x_2 + 2.23x_3 + 0.835x_1^2 + 0.075x_2^2 + 0.057x_3^2 - 0.388x_1x_2 - 0.038x_1x_3 + 0.312x_2x_3 \quad \text{con } R^2 = 89.2\% \quad (4.10)$$

Debido a que en ambos modelos sus coeficientes de determinación fueron mayores a 70%, la metodología de optimización puede ser aplicada a ambas variables de respuesta utilizando los modelos 4.9 y 4.10 en la ecuación 2.42 para la generación de valores.

Fase IV: En cuanto a la cantidad de combinaciones a simular, en este ejemplo al igual que en el anterior se utilizó un arreglo 11^2 , es decir, se trabajó con 121 combinaciones diferentes de los factores de control x_2 y x_3 , utilizando los siguientes niveles para cada variable:

{-1.682 -1.345 -1.009 -0.672 -0.336 0 0.336 0.672 1.009 1.345 1.682}

Como puede observarse se consideró desde el nivel más pequeño para cada variable, hasta su nivel más grande, con saltos de 0.3364 para que quedaran espacios de la misma dimensión entre cada nivel de las variables. Las combinaciones resultantes se muestran la tabla 4.12. Una vez que se determinó la cantidad de combinaciones, se requiere determinar la cantidad de corridas a simularse en cada una de ellas utilizando la ecuación 2.42. En este caso, al igual que en el caso de estudio I, se utilizaron 5000 corridas en cada una de las 121 combinaciones.

Tabla 4.11 Datos Tomados para Ilustrar la Metodología: Caso de Estudio II

x_1	x_2	x_3	y_1	y_2
-1.000	-1.000	-1.000	74	53.2
1.000	-1.000	-1.000	51	62.9
-1.000	1.000	-1.000	88	53.4
1.000	1.000	-1.000	70	62.6
-1.000	-1.000	1.000	71	57.3
1.000	-1.000	1.000	90	67.9
-1.000	1.000	1.000	66	59.8
1.000	1.000	1.000	97	67.8
-1.682	0.000	0.000	76	59.1
1.682	0.000	0.000	79	65.9
0.000	-1.682	0.000	85	60.0
0.000	1.682	0.000	97	60.7
0.000	0.000	-1.682	55	57.4
0.000	0.000	1.682	81	63.2
0.000	0.000	0.000	81	59.2
0.000	0.000	0.000	75	60.4
0.000	0.000	0.000	76	59.1
0.000	0.000	0.000	83	60.6
0.000	0.000	0.000	80	60.8
0.000	0.000	0.000	91	58.9

Fase V: Cuando se han generado las iteraciones necesarias se procede a hacer los cálculos de la media geométrica con las ponderaciones que se crean convenientes de acuerdo a la relevancia en la funcionalidad o costos del producto final de cada una de las variables de respuesta analizadas. En este ejemplo se consideraron cuatro escenarios distintos, en cada uno de los cuales las medias geométricas se calcularon con las ponderaciones que se indican en la tabla 4.13.

En la tabla 4.14 se muestran los resultados obtenidos por medio de simulación para las 121 combinaciones de las variables de control x_2 y x_3 . Para realizar los cálculos de las columnas p_1 y p_2 se consideraron las siguientes especificaciones para cada característica de calidad: $\{y_1 \geq 80 \text{ y } 55 \leq y_2 \leq 60\}$. Por lo que cada vector generado se compara contra estos valores, para calcular las probabilidades de concordancia.

Tabla 4.12 Combinaciones de los Factores de Control Utilizadas para Formar la Malla en la Simulación: Caso de Estudio II

Corrida	1	2	3	4	5	6	7	8	9	10	11	12	13
X ₂	-1.682	-1.682	-1.682	-1.682	-1.682	-1.682	-1.682	-1.682	-1.682	-1.682	-1.682	-1.3456	-1.3456
X ₃	-1.682	-1.3456	-1.0092	-0.6728	-0.3364	0	0.3364	0.6728	1.0092	1.3456	1.682	-1.682	-1.3456
Corrida	14	15	16	17	18	19	20	21	22	23	24	25	26
X ₂	-1.3456	-1.3456	-1.3456	-1.3456	-1.3456	-1.3456	-1.3456	-1.3456	-1.3456	-1.0092	-1.0092	-1.0092	-1.0092
X ₃	-1.0092	-0.6728	-0.3364	0	0.3364	0.6728	1.0092	1.3456	1.682	-1.682	-1.3456	-1.0092	-0.6728
Corrida	27	28	29	30	31	32	33	34	35	36	37	38	39
X ₂	-1.0092	-1.0092	-1.0092	-1.0092	-1.0092	-1.0092	-1.0092	-0.6728	-0.6728	-0.6728	-0.6728	-0.6728	-0.6728
X ₃	-0.3364	0	0.3364	0.6728	1.0092	1.3456	1.682	-1.682	-1.3456	-1.0092	-0.6728	-0.3364	0
Corrida	40	41	42	43	44	45	46	47	48	49	50	51	52
X ₂	-0.6728	-0.6728	-0.6728	-0.6728	-0.6728	-0.3364	-0.3364	-0.3364	-0.3364	-0.3364	-0.3364	-0.3364	0.3364
X ₃	0.3364	0.6728	1.0092	1.3456	1.682	-1.682	-1.3456	-1.0092	-0.6728	-0.3364	0	0.3364	0.6728
Corrida	53	54	55	56	57	58	59	60	61	62	63	64	65
X ₂	-0.3364	-0.3364	-0.3364	0	0	0	0	0	0	0	0	0	0
X ₃	1.0092	1.3456	1.682	-1.682	-1.3456	-1.0092	-0.6728	-0.3364	0	0.3364	0.6728	1.0092	1.3456
Corrida	66	67	68	69	70	71	72	73	74	75	76	77	78
X ₂	0	0.3364	0.3364	0.3364	0.3364	0.3364	0.3364	0.3364	0.3364	0.3364	0.3364	0.3364	0.6728
X ₃	1.682	-1.682	-1.3456	-1.0092	-0.6728	-0.3364	0	0.3364	0.6728	1.0092	1.3456	1.682	-1.682
Corrida	79	80	81	82	83	84	85	86	87	88	89	90	91
X ₂	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	1.0092	1.0092	1.0092
X ₃	-1.3456	-1.0092	-0.6728	-0.3364	0	0.3364	0.6728	1.0092	1.3456	1.682	-1.682	-1.3456	-1.0092
Corrida	92	93	94	95	96	97	98	99	100	101	102	103	104
X ₂	1.0092	1.0092	1.0092	1.0092	1.0092	1.0092	1.0092	1.0092	1.3456	1.3456	1.3456	1.3456	1.3456
X ₃	-0.6728	-0.3364	0	0.3364	0.6728	1.0092	1.3456	1.682	-1.682	-1.3456	-1.0092	-0.6728	-0.3364
Corrida	105	106	107	108	109	110	111	112	113	114	115	116	117
X ₂	1.3456	1.3456	1.3456	1.3456	1.3456	1.3456	1.682	1.682	1.682	1.682	1.682	1.682	1.682
X ₃	0	0.3364	0.6728	1.0092	1.3456	1.682	-1.682	1.3456	1.0092	0.6728	0.3364	0	0.3364
Corrida	118	119	120	121									
X ₂	1.682	1.682	1.682	1.682									
X ₃	0.6728	1.0092	1.3456	1.682									

Tabla 4.13 Ponderaciones Consideradas para las Variables de Respuesta

Variable	Ponderaciones			
	Mgp_1	Mgp_2	Mgp_3	Mgp_4
y_1	0.5	0.6	0.8	0.2
y_2	0.5	0.4	0.2	0.8

Fase VI: Toda vez que se ha simulado la malla con las diferentes combinaciones de los factores de interés, lo siguiente es ajustar un modelo logístico para Mgp . En las ecuaciones 4.11, 4.12, 4.13 y 4.14, se muestran los modelos ajustados para la Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4 respectivamente considerando las variables de control, todos construidos utilizando el software minitab 15.

$$\begin{aligned} \log it(\hat{p}(x_c)) = & 0.158 + 0.523x_2 + 0.316x_3 - 1.204x_2x_3 + 0.297x_2^2 - 1.485x_3^2 - 0.134x_2^2x_3 \\ & + 0.529x_2x_3^2 + 0.155x_2^2x_3^2 - 0.180x_2^3 + 0.350x_3^3 + 0.154x_2^3x_3 + 0.201x_2x_3^3 - 0.034x_2^4 \\ & + 0.187x_3^4 - 0.054x_2^4x_3 - 0.261x_2x_3^4 + 0.014x_2^5 - 0.070x_3^5 - 0.008x_2^5x_3 - 0.019x_2x_3^5 \\ & - 0.006x_2^6 - 0.012x_3^6 + 0.017x_2^6x_3 + 0.038x_2x_3^6 \end{aligned} \quad (4.11)$$

$$\begin{aligned} \log it(\hat{p}(x_c)) = & 0.240 + 0.814x_2 + 0.982x_3 - 1.414x_2x_3 + 0.479x_2^2 - 1.651x_3^2 - 0.178x_2^2x_3 \\ & + 0.457x_2x_3^2 + 0.168x_2^2x_3^2 - 0.206x_2^3 + 0.264x_3^3 + 0.166x_2^3x_3 + 0.254x_2x_3^3 - 0.054x_2^4 \\ & + 0.167x_3^4 - 0.049x_2^4x_3 - 0.210x_2x_3^4 + 0.012x_2^5 - 0.064x_3^5 - 0.008x_2^5x_3 - 0.029x_2x_3^5 \\ & - 0.005x_2^6 - 0.006x_3^6 + 0.014x_2^6x_3 + 0.027x_2x_3^6 \end{aligned} \quad (4.12)$$

$$\begin{aligned} \log it(\hat{p}(x_c)) = & 0.045 + 0.065x_2 - 0.914x_3 - 0.659x_2x_3 + 0.063x_2^2 - 0.886x_3^2 - 0.018x_2^2x_3 \\ & + 0.335x_2x_3^2 + 0.088x_2^2x_3^2 - 0.089x_2^3 + 0.419x_3^3 + 0.091x_2^3x_3 + 0.064x_2x_3^3 - 0.011x_2^4 \\ & + 0.059x_3^4 - 0.035x_2^4x_3 - 0.187x_2x_3^4 + 0.012x_2^5 - 0.065x_3^5 - 0.006x_2^5x_3 + 0.005x_2x_3^5 \\ & - 0.005x_2^6 - 0.007x_3^6 + 0.014x_2^6x_3 + 0.029x_2x_3^6 \end{aligned} \quad (4.13)$$

$$\begin{aligned} \log it(\hat{p}(x_c)) = & 0.123 + 0.398x_2 + 0.009x_3 - 1.078x_2x_3 + 0.229x_2^2 - 1.370x_3^2 - 0.107x_2^2x_3 \\ & + 0.514x_2x_3^2 + 0.141x_2^2x_3^2 - 0.159x_2^3 + 0.376x_3^3 + 0.139x_2^3x_3 + 0.168x_2x_3^3 - 0.025x_2^4 \\ & + 0.176x_3^4 - 0.054x_2^4x_3 - 0.261x_2x_3^4 + 0.014x_2^5 - 0.071x_3^5 - 0.007x_2^5x_3 - 0.013x_2x_3^5 \\ & - 0.006x_2^6 - 0.011x_3^6 + 0.018x_2^6x_3 + 0.039x_2x_3^6 \end{aligned} \quad (4.14)$$

Tabla 4.14 Resultados de la Simulación del Arreglo 11²: Caso de Estudio II

x_2	x_3	ρ_1	ρ_2	Mgp_1	Mgp_2	Mgp_3	Mgp_4
-1.6820	-1.6820	0.0014	0.5248	0.0271	0.0150	0.0046	0.1604
-1.6820	-1.3456	0.0058	0.5774	0.0579	0.0365	0.0146	0.2301
-1.6820	-1.0092	0.0308	0.6230	0.1385	0.1025	0.0562	0.3414
-1.6820	-0.6728	0.1238	0.6318	0.2797	0.2376	0.1715	0.4560
-1.6820	-0.3364	0.3668	0.5716	0.4579	0.4380	0.4008	0.5231
-1.6820	0.0000	0.6462	0.5234	0.5816	0.5940	0.6195	0.5459
-1.6820	0.3364	0.8210	0.4260	0.5914	0.6315	0.7200	0.4857
-1.6820	0.6728	0.8958	0.3588	0.5669	0.6213	0.7460	0.4308
-1.6820	1.0092	0.9116	0.2754	0.5011	0.5648	0.7175	0.3499
-1.6820	1.3456	0.8904	0.2398	0.4621	0.5269	0.6849	0.3117
-1.6820	1.6820	0.8316	0.1944	0.4021	0.4650	0.6218	0.2600
-1.3456	-1.6820	0.0020	0.5702	0.0338	0.0192	0.0062	0.1841
-1.3456	-1.3456	0.0044	0.6366	0.0529	0.0322	0.0119	0.2354
-1.3456	-1.0092	0.0162	0.6814	0.1051	0.0723	0.0342	0.3226
-1.3456	-0.6728	0.0882	0.6670	0.2425	0.1981	0.1322	0.4450
-1.3456	-0.3364	0.2804	0.6302	0.4204	0.3877	0.3297	0.5360
-1.3456	0.0000	0.5718	0.5232	0.5470	0.5518	0.5617	0.5326
-1.3456	0.3364	0.7662	0.4348	0.5772	0.6108	0.6841	0.4870
-1.3456	0.6728	0.8440	0.3346	0.5314	0.5829	0.7014	0.4026
-1.3456	1.0092	0.8544	0.2506	0.4627	0.5231	0.6685	0.3203
-1.3456	1.3456	0.8168	0.2044	0.4086	0.4693	0.6191	0.2697
-1.3456	1.6820	0.7420	0.1650	0.3499	0.4067	0.5493	0.2229
-1.0092	-1.6820	0.0008	0.6064	0.0220	0.0113	0.0030	0.1610
-1.0092	-1.3456	0.0024	0.6676	0.0400	0.0228	0.0074	0.2166
-1.0092	-1.0092	0.0112	0.7014	0.0886	0.0586	0.0256	0.3066
-1.0092	-0.6728	0.0618	0.6874	0.2061	0.1620	0.1001	0.4246
-1.0092	-0.3364	0.2490	0.6364	0.3981	0.3624	0.3004	0.5275
-1.0092	0.0000	0.4958	0.5296	0.5124	0.5091	0.5024	0.5227
-1.0092	0.3364	0.6950	0.4260	0.5441	0.5714	0.6302	0.4698
-1.0092	0.6728	0.7690	0.3112	0.4892	0.5355	0.6417	0.3729
-1.0092	1.0092	0.7878	0.2216	0.4178	0.4743	0.6113	0.2856
-1.0092	1.3456	0.7216	0.1730	0.3533	0.4076	0.5423	0.2302
-1.0092	1.6820	0.6120	0.1364	0.2889	0.3357	0.4533	0.1842

Tabla 4.14 Continuación

x_2	x_3	p_1	p_2	Mgp_1	Mgp_2	Mgp_3	Mgp_4
-0.6728	-1.6820	0.0006	0.6138	0.0192	0.0096	0.0024	0.1535
-0.6728	-1.3456	0.0018	0.6942	0.0353	0.0195	0.0059	0.2110
-0.6728	-1.0092	0.0164	0.7290	0.1093	0.0748	0.0350	0.3413
-0.6728	-0.6728	0.0754	0.7072	0.2309	0.1846	0.1180	0.4520
-0.6728	-0.3364	0.2320	0.6406	0.3855	0.3483	0.2843	0.5228
-0.6728	0.0000	0.4816	0.5112	0.4962	0.4932	0.4874	0.5051
-0.6728	0.3364	0.6418	0.3980	0.5054	0.5301	0.5833	0.4379
-0.6728	0.6728	0.7100	0.3054	0.4657	0.5066	0.5998	0.3615
-0.6728	1.0092	0.7140	0.1944	0.3726	0.4243	0.5504	0.2522
-0.6728	1.3456	0.6252	0.1456	0.3017	0.3490	0.4671	0.1949
-0.6728	1.6820	0.4892	0.1134	0.2355	0.2726	0.3652	0.1519
-0.3364	-1.6820	0.0014	0.6182	0.0294	0.0160	0.0047	0.1829
-0.3364	-1.3456	0.0040	0.6972	0.0528	0.0315	0.0112	0.2484
-0.3364	-1.0092	0.0230	0.7406	0.1305	0.0922	0.0461	0.3698
-0.3364	-0.6728	0.0988	0.7010	0.2632	0.2163	0.1462	0.4737
-0.3364	-0.3364	0.2924	0.6302	0.4293	0.3975	0.3409	0.5405
-0.3364	0.0000	0.5074	0.5192	0.5133	0.5121	0.5097	0.5168
-0.3364	0.3364	0.6402	0.3914	0.5006	0.5258	0.5802	0.4319
-0.3364	0.6728	0.6902	0.2716	0.4330	0.4753	0.5728	0.3273
-0.3364	1.0092	0.6504	0.1774	0.3397	0.3868	0.5016	0.2300
-0.3364	1.3456	0.5434	0.1114	0.2460	0.2883	0.3958	0.1529
-0.3364	1.6820	0.3822	0.0872	0.1826	0.2116	0.2844	0.1172
0.0000	-1.6820	0.0026	0.6258	0.0403	0.0233	0.0078	0.2090
0.0000	-1.3456	0.0090	0.7070	0.0798	0.0516	0.0215	0.2954
0.0000	-1.0092	0.0482	0.7506	0.1902	0.1445	0.0835	0.4334
0.0000	-0.6728	0.1676	0.7044	0.3436	0.2976	0.2233	0.5286
0.0000	-0.3364	0.3884	0.6302	0.4947	0.4714	0.4279	0.5721
0.0000	0.0000	0.5776	0.5082	0.5418	0.5488	0.5630	0.5214
0.0000	0.3364	0.6826	0.3718	0.5038	0.5353	0.6045	0.4198
0.0000	0.6728	0.6950	0.2504	0.4172	0.4620	0.5666	0.3071
0.0000	1.0092	0.6238	0.1572	0.3131	0.3594	0.4735	0.2071
0.0000	1.3456	0.5046	0.0940	0.2178	0.2576	0.3606	0.1316
0.0000	1.6820	0.3264	0.0732	0.1546	0.1795	0.2420	0.0987

Tabla 4.14 Continuación

x_2	x_3	p_1	p_2	Mgp_1	Mgp_2	Mgp_3	Mgp_4
0.3364	-1.6820	0.0066	0.6022	0.0630	0.0401	0.0163	0.2442
0.3364	-1.3456	0.0238	0.7040	0.1294	0.0923	0.0469	0.3576
0.3364	-1.0092	0.1076	0.7564	0.2853	0.2347	0.1589	0.5121
0.3364	-0.6728	0.2980	0.7078	0.4593	0.4212	0.3543	0.5953
0.3364	-0.3364	0.5416	0.6090	0.5743	0.5676	0.5545	0.5949
0.3364	0.0000	0.6976	0.4836	0.5808	0.6025	0.6483	0.5204
0.3364	0.3364	0.7420	0.3552	0.5134	0.5526	0.6404	0.4116
0.3364	0.6728	0.7416	0.2192	0.4032	0.4554	0.5812	0.2797
0.3364	1.0092	0.6624	0.1288	0.2921	0.3441	0.4774	0.1787
0.3364	1.3456	0.4918	0.0824	0.2013	0.2407	0.3440	0.1178
0.3364	1.6820	0.3026	0.0578	0.1323	0.1561	0.2173	0.0805
0.6728	-1.6820	0.0262	0.5948	0.1248	0.0914	0.0489	0.3185
0.6728	-1.3456	0.0826	0.6946	0.2395	0.1936	0.1265	0.4537
0.6728	-1.0092	0.2628	0.7438	0.4421	0.3984	0.3236	0.6041
0.6728	-0.6728	0.5230	0.6946	0.6027	0.5859	0.5535	0.6563
0.6728	-0.3364	0.7224	0.6076	0.6625	0.6741	0.6978	0.6290
0.6728	0.0000	0.8162	0.4696	0.6191	0.6543	0.7308	0.5245
0.6728	0.3364	0.8534	0.3284	0.5294	0.5824	0.7050	0.3975
0.6728	0.6728	0.8136	0.1986	0.4020	0.4628	0.6137	0.2633
0.6728	1.0092	0.7052	0.1274	0.2997	0.3557	0.5008	0.1794
0.6728	1.3456	0.5464	0.0794	0.2083	0.2526	0.3715	0.1168
0.6728	1.6820	0.3114	0.0554	0.1313	0.1561	0.2205	0.0782
1.0092	-1.6820	0.0946	0.5532	0.2288	0.1917	0.1347	0.3886
1.0092	-1.3456	0.2574	0.6788	0.4180	0.3794	0.3125	0.5591
1.0092	-1.0092	0.5268	0.7128	0.6128	0.5945	0.5596	0.6710
1.0092	-0.6728	0.7714	0.6750	0.7216	0.7313	0.7511	0.6933
1.0092	-0.3364	0.8780	0.5880	0.7185	0.7479	0.8103	0.6371
1.0092	0.0000	0.9192	0.4500	0.6431	0.6908	0.7968	0.5191
1.0092	0.3364	0.9176	0.3036	0.5278	0.5895	0.7355	0.3788
1.0092	0.6728	0.8824	0.1806	0.3992	0.4678	0.6425	0.2480
1.0092	1.0092	0.7938	0.1146	0.3016	0.3660	0.5390	0.1688
1.0092	1.3456	0.6018	0.0724	0.2087	0.2580	0.3940	0.1106
1.0092	1.6820	0.3720	0.0468	0.1319	0.1623	0.2457	0.0708

Tabla 4.14 Continuación

x_2	x_3	p_1	p_2	Mgp_1	Mgp_2	Mgp_3	Mgp_4
1.3456	-1.6820	0.2920	0.5330	0.3945	0.3715	0.3293	0.4726
1.3456	-1.3456	0.5484	0.6370	0.5910	0.5823	0.5651	0.6182
1.3456	-1.0092	0.8026	0.6750	0.7360	0.7489	0.7753	0.6988
1.3456	-0.6728	0.9148	0.6576	0.7756	0.8016	0.8564	0.7025
1.3456	-0.3364	0.9468	0.5526	0.7233	0.7633	0.8501	0.6154
1.3456	0.0000	0.9676	0.4144	0.6332	0.6893	0.8167	0.4910
1.3456	0.3364	0.9592	0.2852	0.5230	0.5905	0.7526	0.3635
1.3456	0.6728	0.9322	0.1698	0.3979	0.4717	0.6631	0.2387
1.3456	1.0092	0.8534	0.1162	0.3149	0.3844	0.5727	0.1731
1.3456	1.3456	0.7006	0.0674	0.2173	0.2746	0.4386	0.1077
1.3456	1.6820	0.4410	0.0604	0.1632	0.1991	0.2963	0.0899
1.6820	-1.6820	0.5844	0.4950	0.5378	0.5468	0.5653	0.5117
1.6820	-1.3456	0.8088	0.5794	0.6846	0.7078	0.7566	0.6194
1.6820	-1.0092	0.9244	0.6252	0.7602	0.7905	0.8549	0.6761
1.6820	-0.6728	0.9688	0.6184	0.7740	0.8096	0.8856	0.6765
1.6820	-0.3364	0.9822	0.5132	0.7100	0.7576	0.8626	0.5843
1.6820	0.0000	0.9804	0.3842	0.6137	0.6740	0.8129	0.4634
1.6820	0.3364	0.9794	0.2780	0.5218	0.5918	0.7613	0.3576
1.6820	0.6728	0.9580	0.1734	0.4076	0.4835	0.6806	0.2441
1.6820	1.0092	0.9120	0.1114	0.3187	0.3933	0.5989	0.1696
1.6820	1.3456	0.7654	0.0834	0.2527	0.3154	0.4913	0.1299
1.6820	1.6820	0.5334	0.0598	0.1786	0.2223	0.3443	0.0926

Fase VII: Para apreciar el ajuste de los modelos a cada una de las medias geométricas para la que fue construido, se procedió a elaborar gráficos en matlab en donde se representarán al mismo tiempo las medias geométricas de las probabilidades de concordancia según las 5000 simulaciones para cada una de las 121 combinaciones de los factores de interés y las medias geométricas pronosticadas por los modelos anteriores en estas mismas combinaciones. Las comparaciones de los resultados simulados contra los pronosticados por el modelo para la Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4 se muestran en las figuras 4.13, 4.14, 4.15 y 4.16, respectivamente. Como puede verse los modelos pronostican con mucha exactitud lo obtenido mediante la simulación ya que el valor pronosticado prácticamente se sobrepone al valor simulado, por lo cual estos modelos pueden utilizarse para la optimización mediante métodos numéricos.

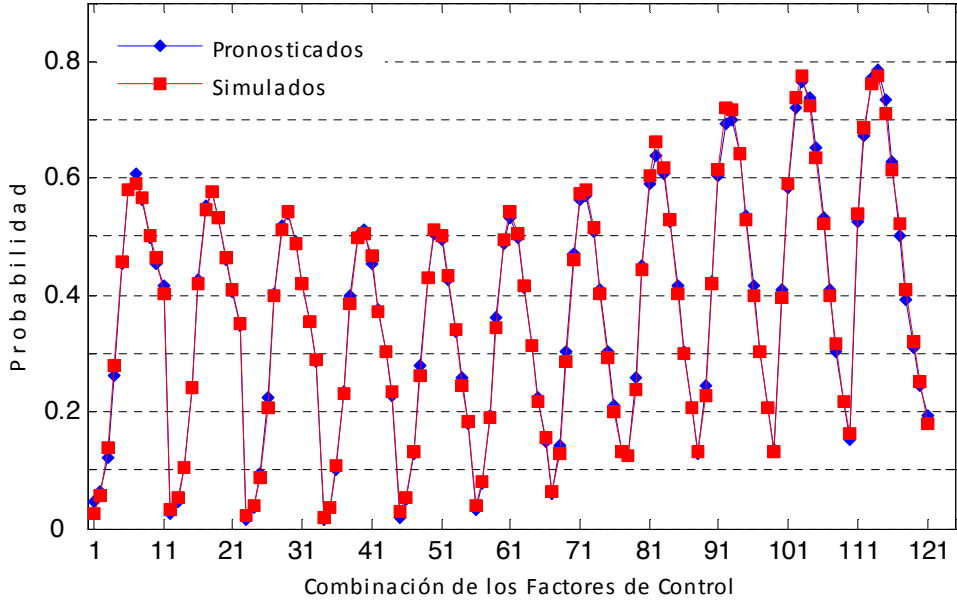


Figura 4.13 Resultados Simulados contra los Pronosticados para Mgp_1 .

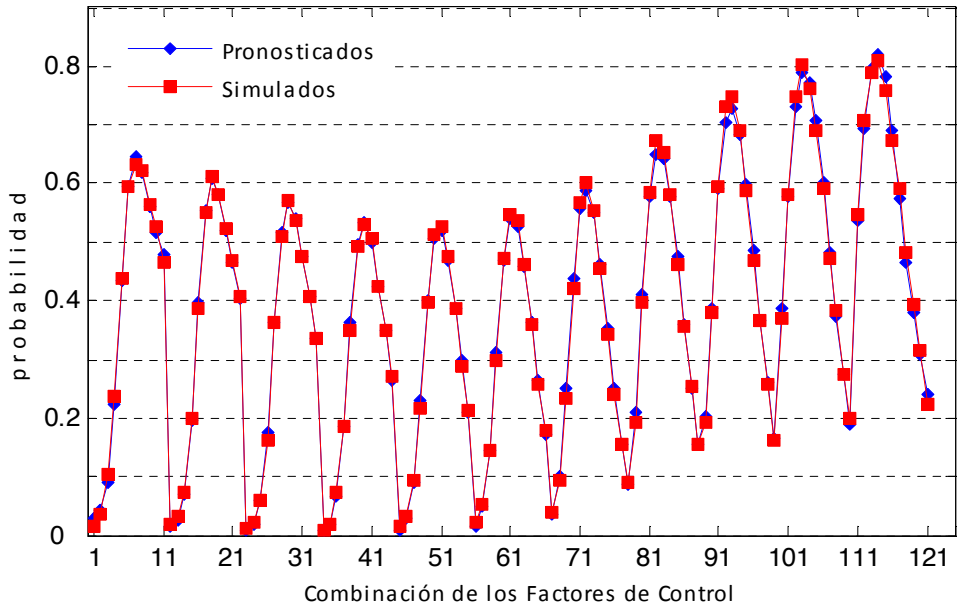


Figura 4.14 Resultados Simulados contra los Pronosticados para Mgp_2 .

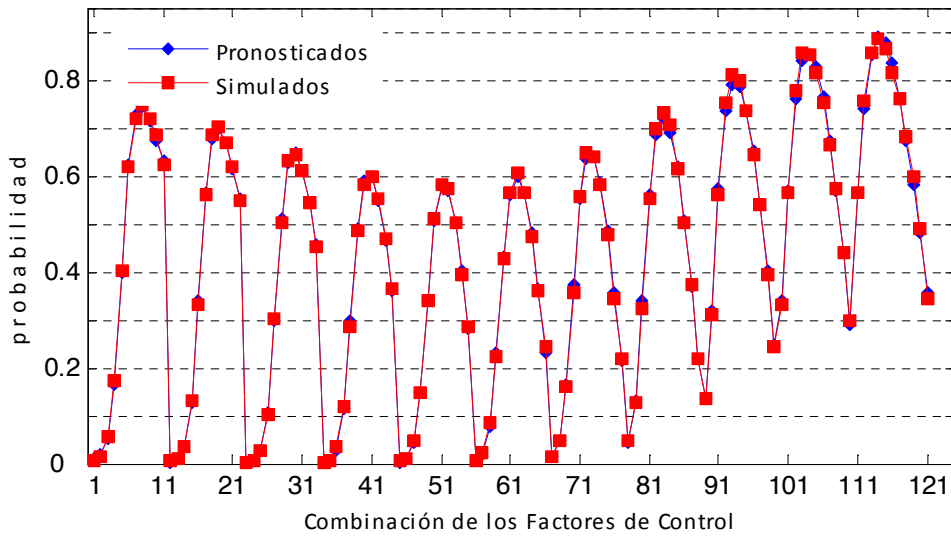


Figura 4.15 Resultados Simulados contra los Pronosticados para Mgp_3 .

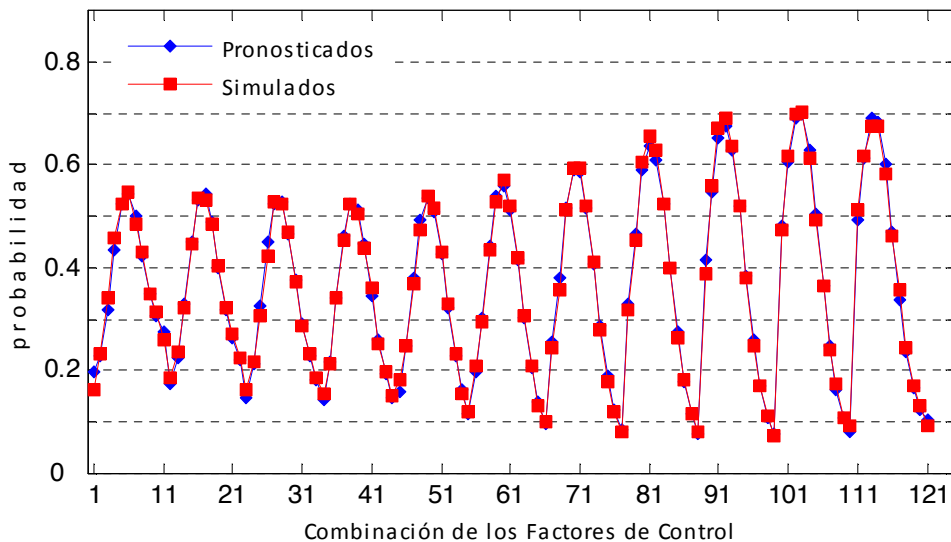


Figura 4.16 Resultados Simulados contra los Pronosticados para Mgp_4 .

Con los modelos logísticos construidos para Mgp_1 , Mgp_2 , Mgp_3 , y Mgp_4 se procede a construir sus respectivas gráficas de superficie y de contornos, las cuales se muestran en las figuras 4.17 y 4.18, 4.19 y 4.20, 4.21 y 4.22, y 4.23 y 4.24 para cada uno de los modelos respectivamente. Como se explicaba en el caso de estudio I, con estas gráficas se puede tener una idea previa de la localización de la región óptima de forma visual y no debe contraponerse a los resultados alcanzados mediante métodos numéricos. De hecho como puede apreciarse en estas figuras, en Mgp_1 , Mgp_2 y Mgp_3 el punto óptimo parece estar un poco por arriba de $x_2 = 1.5$ y alrededor de $x_3 = -0.8$ para Mgp_1 y Mgp_2 , pero en Mgp_3 x_3 parece estar más cercano a -0.5 . Por otra parte en el caso de Mgp_4 al parecer sus mejores resultados se obtienen en una vecindad mas acotada alrededor de $x_2 = 1.5$ y $x_3 = -0.8$. Para conocer la combinación óptima de las variables de control de forma exacta, es que se procede a la optimización de los polinomios mediante métodos numéricos como se ilustra posteriormente.

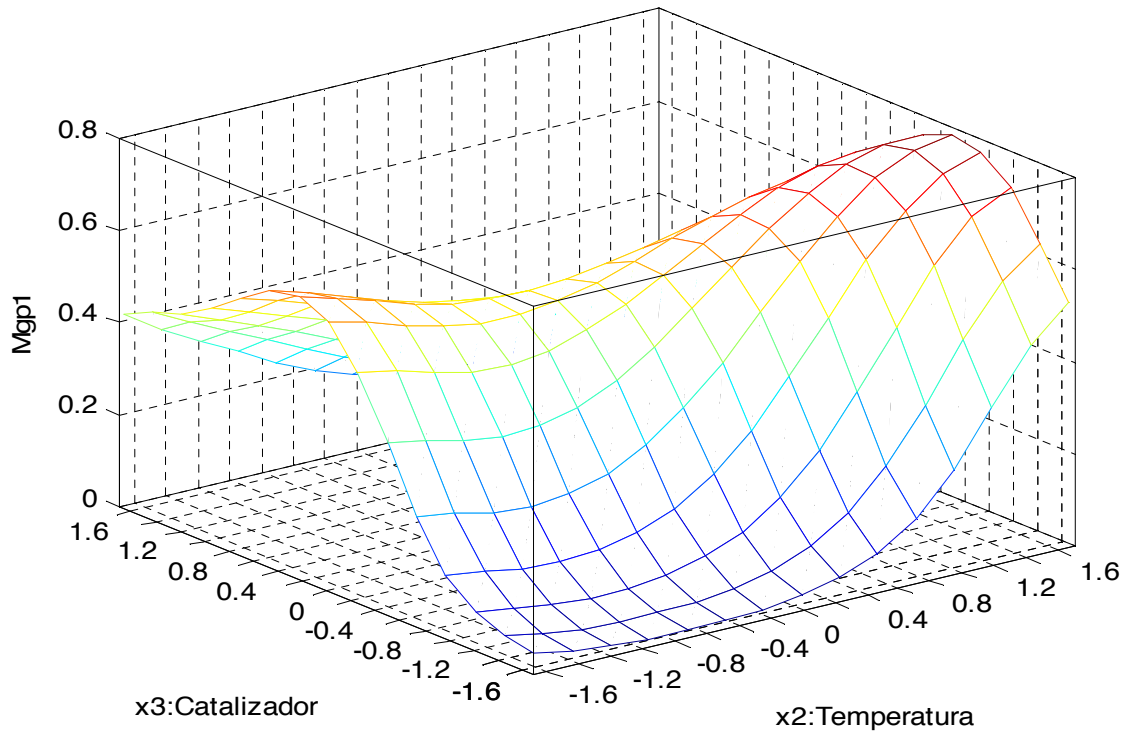


Figura 4.17 Gráfica de Superficie para Mgp_1

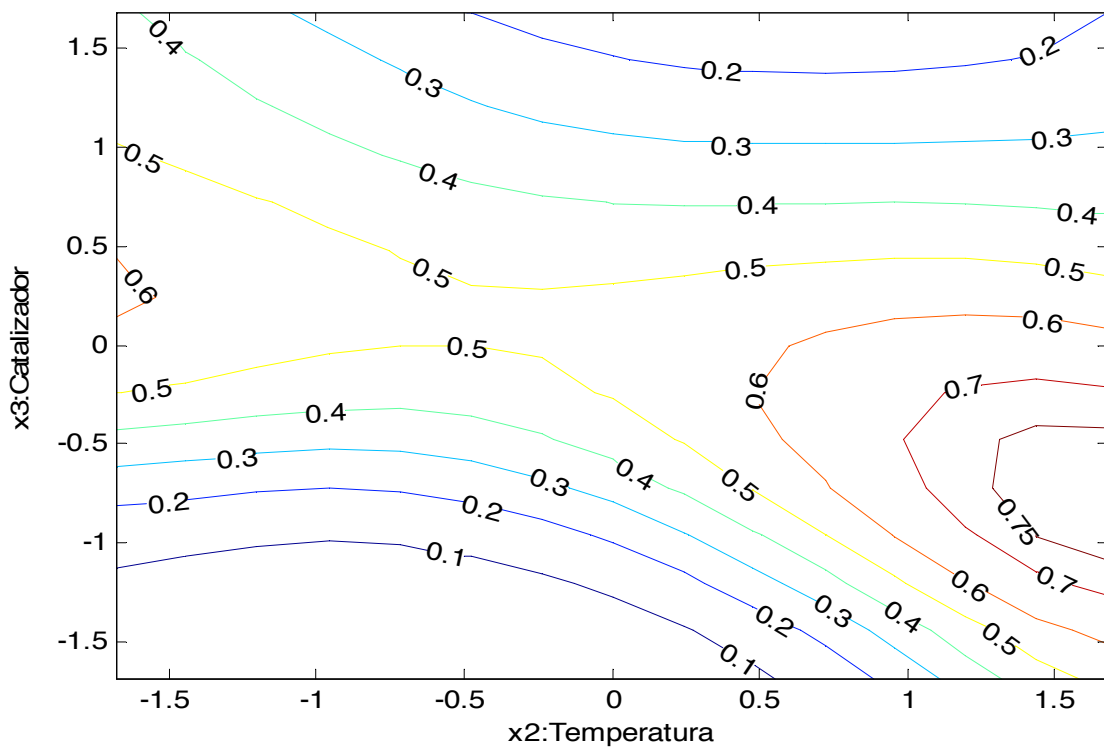


Figura 4.18 Gráfica de Contornos para Mgp_1

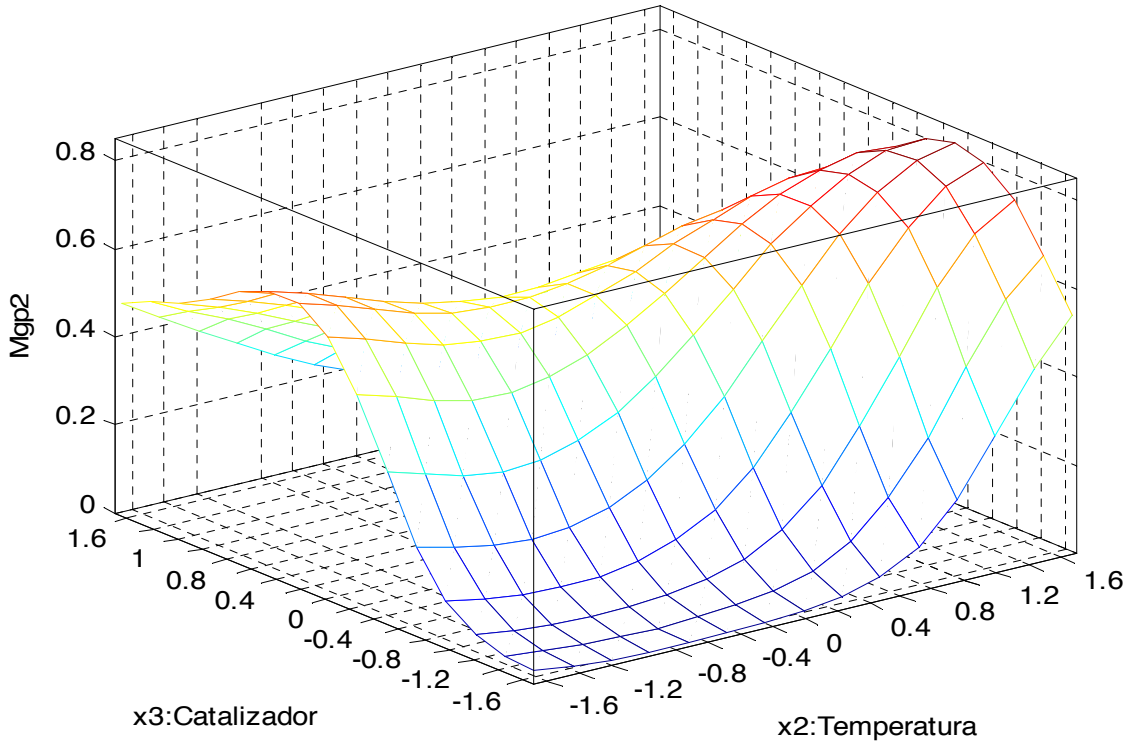


Figura 4.19 Gráfica de Superficie para Mgp_2

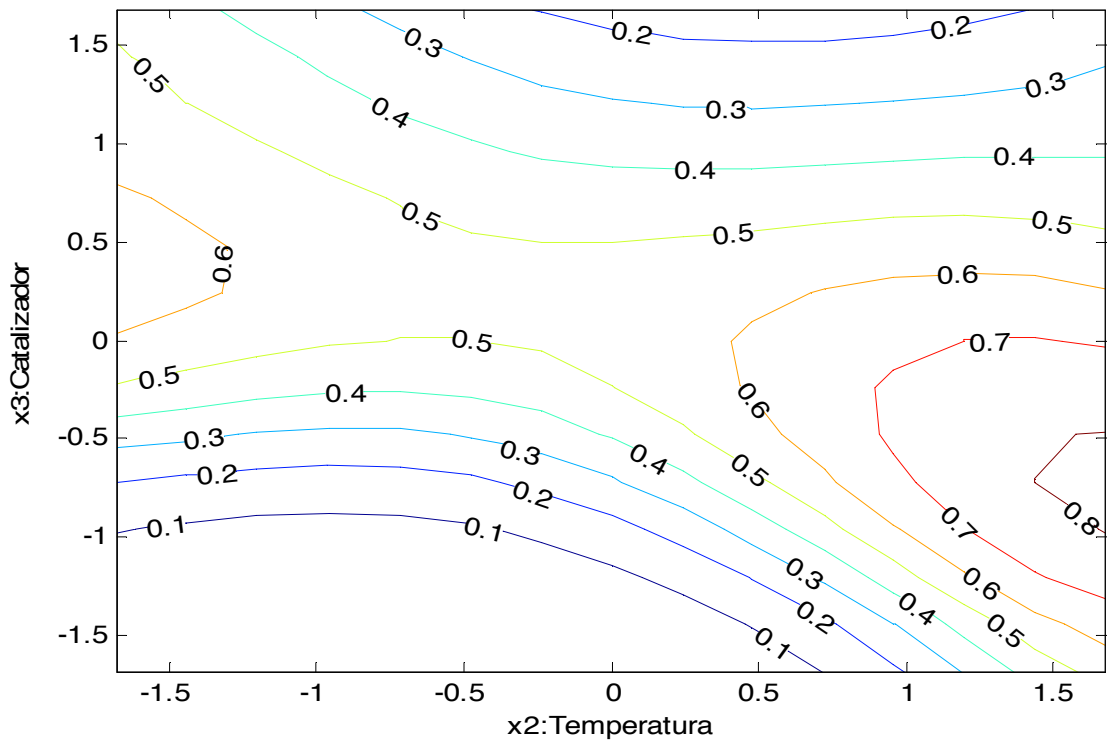


Figura 4.20 Gráfica de Contornos para Mgp_2

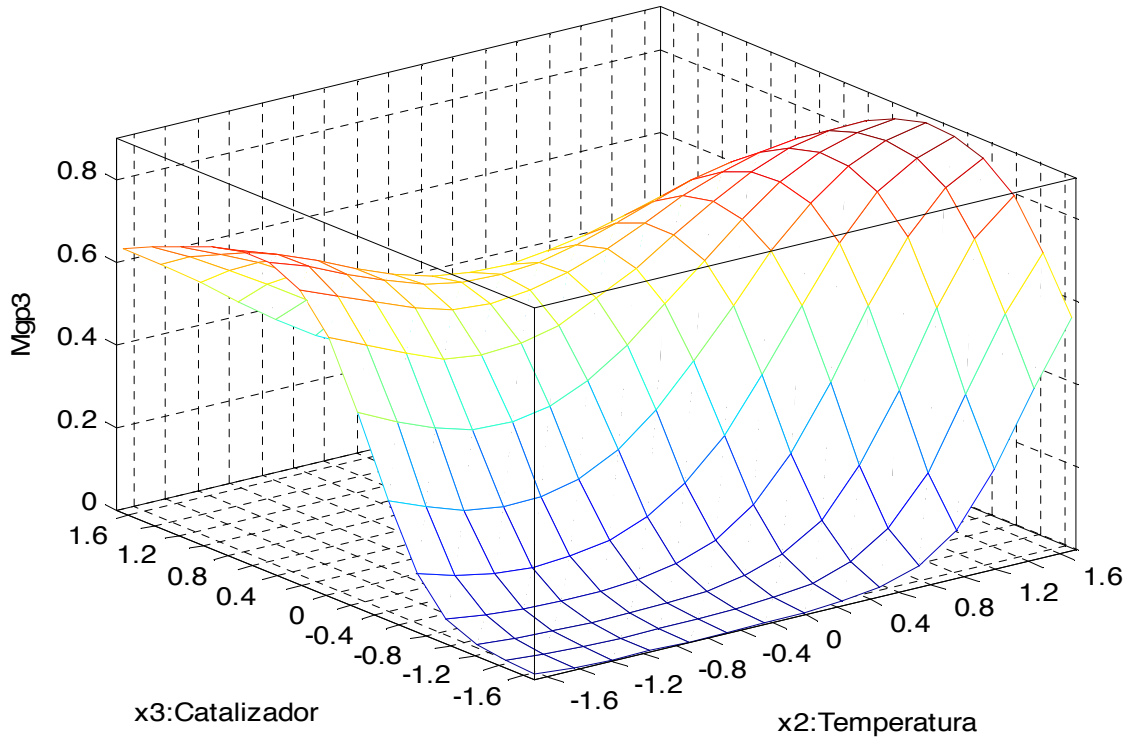


Figura 4.21 Gráfica de Superficie para Mgp_3

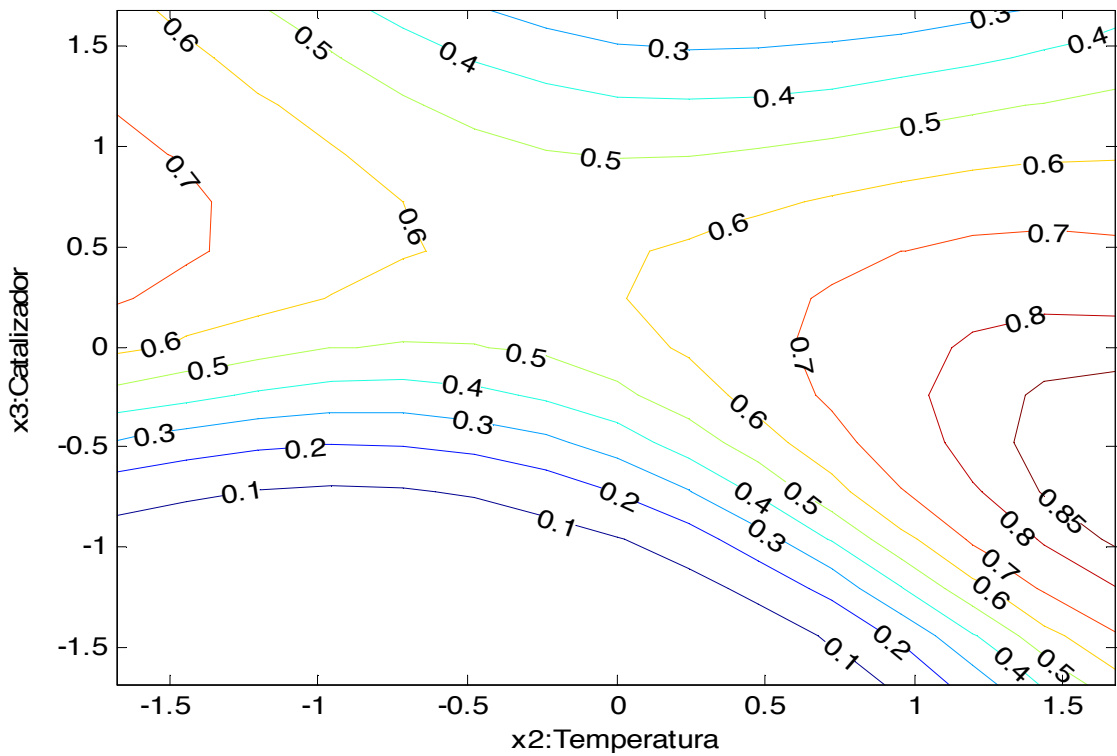


Figura 4.22 Gráfica de Contornos para Mgp_3

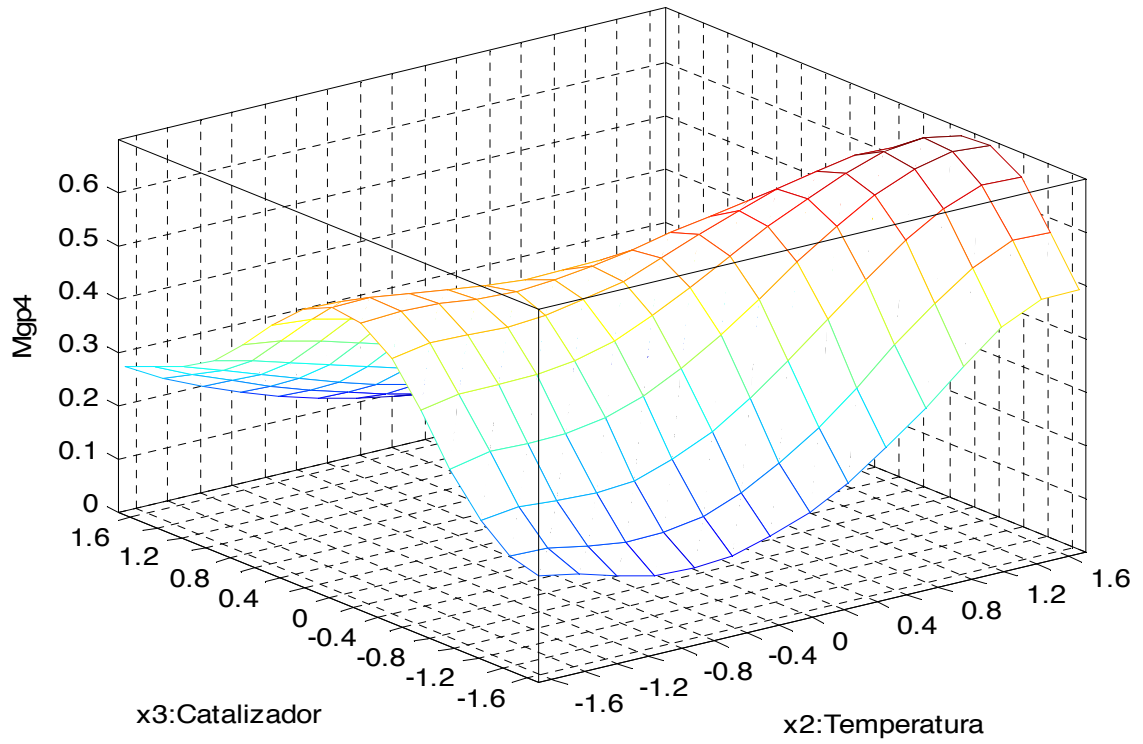


Figura 4.23 Gráfica de Superficie para Mgp_4

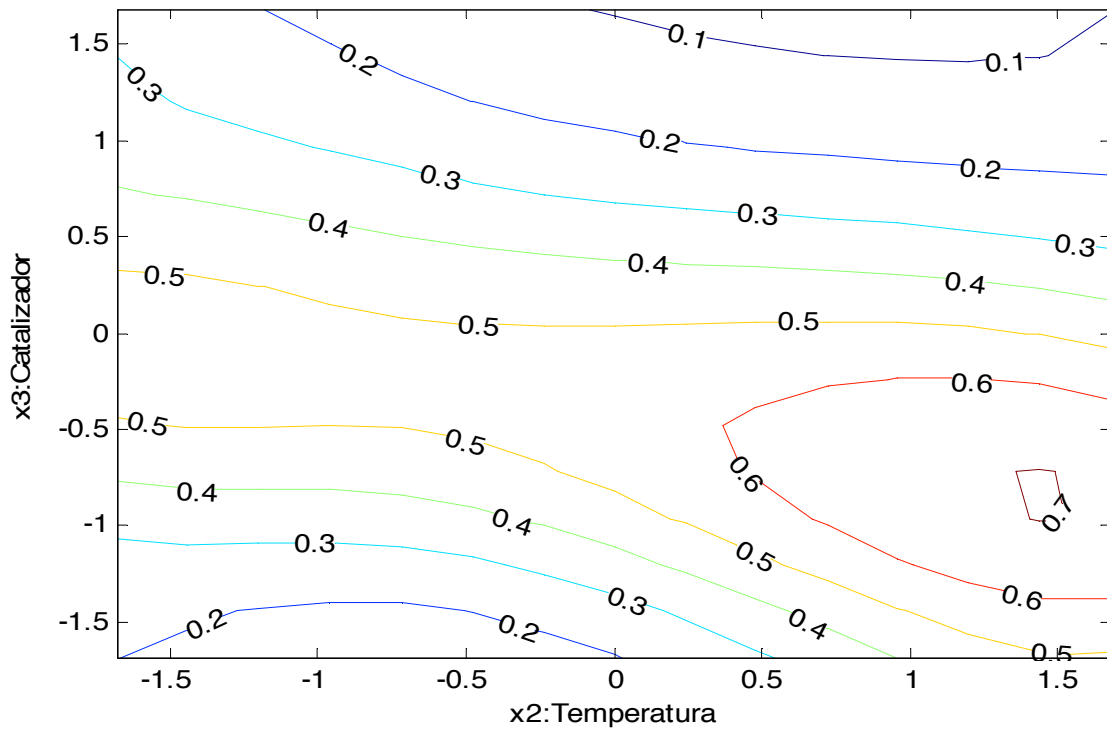


Figura 4.24 Gráfica de Contornos para Mgp_4

Fase VIII: Después de verificar gráficamente que los modelos realmente representen los resultados simulados, así como su significancia, se puede proceder a la fase de optimización con ellos, utilizando el *MGRG*. Debido a que se sugiere que se inicie el algoritmo con distintos valores para evaluar la consistencia de la solución alcanzada, se utilizaron 15 valores iniciales diferentes para la búsqueda de los valores óptimos al igual que en el caso de estudio I. Estos valores de partida fueron los mismos para los cuatro modelos logísticos. Los resultados obtenidos muestran que en los cuatro casos siempre hubo convergencia hacia el mismo resultado, independientemente de los valores iniciales, lo que puede apreciarse en la tabla 4.15.

Fase IX: Una vez obtenidas las condiciones óptimas de operación para el proceso, se procede a hacer en cada condición de operación propuesta para cada una de las medias geométricas ponderadas analizadas, una evaluación de los resultados mediante los modelos logísticos y posteriormente haciendo una simulación como corrida de confirmación de 100000 valores de cada respuesta. La tabla 4.16 muestra los resultados obtenidos para Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4 , tanto con la simulación como con los modelos logísticos y como puede verse estos son bastante cercanos en cada uno de los escenarios.

Después de comprobada la consistencia de los resultados entre los modelos y la simulación se procede a hacer 30 simulaciones de 100000 valores para cada respuesta, con el objeto de observar al igual que en el caso de estudio I, el efecto de las ponderaciones en las probabilidades de concordancia de cada variable.

En las tablas 4.17, 4.18, 4.19 y 4.20 se encuentra el resumen de las 30 simulaciones para Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4 respectivamente. En estas tablas se muestran las probabilidades de concordancia, los promedios y las desviaciones estándares para cada respuesta, así como las medias geométricas obtenidas en cada una de las simulaciones. Posteriormente en la tabla 4.21 se encuentran los promedios de cada uno de los estadísticos de las tablas anteriores y con esta información se procedió a calcular los índices Cpk y sus respectivos $Mcpk$ en cada una de las condiciones óptimas propuestas. Como puede observarse en esta tabla las variables que aumentaron su ponderación de un escenario con respecto a otro también aumentaron su índice Cpk , mientras que en aquellas en las que se redujo también se redujo su Cpk , mostrando consistencia con los resultados alcanzados en el ejemplo 1. Respecto al índice $Mcpk$ se puede notar que los valores más altos para cada ponderación coincidieron con la condición óptima encontrada para tal ponderación, excepto en el caso de $Mcpk_2$ que se obtienen prácticamente los mismos resultados en los primeros dos escenarios, debido a que sus ponderaciones son muy cercanas (0.5, 0.5) y (0.6, 0.4), sin embargo si se revisan los índices Cpk y las probabilidades de concordancia individuales se puede observar que la variable de respuesta con mejor ponderación en el escenario dos sí mejoró sus resultados como se esperaba.

Las comparaciones del comportamiento de las probabilidades de concordancia para y_1 y y_2 en los cuatro escenarios planteados se muestran de forma gráfica en las figuras 4.25 y 4.26 respectivamente. En estas puede verse que en el caso de y_1 su probabilidad de concordancia aumentó para Mgp_2 y Mgp_3 en relación a Mgp_1 ya que sus ponderaciones aumentaron de 0.5 a 0.6 y

0.8, pero en Mgp_4 su probabilidad cayó debido a que en este escenario tuvo su ponderación más baja (0.2). En el caso de y_2 el efecto fue contrario ya que sus probabilidades se redujeron para Mgp_2 y Mgp_3 en relación a Mgp_1 , lo cual se puede explicar ya que su ponderación se redujo de 0.5 a 0.4 y 0.2, sin embargo su probabilidad aumentó en Mgp_4 debido a que su ponderación tuvo el valor más alto (0.8). Por lo anterior se puede decir que en todas las respuestas la metodología reaccionó de acuerdo a las ponderaciones, moviéndose hacia zonas con mejores resultados para las variables más importantes.

Tabla 4.15 Resultados de la Optimización: Caso de Estudio II

Valores iniciales		Resultados óptimos encontrados							
		Mgp ₁		Mgp ₂		Mgp ₃		Mgp ₄	
X ₂	X ₃	X ₂	X ₃	X ₂	X ₃	X ₂	X ₃	X ₂	X ₃
0.40	-0.40	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.20	-1.00	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.70	0.70	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
-0.30	-0.70	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.90	-0.90	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.80	0.25	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.00	0.00	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.50	-0.50	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.30	0.50	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.25	0.25	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.47	0.47	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
0.94	0.80	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
1.50	-1.50	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
-0.10	0.10	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848
1.60	1.60	1.6633	-0.7668	1.682	-0.7307	1.682	-0.5969	1.4817	-0.848

Tabla 4.16 Evaluación por Simulación y según los Modelos Logísticos en las Condiciones Óptimas

Escenario	Variables (x ₂ , x ₃)	Resultados	
		Modelo	Simulación
Mgp_1	(1.6630, -0.7670)	0.7887	0.7714
Mgp_2	(1.6820, -0.7310)	0.8204	0.8075
Mgp_3	(1.6820, -0.5970)	0.8890	0.8821
Mgp_4	(1.4817, -0.8480)	0.7086	0.7042

Tabla 4.17 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp_1 ($x_2 = 1.663$, $x_3 = -0.767$)

Simulación	1	2	3	4	5	6	7	8	9	10	
Y_1	93.049	93.099	93.095	93.036	93.093	93.065	93.038	93.053	93.063	92.999	
Y_2	58.419	58.390	58.412	58.424	58.425	58.418	58.408	58.393	58.395	58.405	
S_1	7.618	7.583	7.592	7.550	7.620	7.594	7.603	7.609	7.565	7.595	
S_2	2.852	2.857	2.850	2.864	2.864	2.861	2.849	2.835	2.873	2.843	
P_1	0.958	0.959	0.959	0.959	0.958	0.959	0.958	0.958	0.960	0.958	
P_2	0.623	0.624	0.623	0.622	0.623	0.624	0.624	0.626	0.623	0.625	
(0.5,0.5)	0.772	0.773	0.773	0.772	0.773	0.774	0.773	0.774	0.773	0.773	
Simulación	11	12	13	14	15	16	17	18	19	20	
Y_1	93.070	93.032	93.069	93.033	93.050	93.027	93.016	93.052	93.060	93.081	
Y_2	58.416	58.404	58.426	58.419	58.415	58.407	58.413	58.402	58.403	58.411	
S_1	7.596	7.576	7.581	7.580	7.606	7.608	7.636	7.613	7.600	7.612	
S_2	2.855	2.859	2.857	2.856	2.848	2.860	2.870	2.855	2.865	2.858	
P_1	0.958	0.958	0.958	0.959	0.958	0.958	0.958	0.958	0.959	0.958	
P_2	0.625	0.621	0.621	0.624	0.625	0.622	0.622	0.625	0.624	0.623	
(0.5,0.5)	0.774	0.771	0.771	0.774	0.774	0.772	0.772	0.774	0.774	0.773	
Simulación	21	22	23	24	25	26	27	28	29	30	Promedios
Y_1	92.989	93.048	93.091	93.050	93.008	93.024	93.043	93.056	93.065	93.010	93.049
Y_2	58.406	58.414	58.400	58.410	58.397	58.409	58.432	58.412	58.415	58.428	58.411
S_1	7.596	7.613	7.615	7.590	7.612	7.646	7.599	7.647	7.624	7.601	7.603
S_2	2.857	2.849	2.879	2.862	2.862	2.864	2.840	2.862	2.865	2.840	2.857
P_1	0.958	0.958	0.959	0.959	0.957	0.958	0.958	0.958	0.959	0.958	0.958
P_2	0.623	0.623	0.621	0.622	0.625	0.624	0.624	0.623	0.621	0.624	0.623
(0.5,0.5)	0.773	0.773	0.772	0.772	0.773	0.773	0.773	0.772	0.771	0.774	0.773

Tabla 4.18 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp_2 ($x_2 = 1.682$, $x_3 = -0.731$)

Simulación	1	2	3	4	5	6	7	8	9	10	
Y_1	93.646	93.672	93.681	93.623	93.701	93.642	93.618	93.648	93.660	93.588	
Y_2	58.518	58.492	58.519	58.513	58.527	58.532	58.504	58.492	58.502	58.506	
S_1	7.633	7.588	7.623	7.545	7.638	7.581	7.626	7.605	7.568	7.602	
S_2	2.859	2.858	2.858	2.858	2.873	2.866	2.852	2.841	2.868	2.852	
P_1	0.963	0.964	0.964	0.965	0.964	0.964	0.964	0.964	0.965	0.963	
P_2	0.615	0.617	0.615	0.617	0.615	0.616	0.618	0.619	0.616	0.617	
(0.6,0.4)	0.805	0.806	0.805	0.807	0.805	0.806	0.807	0.807	0.806	0.806	
Simulación	11	12	13	14	15	16	17	18	19	20	
Y_1	93.660	93.616	93.653	93.619	93.648	93.625	93.590	93.652	93.632	93.678	
Y_2	58.515	58.508	58.519	58.525	58.506	58.519	58.508	58.508	58.507	58.509	
S_1	7.592	7.587	7.596	7.583	7.609	7.623	7.636	7.621	7.619	7.604	
S_2	2.854	2.866	2.860	2.859	2.853	2.864	2.874	2.851	2.872	2.865	
P_1	0.964	0.964	0.963	0.964	0.963	0.964	0.963	0.964	0.964	0.964	
P_2	0.618	0.615	0.614	0.616	0.617	0.614	0.616	0.617	0.617	0.616	
(0.6,0.4)	0.807	0.805	0.804	0.806	0.806	0.805	0.805	0.807	0.806	0.806	
Simulación	21	22	23	24	25	26	27	28	29	30	Promedios
Y_1	93.584	93.632	93.681	93.642	93.587	93.603	93.645	93.650	93.634	93.621	93.638
Y_2	58.502	58.517	58.510	58.504	58.499	58.510	58.532	58.512	58.516	58.526	58.512
S_1	7.604	7.624	7.607	7.597	7.627	7.648	7.612	7.644	7.629	7.612	7.609
S_2	2.856	2.854	2.875	2.871	2.867	2.864	2.848	2.867	2.859	2.851	2.860
P_1	0.964	0.963	0.964	0.964	0.963	0.963	0.963	0.963	0.964	0.964	0.964
P_2	0.617	0.617	0.615	0.615	0.617	0.616	0.616	0.616	0.615	0.616	0.616
(0.6,0.4)	0.807	0.806	0.805	0.806	0.805	0.806	0.806	0.806	0.805	0.806	0.806

Tabla 4.19 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp_3 ($x_2 = 1.682$, $x_3 = -0.597$)

Simulación	1	2	3	4	5	6	7	8	9	10	
Y_1	94.534	94.528	94.535	94.568	94.517	94.519	94.523	94.535	94.523	94.536	
Y_2	58.879	58.868	58.874	58.875	58.887	58.880	58.866	58.876	58.879	58.887	
S_1	7.484	7.447	7.441	7.442	7.449	7.479	7.456	7.480	7.464	7.438	
S_2	2.805	2.798	2.800	2.801	2.812	2.802	2.803	2.802	2.786	2.813	
P_1	0.973	0.973	0.973	0.974	0.974	0.972	0.973	0.972	0.973	0.973	
P_2	0.594	0.594	0.594	0.596	0.592	0.591	0.594	0.594	0.594	0.593	
(0.8,0.2)	0.881	0.881	0.881	0.882	0.881	0.880	0.882	0.881	0.882	0.881	
Simulación	11	12	13	14	15	16	17	18	19	20	
Y_1	94.532	94.462	94.512	94.580	94.573	94.555	94.542	94.571	94.542	94.510	
Y_2	58.866	58.877	58.885	58.886	58.878	58.883	58.871	58.879	58.883	58.873	
S_1	7.476	7.412	7.429	7.418	7.452	7.393	7.486	7.430	7.457	7.463	
S_2	2.811	2.785	2.795	2.805	2.796	2.801	2.799	2.783	2.798	2.801	
P_1	0.973	0.972	0.973	0.973	0.974	0.974	0.972	0.974	0.973	0.972	
P_2	0.593	0.595	0.595	0.595	0.594	0.595	0.594	0.596	0.595	0.595	
(0.8,0.2)	0.881	0.881	0.881	0.882	0.882	0.882	0.881	0.883	0.882	0.881	
Simulación	21	22	23	24	25	26	27	28	29	30	Promedios
Y_1	94.496	94.529	94.486	94.483	94.520	94.476	94.528	94.516	94.523	94.561	94.527
Y_2	58.870	58.874	58.851	58.884	58.882	58.876	58.869	58.851	58.872	58.886	58.876
S_1	7.453	7.464	7.439	7.440	7.502	7.473	7.486	7.439	7.432	7.428	7.452
S_2	2.794	2.811	2.808	2.791	2.802	2.801	2.827	2.803	2.801	2.799	2.801
P_1	0.973	0.973	0.972	0.972	0.972	0.972	0.973	0.973	0.973	0.974	0.973
P_2	0.595	0.593	0.596	0.596	0.595	0.594	0.592	0.595	0.595	0.594	0.594
(0.8,0.2)	0.882	0.881	0.882	0.881	0.881	0.881	0.881	0.882	0.882	0.882	0.881

Tabla 4.20 Resumen de las 30 Iteraciones en la Condición Óptima para Mgp_4 ($x_2 = 1.4817$, $x_3 = -0.848$)

Simulación	1	2	3	4	5	6	7	8	9	10	
Y_1	89.385	89.398	89.411	89.355	89.358	89.405	89.393	89.374	89.346	89.392	
Y_2	58.183	58.166	58.141	58.151	58.158	58.161	58.140	58.160	58.161	58.148	
S_1	7.250	7.235	7.285	7.204	7.240	7.217	7.232	7.261	7.225	7.270	
S_2	2.710	2.709	2.711	2.715	2.736	2.725	2.722	2.722	2.719	2.721	
P_1	0.912	0.913	0.912	0.912	0.910	0.913	0.913	0.911	0.912	0.913	
P_2	0.656	0.660	0.659	0.659	0.656	0.657	0.660	0.658	0.661	0.658	
(0.2,0.8)	0.701	0.704	0.703	0.703	0.701	0.701	0.704	0.702	0.705	0.702	
Simulación	11	12	13	14	15	16	17	18	19	20	
Y_1	89.387	89.364	89.407	89.410	89.367	89.408	89.376	89.382	89.383	89.398	
Y_2	58.163	58.151	58.156	58.161	58.155	58.160	58.161	58.168	58.157	58.149	
S_1	7.259	7.226	7.260	7.231	7.298	7.237	7.257	7.233	7.256	7.255	
S_2	2.735	2.724	2.725	2.719	2.718	2.720	2.714	2.706	2.710	2.721	
P_1	0.911	0.912	0.912	0.912	0.909	0.912	0.911	0.912	0.912	0.912	
P_2	0.656	0.658	0.657	0.657	0.658	0.658	0.658	0.659	0.660	0.660	
(0.2,0.8)	0.701	0.702	0.702	0.702	0.702	0.702	0.703	0.703	0.704	0.704	
Simulación	21	22	23	24	25	26	27	28	29	30	Promedios
Y_1	89.373	89.381	89.378	89.402	89.436	89.383	89.381	89.373	89.397	89.343	89.385
Y_2	58.163	58.160	58.168	58.169	58.142	58.168	58.162	58.155	58.156	58.163	58.159
S_1	7.233	7.278	7.207	7.242	7.232	7.236	7.230	7.270	7.246	7.268	7.246
S_2	2.707	2.728	2.720	2.716	2.715	2.723	2.711	2.725	2.728	2.715	2.719
P_1	0.912	0.912	0.912	0.913	0.913	0.912	0.913	0.911	0.912	0.911	0.912
P_2	0.659	0.659	0.656	0.658	0.660	0.659	0.660	0.657	0.658	0.657	0.658
(0.2,0.8)	0.703	0.703	0.701	0.703	0.704	0.703	0.705	0.701	0.702	0.702	0.703

Tabla 4.21 Estadísticos Resumen para las Condiciones Óptimas Propuestas según Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4

Variables de control		Probabilidades de concordancia		Promedios y desviaciones		Cpks		Mcpks			
X_1	X_2	P_1	P_2	(\bar{Y}_1, S_1)	(\bar{Y}_2, S_2)	Cpk_{y1}	Cpk_{y2}	$Mcpk_1$ (0.5,0.5)	$Mcpk_2$ (0.6,0.4)	$Mcpk_3$ (0.8,0.2)	$Mcpk_4$ (0.2,0.8)
1.663	-0.767	0.958	0.623	(93.049, 7.603)	(58.411, 2.857)	0.572	0.185	0.326	0.365	0.457	0.232
1.682	-0.731	0.964	0.616	(93.638, 7.609)	(58.512, 2.860)	0.597	0.173	0.322	0.364	0.466	0.222
1.682	-0.597	0.973	0.594	(94.527, 7.452)	(58.876, 2.801)	0.650	0.134	0.295	0.345	0.474	0.184
1.4817	-0.848	0.912	0.658	(89.385, 7.246)	(58.159, 2.719)	0.432	0.226	0.312	0.333	0.379	0.257

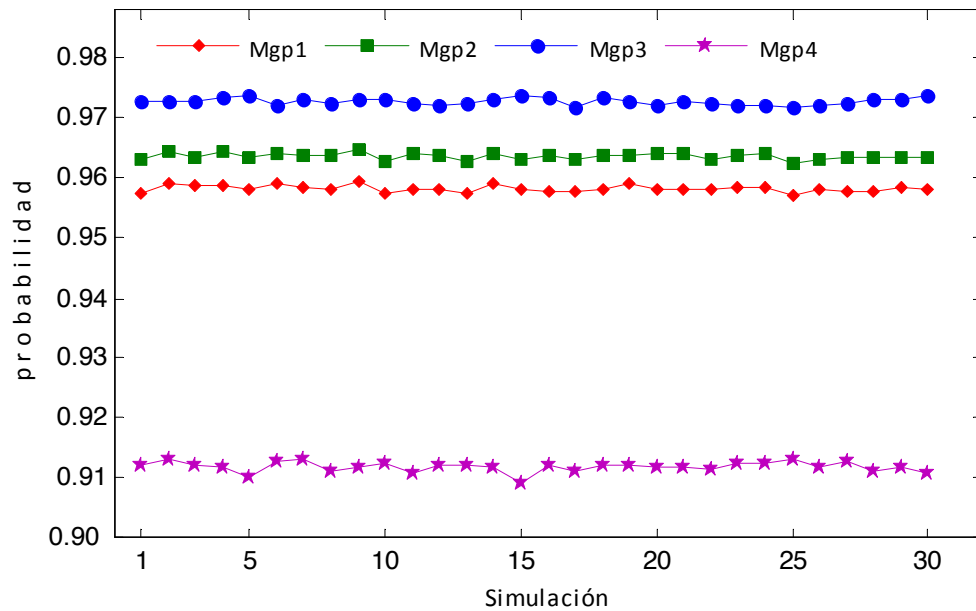


Figura 4.25 Comportamiento de la Probabilidad de Concordancia para y_1 en las Condiciones Óptimas Propuestas para los Cuatro Escenarios Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4 , con Ponderaciones 0.5, 0.6, 0.8 y 0.2 Respectivamente.

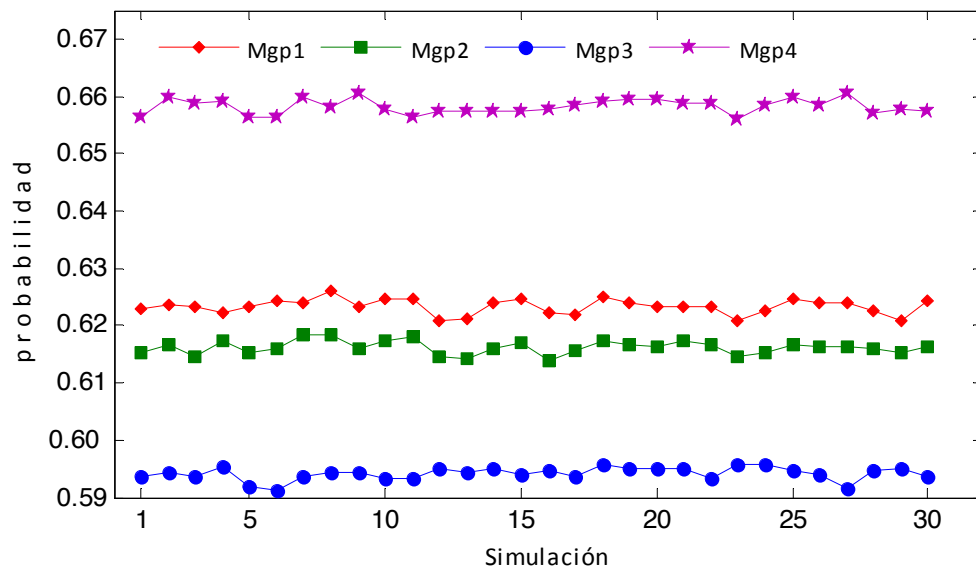


Figura 4.26 Comportamiento de la Probabilidad de Concordancia para y_2 en las Condiciones Óptimas Propuestas para los Cuatro Escenarios Mgp_1 , Mgp_2 , Mgp_3 y Mgp_4 , con Ponderaciones 0.5, 0.4, 0.2 y 0.8 Respectivamente.

Fase X: Esta fase se explica en el siguiente apartado.

4.3 Discusión de Resultados

4.3.1 Conclusiones Fase X: Análisis de Sistema HPLC

En el caso de estudio I utilizado para ilustrar la modificación planteada a la metodología propuesta por Miró-Quesada et al. (2004), se puede apreciar el efecto de utilizar la media geométrica ponderada como la variable de respuesta a optimizar, ya que, por ejemplo, si solo se utiliza la media geométrica sin ponderar la mejor condición de operación resultó en (0.514, -1) que está en la región de x_1 entre 0.4 y 0.6, en la cual la media geométrica tiene un mejor desempeño, de hecho muy similar, pero ligeramente mejor para $x_1 = 0.4$. Una vez que la *Mg* se sustituyó por la *Mgp*, se pudo observar que esta última penaliza a la región de $x_1 = 0.4$, mientras que aumenta en $x_1 = 0.6$ por tener un mejor desempeño en las variables y_2 y y_3 que fueron las de mayor ponderación, de tal forma que ahora la región $x_1 = 0.6$ tiene un mejor desempeño que $x_1 = 0.4$, cuestión que debiera hacer que la condición óptima de operación se mueva un poco hacia $x_1 = 0.6$, hecho que efectivamente sucedió, ya que la nueva solución sugerida fue (0.552, -1).

Como puede apreciarse en este ejemplo, sí hubo un cambio en la región de operación óptima para el proceso, cargándose hacia las condiciones en donde las variables más importantes se comportan de mejor manera. El desplazamiento observado fue pequeño de $x_1 = 0.514$ a $x_1 = 0.552$, por lo que el cambio pareciera no ser tan evidente. Sin embargo esto sucedió porque las regiones de operación con $x_1 = 0.4$ y 0.6, son las que mejor se desempeñan respecto a la *Mg*, muy por encima de las otras condiciones y además tienen un desempeño muy similar en las cuatro variables analizadas, (de hecho para marcar un poco más estas diferencias es que las especificaciones se modificaron con respecto al trabajo de referencia) provocando que aún en la *Mgp* su desempeño siga siendo muy parecido, por lo que a la hora de optimizar mediante el *MGRG* la solución si se mueve hacia $x_1 = 0.6$ pero no parece ser tan claro.

En la tabla 4.22 se muestran las estimaciones por intervalo con un 95% de confianza de las probabilidades de concordancia individuales, partiendo de las estimaciones puntuales obtenidas en las condiciones óptimas que se muestran en la tabla 4.10. En estos intervalos puede verse que existe una diferencia estadísticamente significativa en las probabilidades de concordancia de las variables de respuesta, dependiendo de los pesos que tengan en los dos escenarios planteados. En general puede observarse que los límites de los intervalos se desplazan hacia arriba conforme las ponderaciones aumentan y que se tiene el efecto contrario cuando las ponderaciones se reducen. En el caso de los intervalos para P_1 (relacionados con y_1), se puede observar que aunque una menor ponderación en *Mgp* jala los valores de los límites hacia abajo, este desplazamiento no es suficiente ya que se alcanzan a presentar algunos valores comunes en ambos intervalos por lo que la diferencia no es significativa para un 95% de confianza; sin embargo, sí lo es para un nivel de confianza del 85%.

Tabla 4.22 Intervalos de Confianza para las Probabilidades de Concordancia de cada Variable de Respuesta en las Condiciones Óptimas para el Caso de Estudio I

	Pesos	Estimación puntual				Estimación por intervalo al 95% de confianza			
		P ₁	P ₂	P ₃	P ₄	P ₁	P ₂	P ₃	P ₄
<i>Mg</i>	(0.25, 0.25, 0.25, 0.25)	0.990	0.939	0.994	0.908	(0.989547, 0.990453)*	(0.937517, 0.940483)	(0.993521, 0.994479)	(0.906209, 0.909791)
<i>Mgp</i>	(0.10, 0.40, 0.40, 0.10)	0.989	0.945	0.995	0.886	(0.988525, 0.989475)*	(0.943587, 0.946413)	(0.994563, 0.995437)	(0.884030, 0.887970)

*Intervalos construidos con un 85% de confianza

4.3.2 Conclusiones Fase X: Porcentaje de Conversión y Actividad Térmica

En los resultados que se muestran en la tabla 4.15 puede apreciarse que los valores óptimos de x_2 y x_3 en Mgp_1 , Mgp_2 y Mgp_3 son similares, sin embargo, también se observa en las tablas 4.17 y 4.18 y de manera resumida en la tabla 4.21 que si se compara Mgp_2 con Mgp_1 , el procedimiento por sí solo busca condiciones en donde se comporte de mejor manera y_1 sin afectar demasiado a y_2 , debido a que sus ponderaciones no son tan diferentes, pero si se comparan los resultados de Mgp_1 contra los de Mgp_3 en las tablas 4.17 y 4.19, las modificaciones a favor de y_1 y en contra de y_2 se empiezan a hacer más evidentes, esto debido a que las diferencias en las ponderaciones se hacen más grandes a favor de y_1 . De la misma manera, al considerar los resultados para Mgp_4 en la tabla 4.20, se puede observar que los valores óptimos propuestos tanto para x_2 como para x_3 , ahora sí son bastante diferentes a los anteriores, sobre todo en x_2 , esto es debido a que en esta ocasión se le dio una ponderación de 0.2 a y_1 y de 0.8 a y_2 , por lo que la metodología busca la zona en donde y_2 tuvo su mejor desempeño a expensas de y_1 . Todo lo anterior puede apreciarse de manera resumida en las probabilidades de concordancia, las medias y los índices Cpk en la tabla 4.21.

Adicionalmente en la tabla 4.14 puede observarse que Mgp_4 tiene su resultado más alto en $x_2 = 1.3456$ y $x_3 = -0.6728$ por lo que el resultado óptimo alcanzado es muy similar a este según la tabla 4.15. Respecto a Mgp_1 también tiene su mejor resultado en $x_2 = 1.3456$ y $x_3 = -0.6728$ de acuerdo con la tabla 4.14, sin embargo su resultado óptimo se encontró en $x_2 = 1.6630$ y $x_3 = -0.7670$, esto fue debido a que Mgp_1 también se comporta de buena manera en $x_2 = 1.682$, de hecho su segundo y tercer resultados más altos se encuentran en esta zona con valores muy cercanos al primero. Es importante señalar también que al resolver el problema presentado como ejemplo en este trabajo, Myers y Montgomery (1995) encuentran como mejor propuesta para solución óptima $x_2 = 1.682$ y $x_3 = -0.5955$, la cual es similar a las encontradas utilizando Mgp_1 y Mgp_2 y es casi idéntica a la obtenida con Mgp_3 .

En la tabla 4.23 se muestran las estimaciones por intervalo con un 95% de confianza de las probabilidades de concordancia individuales, partiendo de las estimaciones puntuales obtenidas en las condiciones óptimas que se muestran en la tabla 4.21. En estos intervalos puede verse que existe una diferencia estadísticamente significativa en las probabilidades de concordancia de las variables de respuesta, dependiendo de los pesos que tengan en los cuatro escenarios planteados. En general,

al igual que en el caso de estudio I, puede observarse que los límites de los intervalos se desplazan hacia arriba conforme las ponderaciones aumentan y que se tiene el efecto contrario cuando las ponderaciones se reducen.

Tabla 4.23 Intervalos de Confianza para las Probabilidades de Concordancia de cada Variable de Respuesta en las Condiciones Óptimas para el Caso de Estudio II

	Pesos	Estimación puntuales		Estimación por intervalo al 95% de confianza	
		P ₁	P ₂	P ₁	P ₂
<i>Mgp₁</i>	(0.5, 0.5)	0.958	0.623	(0.956757, 0.959243)	(0.619996, 0.626004)
<i>Mgp₂</i>	(0.6, 0.4)	0.964	0.616	(0.962845, 0.965155)	(0.612986, 0.619014)
<i>Mgp₃</i>	(0.8, 0.2)	0.973	0.594	(0.971995, 0.974005)	(0.590956, 0.597044)
<i>Mgp₄</i>	(0.2, 0.8)	0.912	0.658	(0.910244, 0.913756)	(0.655060, 0.660940)

Así pues, según los resultados alcanzados en ambos casos de estudio se puede decir que la metodología reaccionó de acuerdo con las ponderaciones dadas a cada variable de respuesta, buscando las condiciones de operación del proceso, en donde las variables con las ponderaciones más altas se comportan de mejor manera, de acuerdo con su probabilidad de concordancia.

5. CONCLUSIONES Y RECOMENDACIONES

En este capítulo se presentan las conclusiones alcanzadas respecto a las hipótesis y los objetivos planteados y posteriormente se muestran las conclusiones generales de este trabajo, seguidas por las recomendaciones.

5.1 Conclusiones de Acuerdo con las Hipótesis y los Objetivos

A continuación se presentan las conclusiones de la presente investigación de acuerdo con las hipótesis y los objetivos planteados.

5.1.1 Conclusiones de Acuerdo con las Hipótesis

Hipótesis 1: La metodología propuesta por Peterson (2004) y complementada por Miró-Quesada et al. (2004) funciona correctamente si las variables de respuesta originales se sustituyen por la media geométrica ponderada de las probabilidades de concordancia individuales.

Respecto a esta hipótesis se puede decir que resultó verdadera, ya que el sustituir las variables de respuesta originales, por las probabilidades de concordancia individuales de cada una de las variables de respuesta, y estas últimas a su vez por su media geométrica ponderada, se tuvieron las siguientes ventajas:

- Se redujo la dimensión del problema original, por lo que se vuelve más simple.
- La nueva variable de respuesta fue fácil de entender, ya que es un promedio ponderado de las probabilidades de que cada variable de respuesta cumpla con su especificación.
- Fue posible introducir diferencias en las ponderaciones de cada una de las variables de respuesta de acuerdo con su relevancia de acuerdo con los costos o a la calidad final del producto.

Hipótesis 2: La media geométrica ponderada se puede modelar utilizando regresión logística binaria, ya que será calculada a partir de probabilidades individuales, las cuales son el resultado de una variable de respuesta con únicamente dos resultados posibles, evaluando si cada valor generado mediante simulación cumple o no con su respectiva especificación.

Respecto a esta hipótesis se puede decir que también resultó verdadera, ya que en los dos casos utilizados para ilustrar la metodología y en los diferentes escenarios considerados en cada uno de ellos fue posible ajustar modelos de regresión logística que se apegaron bastante bien a los datos simulados.

Hipótesis 3: Se puede utilizar el *MGRG* para optimizar el polinomio resultante de la regresión logística, lo que es deseable, ya que este permite la optimización restringida.

Esta hipótesis también fue verdadera, ya que en todos los casos planteados en este trabajo se utilizó el *MGRG* con el cual se alcanzaron resultados óptimos siempre dentro de la región de diseño, los cuales fueron consistentes con la información contenida en las gráficas de superficie y de contornos, así como con las gráficas de líneas en las que se contrastó la información simulada contra la obtenida por los modelos de regresión, ya que en todas ellas se puede ubicar de forma visual una región óptima aproximada.

Hipótesis 4: La condición óptima resultante para operar el proceso será diferente si se consideran diferentes ponderaciones y tenderá hacia donde las variables de respuesta más importantes se comportan de la mejor manera, si se utiliza como variable de respuesta a optimizar la media geométrica ponderada de las probabilidades de concordancia individuales, dando ponderaciones más altas a las variables de respuesta más importantes.

Respecto a esta hipótesis se pudo observar en todos los casos considerados en este trabajo que las condiciones óptimas propuestas se desplazaban hacia las zonas en donde las variables de respuesta más importantes ofrecían un mejor desempeño a expensas de las menos importantes y los resultados alcanzados diferían dependiendo de las ponderaciones consideradas para las variables de respuesta, por lo cual, es posible decir que esta hipótesis también resultó verdadera.

5.1.2 Conclusiones de Acuerdo con los Objetivos

En cuanto a los objetivos propuestos al inicio de este trabajo, se puede decir que se cumplieron tanto el general como los específicos en ambos ejemplos ilustrativos.

5.1.3 Conclusiones Generales

Con los resultados alcanzados en ambos casos de estudio se puede concluir que la metodología propuesta por Peterson (2004) y complementada por Miró-Quesada et al. (2004), funciona adecuadamente con las modificaciones propuestas y que dar ponderaciones diferentes a las variables de respuesta, sí puede modificar los resultados óptimos alcanzados, buscando configuraciones del proceso con mejores comportamientos en las variables con los mayores pesos. Esto último hace posible analizar casos en los que no todas las respuestas son igualmente importantes, con lo que la metodología se vuelve más flexible y atractiva en muchos casos prácticos, ya que en su diseño original no permitía incorporar estas consideraciones. Una ventaja adicional es que con la modificación propuesta acerca del uso de ponderaciones, el problema se puede resolver utilizando diferentes combinaciones de pesos, lo que proporcionaría a los ingenieros o a los investigadores diferentes alternativas óptimas de operación del proceso para escoger, tomando en consideración algunos otros aspectos del propio proceso, lo que puede facilitar y enriquecer la toma de decisiones.

Como conclusión general de este trabajo, se puede decir que la inferencia bayesiana, es una opción que permite ser aplicada en todas las situaciones en las que el análisis de datos sea requerido. A favor de la opción bayesiana se puede decir que es muy flexible y permite considerar aspectos que dentro de la estadística clásica no es posible manejar, por ejemplo la incertidumbre en los parámetros de los modelos de regresión.

En contra de la estadística bayesiana se puede decir, que todavía no es muy conocida y por lo tanto poco aplicada en casos prácticos dentro de la ingeniería industrial, pero poco a poco va ganando terreno y cada vez es más común encontrar publicaciones científicas en donde se ha utilizado. Por otra parte también existe un mayor grado de complejidad para su implementación, ya que en ocasiones puede ser necesario generar programas de cómputo que realicen tal tarea, aunque se puede obtener software de libre acceso como es el caso del WinBUGS y el paquete estadístico R, los cuales están contruidos con paqueterías propias de la inferencia bayesiana. La complejidad matemática derivada en muchas ocasiones de la integración multidimensional, es un aspecto que ya se ha superado debido a técnicas como las CMMC, que mediante la simulación hacen posible su solución.

5.2 Recomendaciones

Como ya se comentó previamente existen distintas opciones para trabajar con la optimización simultánea de múltiples respuestas, dentro de estas se encuentran las que utilizan la sustitución de las variables de respuesta originales por una sola variable, en las que es necesario modelar esta última mediante un polinomio para posteriormente optimizarlo mediante técnicas de métodos numéricos, pretendiendo encontrar la mejor combinación de los factores de control para la futura operación del proceso. Sin embargo según mencionan Ortiz et al. (2004) conforme el problema crece moderadamente ya sea en el número de factores o de respuestas, los algoritmos de optimización convencionales pueden fallar al tratar de encontrar el óptimo global, razón por la cual se ha puesto interés en las técnicas que utilizan la inteligencia artificial, en su mayoría algoritmos genéticos para la optimización de los polinomios resultantes.

Sin embargo se debe de tener cuidado al incorporarle cada vez mas variantes a las metodologías actuales, no se debe perder de vista que el objetivo último es encontrar opciones que puedan ser utilizadas por los ingenieros en la solución de problemas reales, los cuales pueden no tener una formación estadística tan completa para tomar las nuevas propuestas que están surgiendo y terminan por aplicar las metodologías incorporadas en el software comercial, llegando posiblemente a soluciones subóptimas. Sin embargo debido a que en la mayoría de las ocasiones se deben entregar resultados en un periodo de tiempo lo suficientemente corto, no se dispone de este recurso para utilizarlo descifrando herramientas más precisas, pero que tienen el inconveniente de ser también más complicadas. El reto entonces por parte de los investigadores en este campo es cerrar el ciclo, es decir, desarrollar herramientas más precisas, pero posteriormente ayudar a que su aplicación sea amigable, posiblemente desarrollando software en el que solo se requiera la alimentación de información y un posterior análisis de resultados.

En la actualidad se está poniendo mucho énfasis en la investigación aplicada. Se habla de la triple hélice, universidad – empresa – gobierno como la respuesta a los problemas de competitividad de las organizaciones en la que se propone trabajo conjunto de estos tres entes, uno desarrollando investigación (universidad), el otro apoyando con recursos para que esto se dé (gobierno), pero procurando que ésta al aplicarse pueda tener repercusiones positivas en la productividad de las organizaciones (empresa), es decir, los desarrollos conseguidos deben de ser aplicables. El hacer esto posible, pareciera ser más factible si la complejidad de las herramientas propuestas derivadas de la actividad de investigación no es tal que los usuarios potenciales prefieran opciones menos complejas para atender sus problemáticas, es decir, se pudiera tener una herramienta extremadamente precisa, pero también extremadamente compleja por lo que pudiera ser poco utilizada. El desarrollo de programas de cómputo amigables puede ser el puente que haga que el usuario final pueda acceder a su implementación, como ya se comentaba con anterioridad.

5.2.1 Futuras Líneas de Investigación

Como una línea de investigación se propone implementar la modificación planteada en este estudio, al trabajo desarrollado por Peterson et al. (2009), en el cual utilizan los métodos SUR (Seemingly Unrelated Regression) para ajustar los modelos de regresión que se utilizan en la simulación mediante la distribución predictiva, sustituyendo al método de mínimos cuadrados, consiguiendo con ello considerar la correlación entre las variables de respuesta desde el modelado en la fase III, con lo cual quedaría mejor cubierto este relevante aspecto en la optimización simultánea de múltiples respuestas.

6. REFERENCIAS

1. Ames, A. E., Matucci, N., Macdonald S., Szonyi, G., y Hawkins, D. M. (1997). Quality Loss Function for Optimization Across Multiple Response Surfaces. *Journal of Quality Technology* , 339-346.
2. Berger, J. O. (2000). Bayesian Analysis: A Look at Today and Thoughts of Tomorrow. *Journal of the American Statistical Association*, 1269-1276.
3. Box, G. E. P., y Tiao G. C. (1992). *Bayesian Inference in Statistical Analysis*. New York: Wiley Interscience.
4. Casella, G., y George, E. I. (1992). Explaining the Gibbs Sampler. *The American Statistician*, 167-174.
5. Ch'ng, C. K., Quah, S.H., y Low, H.C. (2005). Index C_{pm}^* in Multiple Response Optimization. *Quality Engineering*, 165-171.
6. Chapra, S. C., y Canale, R. P. (2011). *Métodos Numéricos Para Ingenieros*. México: Mc Graw Hill.
7. Chiao, C. H., y Hamada, M. (2001). Analyzing Experiments With Correlated Multiple Responses. *Journal of Quality Technology*, 451-465.
8. Chib, S., y Greenberg, E. (1995). Understanding the Metropolis – Hastings Algorithm. *The American Statistician*, 327-335.
9. Colosimo, B. M., y Del Castillo, E. (2007). *Bayesian Process Monitoring, Control and Optimization*. Boca Raton, Florida: Chapman and Hall/CRC.
10. Copeland, K. A. F., y Nelson, P.R. (1996). Dual Response Optimization Via Direct Function Minimization. *Journal of Quality Technology*, 331-336.
11. Costa, N., Lopes Pereira, Z., y Lourenco, J. (2011). Mean and Standard Deviation Optimization of Multiple Responses. *Proceedings of the International MultiConference of Engineers and Computer Scientists* (págs. 1225-1230). Hong Kong: International Association of Engineers.
12. Del Castillo, E. (27 de Enero de 2006). *Página Web de Enrique del Castillo*. Recuperado el 23 de Mayo de 2011, de <http://www2.ie.psu.edu/Castillo/Personal%20Web%20Page/Castillo.htm>.
13. Del Castillo, E., Montgomery, D.C., y McCarville, D. R. (1996). Modified Desirability Functions for Multiple Response Optimization. *Journal of quality technology*, 337-345.
14. Derringer, G. C. y Suich, R. (1980). Simultaneous Optimization of Several Response Variables. *Journal of Quality Technology*, 214-219.
15. Ding, R., Lin, D. K. J., y Wei, D. (2004). Dual-Response Surface Optimization: A weighted MSE Approach. *Quality engineering*, 377-385.
16. Gelman, A., Carlin, J. B., Stern, H. S., y Rubin, D. B. (1995). *Bayesian Data Analysis*. Londres: Chapman and Hall.
17. Geyer, C. J. (1992). Practical Markov Chain Monte Carlo. *Statistical Science*, 473-483.
18. Gutiérrez, H. y De la Vara, R. (2012). *Análisis y Diseño de Experimentos*. México: Mc Graw Hill.
19. Hamada, M. S, Wilson, A. G, Reese, C. S., y Martz, H. F. (2010). *Bayesian Reliability*. New York: Springer.

20. Jackman, S. (2004). Bayesian Analysis for Political Research. *Annual Review of Political Science*, 483-505.
21. Ko, Y. H., Kim, K. J., y Jun, C. H. (2005). A New Loss Function-Based Method for Multiresponse Optimization. *Journal of Quality Technology*, 50-59.
22. Lecoutre, B. (2006). And if You Were a Bayesian Without Knowing it? *AIP Conference Proceedings, 26th International workshop*. (págs. 15-22). Paris: American Institute of Physics.
23. Lee, M. S., y Kim, K. J. (2007). Expected Desirability Function: Consideration of Both Location and Dispersion Effects in Desirability Function Approach. *Quality Technology and Quantitative Management*, 365-377.
24. Lin, D. K. J. y Tu, W. (1995). Dual Response Surface Optimization. *Journal of Quality Technology*, 34-39.
25. Miró-Quesada, G., Del Castillo, E., y Peterson, J. J. (2004). A Bayesian Approach for Multiple Response Surface Optimization in the Presence of Noise Variables. *Journal of Applied Statistics*, 251-270.
26. Montgomery, D. C. (2005). *Diseño y Análisis de Experimentos*. México: Limusa Wiley.
27. Myers, R. H., y Montgomery, D. C. (1995). *Response Surface Methodology Process and Product Optimization Using Designed Experiments*. New York: Wiley Interscience.
28. Najafi, S., Salmasina, A., y Kazemzadeh, R. B. (2011). Optimization of Robust Design for Multiple Response Problem. *Australian Journal of Basic and Applied Sciences*, 1566-1577.
29. Ortiz, F., Simpson, J. R., Pignatiello, J. J., y Heredia-Lagner, A. (2004). A Genetic Algorithm Approach to Multiple-Response Optimization. *Journal of Quality Technology*, 432-450.
30. Pal, S., y Gauri, S. K. (2010). Multi-Response Optimization Using Multiple Regression-Based Weighted Signal-to-Noise Ratio (MRWSN). *Quality Engineering*, 336-350.
31. Peterson, J. J. (2004). A Posterior Predictive Approach to Multiple Response Surface Optimization. *Journal of Quality Technology*, 139-153.
32. Peterson, J. J., Miró-Quesada, G., y Del Castillo, E. (2009). A Bayesian Reliability Approach to Multiple Response Optimization with Seemingly Unrelated Regression Models. *Quality Technology and Quantitative Management*, 353-369.
33. Plante, R. D. (2001). Process Capability: A Criterion for Optimizing Multiple Response Product and Process Design. *IEE Transactions*, 497-509.
34. Retzer, J. J. (2006). The Century of Bayes. *International Journal of Market Research*, 49-59.
35. Romano, D., Varetto, M., y Vicario, G. (2004). Multiresponse Robust Design: A General Framework Based on Combined Array. *Journal of Quality Technology*, 27-37.
36. Ross, S. M. (2002). *Probabilidad y Estadística para Ingenieros*. México: Mc Graw Hil.
37. Spiegel, M. R., y Stephens, L. J. (2002). *Estadística*. México: Mc Graw Hill Interamericana.
38. Taleghani, M. (2009). Important Issues in Multiple Response Optimization. *Iranian Journal of Optimization*, 230-242.
39. Tuck, M. G., Lewis, S. M., y Cottrell, J. I. L. (1993). Response Surface Methodology and Taguchi: a Quality Improvement Study from the Milling Industry. *Royal Statistical Society*, 671-681.
40. Vining, G. G., y Myers, R. H. (1990). Combining Taguchi and Response Surface Philosophies: A Dual Response Approach. *Journal of Quality Technology*, 38-45.

41. Withers, S. D. (2002). Quantitative methods: Bayesian Inference, Bayesian Thinking. *Progress in Human Geography*, 553-566.
42. Woolley, T. W. (2004). Classical Versus Bayesian Inference: A Classroom Illustration. *Teaching statistics*, 42-45.

7. ANEXOS

7.1 ANEXO 1

%Programa en matlab para realizar las simulaciones en el caso de estudio I

```
clear all;
```

```
clc;
```

```
%Matriz de diseño
```

```
X=[1 -1 -1 0 1 0 0 1 1 0  
 1 -1 1 0 -1 0 0 1 1 0  
 1 -1 0 -1 0 1 0 1 0 1  
 1 -1 0 1 0 -1 0 1 0 1  
 1 0 0 0 0 0 0 0 0 0  
 1 0 1 -1 0 0 -1 0 1 1  
 1 0 -1 1 0 0 -1 0 1 1  
 1 0 1 1 0 0 1 0 1 1  
 1 0 -1 -1 0 0 1 0 1 1  
 1 0 0 0 0 0 0 0 0 0  
 1 1 0 1 0 1 0 1 0 1  
 1 1 -1 0 -1 0 0 1 1 0  
 1 1 1 0 1 0 0 1 1 0  
 1 1 0 -1 0 -1 0 1 0 1  
 1 0 0 0 0 0 0 0 0 0];
```

```
%Respuestas articulo del Castillo
```

```
Y=[2.14 22 172 0.76;  
 1.73 12 311 0.88;  
 1.93 16 251 0.8;  
 1.95 16 241 0.8;  
 2.17 14 278 0.79;  
 1.97 11 371 0.86;  
 2.38 19 194 0.74;  
 1.98 11 360 0.86;  
 2.37 18 204 0.74;  
 2.20 14 280 0.78;  
 2.42 13 314 0.78;  
 2.61 17 223 0.73;  
 2.14 10 410 0.85;  
 2.42 12 324 0.78;  
 2.20 14 281 0.79];
```

```
[N p]=size(X); %Tamaño de la matriz X
```

```
[N q]=size(Y); %Tamaño de la matriz Y
```

```
v=N-p-q+1; %Cálculo de los grados de libertad
```

```
%Estimación por mínimos cuadrados de los parámetros de los modelos de regresión para cada variable de respuesta
```

```
Bhat=((X'*X)^(-1)*X'*Y);
```

```
S=(Y-X*Bhat)'*(Y-X*Bhat);
```

```
%Con este ciclo se define la cantidad de veces que se va a repetir la simulación
```

```
for k=1:30;
```

```
ccontador=0;
```

```
%Con los ciclos siguientes se generan las n iteraciones para cada combinación de variables de control
```

```
for x2=-1:0.2:1; %Se cancela para corrida de validación y se le da el valor óptimo encontrado
```

```
for x3=-1:0.2:1; %Se cancela para corrida de validación y se le da el valor óptimo encontrado
```

```
% Para la corrida de validación se eliminan los dos ciclos anteriores y se le dan valores fijos a las variables (valores óptimos)
```

```
x6=x2*x3;
```

```
x8=x2*x2;
```

```
x9=x3*x3;
```

```
contador1=0;
```

```
contador2=0;
```

```
contador3=0;
```

```
contador4=0;
```

```
cont1=0;
```

```
cont2=0;
```

```
cont3=0;
```

```
cont4=0;
```

```
%Cantidad de vectores de respuesta a generar
```

```
for n=1:100000
```

```
%Genera una x1 en cada iteración, ya que esta es la variable de ruido
```

```
x1= normrnd(0, 0.1);
```

```
%Variables acomodadas según artículo del Castillo
```

```
x4=x1*x2;
```

```
x5=x1*x3;
```

```
x7=x1*x1;
```

```
%Modelo de segundo orden completo
```

```
x=[1 x1 x2 x3 x4 x5 x6 x7 x8 x9];
```

```
H=v*S^(-1)/(1+x*(X'*X)^(-1)*x');
```

```

u=mvnrnd(zeros(1,q),H^(-1));
c = chi2rnd(ones(1,1)*v);
mean=(Bhat'*x');
% Función predictiva
Ytilde = (u*sqrt(v/c))'+mean;

```

%Con este segmento calcula la probabilidad de concordancia individual comparando cada valor
%generado contra su respectiva especificación

```

if (Ytilde(1,1) >= 1.8)
    cont1=1;
elseif (Ytilde(1,1) < 1.8)
    cont1=0;
end

if (Ytilde(2,1)<=13.5)
    cont2=1;
elseif (Ytilde(2,1)>13.5)
    cont2=0;
end

if (Ytilde(3,1)>=302)
    cont3=1;
elseif (Ytilde(3,1)<302);
    cont3=0;
end

if (Ytilde(4,1)>=0.75) & (Ytilde(4,1)<=0.84)
    cont4=1;
elseif (Ytilde(4,1)<0.75)|(Ytilde(4,1)>0.84)
    cont4=0;
end

```

%Contadores que indican la cantidad de veces que cada variable generada cumplió con su
%especificación, para posteriormente calcular las respectivas probabilidades de concordancia

```

contador1=contador1+cont1;
contador2=contador2+cont2;
contador3=contador3+cont3;
contador4=contador4+cont4;

```

%La siguiente instrucción acumula los valores de las p respuestas en una columna para cada una
%de las n iteraciones de tal forma que al final se tenga una matriz de pxn (en este caso de
%4x100000) encontrándose en los renglones los valores para cada variable de respuesta

```

resp(:,n)=Ytilde;
end

```

%Instrucciones para calcular los promedios y las desviaciones estándar para cada variable de
%respuesta partiendo de los valores de la matriz **resp** (esto para validación)

```

suma1=sum(resp(1,:));
media1=suma1/n ;%Promedio de la respuesta y1
suma2=sum(resp(2,:));
media2=suma2/n ; %Promedio de la respuesta y2

```

```

suma3=sum(resp(3,:));
media3=suma3/n ; %Promedio de la respuesta y3
suma4=sum(resp(4,:));
media4=suma4/n ; %Promedio de la respuesta y4

desv1=std(resp(1,:)); %Desviación estándar de la respuesta y1
desv2=std(resp(2,:)); %Desviación estándar de la respuesta y2
desv3=std(resp(3,:)); %Desviación estándar de la respuesta y3
desv4=std(resp(4,:)); %Desviación estándar de la respuesta y4

resp;

%Cálculo de las probabilidades de concordancia individuales dividiendo cada contador entre la
cantidad de variables generadas (100 000 en este caso)
p1=contador1/n;
p2=contador2/n;
p3=contador3/n;
p4=contador4/n;

%Cálculo de las medias geométricas partiendo de las probabilidades anteriores
mgeomp=((p1^0.1)*(p2^0.4)*(p3^0.4)*(p4^0.1));
mgeom=((p1)*(p2)*(p3)*(p4))^(1/4);

end %Para for x2=-1:0.2:1; Se cancelan para corrida de validación
end %Para for x3=-1:0.2:1; Se cancelan para corrida de validación

%Escribe los estadísticos calculados en cada una de las k=30 iteraciones en una columna de la matriz
%resultados (esto para validación)
resultado(1,k)=media1;
resultado(2,k)=media2;
resultado(3,k)=media3;
resultado(4,k)=media4;
resultado(5,k)=desv1;
resultado(6,k)=desv2;
resultado(7,k)=desv3;
resultado(8,k)=desv4;
resultado(9,k)=p1;
resultado(10,k)=p2;
resultado(11,k)=p3;
resultado(12,k)=p4;
resultado(13,k)=mgeom;
resultado(14,k)=mgeomp;

end %Es para el contador k, para hacer la simulación varias veces (validación)

%Escribe la matriz resultado (14x30)
resultado

```

%Cálculos para obtener un promedio de los estadísticos obtenidos en la matriz **resultado**

```
sumay1=sum(resultado(1,:));
mediay1=sumay1/k ; %Promedio de los 30 promedios de  $y_1$ 
sumay2=sum(resultado(2,:));
mediay2=sumay2/k ; %Promedio de los 30 promedios de  $y_2$ 
sumay3=sum(resultado(3,:));
mediay3=sumay3/k; %Promedio de los 30 promedios de  $y_3$ 
sumay4=sum(resultado(4,:));
mediay4=sumay4/k ; %Promedio de los 30 promedios de  $y_4$ 
sumadesy1=sum(resultado(5,:));
mediadesy1=sumadesy1/k; %Promedio de las 30 desviaciones estándar en  $y_1$ 
sumadesy2=sum(resultado(6,:));
mediadesy2=sumadesy2/k; %Promedio de las 30 desviaciones estándar en  $y_2$ 
sumadesy3=sum(resultado(7,:));
mediadesy3=sumadesy3/k; %Promedio de las 30 desviaciones estándar en  $y_3$ 
sumadesy4=sum(resultado(8,:));
mediadesy4=sumadesy4/k; %Promedio de las 30 desviaciones estándar en  $y_4$ 
sumapy1=sum(resultado(9,:));
mediapy1=sumapy1/k; %Promedio de las 30 probabilidades de concordancia  $p_1$ 
sumapy2=sum(resultado(10,:));
mediapy2=sumapy2/k; %Promedio de las 30 probabilidades de concordancia  $p_2$ 
sumapy3=sum(resultado(11,:));
mediapy3=sumapy3/k; %Promedio de las 30 probabilidades de concordancia  $p_3$ 
sumapy4=sum(resultado(12,:));
mediapy4=sumapy4/k; %Promedio de las 30 probabilidades de concordancia  $p_4$ 
sumamg=sum(resultado(13,:));
mediamg=sumamg/k; %Promedio de las 30 medias geométricas
sumamgp=sum(resultado(14,:));
mediamgp=sumamgp/k; %Promedio de las 30 medias geométricas ponderadas
```

%Imprime en un vector columna los promedios anteriores

```
promedios(1,1)=mediay1;
promedios(2,1)=mediay2;
promedios(3,1)=mediay3;
promedios(4,1)=mediay4;
promedios(5,1)=mediadesy1;
promedios(6,1)=mediadesy2;
promedios(7,1)=mediadesy3;
promedios(8,1)=mediadesy4;
promedios(9,1)=mediapy1;
promedios(10,1)=mediapy2;
promedios(11,1)=mediapy3;
promedios(12,1)=mediapy4;
promedios(13,1)=mediamg;
promedios(14,1)=mediamgp;
```

promedios