



CIATEC

ESTUDIO COMPARATIVO ENTRE DOS ALGORITMOS BASADOS
EN POBLACIONES MEDIANTE DISEÑO DE EXPERIMENTOS
ANIDADO

Tesis

QUE PARA OBTENER EL GRADO ACADEMICO DE

Maestro en Ciencia y
Tecnología
en la Especialidad de
Ingeniería Industrial y de
Manufactura

PRESENTA

Luis Francisco Villalobos González

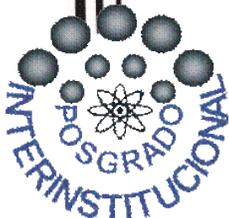
Director

Dra. Anayansi Estrada Monje

Co director

Dr. Jöns Sánchez Aguilar

León, Guanajuato, México, Enero 2011.



RESUMEN

En este trabajo se comparan dos algoritmos de búsqueda basados en poblaciones (Genetic algorithm y Particle Swarm Optimization). Los trabajos que proponen variaciones de algoritmos de búsqueda heurística, comparan los resultados del algoritmo modificado contra el algoritmo original al resolver el mismo problema. El uso de estadísticos de tendencia central y dispersión resultan útiles para este propósito (media y desviación estándar). Sin embargo, estos estadísticos están relacionados con distribuciones de probabilidad. Este hecho permite el uso de herramientas estadísticas para realizar comparaciones objetivas.

Utilizando los conceptos de precisión y exactitud referentes a la calidad en las mediciones, se busca sintonizar los algoritmos utilizando funciones de benchmarking para optimización continua. La sintonización busca que la media de varias corridas tienda a un valor (solución) y la desviación estándar sea mínima. Mediante el diseño de experimentos 2^k , se sintonizan los algoritmos, utilizando factores como, el número de variables y el tamaño de la población, además de otros parámetros inherentes a cada algoritmo.

El objetivo principal de este trabajo es realizar una comparación objetiva (precisión y exactitud) entre los dos algoritmos mediante herramientas estadísticas diferentes a la prueba de hipótesis para dos poblaciones. Utilizando un diseño de experimentos anidado, se consideran cinco casos de estudio, y las soluciones encontradas por los algoritmos, para determinar si existe diferencia entre dichas soluciones.

Contenido

CAPÍTULO 1.....	8
GENERALIDADES	8
1.1 ANTECEDENTES	9
1.2 HIPÓTESIS	12
1.3 JUSTIFICACIÓN	12
1.4 OBJETIVOS	14
1.4.1 <i>Objetivo general</i>	14
1.4.2 <i>Objetivos particulares</i>	14
CAPÍTULO 2. MARCO TEÓRICO	15
2.1 VARIABLES ALEATORIAS.	16
2.2 PRECISIÓN Y EXACTITUD.....	17
2.3 DISEÑO FACTORIAL GENERAL 2^k	20
2.4 DISEÑO ANIDADO.....	24
2.5 ANÁLISIS DE REGRESIÓN.	26
2.5.1 <i>Estimación de los Parámetros del Modelo de Regresión</i>	32
2.5.2 <i>Prueba de Hipótesis para la Significancia del Modelo</i>	32
2.6 OPTIMIZACIÓN DE RESPUESTA MÚLTIPLE.....	34
2.6.1 <i>Superposición de Gráficos de contorno.</i>	35
2.6.2 <i>Funciones de Utilidad.</i>	37
2.7 BÚSQUEDA HEURÍSTICA ALEATORIA	42
2.7.1 <i>Algoritmo Genético (GA)</i>	42
2.7.2 <i>Optimización por Enjambre de Partículas (PSO)</i>	43
CAPÍTULO 3.....	45
METODOLOGÍA.....	45
3.1 METODOLOGÍA PROPUESTA.	46
3.2 SINTONIZACIÓN DE AG Y PSO.....	47
3.3 FUNCIONES DE PRUEBA.	48
3.4 DISEÑO ANIDADO PARA COMPARAR AG VS PSO.....	52

3.4.1	<i>Optimización de Respuesta Múltiple, Casos de Estudio.</i>	52
CAPÍTULO 4.		57
RESULTADOS		57
4.1	SINTONIZACIÓN DEL ALGORITMO GENÉTICO	58
4.1.1	<i>Función de prueba Ackley (ecuación 3-1).</i>	59
4.1.2	<i>Función de prueba Griewank (ecuación 3-2).</i>	62
4.1.3	<i>Función Rastrigin (ecuación 3-3).</i>	65
4.1.4	<i>Función de prueba Rosenbrock (ecuación 3-4)</i>	68
4.1.5	<i>Función de prueba Schwefel (ecuación 3-5).</i>	71
4.1.6	<i>Función de prueba Spherical (ecuación 3-6).</i>	74
4.1.7	<i>Optimización.</i>	77
4.2	SINTONIZACIÓN DE PSO	82
4.2.1	<i>Función de prueba Ackley (ecuación 3-1).</i>	83
4.2.2	<i>Función de prueba Griewank (ecuación 3-2).</i>	86
4.2.3	<i>Función de prueba Rastrigin (ecuación 3-3)</i>	89
4.2.4	<i>Función de prueba Rosenbrock (ecuación 3-4)</i>	92
4.2.5	<i>Función de prueba Schwefel (ecuación 3-5).</i>	95
4.2.6	<i>Función de prueba Spherical (ecuación 3-6).</i>	98
4.2.7	<i>Optimización.</i>	101
4.3	COMPARACIÓN DE ALGORITMOS (GA Y PSO)	107
CAPÍTULO 5.	CONCLUSIONES	110
	BIBLIOGRAFIA	117

INDICE DE TABLAS

Tabla 2-2 matriz de diseño 2^k para $k \leq 5$	20
Tabla 2-3 Análisis de varianza del diseño 2^k	21
Tabla 2-4 Diseño 2^k y efectos de interacción dobles para $k \leq 5$	23
Tabla 2-5 Hipótesis estadística y estadístico de prueba para los tipos de factores	25
Tabla 2-6 Análisis de varianza para el diseño anidado en dos etapas	26
Tabla 2-1 Análisis de varianza para la prueba de significancia del modelo de regresión	34
Tabla 3-1 Operadores del GA empleado	47
Tabla 3-2 Factores del DDE 2^5 para sintonización del GA	47
Tabla 3-3 Factores del DDE 2^6 para sintonización del PSO	48
Tabla 3-4 Diseño anidado para comparar AG Vs PSO	52
Tabla 3-5 Variables del problema de Derringer	53
Tabla 3-6 Coeficientes de regresión para el problema de Derringer	53
Tabla 3-7 Variables del problema de Del Castillo	54
Tabla 3-8 Variables del problema de Hendricks	55
Tabla 3-9 Coeficientes de regresión para el problema de Hendriks	55
Tabla 3-10 Variables del problema de Contreras	55
Tabla 3-11 Variables del problema de Wold	56
Tabla 4-1 ANOVA para <i>Xy ackley</i>	59
Tabla 4-2 ANOVA para <i>Sy ackley</i>	60
Tabla 4-3 Coeficientes de regresión para <i>Xy ackley</i>	61
Tabla 4-4 Coeficientes de regresión para <i>Sy ackley</i>	61
Tabla 4-5 Estadísticos para los modelos de regresión (Ackley)	61
Tabla 4-6 ANOVA para <i>ln(Xy griewank)</i>	62
Tabla 4-7 ANOVA para <i>ln(Sy griewank)</i>	63
Tabla 4-8 Coeficientes de regresión para <i>ln(Sy griewank)</i>	64
Tabla 4-9 Coeficientes de regresión para <i>ln(Xy griewank)</i>	64
Tabla 4-10 Estadísticos para los modelos de regresión (Griewank)	64
Tabla 4-11 ANOVA para <i>Xy rastrigin</i>	65
Tabla 4-12 ANOVA para <i>Sy rastrigin</i>	66
Tabla 4-13 Coeficientes de regresión para <i>Sy rastrigin</i>	67

Tabla 4-14 Coeficientes de regresión para <i>Xy rastrigin</i>	67
Tabla 4-15 Estadísticos para los modelos de regresión (Rastrigin).....	67
Tabla 4-16 ANOVA para <i>Xy rosenbrock</i>	68
Tabla 4-17 ANOVA para <i>Sy rosenbrock</i>	69
Tabla 4-18 Coeficientes de regresión para <i>Sy rosenbrock</i>	70
Tabla 4-19 Coeficientes de regresión para <i>Xy rosenbrock</i>	70
Tabla 4-20 Estadísticos para los modelos de regresión (Rosenbrock)	70
Tabla 4-21 ANOVA para <i>Xy schwefel</i>	71
Tabla 4-22 ANOVA para <i>Sy schwefel</i>	72
Tabla 4-23 Coeficientes de regresión para <i>Sy schwefel</i>	73
Tabla 4-24 Coeficientes de regresión para <i>Xy schwefel</i>	73
Tabla 4-25 Estadísticos para los modelos de regresión (Schwefel)	73
Tabla 4-26 ANOVA para <i>Xy spherical</i>	74
Tabla 4-27 ANOVA para <i>Sy spherical</i>	75
Tabla 4-28 Coeficientes de regresión para <i>Sy spherical</i>	76
Tabla 4-29 Coeficientes de regresión para <i>Xy spherical</i>	76
Tabla 4-30 Estadísticos para los modelos de regresión (Spherical)	76
Tabla 4-31 límites para el cálculo de la función de deseabilidad (GA)	77
Tabla 4-32 Deseabilidad para cada punto experimental (GA)	78
Tabla 4-33 deseabilidades para prueba de hipótesis.....	79
4-34 Prueba de Shapiro-Wilk W	80
Tabla 4-35 intervalos de confianza del 95% para la media y desviación estándar	80
Tabla 4-36 ANOVA para <i>ln(Xx ackley)</i>	83
Tabla 4-37 ANOVA para <i>ln(Sy ackley)</i>	84
Tabla 4-38 coeficientes de regresión para <i>ln(Sy ackley)</i>	85
Tabla 4-39 coeficientes de regresión para <i>ln(Xy ackley)</i>	85
Tabla 4-40 Estadísticos para los modelos de regresión (Ackley).....	85
Tabla 4-41 ANOVA para <i>ln(Xy griewnk)</i>	86
Tabla 4-42 ANOVA para <i>ln(Sy griewnk)</i>	87
Tabla 4-43 Coeficientes de regresión para <i>ln(Xy griewnk)</i>	88
Tabla 4-44 Coeficientes de regresión para <i>ln(Sy griewnk)</i>	88

Tabla 4-45 Estadísticos para los modelos de regresión (Griewank)	88
Tabla 4-46 ANOVA para <i>Xy rastrigin</i>	89
Tabla 4-47 ANOVA para <i>Sy rastirgin</i>	90
Tabla 4-48 Coeficientes de regresión para <i>Sy rastirgin</i>	91
Tabla 4-49 Coeficientes de regresión para <i>Xy rastrigin</i>	91
Tabla 4-50 Estadísticos para los modelos de regresión (Rastrigin).....	91
Tabla 4-51 ANOVA para <i>Xy rosenbrock</i>	92
Tabla 4-52 ANOVA para <i>Sy rosenbrock</i>	93
Tabla 4-53 coeficientes de regresión para <i>Xy rosenbrock</i>	94
Tabla 4-54 coeficientes de regresión para <i>Sy rosenbrock</i>	94
Tabla 4-55 Estadísticos para los modelos de regresión (Rosenbrock)	94
Tabla 4-56 ANOVA para <i>Xy schwefel</i>	95
Tabla 4-57 ANOVA para <i>Sy schwefel</i>	96
Tabla 4-58 Coeficientes de regresión para <i>Sy schwefel</i>	97
Tabla 4-59 Coeficientes de regresión para <i>Xy schwefel</i>	97
Tabla 4-60 Estadísticos para los modelos de regresión (Schwefel)	97
Tabla 4-61 ANOVA para <i>ln(Xy spherical)</i>	98
Tabla 4-62 ANOVA para <i>ln(Sy spherical)</i>	99
Tabla 4-63 coeficientes de regresión para <i>ln(Sy spherical)</i>	100
Tabla 4-64 Coeficientes de regresión para <i>ln(Xy spherical)</i>	100
Tabla 4-65 Estadísticos para los modelos de regresión (Spherical)	100
Tabla 4-66 limites para el cálculo del a función de deseabilidad (PSO).....	101
Tabla 4-67 Deseabilidad para cada punto experimental (PSO).....	102
Tabla 4-68 deseabilidades para prueba de hipótesis.....	104
Tabla 4-69 Prueba de Shapiro-Wilk W	105
Tabla 4-70 intervalos de confianza del 95% para la media y desviación estándar	105
Tabla 4-71 valores óptimos para el PSO	106
Tabla 4-72 Resultados del experimento anidado.....	107
Tabla 4-73 ANOVA para la significancia del modelo	107
Tabla 4-74 ANOVA para el diseño anidado	107
Tabla 4-75 Estadísticos para el modelo de regresión	108

Tabla 4-76 Componentes de varianza	108
Tabla 5-1 Resultados para la funciones Rosenbrock y Schwefel en el tratamiento optimo	111
Tabla 5-2 Parámetros de los algoritmos no sintonizados	112

INDICE DE FIGURAS

Figura 1-1 Diagramas de flujo de GA y PSO.....	11
Figura 2-2 Precisión y exactitud, representadas como un tiro al blanco	18
Figura 2-3 Precisión y exactitud atreves de la curva normal.....	18
Figura 2-1 Diseño anidado en dos etapas	24
Figura 3-1 Esquema de la metodología	46

INDICE DE GRAFICOS

Gráfico 2-1 Diagrama de dispersión.....	27
Gráfico 2-2 Distribución de Y dado un valor x	28
Gráfico 2-3 a) plano de regresión del modelo $EY = 40 + 35x_1 + 15x_2$ b) gráfico de contorno.....	30
Gráfico 2-4 $Y = 2 + 4x + 1.5x_2 + 3x_3$	31
Gráfico 2-5 Híper planos generados por $Y = 2 + 4x_1 + 1.5x_2 + 3x_3$	31
Gráfico 2-6. Superficie de respuesta (a), y gráfico de contorno (b) para y_1	36
Gráfico 2-7. Superficie de respuesta (a), y gráfico de contorno (b) para y_2	36
Gráfico 2-8. Superficie de respuesta (a), y gráfico de contorno (b) para y_3	36
Gráfico 2-9. Región optima para el rendimiento (y_1) la viscosidad (y_2) y el peso molecular (y_3)	37
Gráfico 2-10. Deseabilidad bilateral.....	38
Gráfico 2-11 deseabilidad unilateral	39
Gráfico 3-1 Función Spherical para $n=2$ en el cubo de inspección	49
Gráfico 3-3 Función Rosenbrock para $n=2$ en el cubo de inspección	49
Gráfico 3-4 Función Griewank para $n=2$ en el cubo de inspección	50
Gráfico 3-5 Función Ackley para $n=2$ en el cubo de inspección	50
Gráfico 3-6 Función Rastrigin para $n=2$ en el cubo de inspección	51

Gráfico 3-7 Función Schwefel para n=2 en el cubo de inspección	51
Gráfico 4-1 residuales Vs predichos <i>Xy ackley</i>	59
Gráfico 4-2 Probabilidad normal para los residuales de <i>Xy ackley</i>	59
Gráfico 4-3 Residuales Vs predichos <i>Sy ackley</i>	60
Gráfico 4-4 Probabilidad normal de residuales para <i>Sy ackley</i>	60
Gráfico 4-5 Residuales Vs predichos para <i>ln(Xy griewank)</i>	62
Gráfico 4-6 Probabilidad normal de residuales para <i>ln(Xy griewank)</i>	62
Gráfico 4-7 Residuales Vs predichos para <i>ln(Sy griewank)</i>	63
Gráfico 4-8 Probabilidad normal de residuales para <i>ln(Sy griewank)</i>	63
Gráfico 4-9 Residuales Vs predichos para <i>Xy rastrigin</i>	65
Gráfico 4-10 Probabilidad normal de residuales para <i>Xy rastrigin</i>	65
Gráfico 4-11 Residuales Vs predichos para <i>Sy rastrigin</i>	66
Gráfico 4-12 Probabilidad normal de residuales para <i>Sy rastrigin</i>	66
Gráfico 4-13 Residuales Vs predichos para <i>Xy rosenbrock</i>	68
Gráfico 4-14 Probabilidad normal de residuales para <i>Xy rosenbrock</i>	68
Gráfico 4-15 Residuales Vs predichos para <i>Sy rosenbrock</i>	69
Gráfico 4-16 Probabilidad normal de residuales para <i>Sy rosenbrock</i>	69
Gráfico 4-17 Residuales Vs predichos para <i>Xy schwefel</i>	71
Gráfico 4-18 Probabilidad normal de residuales para <i>Xy schwefel</i>	71
Gráfico 4-19 Residuales Vs predichos para <i>Sy schwefel</i>	72
Gráfico 4-20 Probabilidad normal de residuales para <i>Sy schwefel</i>	72
Gráfico 4-21 Residuales Vs predichos para <i>Xy spherical</i>	74
Gráfico 4-22 Probabilidad normal de residuales para <i>Xy spherical</i>	74
Gráfico 4-23 Residuales Vs predichos para <i>Sy spherical</i>	75
Gráfico 4-24 Probabilidad normal de residuales para <i>Sy spherical</i>	75
Gráfico 4-25 Cuantil-cuantil para Dag5 Vs distribución normal	79
Gráfico 4-26 Cuantil-cuantil para Dag9 Vs distribución normal	80
Gráfico 4-27 Residuales Vs predichos para <i>ln(Xx ackley)</i>	83
Gráfico 4-28 Probabilidad normal de residuales para <i>ln(Xx ackley)</i>	83
Gráfico 4-29 residuales Vs predichos para <i>ln(Sy ackley)</i>	84

Gráfico 4-30 Probabilidad normal de residuales para <i>ln(Sy ackley)</i>	84
Gráfico 4-31 residuales Vs orden de corrida para <i>ln(Xy ackley)</i> (a) y <i>ln(Sy ackley)</i> (b)	85
Gráfico 4-32 Residuales Vs predichos para <i>ln(Xy griewnk)</i>	86
Gráfico 4-33 Probabilidad normal de residuales para <i>ln(Xy griewnk)</i>	86
Gráfico 4-34 Residuales Vs predichos para <i>ln(Sy griewnk)</i>	87
Gráfico 4-35 Probabilidad normal de residuales para <i>ln(Sy griewnk)</i>	87
Gráfico 4-36 Residuales Vs predichos para <i>Xy rstrigin</i>	89
Gráfico 4-37 Probabilidad normal de residuales para <i>Xy rstrigin</i>	89
Gráfico 4-38 Residuales Vs predichos para <i>Sy rastirgin</i>	90
Gráfico 4-39 Probabilidad normal de residuales para <i>Sy rastirgin</i>	90
Gráfico 4-40 Residuales Vs predichos para <i>Xy rosenbrock</i>	92
Gráfico 4-41 Probabilidad normal de residuales para <i>Xy rosenbrock</i>	92
Gráfico 4-42 Residuales Vs predichos para <i>Sy rosenbrock</i>	93
Gráfico 4-43 Probabilidad normal de residuales para <i>Sy rosenbrock</i>	93
Gráfico 4-44 Residuales Vs predichos para <i>Xy schwefel</i>	95
Gráfico 4-45 Probabilidad normal de residuales para <i>Xy schwefel</i>	95
Gráfico 4-46 Residuales Vs predichos para <i>Sy schwefel</i>	96
Gráfico 4-47 Probabilidad normal de residuales para <i>Sy schwefel</i>	96
Gráfico 4-48 Residuales Vs predichos para <i>ln(Xy spherical)</i>	98
Gráfico 4-49 Probabilidad normal de residuales para <i>ln(Xy spherical)</i>	98
Gráfico 4-50 Residuales Vs predichos para <i>ln(Sy spherical)</i>	99
Gráfico 4-51 Probabilidad normal de residuales para <i>ln(Sy spherical)</i>	99
Gráfico 4-52 Cuantil-cuantil para Dpso5 Vs distribución normal	104
Gráfico 4-53 Cuantil-cuantil para Dpso9 Vs distribución normal	105
Gráfico 4-54 residuales Vs Factor (algoritmo).....	108
Gráfico 4-55 Residuales Vs predichos	108
Gráfico 4-56 Residuales Vs factor excluyendo puntos atípicos	109
Gráfico 5-1 Diagrama de caja para el experimento del caso Derringer	113
Gráfico 5-2 Medias e intervalo de diferencia mínima significativa (no sintonizado)	113
Gráfico 5-3 Medias e intervalo de diferencia mínima significativa (sintonizado)	114

Gráfico 5-4 D Vs X1 Vs X2 para el caso Derringer.....	114
Gráfico 5-5 D Vs X1 Vs X2 para el caso Del Castillo.....	115
Gráfico 5-6 D Vs X1 Vs X2 para el caso Hendricks.....	115
Gráfico 5-7 D Vs X1 Vs X2 para el caso Contreras.....	115
Gráfico 5-8 D Vs X1 Vs X3 para el caso Wold	116

CAPÍTULO 1.

GENERALIDADES

1.1 Antecedentes

En muchos experimentos es importante evaluar más de una variable de salida para cada unidad experimental. Este tipo de experimentos se aplica en el desarrollo de procesos y productos industriales, donde el desempeño está relacionado con múltiples variables de salida. Es común que algunas de estas variables de salida estén en conflicto (una crece mientras la otra decrece) es decir, son antagónicas, y el analista debe decidir cuales respuestas son más importantes. Para evaluar las respuestas, el analista puede, determinar el conjunto de condiciones operativas para la elaboración del producto con la mejor respuesta global. Este conjunto de condiciones de operación se llama condición óptima para el proceso. Puede encontrarse una amplia gama de técnicas para la optimización respuesta múltiple (ORM). Sobresalen la superposición de gráficos de contorno y las funciones de conveniencia, debido a su facilidad de uso y la capacidad para localizar una solución óptima ⁽¹⁾.

La optimización de respuesta múltiple implica determinar un punto en la región experimental, en el cual las variables de respuestas presentan algunas características deseadas, particularmente próximas a un valor específico y con varianza pequeña. En esto consiste la optimización de respuesta múltiple, donde todas las respuestas se consideran simultáneamente y es improbable que alcancen el óptimo en el mismo punto.

Los procedimientos de optimización de respuesta múltiple, generalmente se desarrollan en cuatro pasos:

- i. Desarrollo de modelos de regresión individuales para cada respuesta, en función de los factores de control, o variables independientes.
- ii. Calcular medidas de desempeño predichas, como distancia al objetivo y varianza, utilizando los estimadores para cada respuesta.
- iii. Combinar las medidas de desempeño predichas de todas las respuestas, utilizando alguna función de utilidad (sección 2.6.2) o procedimiento auxiliar.
- iv. Optimizar la función para identificar el mejor conjunto solución.

De los pasos anteriores, resulta evidente que un modelado apropiado de las variables de respuesta es crucial para que la técnica de optimización provea resultados confiables. Un modelo apropiado de las variables de respuesta en función de los factores de control, conduce a predicciones con varianza pequeña. Si la predicción es precisa, entonces los valores observados de las respuestas estarán cerca de los valores predichos cuando los niveles óptimos se estudien en una prueba de confirmación ⁽²⁾.

La superposición de de gráficos de contorno (SGC) es una herramienta útil, para encontrar la región de interés. Los gráficos de contorno se obtienen de modelos de regresión desarrollados para describir la variación de las respuestas. La región de interés es aquella en donde todas las variables de respuesta son satisfechas simultáneamente. Sin embargo, existen dos limitaciones en la aplicación de la superposición de gráficos de contorno ⁽³⁾. Primero: este método no permite reconocer la solución más dominante en el espacio experimental, y segundo: su uso puede dificultar el análisis en problemas con tres o más factores de control espacio multidimensional (sección 2.6.1).

Las limitaciones de la superposición de gráficos de contorno, abren paso a técnicas más adecuadas para problemas en el espacio n-dimensional. Las funciones de utilidad (sección 2.6.2) permiten el manejo de varias dimensiones, “calificando” cada punto en el espacio en función de todas las variables de respuesta asociadas. La función de deseabilidad (sección 2.6.2.1) propuesta por Derringer y Suich ⁽⁴⁾, es la función de utilidad más difundida en la optimización de respuesta múltiple, con la limitación debida al uso de un estimador \hat{Y}_j , ya que la función de deseabilidad no considera que existe un intervalo de confianza asociado a la aleatoriedad de la estimación ⁽⁵⁾. Además, con el aumento del número de variables independientes, la función de deseabilidad se vuelve altamente no lineal y multimodal, y el uso de métodos convencionales de búsqueda se dificulta. En estos casos, una alternativa es el uso de métodos heurísticos de búsqueda ⁽⁶⁾. La aplicación de estos algoritmos (como el algoritmo genético, optimización por colonia de hormigas, optimización por enjambre de partículas, u optimización por bacterias forrajeras) se realiza en el paso (iv) del desarrollo de la optimización de respuesta múltiple.

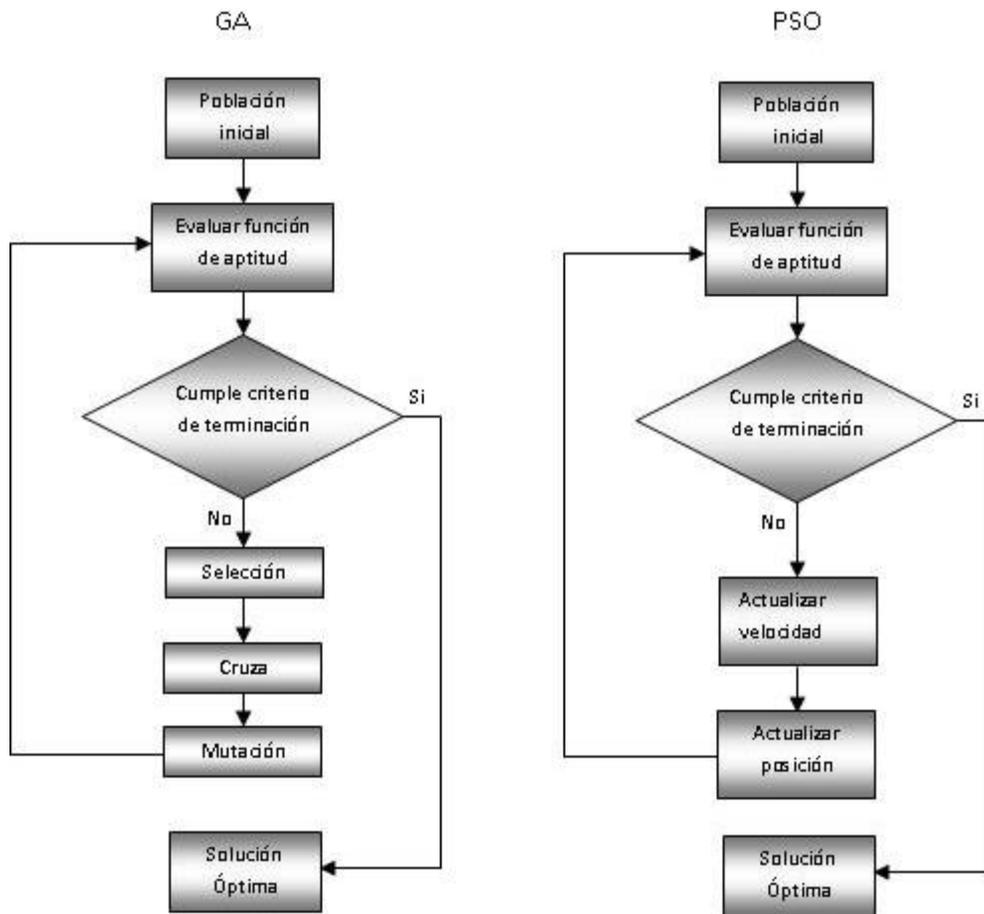


Figura 0-1 Diagramas de flujo de GA y PSO.

La optimización por enjambre de partículas o PSO, (Particle Swarm Optimization) es un método heurístico de búsqueda relativamente reciente (1998), cuya mecánica se inspiran en el comportamiento de enjambres o en el comportamiento colectivo de poblaciones biológicas. PSO es similar al algoritmo genético o GA, (Genetic Algorithm) en el sentido de que estas dos heurísticas son métodos de búsqueda basados en poblaciones. En otras palabras, los algoritmos PSO y GA (Figura 0-1) se mueven de un conjunto de puntos (población) a otro conjunto de puntos en una sola iteración con posibles mejoras utilizando una combinación de reglas deterministas y probabilísticos. El GA y sus muchas variaciones han sido populares en el ámbito académico y el industrial, debido principalmente a la facilidad de implementación y a la capacidad de resolver con eficacia problemas altamente no lineales, mezclados con problemas de optimización entera que son típicos de los sistemas complejos de ingeniería ⁽⁷⁾.

Para comparar un algoritmo básico contra uno mejorado, se ha utilizado la media de varias corridas, y su desviación estándar ⁽⁸⁾ ⁽⁹⁾, o métricas de eficiencia como el porcentaje de éxitos ⁽¹⁰⁾. Sin embargo el PSO y el GA utilizan un conjunto de soluciones iniciales generadas aleatoriamente y esto representa la principal fuente de variación en las soluciones obtenidas con estos métodos de búsqueda, es decir, que las soluciones obtenidas con PSO y GA representan variables aleatorias (sección 2.1). Esta consideración permite utilizar estadística inferencial, como las pruebas paramétricas ⁽⁷⁾ ⁽¹¹⁾ ⁽¹²⁾ y pruebas no paramétricas ⁽¹³⁾, para comparar el desempeño de los diferentes tipos de algoritmos.

El diseño de experimentos es un proceso de planeación del experimento que permite recolectar datos, analizarlos con métodos estadísticos, resultando una solución *válida* y *objetiva*. En algunos tipos de diseños de experimentos (diseños anidados o jerárquicos), los niveles de un factor B son similares, pero no idénticos para los niveles del otro factor (A).

Cada nivel del factor B anidado en A es único ⁽¹¹⁾. Utilizando este tipo de diseño, puede compararse los algoritmos en términos de parámetros en común, por ejemplo, tamaño de población inicial, o tipo de función de aptitud evaluada.

1.2 Hipótesis

- Existe igualdad en las medias de las soluciones encontradas por PSO y GA al resolver problemas de optimización de respuesta múltiple, con un nivel de confianza del 95%.
- Existe igualdad en las desviaciones estándar de soluciones encontradas por PSO y GA al resolver problemas de optimización de respuesta múltiple, con un nivel de confianza del 95%.

1.3 Justificación

Las soluciones obtenidas con métodos heurísticos de búsqueda son descritas por distribuciones de probabilidad con parámetros μ y σ (variable aleatoria). La prueba t es una

herramienta que permite comparar dos poblaciones a la vez, suponiendo que las poblaciones son distribuidas normalmente con varianzas iguales. Cualquier violación de estos supuestos invalida los resultados de la prueba, y las pruebas no paramétricas resultan adecuadas.

El análisis de varianza (ANOVA) supone normalidad entre las poblaciones y varianzas iguales. Debido al uso del estadístico F, el ANOVA es una prueba robusta ante la violación de los supuestos, además que permite la comparación de múltiples poblaciones a la vez. Por esto, el diseño de experimentos anidado es una herramienta adecuada para comparar algoritmos basados en poblaciones.

1.4 Objetivos

1.4.1 Objetivo general

Comparar y analizar la precisión y exactitud del PSO y GA sintonizados, en la optimización simultánea de varias respuestas a través de un diseño anidado, que permita formular un método para **comparar objetivamente** la precisión y exactitud de algoritmos basados en poblaciones.

1.4.2 Objetivos particulares

- Sintonizar los parámetros para el GA canónico y el PSO básico para ser precisos y exactos.
- Utilizar funciones de benchmarking para sintonizar los algoritmos.
- Seleccionar los métodos y/o diseños experimentales para el análisis.
- Utilizar casos de estudio de optimización simultánea de varias respuestas para comparar los algoritmos sintonizados.

CAPÍTULO 2.

MARCO TEÓRICO

2.1 Variables aleatorias.

Frecuentemente, en los experimentos aleatorios, los resultados se resumen a un solo número. Para algunos experimentos aleatorios, el espacio muestral se define como el total de posibles resultados. En otros casos resulta útil asociar un número a cada uno de los posibles resultados, es decir, se define una función sobre los puntos de un espacio muestral. Debido a que el resultado concreto del experimento no se conoce de antemano, el valor resultante de nuestra variable tampoco se conoce de antemano. Por esta razón, la variable que asocia un número al resultado de un experimento se conoce como variable aleatoria.

Definición formal.- Si S es un espacio muestral con una medida de probabilidad, y X es una función de valor real definida sobre los puntos de S , entonces X se llama variable aleatoria.

Es común que una medición (por ejemplo, la corriente eléctrica en un alambre de cobre o las dimensiones de una pieza mecánica) pueda tomar cualquier valor en un intervalo de números reales (al menos teóricamente). Entonces la precisión arbitraria en la medición es posible. Por supuesto que, en la práctica, podemos redondear con precisión de décimas o centésimas de unidad. La variable aleatoria que representa esta medida se dice que es una variable aleatoria continua. El rango de la variable aleatoria incluye todos los valores en un intervalo de números reales, es decir, el rango puede ser considerado como continuo.

La definición de probabilidad en el caso continuo, supone para cada variable aleatoria la existencia de una función, llamada función de densidad de probabilidad, tal que el área bajo la curva es la probabilidad asociada para el intervalo correspondiente. En otras palabras, la función de densidad integrada en el intervalo (a, b) con $a \leq b$, da la probabilidad de que la variable aleatoria correspondiente, asumirá el valor en el intervalo (a, b) .

Definición.- Una función con valores $f(x)$, definida sobre el conjunto de todos los números reales, se llama **función de densidad de probabilidad** sí y solo sí:

$$P(a \leq X \leq b) = \int_a^b f(x) dx \quad 2-1$$

Para cualesquiera constantes reales a y b con $a \leq b$.

Adviértase que $f(c)$, el valor de la función de densidad de X en c , no da la probabilidad de $X=c$. Las variables aleatorias continuas siempre están asociadas a un intervalo, por lo que $P(X=c)=0$, para cualquier constante real c .

Definición.- Si X es una variable aleatoria continua, y su densidad de probabilidad en t es $f(t)$, entonces la función está dada por:

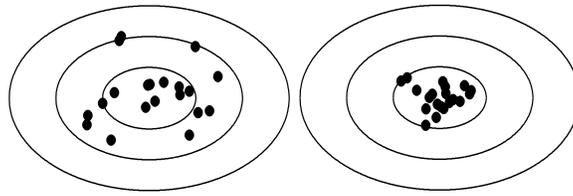
$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \text{ para } -\infty < x < \infty \quad 2-2$$

Se llama **función de distribución** o la **distribución acumulativa** de X .

2.2 Precisión y exactitud.

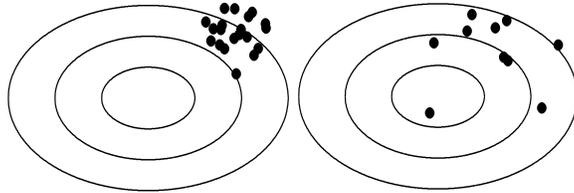
Estos dos conceptos son manifestaciones de la variabilidad de cualquier proceso de medición. La *precisión* es la variabilidad que presentan los resultados al medir repetidamente el mismo objeto o mensurado con el mismo instrumento de medición, en pocas palabras, la precisión es la habilidad de un proceso de medición de repetir y reproducir sus mediciones, independientemente de si la medición es correcta o no.

Por otro lado, la *exactitud* o sesgo, se refiere al desplazamiento de un valor estándar o verdadero que se considera conocido. Para mostrar estos conceptos, la Figura 0-1 muestra una representación de la precisión y exactitud por medio de un tiro al blanco. En donde el centro de la diana es el valor nominal (N) y los puntos son el resultado de varias mediciones.



a) Impreciso y exacto

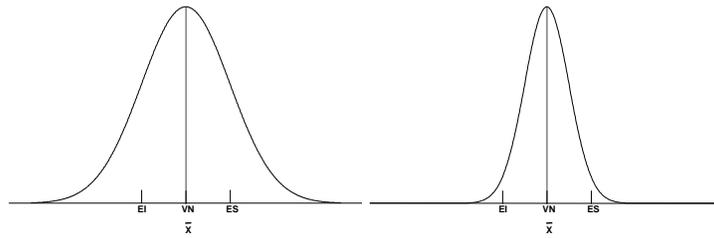
b) Preciso y exacto



a) Preciso e inexacto

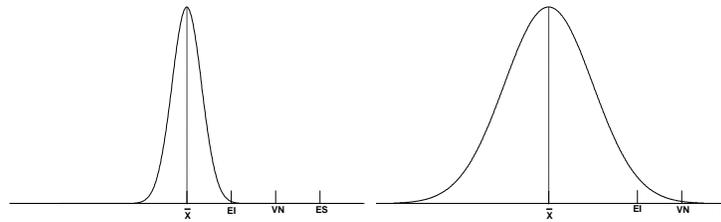
b) Impreciso e inexacto

Figura 2-1 Precisión y exactitud, representadas como un tiro al blanco



a) Impreciso y exacto

b) Preciso y exacto



c) Preciso e inexacto

d) Impreciso e inexacto

Figura 2-2 Precisión y exactitud a través de la curva normal

La Figura 2-2 muestra la precisión y exactitud de las mediciones utilizando la curva de la distribución normal donde se considera la variabilidad o amplitud de la curva, y el desfase de la media respecto al valor nominal. En ambos casos, el inciso a) muestra un proceso de medición impreciso y exacto, tiene una exactitud adecuada pues porque el promedio de las mediciones es el valor nominal (N), pero tiene una dispersión elevada. En b) el promedio es el valor nominal, con una baja dispersión. El inciso c) las mediciones tienen buena exactitud (baja dispersión), pero la media se aleja del valor nominal, por último, el inciso d) las mediciones están dispersas y el promedio de las mediciones se aleja del valor nominal.

Lo más recomendable es que los sistemas de medición sean precisos y exactos como en el inciso b). Estos conceptos son aplicables a los métodos heurísticos de búsqueda, es deseable que los resultados de estos algoritmos sean precisos y exactos, es decir, que una serie de soluciones encontradas por el algoritmo tengan poca dispersión, y la media sea el valor de la solución verdadera.

2.3 Diseño factorial general 2^k .

En el diseño factorial 2^k , en el que se consideran k factores con dos niveles cada uno, y tiene 2^k tratamientos o puntos de diseño. Las k columnas y los 2^k renglones que componen la matriz de diseño considerando una réplica, se construyen de la siguiente manera: en la primera columna, que corresponde a los niveles del factor A se alternan los signos + y -, comenzando con -, hasta completar los 2^k renglones. En la segunda columna se alternan dos signos menos con dos signos más, en la tercera columna se alternan 4 signos menos con cuatro signos más, y sucesivamente hasta la columna k , donde se alternan 2^{k-1} signos menos, con 2^{k-1} signos más (Tabla 2-1). La notación de Yates representa el total o la suma de las observaciones en cada tratamiento, en específico (1) es la suma de los datos obtenidos en el tratamiento en el que todos los factores se encuentran en un nivel bajo; a , es la suma de los datos del tratamiento donde el factor A esta en un nivel alto y los factores restantes corresponden a un nivel bajo, y así sucesivamente.

Tratamiento	Notación de Yates	A	B	C	D	E	Tratamiento	Notación de Yates	A	B	C	D	E
1	(1)	-	-	-	-	-	17	e	-	-	-	-	+
2	A	+	-	-	-	-	18	ae	+	-	-	-	+
3	B	-	+	-	-	-	19	be	-	+	-	-	+
4	ab	+	+	-	-	-	20	abe	+	+	-	-	+
5	C	-	-	+	-	-	21	ce	-	-	+	-	+
6	ac	+	-	+	-	-	22	ace	+	-	+	-	+
7	bc	-	+	+	-	-	23	bce	-	+	+	-	+
8	abc	+	+	+	-	-	24	$abce$	+	+	+	-	+
9	D	-	-	-	+	-	25	De	-	-	-	+	+
10	ad	+	-	-	+	-	26	Ade	+	-	-	+	+
11	bd	-	+	-	+	-	27	Bde	-	+	-	+	+
12	abd	+	+	-	+	-	28	$abde$	+	+	-	+	+
13	cd	-	-	+	+	-	29	Cde	-	-	+	+	+
14	acd	+	-	+	+	-	30	$acde$	+	-	+	+	+
15	bcd	-	+	+	+	-	31	$bcde$	-	+	+	+	+
16	$abcd$	+	+	+	+	-	32	$abcde$	+	+	+	+	+

Tabla 2-1 matriz de diseño 2^k para $k \leq 5$

Con el diseño factorial completo 2^k se pueden estudiar los 2^{k-1} efectos siguientes:

$$\binom{k}{1} = k \quad \text{Efectos principales}$$

$$\binom{k}{2} = \frac{k!}{2(k-2)!} \quad \text{Interacciones dobles}$$

$$\binom{k}{3} = \frac{k!}{3(k-3)!} \quad \text{Interacciones triples}$$

$$\vdots$$

$$\binom{k}{k-1} = \frac{k!}{(k-1)!} \quad \text{Interacción de k-1 factores}$$

$$\binom{k}{k} = 1 \quad \text{Interacción de k factores}$$

Donde $\binom{k}{r}$ corresponde a las combinaciones de k en r .

Fuente de variación	Suma de cuadrados	Grados de libertad
k efectos principales		
A	SS_A	1
B	SS_B	1
\vdots	\vdots	\vdots
K	SS_K	1
$\binom{k}{2}$ interacciones dobles		
AB	SS_{AB}	1
AC	SS_{AC}	1
\vdots	\vdots	\vdots
JK	SS_{JK}	1
$\binom{k}{3}$ interacciones triples		
ABC	SS_{ABC}	1
ABD	SS_{ABD}	1
\vdots	\vdots	\vdots
IJK	SS_{IJK}	1
\vdots	\vdots	\vdots
$\binom{k}{k} = 1$ interacción de k factores:		
ABC ... K	$SS_{ABC \dots K}$	1
Error	SS_E	$2^k(n-1)$
Totales	SS_T	$n2^k - 1$

Tabla 2-2 Análisis de varianza del diseño 2^k

Cada uno de los efectos se obtiene a partir de los contrastes¹, los que se obtienen a su vez de la tabla de signos del diseño. La tabla de signos consiste en determinar la columna de

¹ El contraste es una combinación lineal de la forma $C = \sum_{i=1}^{2^k} c_i Y_i$, con $\sum_{i=1}^{2^k} c_i = 0$

signos que llevan los totales en la notación de Yates para formar el contraste de cada efecto. Las columnas de signos para los contrastes que definen a los efectos principales están dadas por la matriz de diseño, mientras que los efectos de interacción se obtienen de la multiplicación de las columnas que señala dicho efecto de interacción. Para calcular la columna de efectos que define el contraste del efecto ABC en un diseño factorial 2^3 se multiplican las columnas de signos A×B×C. en la Tabla 2-3 se muestran los contrastes para los efectos principales y los efectos de interacción doble para un diseño con $k \leq 5$ factores

El contraste de cada efecto se determina multiplicando la columna de signos por la columna de totales (expresado por la notación de Yates). Con los contrastes se calcula el efecto y la suma de cuadrados (con un grado de libertad) mediante:

$$\text{Efecto } ABC \dots K = \frac{1}{n2^{k-1}} [\text{contraste } ABC \dots K] \quad 2-3$$

$$SS_{ABC \dots K} = \frac{1}{n2^k} [\text{contraste } ABC \dots K]^2 \quad 2-4$$

Donde n es el número de replicas del experimento. La suma de cuadrados totales está dada

$$SST = \sum_{i=1}^{n2^k} Y_i^2 - \frac{1}{n2^k} \left(\sum_{i=1}^{n2^k} Y_i \right)^2 \quad 2-5$$

La suma de cuadrados totales tiene $n2^k$ grados de libertad. La suma de cuadrados del error se obtiene por diferencia, y tiene $2^k(n-1)$ grados de libertad.

$$SS_T = \sum_{i=1}^{n2^k} Y_i^2 - \left(\frac{1}{n2^k} \sum_{i=1}^{n2^k} Y_i \right)^2 \quad 2-5$$

Tratamiento	Notación de Yates	A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
1	(1)	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+
2	<i>a</i>	+	-	-	-	-	-	-	-	-	+	+	+	+	+	+
3	<i>b</i>	-	+	-	-	-	-	+	+	+	-	-	-	+	+	+
4	<i>ab</i>	+	+	-	-	-	+	-	-	-	-	-	-	+	+	+
5	<i>c</i>	-	-	+	-	-	+	-	+	+	-	+	+	-	-	+
6	<i>ac</i>	+	-	+	-	-	-	+	-	-	-	+	+	-	-	+
7	<i>bc</i>	-	+	+	-	-	-	-	+	+	+	-	-	-	-	+
8	<i>abc</i>	+	+	+	-	-	+	+	-	-	+	-	-	-	-	+
9	<i>d</i>	-	-	-	+	-	+	+	-	+	+	-	+	-	+	-
10	<i>ad</i>	+	-	-	+	-	-	-	+	-	+	-	+	-	+	-
11	<i>bd</i>	-	+	-	+	-	-	+	-	+	-	+	-	-	+	-
12	<i>abd</i>	+	+	-	+	-	+	-	+	-	-	+	-	-	+	-
13	<i>cd</i>	-	-	+	+	-	+	-	-	+	-	-	+	+	-	-
14	<i>acd</i>	+	-	+	+	-	-	+	+	-	-	-	+	+	-	-
15	<i>bcd</i>	-	+	+	+	-	-	-	-	+	+	+	-	+	-	-
16	<i>abcd</i>	+	+	+	+	-	+	+	+	-	+	+	-	+	-	-
17	<i>e</i>	-	-	-	-	+	+	+	+	-	+	+	-	+	-	-
18	<i>ae</i>	+	-	-	-	+	-	-	-	+	+	+	-	+	-	-
19	<i>be</i>	-	+	-	-	+	-	+	+	-	-	-	+	+	-	-
20	<i>abe</i>	+	+	-	-	+	+	-	-	+	-	-	+	+	-	-
21	<i>ce</i>	-	-	+	-	+	+	-	+	-	-	+	-	-	+	-
22	<i>ace</i>	+	-	+	-	+	-	+	-	+	-	+	-	-	+	-
23	<i>bce</i>	-	+	+	-	+	-	-	+	-	+	-	+	-	+	-
24	<i>abce</i>	+	+	+	-	+	+	+	-	+	+	-	+	-	+	-
25	<i>de</i>	-	-	-	+	+	+	+	-	-	+	-	-	-	-	+
26	<i>ade</i>	+	-	-	+	+	-	-	+	+	+	-	-	-	-	+
27	<i>bde</i>	-	+	-	+	+	-	+	-	-	-	+	+	-	-	+
28	<i>abde</i>	+	+	-	+	+	+	-	+	+	-	+	+	-	-	+
29	<i>cde</i>	-	-	+	+	+	+	-	-	-	-	-	-	+	+	+
30	<i>acde</i>	+	-	+	+	+	-	+	+	+	-	-	-	+	+	+
31	<i>bcde</i>	-	+	+	+	+	-	-	-	-	+	+	+	+	+	+
32	<i>abcde</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Tabla 2-3 Diseño 2^k y efectos de interacción dobles para $k \leq 5$

2.4 Diseño anidado.

Los diseños anidados o jerárquicos se utilizan cuando es difícil o imposible cruzar todos los factores de interés en un diseño experimental. Si cada nivel de un factor está asociado con un solo nivel de otro factor, el primer factor se dice que está anidado dentro del segundo factor. Considerando un diseño anidado de dos etapas, el factor A en el primer nivel, y el factor B en el segundo (Figura 2-3). En este diseño, B₁ en A₁ no tiene relación alguna con B₁ en A₂, esto es porque Cada nivel de B es único.

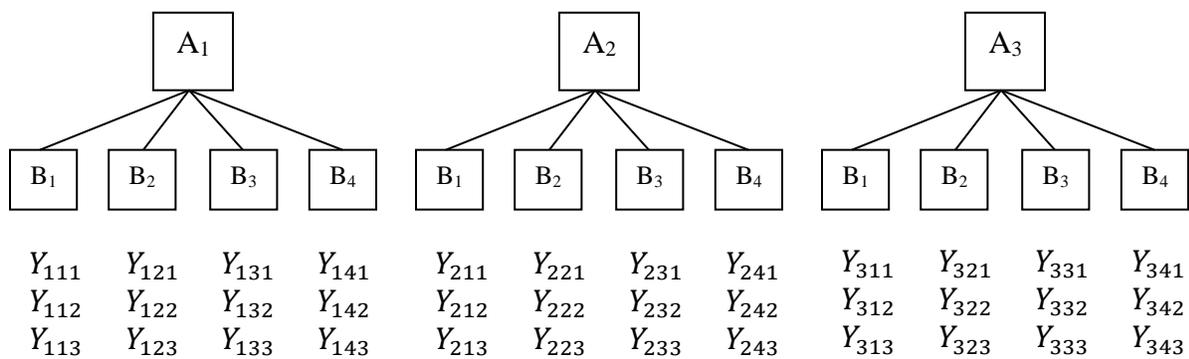


Figura 2-3 Diseño anidado en dos etapas

El modelo lineal para el diseño anidado en dos etapas es:

$$Y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{(ij)k} \quad \begin{cases} i = 1, 2 \dots a \\ j = 1, 2 \dots b \\ k = 1, 2 \dots n \end{cases} \quad 2-6$$

Con a niveles en el factor A, b niveles en B, y n replicas. El subíndice $j(i)$ se refiere a que le j -ésimo nivel del factor B esta anidado en A. Es útil considerar las replicas anidadas dentro de la combinación de niveles de A y B, por eso se utiliza el subíndice $(ij)k$ para el termino del error. Este es un diseño balanceado, porque hay igual número de niveles de B dentro de A e igual número de replicas. Porque cada nivel del factor B no aparece en cada nivel del factor A, no puede haber interacción del factor A y B. La suma de cuadrados corregida puede escribirse como:

$$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{...})^2 = bn \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2 + n \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_{ij.} - \bar{y}_{i..})^2 + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{ij.})^2 \quad 2-7$$

Debido a que tres términos de producto cruzado son cero. La ecuación

$$\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{...})^2 = bn \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2 + \quad 2-7 \text{ indica}$$

que la suma de cuadrados puede partirse en la suma de cuadrados debida al factor A, a la suma de cuadrados debida al factor B anidado en A, y a la suma de cuadrados debida al error y la ecuación $\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (y_{ijk} - \bar{y}_{...})^2 = bn \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{...})^2 +$

2-7 queda:

$$SS_T = SS_A + SS_{B(A)} + SS_E \quad 2-8$$

Donde SS_T tiene $abn-1$ grados de libertad, SS_A $a-1$ grados de libertad, $a(b-1)$ grados de libertad para $SS_{B(A)}$, y $ab(n-1)$ grados de libertad para SS_E si $e \sim N(0, \sigma^2)$, entonces las sumas de cuadrados del lado derecho de la ecuación 2-8 se dividen entre los grados de libertad, entonces se tiene una suma de cuadrados medios independientemente distribuidos, talque la relación entre dos sumas de cuadrados medios tiene distribución F .

El estadístico apropiado para probar los efectos de los factores A y B dependerá de si es tos son fijos o aleatorios. Si A y B son fijos, entonces se asume $\sum_{i=1}^a \tau_i = 0$, y $\sum_{j=1}^b \beta_{j(i)} = 0$, es decir, que la suma de los efectos de los tratamientos de A son cero, y la suma de los efectos de los tratamientos de B dentro de A son cero. Si los efectos son aleatorios entonces $\tau_i \sim N(0, \sigma_\tau^2)$ y $\beta_{j(i)} \sim N(0, \sigma_\beta^2)$. Los modelos mixtos son muy comunes, los cuadrados medios esperados pueden encontrarse con la reglas para los diseños factoriales con factores aleatorios.

	H_0	F_0
A y B fijos	$H_0: \tau_i = 0$	$\frac{MS_A}{MS_E}, \frac{MS_{B(A)}}{MS_E}$
	$H_0: \beta_{j(i)} = 0$	$\frac{MS_A}{MS_E}, \frac{MS_{B(A)}}{MS_E}$
A fijo B aleatorio	$H_0: \tau_i = 0$	$\frac{MS_A}{MS_{B(A)}}, \frac{MS_E}{MS_{B(A)}}$
	$H_0: \sigma_\beta = 0$	$\frac{MS_{B(A)}}{MS_E}, \frac{MS_E}{MS_{B(A)}}$
A y B aleatorio	$H_0: \sigma_\tau = 0$	$\frac{MS_A}{MS_{B(A)}}, \frac{MS_{B(A)}}{MS_E}$
	$H_0: \sigma_\beta = 0$	$\frac{MS_{B(A)}}{MS_E}, \frac{MS_E}{MS_{B(A)}}$

Tabla 2-4 Hipótesis estadística y estadístico de prueba para los tipos de factores

El análisis de varianza se muestra en la Tabla 2-5. Las expresiones para las sumas de cuadrados son:

$$SS_A = \frac{1}{bn} \sum_{i=1}^a y_{i..}^2 - \frac{y_{...}^2}{abn} \quad 2-9$$

$$SS_{B(A)} = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b y_{ij.}^2 - \frac{y_{...}^2}{abn} \quad 2-10$$

$$SS_E = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 - \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b y_{ij.}^2 - \frac{y_{...}^2}{abn} \quad 2-11$$

$$SS_T = \frac{1}{n} \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n y_{ijk}^2 - \frac{y_{...}^2}{abn} \quad 2-12$$

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio
A	SS_A	a-1	MS_A
B dentro de A	$SS_{B(A)}$	a(b-1)	$MS_{B(A)}$
Error	SS_E	ab(n-1)	MS_E
Total	SS_T	abn-1	

Tabla 2-5 Análisis de varianza para el diseño anidado en dos etapas

2.5 Análisis de Regresión.

El análisis de regresión tiene por objeto modelar en forma matemática el comportamiento de una variable de respuesta en función de una o más variables independientes o factores.

De manera general, supongamos que hay una variable dependiente o de respuesta (y) que depende de k variables independientes o regresoras x_1, x_2, \dots, x_k . La relación entre estas variables puede describirse por un modelo matemático llamado modelo de regresión. El modelo de regresión es apto para un conjunto de datos de una muestra. En algunos casos, el experimentador conoce la forma exacta de la verdadera relación funcional entre y y x_1, x_2, \dots, x_k , por ejemplo $y = \varphi(x_1, x_2, \dots, x_k)$. Sin embargo, en la mayoría de los casos, la relación funcional verdadera es desconocida, y el experimentador elige una función adecuada para aproximar φ . Los modelos de bajo orden polinomial se utilizan ampliamente como la

aproximación de funciones. En muchas ocasiones el modelo de regresión representa los resultados de un experimento.

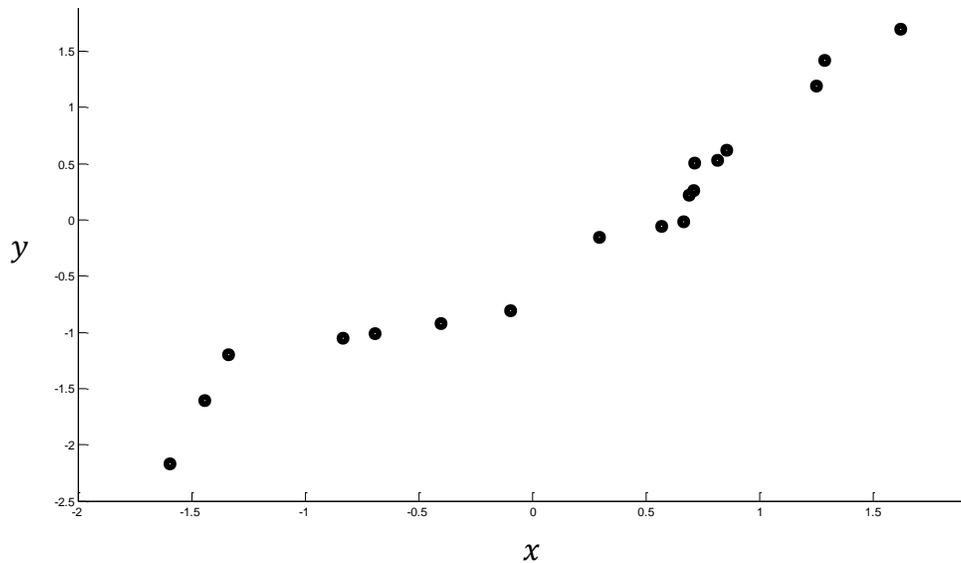


Gráfico 2-1 Diagrama de dispersión

Dado un conjunto de datos x, y , mostrados en el diagrama de dispersión (Gráfico 2-1), donde cada uno de los puntos representa un par (x_i, y_i) . Una inspección del diagrama, muestra que ninguna curva simple pasa exactamente por todos los puntos, pero existe un fuerte indicio de que los puntos se colocan aleatoriamente alrededor de una línea recta. Por eso es posible asumir que la media de la variable Y , se relaciona con la variable x mediante la siguiente expresión:

$$E(Y|x) = \mu_{Y|x} = \beta_0 + \beta_1 x \quad 2-13$$

Donde la pendiente y la ordenada al origen se llaman coeficientes de regresión. Mientras que Y es una función lineal de x , el valor verdadero de Y no está exactamente sobre una línea recta. La forma apropiada de generalizar a un modelo lineal probabilístico es asumir que $E(Y)$ es una función lineal de x , pero para conocer el valor verdadero de Y dado un valor de x , se calcula con el valor promedio de la función modelo de regresión más un error aleatorio y se dice que:

$$Y = \beta_0 + \beta_1 x + \epsilon \quad 2-14$$

Donde ϵ representa el término del error aleatorio. A la expresión anterior se le conoce como modelo de regresión simple, debido a que solo incluye una variable regresora. Algunas veces el modelo anterior se deriva de una relación teórica entre las variables. En otras ocasiones no existe conocimiento previo de la relación entre las variables, y la elección del modelo se apoya en la inspección del diagrama de dispersión, en este caso, se dice que el modelo es empírico.

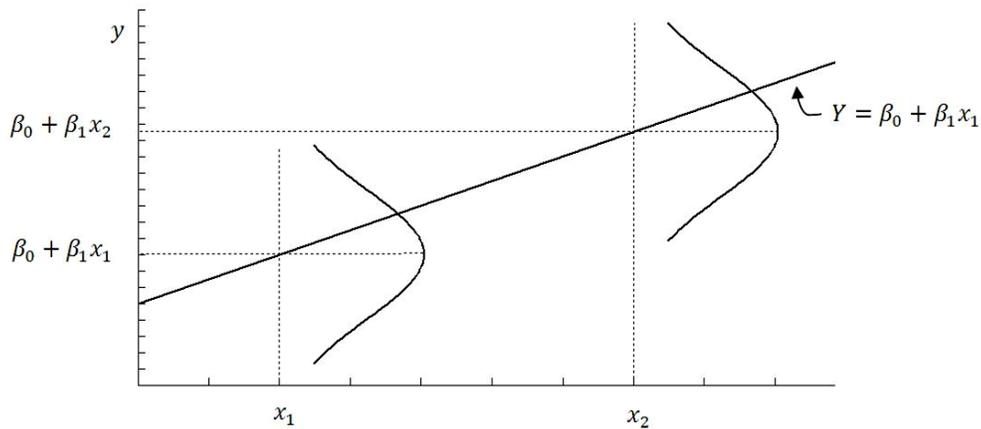


Gráfico 2-2 Distribución de Y dado un valor x

Para profundizar más en el modelo de regresión, suponiendo que el valor de x queda fijo, entonces se puede observar la variación de Y . Entonces el componente aleatorio ϵ del lado derecho determina las propiedades de Y , suponiendo que la media y varianza de ϵ son 0 y σ^2 , entonces:

$$E(Y) = E(\beta_0 + \beta_1 x + \epsilon) = E(\beta_0 + \beta_1 x) + E(\epsilon) = \beta_0 + \beta_1 x \quad 2-15$$

Y la varianza de Y es:

$$V(Y) = V(\beta_0 + \beta_1 x + \epsilon) = V(\beta_0 + \beta_1 x) + V(\epsilon) = 0 + \sigma^2 = \sigma^2 \quad 2-16$$

Por esta razón, se dice que el modelo verdadero de regresión $\mu_{Y|x} = \beta_0 + \beta_1 x$ es una línea de valores promedio, el valor de la ordenada para un valor cualquiera de x es el valor esperado de Y . El valor de la pendiente β_1 , representa el cambio en la media por cada unidad de x . Además, la variación de Y está determinada por la varianza del error σ^2 . Esto

implica que hay una distribución de valores de Y en cada valor de x , y la varianza es la misma en cada valor de x ().

En la práctica, la mayoría de los problemas involucran más de una variable independiente, y los modelos que relacionan a la variable dependiente con las variables independientes son los modelos de regresión lineal múltiple. Suponiendo un proceso donde la variable dependiente Y depende de dos variables x_1 y x_2 . Un modelo de regresión lineal múltiple que pudiese describir la relación entre las variables sería:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon \quad 2-17$$

La expresión anterior representa un modelo de regresión lineal múltiple con dos variables regresoras, la superficie descrita corresponde a un plano. El Gráfico 2-3 a) muestra la superficie generada por el modelo de regresión:

$$E(Y) = 40 + 35x_1 + 15x_2 \quad 2-18$$

Donde se asume que el valor esperado del término del error es cero ($E(\epsilon) = 0$), la intercepción del plano es β_0 . En ocasiones β_1 y β_2 son llamados coeficientes de regresión parciales, debido a que β_1 indica el cambio de la variable Y respecto al incremento unitario de x_1 manteniendo x_2 constante, y β_2 indica el cambio de la variable Y respecto al incremento unitario de x_2 manteniendo x_1 constante. El gráfico de contorno del modelo de regresión (Gráfico 2-3 b)), muestra líneas de valores constantes de $E(Y)$ en función de las variables x_1 y x_2 .

Generalizando, una variable dependiente o de respuesta Y , puede relacionarse a k variables independientes o regresoras mediante:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon \quad 2-19$$

Denominado modelo de regresión lineal múltiple con k variables regresoras, donde $\beta_j, j=0, 1, 2, \dots, k$, son los coeficientes de regresión. Este modelo representa un hiper plano en el espacio k -dimensional de las variables regresoras $\{x_j\}$. El parámetro β_j representa el cambio

en Y por unidad de x_j , mientras todas las demás variables regresoras ($x_i \ i \neq j$) se mantienen constantes.

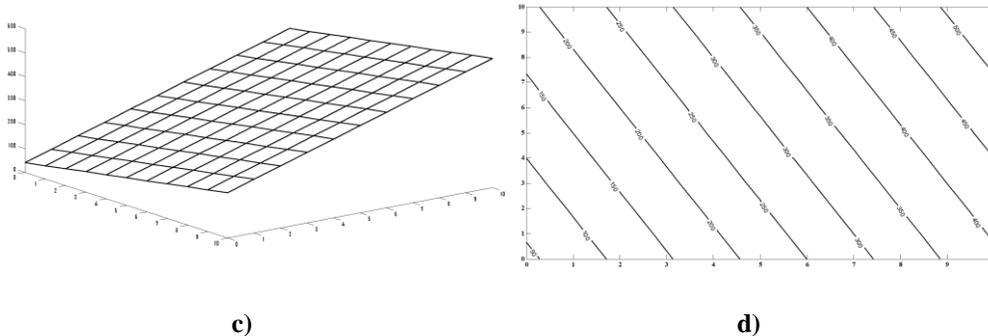


Gráfico 2-3 a) plano de regresión del modelo $E(Y) = 40 + 35x_1 + 15x_2$ b) gráfico de contorno

Cuando se desconoce la función verdadera que relaciona Y con $x_1, x_2 \dots x_k$, los modelos de regresión lineal múltiple son una buena aproximación, pero solo en ciertos intervalos de las variables independientes la aproximación es adecuada. Los modelos polinomiales en una variable, o con interacciones, pueden analizarse con técnicas de regresión lineal múltiple. Para un modelo de regresión cubico con una variable:

$$Y = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3x^3 + \epsilon \quad 2-20$$

Haciendo $x_1 = x$; $x_2 = x^2$ y $x_3 = x^3$ el modelo de regresión lineal múltiple queda:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \epsilon \quad 2-21$$

Que es un modelo de regresión lineal múltiple con tres variables regresoras. Para un modelo de regresión con algún efecto de interacción entre dos variables como:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \epsilon \quad 2-22$$

Haciendo $x_3 = x_1x_2$ y $\beta_3 = \beta_{12}$ la (ecuación $Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \epsilon$

2-22) puede expresarse como un modelo de regresión lineal múltiple con tres variables regresoras

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 \quad 2-23$$

Considerando el modelo de regresión:

$$Y = 2 + 4x + 1.5x^2 + 3x^3 \quad 2-24$$

Que es una curva en dos dimensiones (Gráfico 2-4), y al hacer la siguiente sustitución: $x_1 = x$; $x_2 = x^2$ y $x_3 = x^3$ el modelo de regresión lineal múltiple queda:

$$Y = 2 + 4x_1 + 1.5x_2 + 3x_3 \quad 2-25$$

El cual representa un hiper plano en tres dimensiones. Nótese que a pesar de que la (ecuación $Y = 2 + 4x + 1.5x^2 + 3x^3$ 2-24) representa un modelo de regresión lineal, el grafico generado corresponde a una función no lineal en dos dimensiones. En general, un modelo que contiene coeficientes lineales (β_j) es un modelo de regresión lineal, independientemente de la forma del grafico que genera.

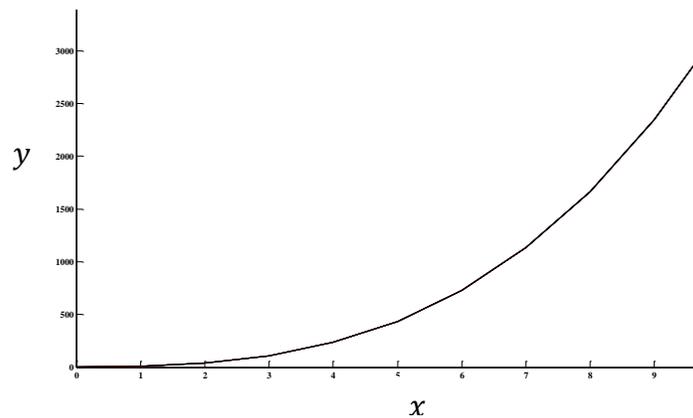


Gráfico 2-4 $Y = 2 + 4x + 1.5x^2 + 3x^3$

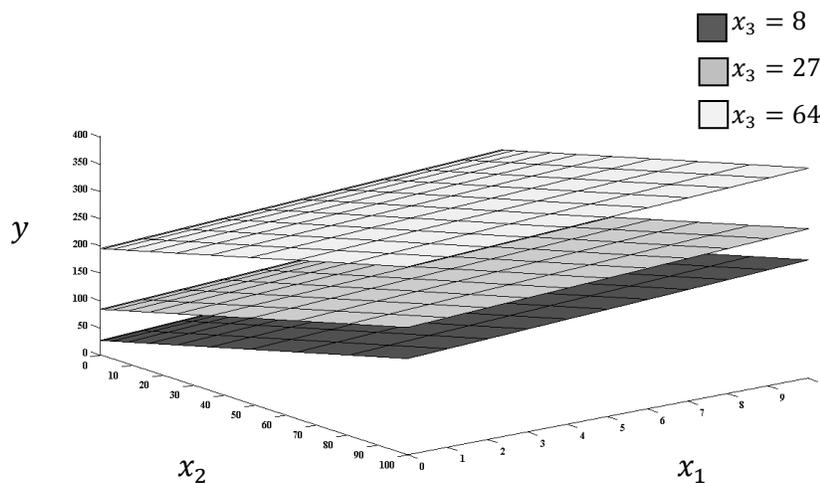


Gráfico 2-5 Híper planos generados por $Y = 2 + 4x_1 + 1.5x_2 + 3x_3$

2.5.1 Estimación de los Parámetros del Modelo de Regresión

Para ajustar el modelo de regresión lineal, es decir, estimar los valores de los parámetros o coeficientes de regresión, resulta muy conveniente utilizar la notación matricial. Suponiendo un modelo con k variables regresoras y n observaciones, el modelo que relaciona la variable independiente con las variables regresoras es:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \dots + \beta_k x_{ik} + \epsilon; \quad i = 1, 2, 3 \dots n \quad 2-26$$

Que puede escribirse en notación matricial

$$y = \beta X + \epsilon \quad 2-27$$

Donde:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_k \end{bmatrix} \quad 2-28$$

El objetivo es encontrar el vector de coeficientes $\hat{\beta}$ que minimizan:

$$L = \sum_{i=1}^n \epsilon_i^2 = \epsilon' \epsilon = y - (\beta X)' y - \beta X \quad 2-29$$

El estimador de mínimos cuadrados $\hat{\beta}$ está dado por

$$\hat{\beta} = (X'X)^{-1} X'y \quad 2-30$$

El modelo ajustado es $\hat{y} = X\hat{\beta}$, la diferencia entre la observación y y el valor ajustado \hat{y} es un residual, decimos que $e_i = y_i - \hat{y}_i$ ($i=1, 2, 3 \dots n$), o en forma matricial:

$$e = y - \hat{y} \quad 2-31$$

2.5.2 Prueba de Hipótesis para la Significancia del Modelo

Seguido a la estimación de los parámetros de regresión, debe analizarse la calidad del modelo ajustado. En la regresión lineal múltiple, ciertas pruebas de hipótesis sobre los parámetros del modelo son útiles para medir la adecuación del modelo de regresión. Estas pruebas de hipótesis se realizan bajo los supuestos de que los errores son independientes, y distribuidos normalmente con media cero y varianza constante ($e \sim N(0, \sigma^2)$).

La prueba para la significancia del modelo de regresión determina el grado de relación lineal entre la variable de respuesta (y) y el conjunto de variables regresoras ($x_1, x_2, x_3 \dots x_k$). La hipótesis apropiada para esta prueba es:

$$H_0: \beta_0 = \beta_1 = \beta_2 = \dots \beta_k = 0 \quad 2-32$$

$$H_1: \beta_j \neq 0 \text{ para al menos una } j$$

La no aceptación de H_0 implica que al menos una de las variables regresoras contribuye significativamente al modelo. Para realizar esta prueba, la suma de cuadrados totales es la suma de dos particiones: la suma de cuadrados debida a la regresión, y la suma de cuadrados debida al error (ecuación $SS_T = SS_R + SS_E$

2-33).

$$SS_T = SS_R + SS_E \quad 2-33$$

Ahora bien, si H_0 es verdadera, entonces SS_R/σ^2 es una variable aleatoria con distribución χ^2 con k (numero de variables regresoras) grados de libertad. También SS_E/σ^2 es una variable aleatoria con distribución χ^2 con $n-p$ (observaciones menos parámetros del modelo) grados de libertad, además SS_R y SS_E son independientes. El estadístico de prueba para H_0 es:

$$F_0 = \frac{SS_R/k}{SS_E/(n-p)} = \frac{MS_R}{MS_E} \quad 2-34$$

No se acepta H_0 si $F_0 > F_{\alpha, k, n-p}$. Las sumatorias de cuadrados se muestran en notación matricial como sigue:

$$SS_E = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}$$

$$SS_R = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{1}{n}(\sum_{i=1}^n y_i)^2 \quad 2-35$$

$$SS_T = \mathbf{y}'\mathbf{y} - \frac{1}{n}(\sum_{i=1}^n y_i)^2$$

Otro criterio para evaluar la calidad del modelo de regresión es el coeficiente de de determinación R^2 .

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T} \quad 2-36$$

El R^2 es una medida de la reducción en la variación de y utilizando las variables regresoras x_1, x_2, \dots, x_k en el modelo. Sin embargo, un valor grande de R^2 no implica necesariamente que el modelo de regresión sea el adecuado. Con la adición de una variable en el modelo, R^2 se incrementa, independientemente de si la variable adicional es estadísticamente significativa o no. Por lo tanto, es posible que los modelos que tienen con valores de R^2 elevados, resulten poco adecuados en la predicción de nuevas observaciones o estimaciones de la respuesta media. Debido al incremento de R^2 conforme añadimos términos al modelo, algunos analistas prefieren utilizar un estadístico R^2 ajustado, definido como:

$$R_{adj}^2 = 1 - \frac{SS_E/(n-p)}{SS_T/(n-1)} = 1 - \left(\frac{n-1}{n-p}\right) (1 - R^2) \quad 2-37$$

Generalmente, el R_{adj}^2 no siempre aumentará a medida que se agregan variables al modelo. De hecho, si los términos agregados son innecesarios, el valor de R_{adj}^2 con frecuencia disminuirá.

Fuente de variación	Suma de cuadrados	Grados de libertad	Cuadrados medios	F_0
Regresión	SS_R	k	SS_R/k	MS_R/MS_E
Error o residuales	SS_E	$n-p$	$SS_E/(n-p)$	
Totales	SS_T	$n-1$		

Tabla 2-6 Análisis de varianza para la prueba de significancia del modelo de regresión

2.6 Optimización de Respuesta Múltiple.

La mayoría de los problemas de calidad, consideran el análisis de más de una variable de calidad. La calidad del producto o proceso depende del valor que toman varias propiedades. En la mayoría de los casos, optimizar una sola de las características, provoca que las demás propiedades se vean afectadas y el resultado es un producto de mala calidad global. De aquí se deriva la importancia de contar con técnicas que permitan optimizar de manera simultáneas a todas las variables de respuesta de interés, o encontrar un punto de operación del proceso donde todas las variables tengan un valor deseado.

Debido a que es improbable que el óptimo global coincida con el óptimo de alguna respuesta individual, la solución del problema de respuesta múltiple suele ser satisfactoria para cada respuesta individual; puede esperarse que exista un punto de operación del proceso en el que todas las variables de respuesta cumplen con los requerimientos de la mejor manera posible.

Considerar simultáneamente varias respuestas, implica la construcción de modelos de regresión adecuados (errores independientes, distribuidos normalmente con media cero y varianza constante, $R_{aj}^2 \geq 70\%$). La optimización por métodos gráficos o analíticos requiere de especificaciones unilaterales o bilaterales.

2.6.1 Superposición de Gráficos de contorno.

Este método consiste en construir gráficos de contorno (Gráfico 2-3 a), para cada una de las variables de respuesta en función de las variables independientes (factores). Como se menciona en la sección 1.1, este método está limitado por el número de factores, (dimensión del problema). Cuando se trata de dos factores de control, resulta fácil superponer los gráficos de contorno de cada variable de respuesta, e identificar la región donde las variables de respuesta toman valores factibles. Tomando el ejemplo de Montgomery (2001) para un proceso químico con dos factores, tiempo (x_1) y temperatura (x_2) y tres variables de respuesta, rendimiento (y_1), viscosidad (y_2) y peso molecular (y_3). Los valores deseados para las variables de respuesta son: $y_1 \geq 78.5$, $62 \leq y_2 \leq 68$, $y_3 \leq 3400$. Los modelos de regresión ajustados se muestran a continuación:

$$\begin{aligned}\hat{y}_1 &= 79.94 + 0.99x_1 + 0.52x_2 - 1.38x_1^2 - x_2^2 + 0.25x_1x_2 \\ \hat{y}_2 &= 70 - 0.16x_1 - 0.95x_2 - 0.69x_1^2 - 6.69x_2^2 - 1.25x_1x_2 \\ \hat{y}_3 &= 3386.2 + 205.1x_1 + 17.4x_2\end{aligned}\tag{2-38}$$

Las superficies y gráficos de contorno se muestran en los gráficos Gráfico 2-6 a Gráfico 2-8. Las restricciones para cada variable de respuesta, representan condiciones que el proceso debe cumplir, en el

Gráfico 2-9, la región sombreada corresponde a combinaciones de tiempo y temperatura que corresponden a valores satisfactorios para las respuestas. El experimentador puede inspeccionar el gráfico de contorno para determinar las condiciones de operación apropiadas.

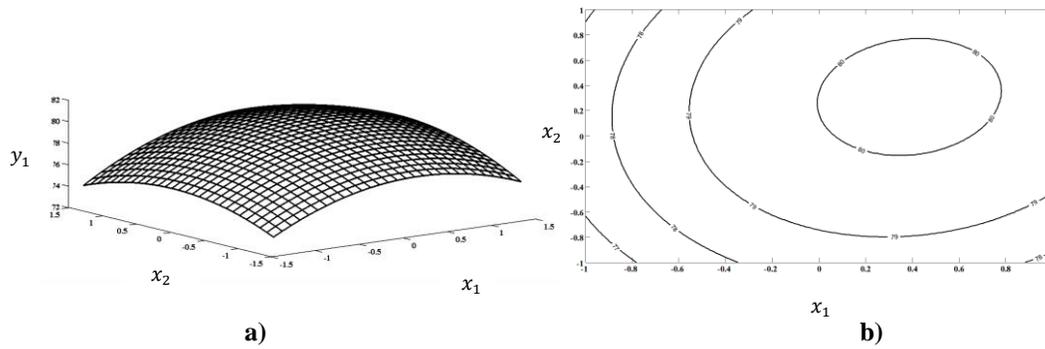


Gráfico 2-6. Superficie de respuesta (a), y gráfico de contorno (b) para y_1

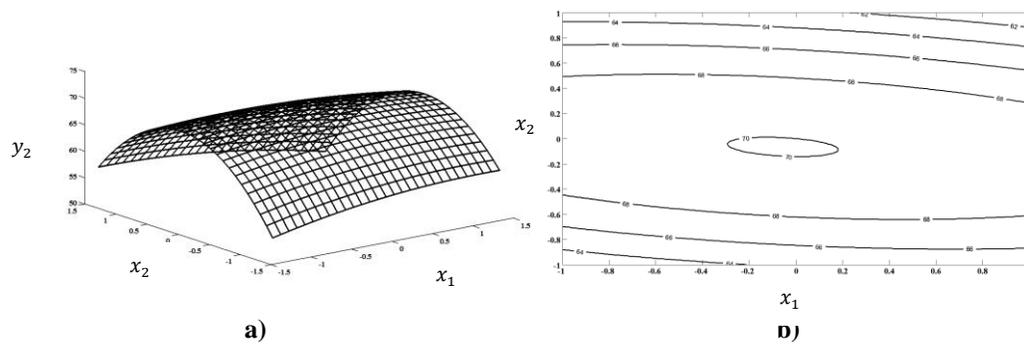


Gráfico 2-7. Superficie de respuesta (a), y gráfico de contorno (b) para y_2

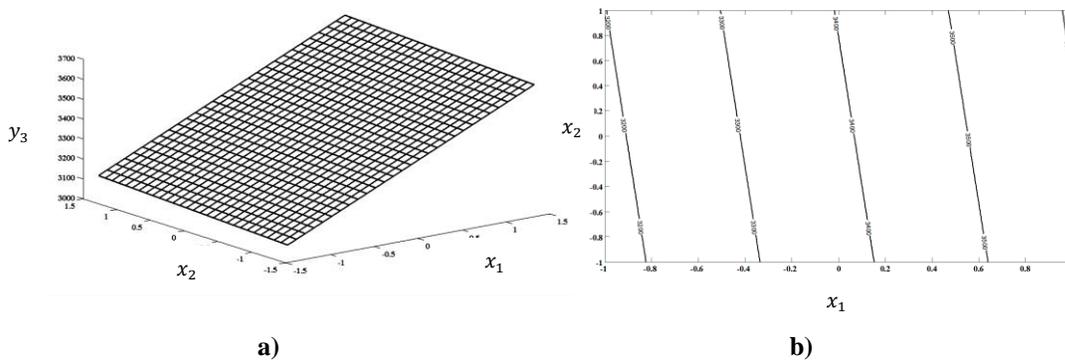


Gráfico 2-8. Superficie de respuesta (a), y gráfico de contorno (b) para y_3

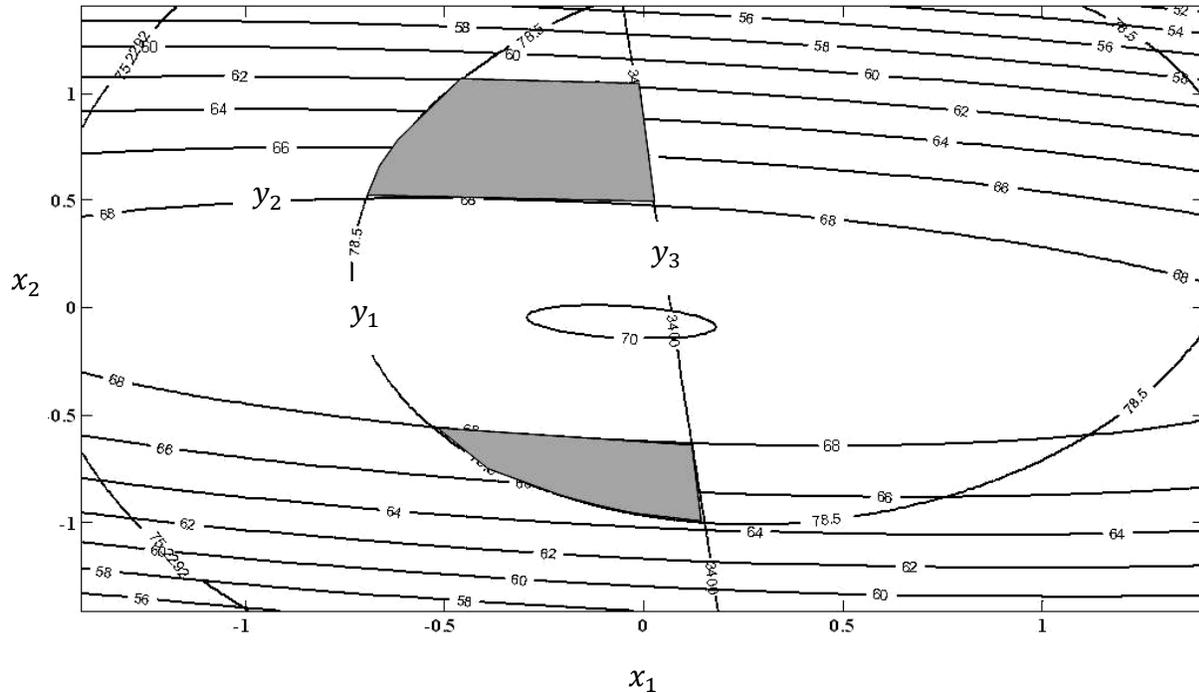


Gráfico 2-9. Región óptima para el rendimiento (y_1) la viscosidad (y_2) y el peso molecular (y_3)

Cuando se tiene más de tres factores, resulta complicado el uso de la superposición de gráficos de contorno, debido a que es una representación en dos dimensiones, esto implica que deben mantenerse $k-2$ factores constantes para la construcción del gráfico de contorno. Esto requiere de varios ensayos a prueba y error, y es en este tipo de problemas, que las funciones de utilidad cobran importancia.

2.6.2 Funciones de Utilidad.

Las funciones mencionadas en este apartado, combinan todas las variables de respuesta en un solo valor cuantitativo, también llamado métrica de desempeño, es decir, las funciones de utilidad son transformaciones del conjunto de variables de respuesta. La optimización de la función de utilidad en función de los factores de control, conduce a las condiciones de proceso óptimas.

2.6.2.1 Derringer y Suich

La función de deseabilidad involucra la transformación del estimador \hat{Y}_{ij} en un valor deseado d_{ij} , donde $0 \leq d_{ij} \leq 1$, después, las deseabilidades individuales se combinan mediante la media geométrica (3) para obtener una calificación de la i -ésima observación para la j -ésima respuesta.

$$D_{gi}(x) = [\prod_{j=1}^k d_{ij}]^{\frac{1}{k}} \quad 2-39$$

La ecuación anterior puede definirse como el antilogaritmo del promedio de los logaritmos naturales de las deseabilidades individuales utilizando:

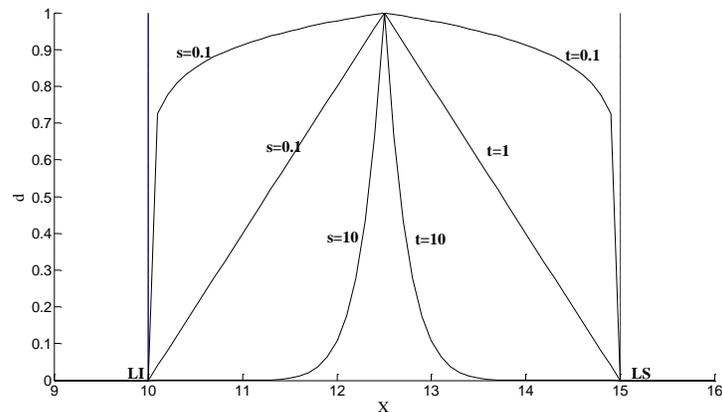
$$D_{gi}(x) = e^{\left[\frac{1}{k} \sum_{j=1}^k \ln(d_{ij})\right]} \quad 2-40$$

Las transformaciones del estimador \hat{Y}_j , dependerán de la característica de calidad para cada

$$d_{ij} = \begin{cases} \frac{Y_{ij} - Y_{j^*}}{c_j - Y_{j^*}} & Y_{j^*} < Y_{ij} < c_j \\ \frac{Y_{j^*} - Y_{ij}}{Y_{j^*} - c_j} & c_j < Y_{ij} < Y_{j^*} \\ 0 & Y_{j^*} > Y_{ij} > Y_{j^*} \end{cases} \quad 2-41. \text{ Para variables del tipo "menor es mejor"} - Y_j \text{ se}$$

trata como "mayor es mejor".

$$d_{ij} = \begin{cases} \left(\frac{Y_{ij} - Y_{j^*}}{c_j - Y_{j^*}}\right)^s & Y_{j^*} < \hat{Y}_{ij} < c_j \\ \left(\frac{Y_{j^*} - \hat{Y}_{ij}}{Y_{j^*} - c_j}\right)^t & c_j < \hat{Y}_{ij} < Y_{j^*} \\ 0 & Y_{j^*} > \hat{Y}_{ij} > Y_{j^*} \end{cases} \quad 2-41$$



Gráfico

$$d_{ij}=0 \quad Y_{ij} \leq Y_j^* Y^t - Y_j^* Y_j^* - Y_j^* r \quad Y_j^* < Y_{ij} < Y_j^* 1 \quad Y_{ij} \geq Y_j^*$$

2-42 es la transformación utilizada. Los exponentes r , s y t , llamados pesos, que permiten un comportamiento lineal o no lineal de la función de deseabilidad. Por ejemplo, si los pesos son iguales a uno, la deseabilidad tiene comportamiento lineal (

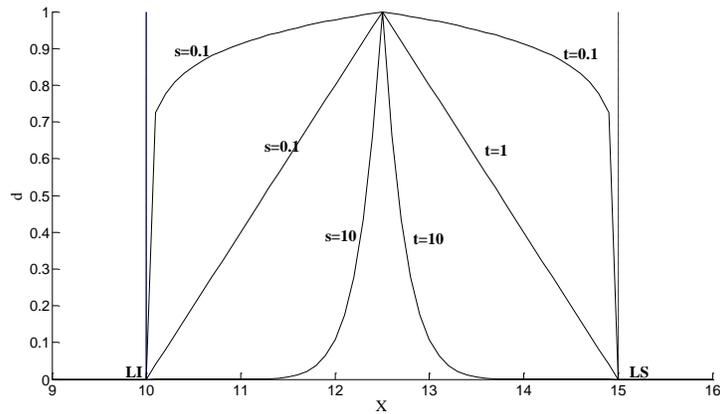


Gráfico 2-10 y Gráfico 2-11).

$$d_{ij} = \begin{cases} 0 & \hat{Y}_{ij} \leq Y_j^* \\ \left(\frac{Y_{ij} - Y_j^*}{Y_j^* - Y_j^*} \right)^r & Y_j^* < \hat{Y}_{ij} < Y_j^* \\ 1 & \hat{Y}_{ij} \geq Y_j^* \end{cases} \quad 2-42$$

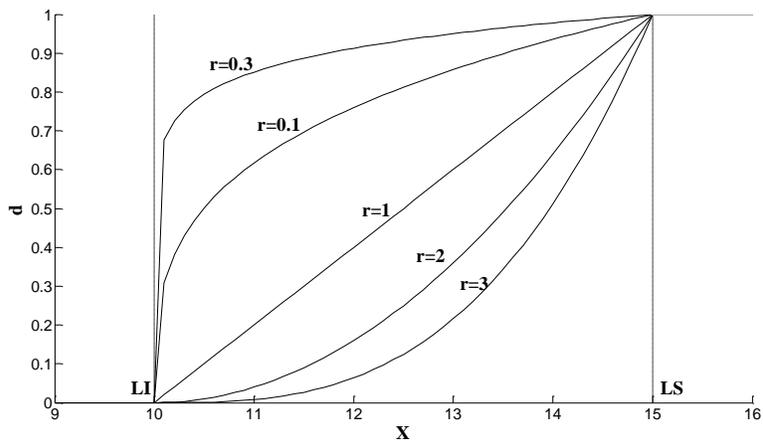


Gráfico 2-11 deseabilidad unilateral

Una inspección de ecuación $D_{gi}(x) = [\prod_{j=1}^k d_{ij}]^{\frac{1}{k}}$ 2-39

revela que, si para alguna de las variables de respuesta la deseabilidad es cero, la deseabilidad global (d_g) es cero. Este hecho permite eliminar soluciones que no cumplen con todas las restricciones (ecuación 2-43):

$$\forall j \in (1, 2, 3 \dots k) \exists d_j = 0 \therefore d_g = 0 \quad 2-43$$

La deseabilidad es una función continua de las variables independientes ($X_1, X_2, \dots X_p$) que transforma el problema multivariable en un problema univariable. Esto permite encontrar el valor máximo de deseabilidad por técnicas de búsqueda analíticas o graficas.

2.6.2.2 Pignatiello

Esta métrica es una variación de la función de pérdida de Taguchi, y se basa en la minimización de la desviación del valor objetivo y maximizar la robustez hacia el ruido. Si alguna variable de respuesta se aleja demasiado del objetivo, se aplica una penalización ponderada de pérdida. La métrica se denota por $\hat{D}_p(x)$.

La base de esta técnica es el error cuadrado de pérdida ponderado, definido por:

$$L_p(x) = [Y(x) - \tau]'C \quad 2-44$$

La pérdida esperada está dada por:

$$E[L_p(x)] = \text{traza}(C\Sigma(x)) + [\eta(x) - \tau]'C[\eta(x) - \tau] \quad 2-45$$

La traza se define como la suma de los elementos de la diagonal de una matriz.

Sustituyendo $\Sigma(x)$ y $\eta(x)$ por sus estimadores $\hat{S}_Y(x)$ y $\hat{Y}(x)$ en la ecuación $E[L_p(x)] = \text{traza}(C\Sigma(x)) + [\eta(x) - \tau]'C[\eta(x) - \tau]$ 2-45 se obtiene el estimador de la perdida esperada, la métrica de desempeño.

$$\hat{D}_p(x) = \text{traza}(C\hat{S}_Y(x)) + [\hat{Y}(x) - \tau]'C[\hat{Y}(x) - \tau] \quad 2-46$$

El método asume que la varianza de una respuesta es distinta para cada conjunto de valores de x , el estimador $\hat{S}_Y(x)$ se obtiene para cada tratamiento y se calcula a partir de las varianzas de la muestra y la covarianzas de $\hat{Y}(x)$. El óptimo de la métrica de desempeño consiste en encontrar el mínimo $\hat{D}_p(x)$.

$\hat{D}_p(x)$ se divide en dos componentes. Uno es el componente debido a la desviación del objetivo (segundo término en la ecuación 2-46), y el componente de varianza (primer término en la ecuación 2-46). La porción de desviación del objetivo es la desviación del objetivo al cuadrado por la matriz de pesos C . esta porción es comparable a la métrica anterior $D_g(x)$. El componente de varianza es la suma de las varianzas ponderadas por C , que penaliza los puntos con varianza de $\hat{Y}(x)$ elevada.

Entre los inconvenientes de esta métrica, es que es aplicable a problemas con variables de salida del tipo “nominal es mejor”, para los casos de maximizar o minimizar debe establecerse un valor objetivo adecuado la matriz de pesos C llamada matriz de costo, con frecuencia no está disponible. Una modificación a este concepto permite utilizar la matriz de costo para representar la importancia relativa de las variables de salida.

2.6.2.3 Khuri y Colon

Esta métrica de desempeño utiliza la diferencia al cuadrado de la respuesta respecto a su objetivo, pero normaliza estas diferencias por medio de la varianza de la predicción de las variables de respuesta.

Esta función de utilidad considera una forma de determinar los objetivos $\tau_i^*(i=1,2,\dots,p)$. Para variables del tipo “menor es menor” o “mayor es mejor”, τ_i se establece como el mínimo (o máximo) valor que toma Y_i en la región experimental.

Debido a la aleatoriedad del estimador de Y_i , el óptimo individual de dicho estimador usado como objetivo, también es una variable aleatoria, la función de utilidad considera dicha

aleatoriedad, y calcula una región de confianza para el objetivo τ_i^* , y se selecciona un nuevo objetivo dentro de la región de confianza, de tal modo que la métrica de desempeño se maximice.

$$\hat{D}_{KC} = \sqrt{[\hat{Y}(x) - \tau]' \text{Var}[\hat{Y}(x)]^{-1} [\hat{Y}(x) - \tau]'} \quad 2-47$$

El término $\text{Var}[\hat{Y}(x)]$ es una matriz de $p \times p$ de las varianzas de los modelos de regresión de para las respuestas Y_i en el tratamiento correspondiente x . un estimador insesgado es:

$$\text{Var}[\hat{Y}(x)] = z'(x)(X'X)^{-1}z'(x)\hat{\Sigma} \quad 2-48$$

En la ecuación 2-48, X es la matriz de variables de utilizada en la estimación de los parámetros del modelo de regresión (ecuación 2-28).

Cada fila de X denotado por $z'(x)$, corresponde a un tratamiento particular de x , los modelos de regresión tienen las mismas variables regresoras. $\hat{\Sigma}$ es el estimador de la matriz de varianza-covarianza. Sustituyendo ecuación 2-48 en la ecuación 2-47, se tiene:

$$\hat{D}_{KC} = \sqrt{\frac{[\hat{Y}(x) - \tau]' \hat{\Sigma}^{-1} [\hat{Y}(x) - \tau]'}{z'(x)(X'X)^{-1}z'(x)}} \quad 2-49$$

Las condiciones óptimas de operación se obtienen minimizando \hat{D}_{KC} . Es importante recalcar que las variables de respuesta deben ser no correlacionadas, en tal caso, deberán

2.7 Búsqueda Heurística Aleatoria

La búsqueda heurística aleatoria se fundamenta en los conceptos de estado y transición entre estados. El mecanismo de búsqueda se construye a partir de la representación del estado y la definición de las transiciones ⁽¹²⁾. Una colección de elementos P_0 , pertenecientes al espacio de búsqueda Ω de n dimensiones, junto con la regla de transición τ , la cual transforma P_i en otra colección P_{i+1} . En general, τ es iterada para producir una serie de poblaciones $P_0 \xrightarrow{\tau} P_1 \xrightarrow{\tau} P_2 \xrightarrow{\tau} P_3$. La generación P_0 es aleatoria, y de regla de

transición τ es determinística, y esta última es la que define algoritmo de búsqueda en particular.

La inteligencia artificial agrupa algoritmos que definen sus reglas de transición a partir de fenómenos biológicos, tales como la selección natural y el movimiento de enjambres. Como los algoritmos presentados a continuación.

2.7.1 Algoritmo Genético (GA).

Esta es una técnica de búsqueda y optimización altamente paralela, inspirada en la selección natural y reproducción. Donde los individuos más aptos son los que pueden transmitir su material genético a la siguiente generación. En este algoritmo, se genera una población inicial de manera aleatoria, cada solución representa el cromosoma de un individuo. Generalmente, la representación de dicha solución es codificada en sistema binario (0,1), esto es para aplicar los operadores genéticos fácilmente.

El GA canónico, utiliza la representación binaria, y tres operadores genéticos, que son: selección, cruzamiento y mutación. Después de que la población es generada, se evalúa el valor de la función objetivo ($f(x)$), en el caso de maximización, al escoger aleatoriamente dos individuos, el más apto será aquel que tenga un valor de $f(x)$ mayor.

El proceso de selección es para escoger individuos para la reproducción, la selección se apoya en la aptitud, individuos más aptos tiene mayor probabilidad de ser seleccionados. La probabilidad de un individuo i ($P_i, i=1,2\dots n$) de ser seleccionado está dada por:

$$P_i = \frac{f(x_i)}{\sum_{i=1}^n f(x_i)} \quad 2-50$$

Los individuos seleccionados pasan al siguiente operador, el cruzamiento, para producir una nueva generación de individuos, la aplicación del operador de reproducción está condicionada a la probabilidad de cruza. El cruzamiento es la combinación de los dos vectores solución para generar una nueva. Las soluciones generadas por el cruzamiento son

susceptibles de cambiar debido al operador de mutación, cuya aplicación está asociada a la probabilidad de mutación. En la Figura 0-1 se muestra el diagrama de flujo del GA.

2.7.2 Optimización por Enjambre de Partículas (PSO).

El algoritmo PSO, fue propuesto por Kennedy y Eberhart ⁽¹³⁾, está basado en dinámica de poblaciones. Comienza con la generación aleatoria de una población inicial de partículas, donde cada partícula representa una solución en el espacio n-dimensional; además cada partícula está influenciada por la mejor posición visitada por la partícula, y la mejor posición global (solución), entre todo el enjambre. El algoritmo básico, en cada iteración las partículas modifican su posición utilizando las siguientes expresiones:

$$V_{k+1} = w_k + C_1 r_1 (LBX - X) + C_2 r_2 (GBX - X) \quad 2-51$$

$$X_{k+1} = X_k + V_k \quad 2-52$$

V_k = velocidad en la iteración k

X_k = posición en la iteración k

w = factor de inercia

C_1 y C_2 = factores de aprendizaje

LBX = mejor posición visitada por la partícula

GBX = mejor posición de todo el enjambre

r_1 y r_2 = números aleatorios $\sim U(0,1)$

C_1 y C_2 afectan a la exploración local y global respectivamente, y junto con w , determinan el desempeño del algoritmo. Después de modificar la velocidad, cada partícula actualiza su mejor posición (asumiendo un problema de minimización):

$$LBX_{k+1} = \begin{cases} LBX_k & f(LBX_k) \leq f(X_{k+1}) \\ X_{k+1} & f(LBX_k) > f(X_{k+1}) \end{cases} \quad 2-53$$

La mejor posición visitada por todo el enjambre corresponde al punto donde $f(LBX_k)$ es mínima:

$$GBX_{k+1} = \arg \min_{LBX_{k+1}} f(LBX_{k+1}) \quad 2-54$$

El proceso es repetido hasta alcanzar el criterio de paro.

CAPÍTULO 3.

METODOLOGÍA

3.1 Metodología propuesta.

Para comparar el desempeño de los algoritmos genético y optimización con enjambre de partículas, es necesario que los algoritmos, con los parámetros adecuados provean soluciones *precisas y exactas*. El uso del diseño de experimentos, permite analizar los datos recolectados con métodos estadísticos, resultando en conclusiones objetivas y validas (Montgomery, 2001). La metodología propuesta (Figura 0-1), consiste en dos etapas, la primera es la sintonización de los algoritmos seguido de la comparación mediante un diseño anidado.

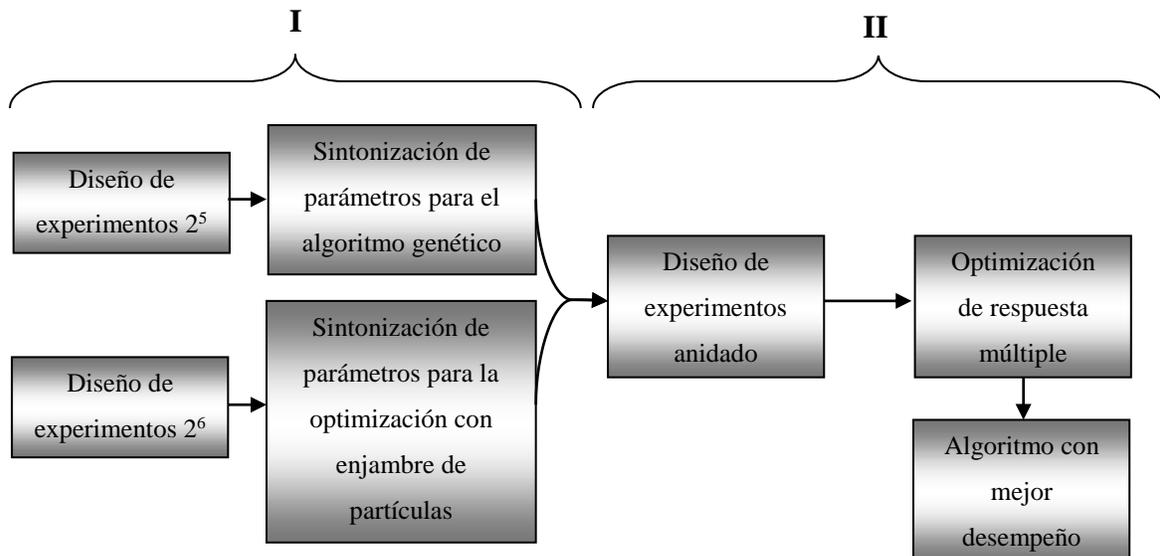


Figura 0-1 Esquema de la metodología

Para sintonizar los algoritmos, se utilizan funciones de benchmarking, estas incluyen funciones unimodales y multimodales, midiendo la media y desviación estándar de la solución obtenida de cien corridas consecutivas (unidad experimental). Con los algoritmos sintonizados para ser precisos y exactos, es posible compararlos solucionando problemas de optimización de respuesta múltiple por medio de un diseño de experimentos anidado, el cual permitirá detectar diferencias estadísticamente significativas entre las soluciones encontradas por los dos algoritmos.

3.2 Sintonización de AG y PSO.

La sintonización del algoritmo genético utilizado, tiene por objetivo encontrar los parámetros los tipos de operadores que minimicen la desviación estándar y el error de las soluciones de múltiples ensayos. El algoritmo genético empleado es el incluido en gatool de Matlab 7.6.0.324, con los operadores mostrados en la Tabla 0-1, el tipo de población es vector doble, y se omite la fracción de migración, debido a que no es un operador del algoritmo genético canónico (Figura 0-1)

Operador	Tipo
Función de creación	Uniforme
Escalamiento	Proporcional
Selección	Ruleta
Cruza	Dos puntos
Mutación	Factible adaptativa

Tabla 0-1 Operadores del GA empleado

La sintonización se realiza mediante un diseño de experimentos 2^5 , que se genera y analiza con el apoyo del software STATGRAPHICS Centurión XV 15.1.02. Los factores utilizados para la sintonización se muestran en la Tabla 0-2.

Factor	Alto	Bajo
A: Dimensión	9	5
B: Tamaño de población	50	30
C: No. De élites	5	3
D: Fracción de cruza	0.9	0.7
E: Máximo de generaciones	1000	500

Tabla 0-2 Factores del DDE 2^5 para sintonización del GA

Las variables de respuesta son la media y desviación estándar de la solución en 100 corridas.

Para el caso del PSO, los factores estudiados en la sintonización, se detallan en la Tabla 0-3. Si alguna solución es infactible, la velocidad es cero, y la partícula se reinicia al límite más cercano:

Para $i=1$ hasta m -partículas

Para $j=1$ hasta n -variables

Si $X_{ij} >$ límite superior

$X_{ij} =$ límite superior

$V_{ij} = 0$

De otra manera Si $X_{ij} <$ límite inferior

$X_{ij} =$ límite inferior

$V_{ij} = 0$

Fin

Fin

Fin

Factor	Alto	Bajo
A: Dimensión	9	5
B: Numero de partículas	50	30
C: W	1.4	0.4
D: C1	2	1.5
E: C2	2.5	2
F: Máximo de iteraciones	1000	500

Tabla 0-3 Factores del DDE 2⁶ para sintonización del PSO

Adicionalmente a las variables de respuesta, se mide el tiempo de cómputo de las 100 corridas y el número de éxitos, siendo este último, cuando la solución es $f(x) \leq 0.01$, dado que el mínimo de todas las funciones de prueba (benchmarking) es cero.

3.3 Funciones de prueba.

Seis funciones de prueba se emplearon para sintonizar los algoritmos, continuas para n dimensiones ($n=1,2,3 \dots \infty$), y se describen a continuación.

i. Spherical (Gráfico 0-1)

$$f(x) = \sum_{d=1}^n x_d^2$$

0-1

Dominio (-100,100)

Mínimo global $x^* = (0, \dots, 0), f(x^*) = 0$.

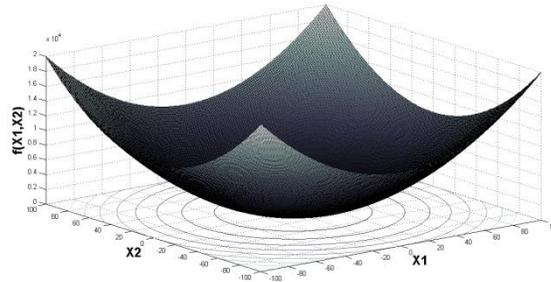


Gráfico 0-1 Función Spherical para n=2 en el cubo de inspección

$[0, 2 \times 10^4]$ por $[-100, 100]$ por $[-100, 100]$

ii. Rosenbrock (Gráfico 0-2)

$$f(x) = \sum_{d=1}^{n-1} 100(x_d^2 - x_{d-1})^2 + (x_d - 1)^2$$

0-2

Dominio (-30,30)

Mínimo global $x^* = (0, \dots, 0), f(x^*) = 0$.

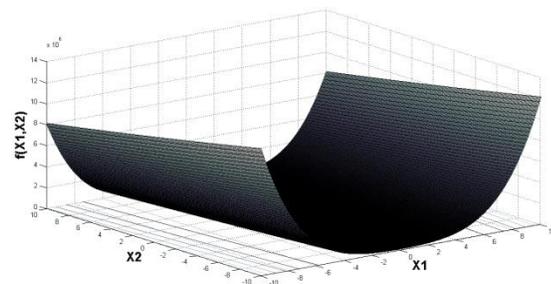


Gráfico 0-2 Función Rosenbrock para n=2 en el cubo de inspección

$[0, 14 \times 10^5]$ por $[-10, 10]$ por $[-10, 10]$

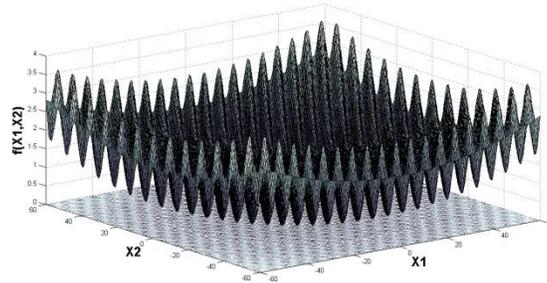
iii. Griewank (Gráfico 0-3)

$$f(x) = \frac{1}{4000} \sum_{d=1}^n x_d^2 - \prod_{d=1}^n \cos\left(\frac{x_d}{\sqrt{d}}\right) + 1$$

0-3

Dominio (-600,600)

Mínimo global $x^* = (0, \dots, 0), f(x^*) = 0$.



**Gráfico 0-3 Función Griewank para n=2 en el cubo de inspección
[0,4] por [-60,60] por [-60,60]**

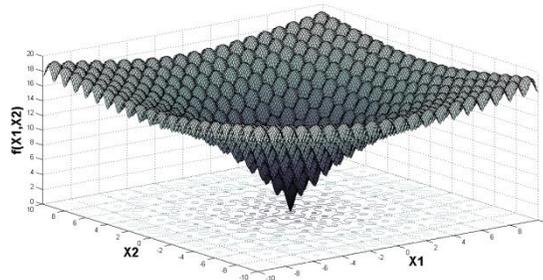
iv. Ackley (Gráfico 0-4)

$$f(x) = 20 + e - 20e\left(-0.2\sqrt{\frac{1}{n}\sum_{d=1}^n x_d^2}\right) - e\left(\frac{1}{n}\sum_{d=1}^n \cos(2\pi x_d)\right)$$

0-4

Dominio (-30,30)

Mínimo global $x^* = (0, \dots, 0), f(x^*) = 0$.



**Gráfico 0-4 Función Ackley para n=2 en el cubo de inspección
[0,20] por [-10,10] por [-10,10]**

v. Rastrigin (Gráfico 0-5)

$$f(x) = \sum_{d=1}^n (x_d^2 - 10 \cos(2\pi x_d) + 10)$$

0-5

Dominio (-5.12,5.12)

Mínimo global $x^* = (0, \dots, 0), f(x^*) = 0$.

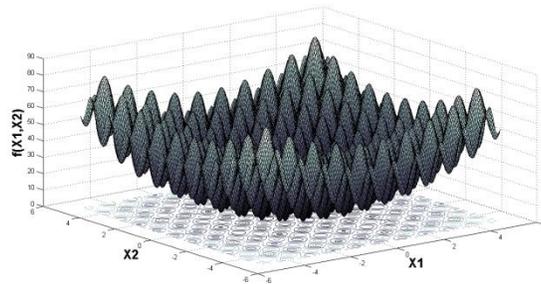


Gráfico 0-5 Función Rastrigin para n=2 en el cubo de inspección
[0,90] por [-6,6] por [-6,6]

vi. Schwefel (Gráfico 0-6)

$$f(x) = 418.9829n \sum_{d=1}^n -x_d \text{sen}(\sqrt{|x_d|})$$

0-6

Dominio (-500,500)

Mínimo global $x^* = (0, \dots, 0), f(x^*) = 420.9687$

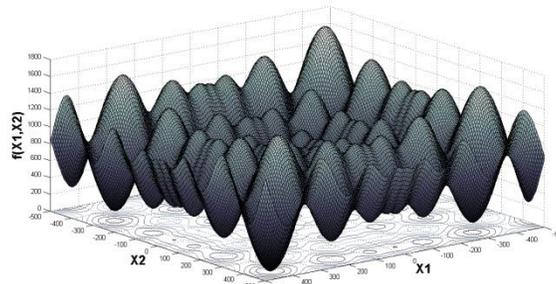


Gráfico 0-6 Función Schwefel para n=2 en el cubo de inspección
[0,1800] por [-500,500] por [-500,500]

3.4 Diseño Anidado para Comparar AG Vs PSO.

El diseño utilizado es el diseño anidado en dos etapas, diferentes problemas de optimización de respuesta múltiple anidados dentro del algoritmo (Tabla 0-4). Los casos de estudio se tomaron de la bibliografía consultada (4; 14; 15; 16; 17). Este diseño permite detectar diferencias entre las soluciones encontradas por los algoritmos. Se consideran cinco casos de estudio, cada uno resuelto por los dos algoritmos.

AG					PSO				
1	2	3	4	5	1	2	3	4	5
Derringer	Del Castillo	Hendriks	Contreras	Wold	Derringer	Del Castillo	Hendriks	Contreras	Wold
Y_{AG11}	Y_{AG21}	Y_{AG31}	Y_{AG41}	Y_{AG51}	Y_{PSO11}	Y_{PSO21}	Y_{PSO31}	Y_{PSO41}	Y_{PSO51}
Y_{AG12}	Y_{AG22}	Y_{AG32}	Y_{AG42}	Y_{AG52}	Y_{PSO12}	Y_{PSO22}	Y_{PSO32}	Y_{PSO42}	Y_{PSO52}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
Y_{AG110}	Y_{AG210}	Y_{AG310}	Y_{AG410}	Y_{AG510}	Y_{PSO110}	Y_{PSO210}	Y_{PSO310}	Y_{PSO410}	Y_{PSO510}

Tabla 0-4 Diseño anidado para comparar AG Vs PSO

La Variable de respuesta es el valor de la deseabilidad del caso estudio.

$$Y_{ijk} = D_{ijk} \quad 0-7$$

$$i = AG, PSO$$

$$j = 1, 2 \dots 5$$

$$k = 1, 2 \dots 10$$

3.4.1 Optimización de Respuesta Múltiple, Casos de Estudio.

Derringer

$$Y_i = \beta_{i0} + \sum_{L=1}^3 \beta_{iL} x_L + \sum_{L=1}^3 \sum_{m=L}^3 \beta_{iLm} x_L x_m; \quad i=1,2,3,4 \quad 0-8), \quad \text{los}$$

coeficientes de los polinomios aparecen en la Tabla 0-6.

$$\hat{Y}_i = \beta_{i0} + \sum_{L=1}^3 \beta_{iL} x_L + \sum_{L=1}^3 \sum_{m=L}^3 \beta_{iLm} x_L x_m; \quad i = 1, 2, 3, 4 \quad 0-8$$

Variable	Descripción	Restricción
X ₁	Sílica hidratada	0.5 ≤ X ₁ ≤ 1.2
X ₂	Silano como agente de acoplamiento	10 ≤ X ₂ ≤ 50
X ₃	Azufre	0.5 ≤ X ₃ ≤ 2.3
Y ₁	Índice de abrasión PICO	120 < Y ₁
Y ₂	Modulo a 200%	1000 < Y ₂
Y ₃	Elongación al rompimiento	400 < Y ₃ < 600
Y ₄	Dureza	60 < Y ₄ < 75

Tabla 0-5 Variables del problema de Derringer

	β_0	β_1	β_2	β_3	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
\hat{Y}_1	139.12	16.49	17.88	10.91	-4.01	-3.45	-1.57	5.13	7.13	7.88
\hat{Y}_2	1261.11	268.15	246.5	139.48	-83.55	-24.79	199.17	69.38	94.13	104.38
\hat{Y}_3	400.38	-99.67	-31.4	-73.92	7.93	17.31	0.43	8.75	6.25	1.25
\hat{Y}_4	68.91	-1.41	4.32	1.63	1.56	0.06	-0.32	-1.63	0.3	-0.25

Tabla 0-6 Coeficientes de regresión para el problema de Derringer

Del Castillo, Montgomery y McCarville (1996), realizan una optimización de respuesta múltiple de un proceso de manufactura de semiconductores, donde se controlan tres factores (X), y optimizan seis variables de respuesta (Y), mostradas en la Tabla 0-7. Los modelos e regresión se ajustaron a partir de un diseño de experimentos Box-Behnken.

Los modelos de regresión para las variables de respuestas están dados por las siguientes ecuaciones:

$$\hat{Y}_1 = 174.93 + 23.38x_2 + 3.6x_3 - 19x_2x_3 \quad 0-9$$

$$\hat{Y}_2 = 141 + 6x_1 + 21.02x_2 + 14.12x_3 \quad 0-10$$

$$\hat{Y}_3 = 139.53 + 7.25x_1 + 16x_2 + 19.75x_3 \quad 0-11$$

$$\hat{Y}_4 = 154.9 + 10.1x_1 + 30.6x_2 + 6.30x_3 - 11.2x_1^2 + 11.3x_1x_2 \quad 0-12$$

$$\hat{Y}_5 = 139.29 + 64.3x_1 + 19.75x_2 + 16.13x_3 - 5.41x_1^2 + 7x_1x_2 \quad 0-13$$

$$\hat{Y}_6 = 146.86 + 4.87x_1 + 15.62x_2 + 27x_3 - 3.98x_1^2 + 4.75x_1x_2 \quad 0-14$$

Variable	Descripción	Restricción
X ₁	Velocidad de flujo	40 ≤ X ₁ ≤ 120
X ₂	Temperatura de flujo	200 ≤ X ₂ ≤ 450
X ₃	Temperatura de bloque	150 ≤ X ₃ ≤ 350
Y ₁	Temperatura máxima en la posición A	185 < Y ₁ < 195, Y ₁ *=190
Y ₂	Temperatura inicial de unión en la posición A	170 < Y ₂ < 195, Y ₂ *=185
Y ₃	Temperatura inicial de unión en la posición A	170 < Y ₃ < 195, Y ₃ *=185
Y ₄	Temperatura máxima en la posición B	185 < Y ₄ < 195, Y ₄ *=190
Y ₅	Temperatura inicial de unión en la posición B	170 < Y ₅ < 195, Y ₅ *=185
Y ₆	Temperatura inicial de unión en la posición B	170 < Y ₆ < 195, Y ₆ *=185

*valor objetivo

Tabla 0-7 Variables del problema de Del Castillo

Hendriks, de Boer, Smilde y Doornbos (1992), utilizaron un diseño de experimentos central compuesto para optimizar simultáneamente tres variables de respuesta para un proceso de recubrimiento sobre gránulos finos.

Los estimadores de las variables de respuestas se ajustaron a un polinomio de segundo orden de la forma:

$$\hat{Y} = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3 + \beta_{33}x_3^2 \quad 0-15$$

Los coeficientes para los modelos de regresión aparecen en la Tabla 0-9.

Variable	Descripción	Restricción
X ₁	Temperatura	-1 ≤ X ₁ ≤ 1*
X ₂	Velocidad de aspersión	-1 ≤ X ₂ ≤ 1*
X ₃	Presión	-1 ≤ X ₃ ≤ 1*
Y ₁	Densidad a granel	0.645 < Y ₁
Y ₂	Tamaño de partícula	105 > Y ₂
Y ₃	% de fármaco	13 < Y ₃

* valor codificado

Tabla 0-8 Variables del problema de Hendricks

β	β_0	β_1	β_2	β_3	β_{12}	β_{13}	β_{23}	β_{33}
\hat{y}_1	0.635	0.004	-0.009	0.016	-0.009	-0.001	0.001	-0.003
\hat{y}_2	104.08	2.02	4.22	-10.64	-1.4	2.3	-3.05	12.44
\hat{y}_3	13.45	-0.02	0.1	-0.18	0.05	0.08	0.05	-0.35

Tabla 0-9 Coeficientes de regresión para el problema de Hendriks.

Contreras, Bruins y Trout (1995), aplican la optimización de respuesta múltiple para un proceso de esterilización con vapor de un producto alimenticio. El diseño de experimentos utilizado fue un central compuesto con cinco factores (Tabla 0-10)

Los modelos de regresión son los siguientes:

$$\hat{Y}_1 = 10.623 + 1.032x_1 + 0.606x_1x_5 + 0.79x_3 + 0.817x_4 - 3.051x_5 + 3.682x_5^2 \quad 0-16$$

$$\hat{Y}_2 = 16.37 - 1.538x_1 - 0.72x_1x_5 + 0.668x_2x_4 - 1.187x_3 - 0.517x_3x_4 - 0.735x_4 + 2.431x_5 - 3.291x_5^2 \quad 0-17$$

Variable	Descripción	Restricción
X ₁	Viscosidad de la salsa	(2000-5000) ≤ X ₁ ≤ (40000-60000)
X ₂	Gas residual	(0-5) ≤ X ₂ ≤ (10-15)
X ₃	Relación sólido/líquido	0.8 ≤ X ₃ ≤ 1.7
X ₄	Peso neto	7.5 ≤ X ₄ ≤ 8.5
X ₅	Velocidad de rotación	0 ≤ X ₅ ≤ 15
Y ₁	Índice de velocidad de calentamiento	7.6 < Y ₁ < 15
Y ₂	Letalidad (microbiológica)	6 < Y ₂ < 11.75

Tabla 0-10 Variables del problema de Contreras

Wold, Carlson, y Skagerberg (1989), estudiaron un proceso químico para reducir la cantidad de subproductos, con cinco factores y cinco variables de respuesta. Utilizando un diseño 2^{5-1} con dos puntos al centro, se ajustan los modelos de regresión correspondientes. Para este caso de estudio se utilizan solo cuatro factores y cuatro variables de respuesta ⁽¹⁸⁾. Los modelos de regresión son:

$$\hat{Y}_1 = 90.889 + 1.8125x_1 + 0.925x_2 - 1.4375x_3 - 2.4875x_1x_2 - 1.15x_1x_4 - 0.9125x_2x_3 - 1.1875x_2x_4 \quad \text{0-18}$$

$$\hat{Y}_2 = 6.4833 + 0.85625x_1 + 1.84375x_2 + 1.91875x_1x_3 + 0.90625x_1x_4 + 0.51875x_1x_3 + 0.70625x_2x_3 \quad \text{0-19}$$

$$\hat{Y}_3 = 2.6722 - 2.61875x_1 - 2.71875x_2 - 1.11875x_4 + 2.49375x_1x_2 + 0.94375x_2x_4 \quad \text{0-20}$$

$$\hat{Y}_4 = 2.7778 + 0.93125x_1 + 1.46875x_2 + 0.98125x_3 + 0.70625x_1x_2 + 0.63125x_2x_3 \quad \text{0-21}$$

Variable	Descripción	Restricción
X ₁	No disponible	-1 ≤ X ₁ ≤ 1*
X ₂	No disponible	-1 ≤ X ₂ ≤ 1*
X ₃	No disponible	-1 ≤ X ₃ ≤ 1*
X ₄	No disponible	-1 ≤ X ₄ ≤ 1*
Y ₁	No disponible	96.2 ≥ Y ₁ ≥ 91
Y ₂	No disponible	2.3 ≤ Y ₂ ≤ 11.5
Y ₃	No disponible	0 ≤ Y ₃ ≤ 6.5
Y ₄	No disponible	0.5 ≤ Y ₃ ≤ 5.5

* valor codificado

Tabla

CAPÍTULO 4.

RESULTADOS

4.1 Sintonización del Algoritmo Genético

Para cada una de las seis funciones de prueba, se corrió un experimento no replicado 2^5 , con 4 puntos centrales. Con excepción de la función Rosenbrock, donde fue necesario añadir puntos estrella. Las variables de respuesta estudiadas son la media de las soluciones

Xy función=1100i=1100yi 0-1) y la desviación

Sy función=199i=1100yi-Xy función2 0-2).

$$1. \quad \bar{X}_{y \text{ función}} = \frac{1}{100} \sum_{i=1}^{100} y_i \quad 0-1$$

$$S_{y \text{ función}} = \sqrt{\frac{1}{99} \sum_{i=1}^{100} (y_i - \bar{X}_{y \text{ función}})^2} \quad 0-2$$

La ANOVA divide la variación de la respuesta, en partes separadas para cada efecto. En esta prueba se mide la significación estadística para cada efecto, comparando la suma de cuadrados medios del error contra un estimado del error experimental. Los efectos con un valor-p menor que 0.05 son significativamente mayores que cero. El estadístico R^2 indica el porcentaje de la variación de la respuesta que se explica con el modelo ajustado. El R^2_{aj} es más adecuado para comparar modelos con diferente número de variables independientes.

El error estándar de estimación, es la desviación estándar de los residuales. Mientras que la media del error absoluto es el promedio de los residuales. El estadístico Durbin-Watson prueba los residuales para determinar si existe alguna correlación significativa en el orden en el que se realizaron las corridas. Un valor-p menor que 0.05 indica la correlación en la secuencia de las corridas.

Cada función se analiza por separado. En dicho análisis se muestran la ANOVA con mayor R^2_{aj} , el grafico de residuales Vs predichos y de probabilidad normal para los residuales, los estadísticos para los modelos de regresión y los coeficientes del modelo. Por último se optimiza el proceso y se validan los resultados.

En los casos donde los residuales se desviaban de la normalidad, se utilizaron transformaciones $(\ln(Y), \frac{1}{\sqrt{Y}}, \sqrt{Y})$ para normalizar los residuales.

4.1.1 Función de prueba Ackley (ecuación 0-1)

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	2.14972	1	2.14972	51.36	0.0000
B:popsize	2.47207	1	2.47207	59.06	0.0000
C:elite	0.0195782	1	0.0195782	0.47	0.4997
D:crossfraction	1.8826	1	1.8826	44.98	0.0000
AC	0.216098	1	0.216098	5.16	0.0310
AD	0.183713	1	0.183713	4.39	0.0453
CD	0.402054	1	0.402054	9.61	0.0044
Total error	1.17204	28	0.0418585		
Total (corr.)	8.49787	35			

Tabla 0-1 ANOVA para $\bar{X}_{y\ ackley}$

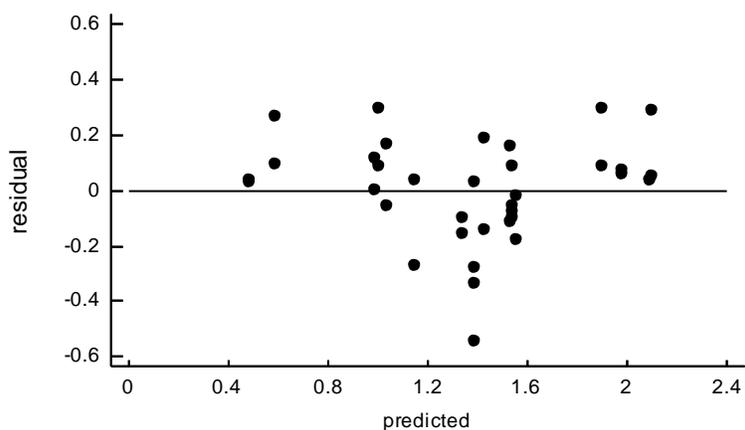


Gráfico 0-1 residuales Vs predichos $\bar{X}_{y\ ackley}$

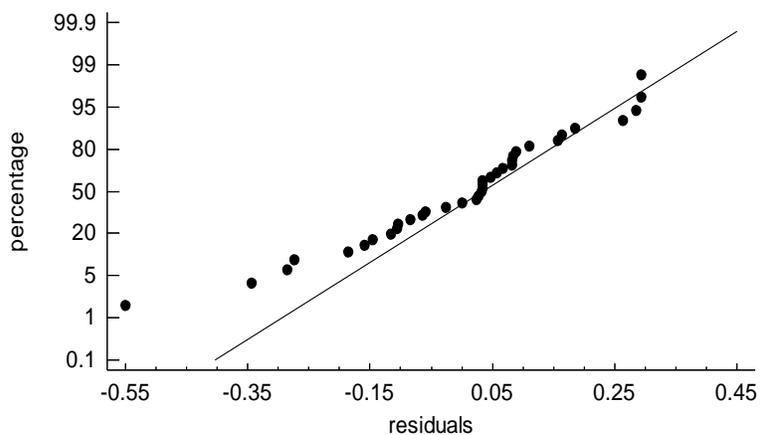


Gráfico 0-2 Probabilidad normal para los residuales de $\bar{X}_{y\ ackley}$

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	0.22459	1	0.22459	49.33	0.0000
B:popsiz	0.00489355	1	0.00489355	1.07	0.3102
C:elite	0.000295774	1	0.000295774	0.06	0.8010
D:crossfraction	0.0334162	1	0.0334162	7.34	0.0122
E:maxit	0.00414049	1	0.00414049	0.91	0.3498
AD	0.0845248	1	0.0845248	18.57	0.0002
CD	0.0108492	1	0.0108492	2.38	0.1357
BCD	0.0483013	1	0.0483013	10.61	0.0033
BDE	0.0138462	1	0.0138462	3.04	0.0940
ACDE	0.0138304	1	0.0138304	3.04	0.0941
BCDE	0.0123367	1	0.0123367	2.71	0.1128
Total error	0.109259	24	0.00455248		
Total (corr.)	0.560284	35			

Tabla 0-2 ANOVA para $S_{y\ ackley}$

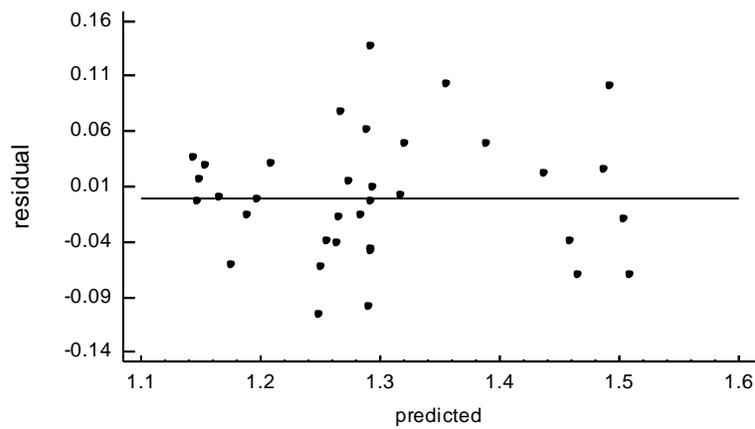


Gráfico 0-3 Residuales Vs predichos $S_{y\ ackley}$

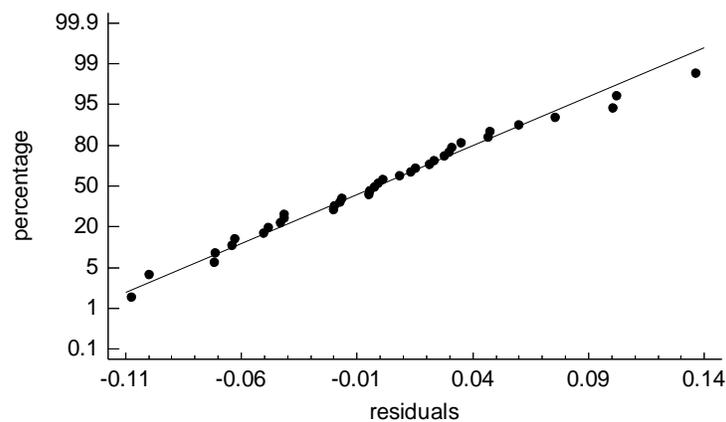


Gráfico 0-4 Probabilidad normal de residuales para $S_{y\ ackley}$

Coefficiente	Estimador
Constante	1.29363
A:dimensión	-0.0837761
B:popsiz	-0.0123662
C:elite	-0.00304022
D:crossfraction	0.0323149
E:maxit	0.011375
AD	-0.0513946
CD	0.0184129
BCD	0.0388512
BDE	0.0208013
ACDE	0.0207894
BCDE	-0.0196347

Coefficiente	Estimador
constante	1.3987
A:dimensión	0.259189
B:popsiz	-0.277943
C:elite	-0.024735
D:crossfraction	0.242551
AC	-0.0821771
AD	-0.0757696
CD	0.11209

Tabla 0-3 Coeficientes de regresión para $\bar{X}_{y\ ackley}$

Tabla 0-4 Coeficientes de regresión para $S_{y\ ackley}$

Estadístico	$\bar{X}_{y\ ackley}$	$S_{y\ ackley}$
R^2	86.2079 %	80.4993 %
R^2_{aj}	82.7598 %	71.5615 %
Error estándar de estimación	0.204594	0.067472
Media del error absoluto	0.137729	0.0433585
Durbin-Watson	1.67053 (P=0.1600)	2.22713 (P=0.7243)

Tabla 0-5 Estadísticos para los modelos de regresión (Ackley)

4.1.2 Función de prueba Griewank (ecuación 0-2)

Fuente de variación	Suma de cuadrados	G.I.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	0.732016	1	0.732016	2.92	0.1023
B:popsiz	0.236975	1	0.236975	0.94	0.3421
C:elite	1.36255	1	1.36255	5.43	0.0298
D:crossfraction	21.4785	1	21.4785	85.64	0.0000
E:maxit	1.39295	1	1.39295	5.55	0.0282
AB	1.95607	1	1.95607	7.80	0.0109
AC	4.97687	1	4.97687	19.84	0.0002
AD	1.09111	1	1.09111	4.35	0.0494
AE	2.03164	1	2.03164	8.10	0.0097
BD	1.85423	1	1.85423	7.39	0.0129
CD	2.10807	1	2.10807	8.41	0.0086
ABE	2.01577	1	2.01577	8.04	0.0099
ACD	2.703	1	2.703	10.78	0.0035
ADE	1.60404	1	1.60404	6.40	0.0195
Total error	5.26668	21	0.250794		
Total (corr.)	50.8104	35			

Tabla 0-6 ANOVA para $\ln(\bar{X}_y \text{ griewank})$

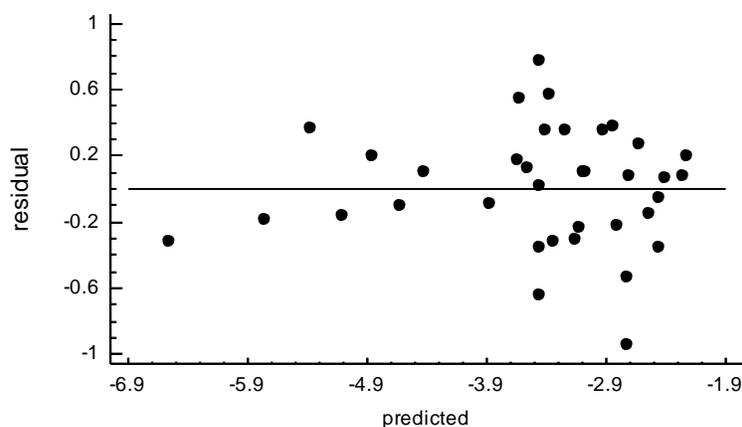


Gráfico 0-5 Residuales Vs predichos para $\ln(\bar{X}_y \text{ griewank})$

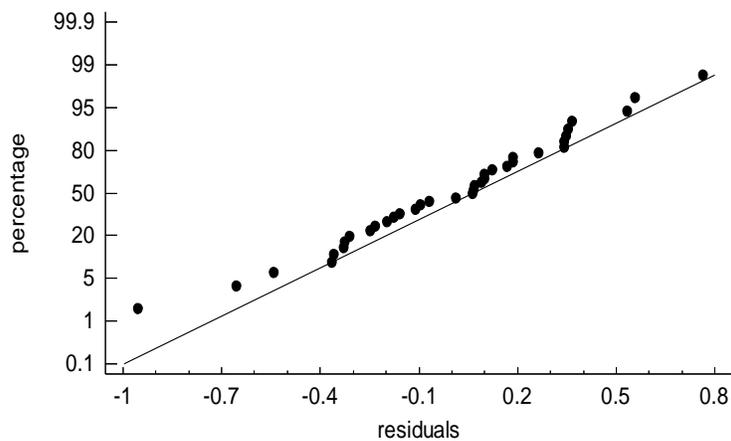


Gráfico 0-6 Probabilidad normal de residuales para $\ln(\bar{X}_y \text{ griewank})$

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	1.41519	1	1.41519	6.68	0.0173
B:popsize	0.940111	1	0.940111	4.44	0.0473
C:elite	1.17237	1	1.17237	5.54	0.0284
D:crossfraction	16.7428	1	16.7428	79.07	0.0000
E:maxit	0.448644	1	0.448644	2.12	0.1603
AC	3.51832	1	3.51832	16.62	0.0005
AD	0.939922	1	0.939922	4.44	0.0473
AE	1.25555	1	1.25555	5.93	0.0239
BD	0.806823	1	0.806823	3.81	0.0644
CD	1.5538	1	1.5538	7.34	0.0131
ABE	2.1612	1	2.1612	10.21	0.0044
ACD	3.29109	1	3.29109	15.54	0.0007
ADE	1.33397	1	1.33397	6.30	0.0203
BDE	1.12458	1	1.12458	5.31	0.0315
Total error	4.44641	21	0.211734		
Total (corr.)	41.1508	35			

Tabla 0-7 ANOVA para $\ln(S_{y \text{ griewank}})$

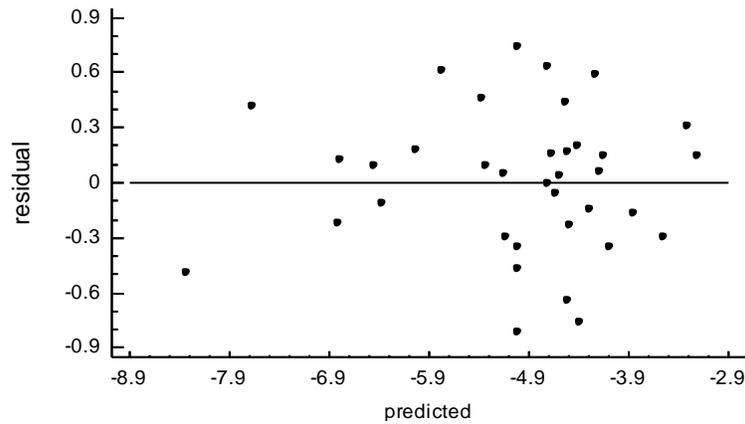


Gráfico 0-7 Residuales Vs predichos para $\ln(S_{y \text{ griewank}})$

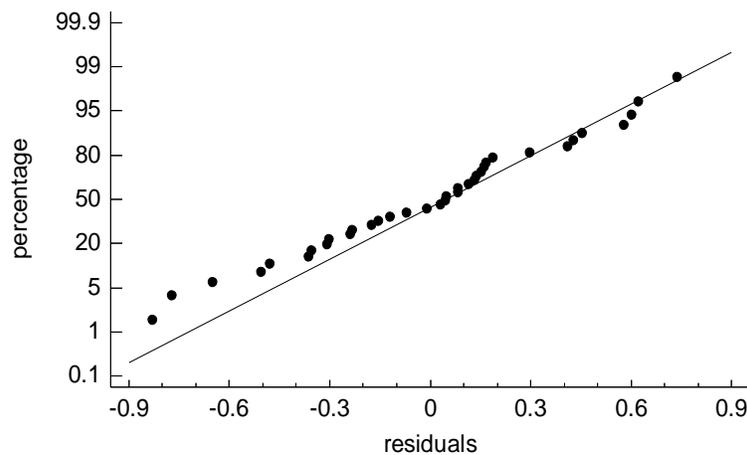


Gráfico 0-8 Probabilidad normal de residuales para $\ln(S_{y \text{ griewank}})$

Coefficiente	Estimador
Constante	-4.99828
A:dimensión	-0.151246
B:popsiz	0.086055
C:elite	-0.206348
D:crossfraction	0.819269
E:maxit	-0.208638
AB	-0.247239
AC	-0.39437
AD	0.184654
AE	-0.25197
BD	-0.240717
CD	0.256666
ABE	0.250984
ACD	0.290635
ADE	0.223889

Tabla 0-9 Coeficientes de regresión para $\ln(\bar{X}_y \text{ griewank})$

Coefficiente	Estimador
constante	-3.45787
A:dimensión	-0.210297
B:popsiz	0.171401
C:elite	-0.191407
D:crossfraction	0.723334
E:maxit	-0.118407
AC	-0.331583
AD	0.171384
AE	-0.19808
BD	-0.158787
CD	0.220355
ABE	0.25988
ACD	0.320697
ADE	0.204173
BDE	0.187465

Tabla 0-8 Coeficientes de regresión para $\ln(S_y \text{ griewank})$

Estadístico	$\ln(\bar{X}_y \text{ griewank})$	$\ln(S_y \text{ griewank})$
R ²	89.6347 %	89.1948 %
R ² _{aj}	82.7244 %	81.9914 %
Error estándar de estimación	0.500793	0.460145
Media del error absoluto	0.306653	0.282202
Durbin-Watson	2.0562 (P=0.6027)	2.03838 (P=0.5393)

Tabla 0-10 Estadísticos para los modelos de regresión (Griewank)

4.1.3 Función Rastrigin (ecuación 0-3)

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	172.597	1	172.597	370.06	0.0000
B:popsiz	51.076	1	51.076	109.51	0.0000
C:elite	0.855344	1	0.855344	1.83	0.1873
D:crossfraction	4.05147	1	4.05147	8.69	0.0067
E:maxit	0.597194	1	0.597194	1.28	0.2681
AB	4.8994	1	4.8994	10.50	0.0033
AD	3.74619	1	3.74619	8.03	0.0088
BC	0.761822	1	0.761822	1.63	0.2125
CD	1.36305	1	1.36305	2.92	0.0993
Total error	12.1263	26	0.466398		
Total (corr.)	252.074	35			

Tabla 0-11 ANOVA para $\bar{X}_{y\text{rastrigin}}$

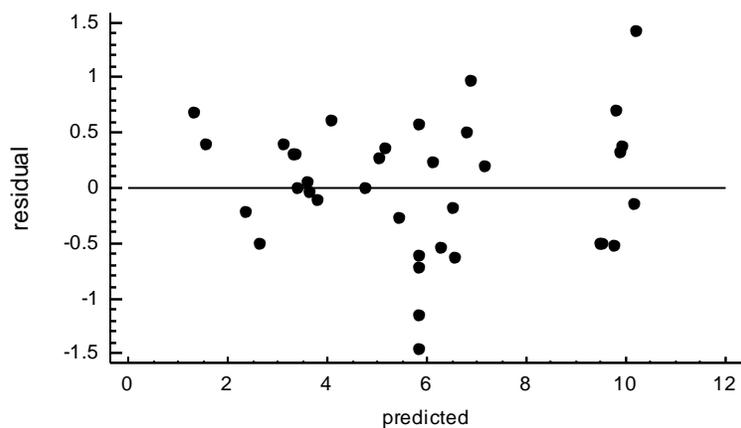


Gráfico 0-9 Residuales Vs predichos para $\bar{X}_{y\text{rastrigin}}$

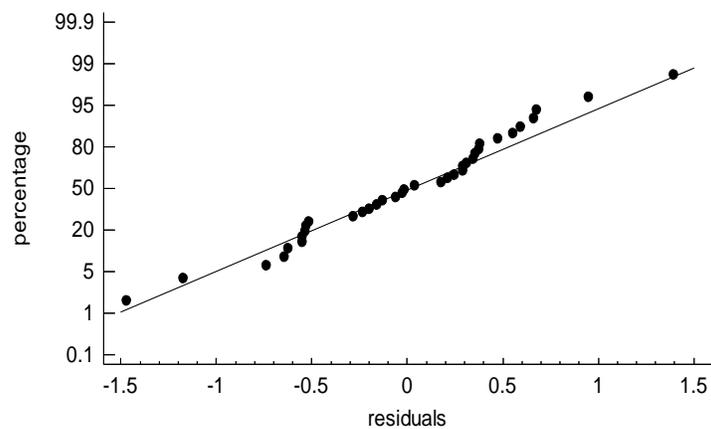


Gráfico 0-10 Probabilidad normal de residuales para $\bar{X}_{y\text{rastrigin}}$

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	67.7236	1	67.7236	394.43	0.0000
B:popsize	14.6604	1	14.6604	85.38	0.0000
C:elite	1.08166	1	1.08166	6.30	0.0181
D:crossfraction	0.00118124	1	0.00118124	0.01	0.9345
E:maxit	0.0725793	1	0.0725793	0.42	0.5209
AD	2.27364	1	2.27364	13.24	0.0011
ACDE	0.952791	1	0.952791	5.55	0.0257
Total error	4.80757	28	0.171699		
Total (corr.)	91.5734	35			

Tabla 0-12 ANOVA para $S_{y\text{rastrigin}}$

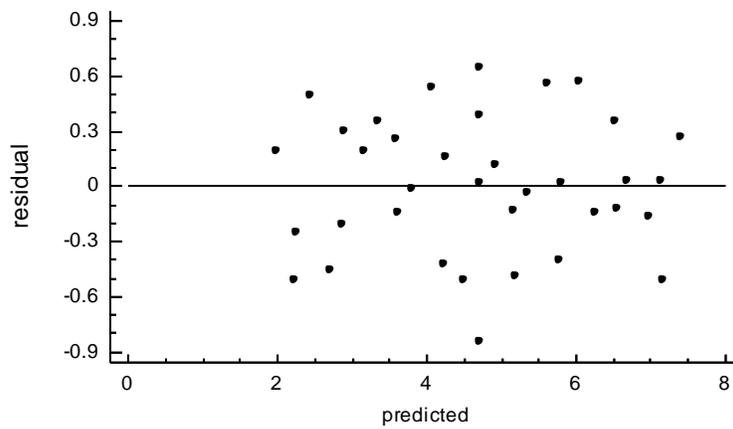


Gráfico 0-11 Residuales Vs predichos para $S_{y\text{rastrigin}}$

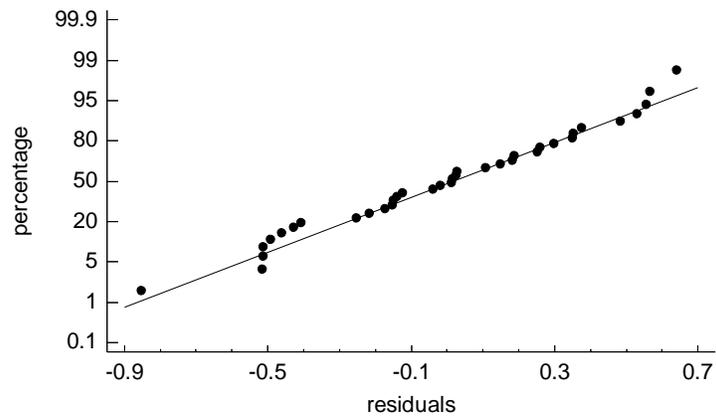


Gráfico 0-12 Probabilidad normal de residuales para $S_{y\text{rastrigin}}$

Coficiente	Estimador
Constante	5.89658
A:dimensión	2.32242
B:popsize	-1.26338
C:elite	-0.163492
D:crossfraction	0.355821
E:maxit	-0.13661
AB	-0.391288
AD	-0.342153
BC	-0.154295
CD	0.206386

Tabla 0-14 Coeficientes de regresión para $\bar{X}_{y\text{rastrigin}}$

Coficiente	Estimador
constante	4.71576
A:dimensión	1.45477
B:popsize	-0.676858
C:elite	-0.183853
D:crossfraction	0.00607566
E:maxit	-0.0476246
AD	-0.266554
ACDE	0.172554

Tabla 0-13 Coeficientes de regresión para

$S_{y\text{rastrigin}}$

Estadístico	$\bar{X}_{y\text{rastrigin}}$	$S_{y\text{rastrigin}}$
R ²	95.1894 %	94.75 %
R ² _{aj}	93.5241 %	93.4376 %
Error estándar de estimación	0.682933	0.414366
Media del error absoluto	0.466193	0.300477
Durbin-Watson	1.52237 (P=0.0943)	1.82991 (P=0.3797)

Tabla 0-15 Estadísticos para los modelos de regresión (Rastrigin)

4.1.4 Función de prueba Rosenbrock (ecuación 0-4)

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	1172.18	1	1172.18	63.13	0.0000
B:popsize	377.529	1	377.529	20.33	0.0001
D:crossfraction	472.158	1	472.158	25.43	0.0000
E:maxit	191.344	1	191.344	10.30	0.0032
AB	127.807	1	127.807	6.88	0.0137
BD	136.191	1	136.191	7.33	0.0112
Total error	538.5	29	18.569		
Total (corr.)	3015.71	35			

Tabla 0-16 ANOVA para $\bar{X}_{y\text{rosenbrock}}$

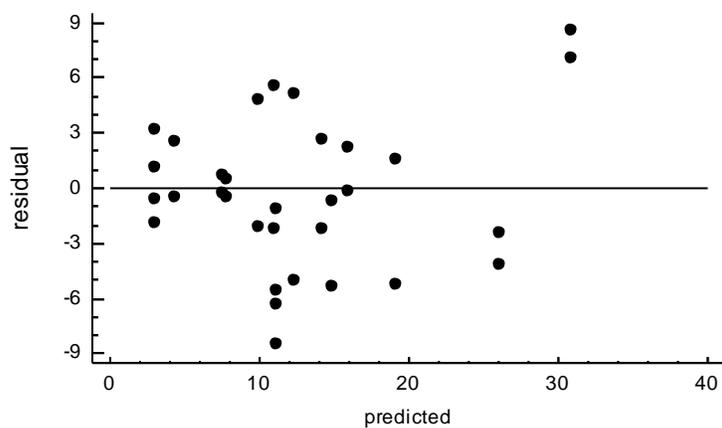


Gráfico 0-13 Residuales Vs predichos para $\bar{X}_{y\text{rosenbrock}}$

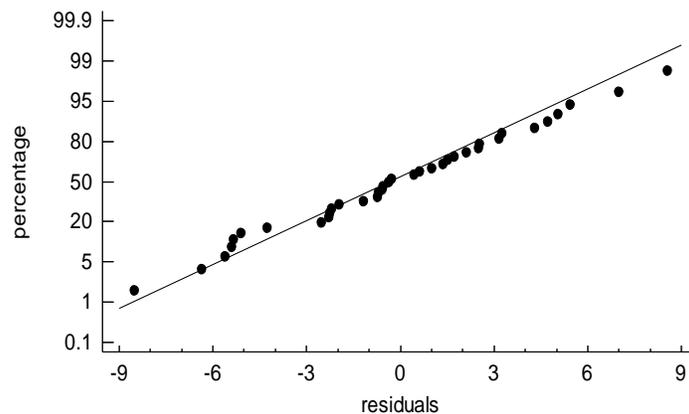


Gráfico 0-14 Probabilidad normal de residuales para $\bar{X}_{y\text{rosenbrock}}$

Fuente de variación	Suma de cuadrados	G.I.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	43.0719	1	43.0719	32.28	0.0000
B:popsize	15.3496	1	15.3496	11.50	0.0022
C:elite	7.74084	1	7.74084	5.80	0.0231
D:crossfraction	25.0458	1	25.0458	18.77	0.0002
E:maxit	16.6991	1	16.6991	12.52	0.0015
BD	5.26527	1	5.26527	3.95	0.0572
BE	6.14147	1	6.14147	4.60	0.0411
CDE	17.8875	1	17.8875	13.41	0.0011
Total error	36.025	27	1.33426		
Total (corr.)	173.226	35			

Tabla 0-17 ANOVA para $\sqrt{S_{y\text{rosenbrock}}}$

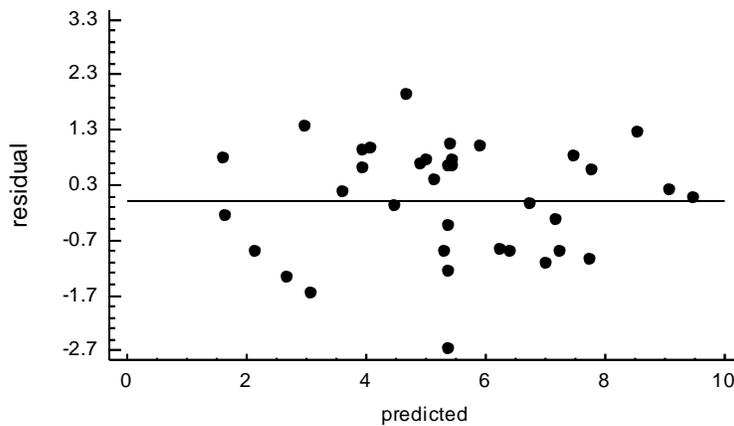


Gráfico 0-15 Residuales Vs predichos para $\sqrt{S_{y\text{rosenbrock}}}$

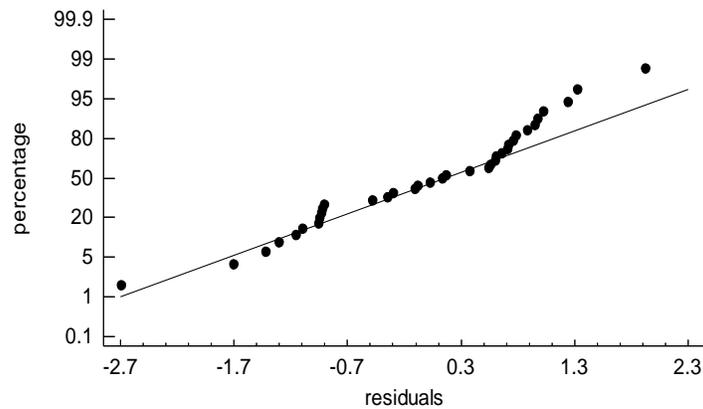


Gráfico 0-16 Probabilidad normal de residuales para $\sqrt{S_{y\text{rosenbrock}}}$

Coefficiente	Estimador
Constante	11.1617
A:dimensión	6.05232
B:popsize	-3.43479
D:crossfraction	3.84122
E:maxit	-2.4453
AB	-1.99849
BD	-2.063

Tabla 0-19 Coeficientes de regresión para

$$\bar{X}_{y \text{ rosenbrock}}$$

Coefficiente	Estimador
Constante	5.41093
A:dimensión	1.16017
B:popsize	-0.692585
C:elite	-0.491834
D:crossfraction	0.884692
E:maxit	-0.722389
BD	-0.405635
BE	-0.438088
CDE	-0.747654

Tabla 0-18 Coeficientes de regresión para

$$\sqrt{S_{y \text{ rosenbrock}}}$$

Estadístico	$\bar{X}_{y \text{ rosenbrock}}$	$\sqrt{S_{y \text{ rosenbrock}}}$
R ²	82.1435 %	79.2035 %
R ² _{aj}	78.449 %	73.0416 %
Error estándar de estimación	4.30917	1.1551
Media del error absoluto	3.09118	0.842819
Durbin-Watson	2.56558 (P=0.9637)	1.46184 (P=0.0634)

Tabla 0-20 Estadísticos para los modelos de regresión (Rosenbrock)

4.1.5 Función de prueba Schwefel (ecuación 0-5)

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	666660.	1	666660.	319.16	0.0000
B:popsiz	214617.	1	214617.	102.75	0.0000
C:elite	13810.6	1	13810.6	6.61	0.0153
E:maxit	7135.2	1	7135.2	3.42	0.0745
AB	46520.3	1	46520.3	22.27	0.0001
Total error	62664.3	30	2088.81		
Total (corr.)	1.01141E6	35			

Tabla 0-21 ANOVA para $\bar{X}_{y\ schwefel}$

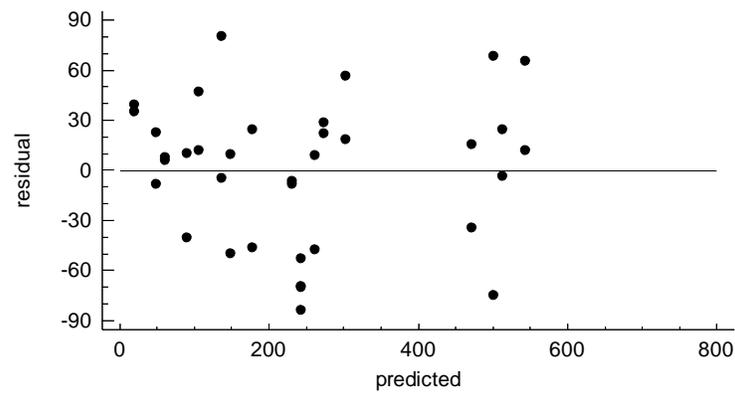


Gráfico 0-17 Residuales Vs predichos para $\bar{X}_{y\ schwefel}$

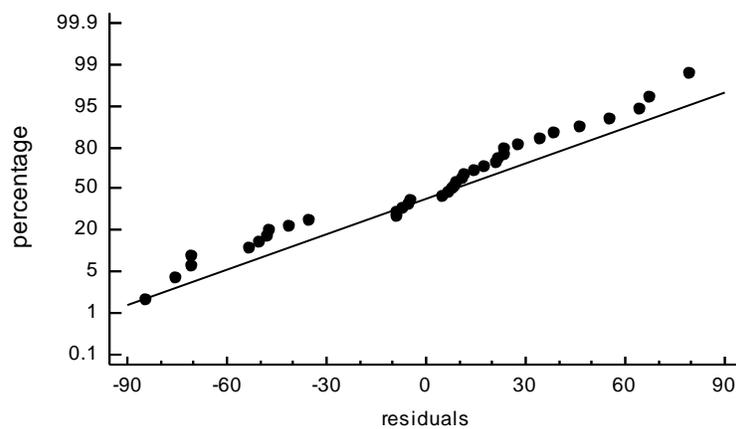


Gráfico 0-18 Probabilidad normal de residuales para $\bar{X}_{y\ schwefel}$

Fuente de variación	Suma de cuadrados	G.L.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	94761.0	1	94761.0	833.39	0.0000
B:popsiz	31733.3	1	31733.3	279.08	0.0000
C:elite	639.189	1	639.189	5.62	0.0251
D:crossfraction	186.201	1	186.201	1.64	0.2115
E:maxit	0.369751	1	0.369751	0.00	0.9549
AB	1277.87	1	1277.87	11.24	0.0024
AD	672.829	1	672.829	5.92	0.0219
CE	514.211	1	514.211	4.52	0.0427
Total error	3070.06	27	113.706		
Total (corr.)	132855.	35			

Tabla 0-22 ANOVA para $S_{y\text{schwefel}}$

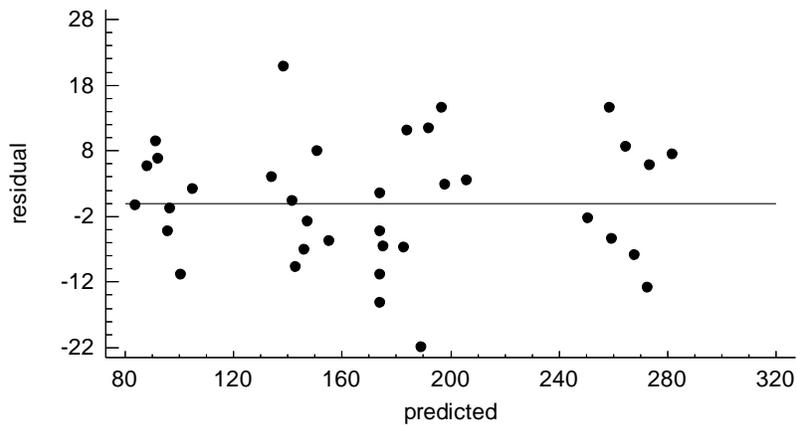


Gráfico 0-19 Residuales Vs predichos para $S_{y\text{schwefel}}$

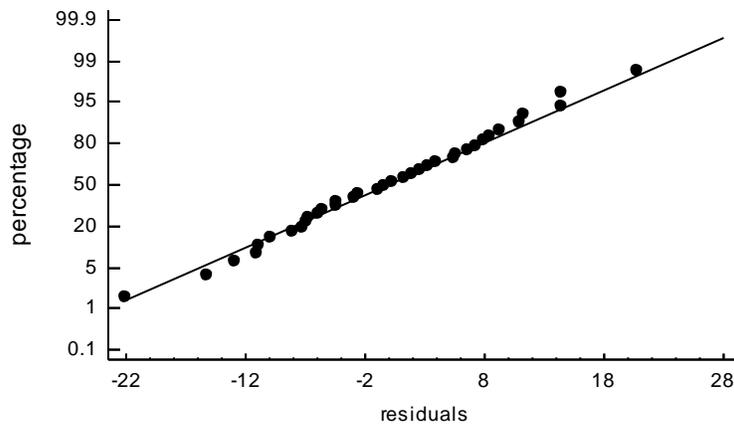


Gráfico 0-20 Probabilidad normal de residuales para $S_{y\text{schwefel}}$

Coefficiente	Estimador
Constante	243.584
A:dimensión	144.337
B:popsize	-81.895
C:elite	-20.7746
E:maxit	-14.9323
AB	-38.1282

Coefficiente	Estimador
Constante	174.099
A:dimensión	54.4177
B:popsize	-31.4907
C:elite	-4.4693
D:crossfraction	-2.41221
E:maxit	-0.107493
AB	-6.3193
AD	-4.5854
CE	4.00863

Tabla 0-24 Coeficientes de regresión para $\bar{X}_{y\text{schwefel}}$

Tabla 0-23 Coeficientes de regresión para $S_{y\text{schwefel}}$

Estadístico	$\bar{X}_{y\text{schwefel}}$	$S_{y\text{schwefel}}$
R ²	97.6892 %	93.8042 %
R ² _{aj}	97.0045 %	92.7716 %
Error estándar de estimación	10.6633	45.7035
Media del error absoluto	7.60956	33.6113
Durbin-Watson	1.62339 (P=0.1566)	1.78076 (P=0.2660)

Tabla 0-25 Estadísticos para los modelos de regresión (Schwefel)

4.1.6 Función de prueba Spherical (ecuación 0-6)

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	0.0255005	1	0.0255005	17.11	0.0005
B:popsiz	0.0173016	1	0.0173016	11.61	0.0027
C:elite	0.0000200376	1	0.0000200376	0.01	0.9088
D:crossfraction	0.0127231	1	0.0127231	8.54	0.0082
E:maxit	0.0205916	1	0.0205916	13.82	0.0013
AB	0.00834047	1	0.00834047	5.60	0.0277
AD	0.00526897	1	0.00526897	3.54	0.0740
AE	0.0110672	1	0.0110672	7.43	0.0127
BD	0.0168153	1	0.0168153	11.28	0.0030
BE	0.00675565	1	0.00675565	4.53	0.0452
DE	0.00812696	1	0.00812696	5.45	0.0295
ABD	0.00800392	1	0.00800392	5.37	0.0307
BCE	0.00640624	1	0.00640624	4.30	0.0506
BDE	0.0130152	1	0.0130152	8.73	0.0076
Total error	0.0312962	21	0.00149029		
Total (corr.)	0.191233	35			

Tabla 0-26 ANOVA para \bar{X}_y spherical

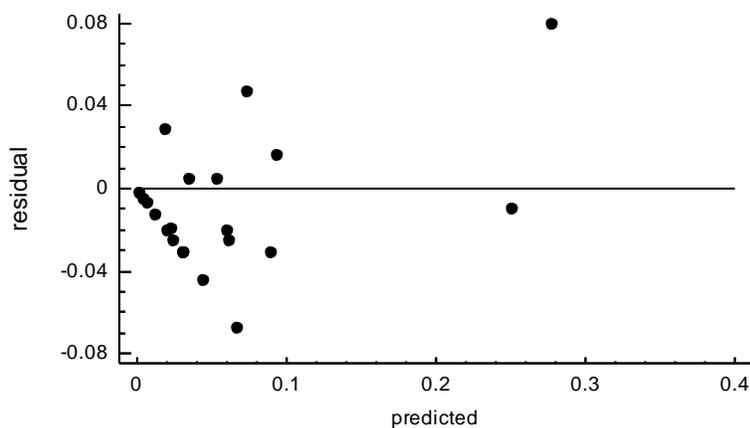


Gráfico 0-21 Residuales Vs predichos para \bar{X}_y spherical

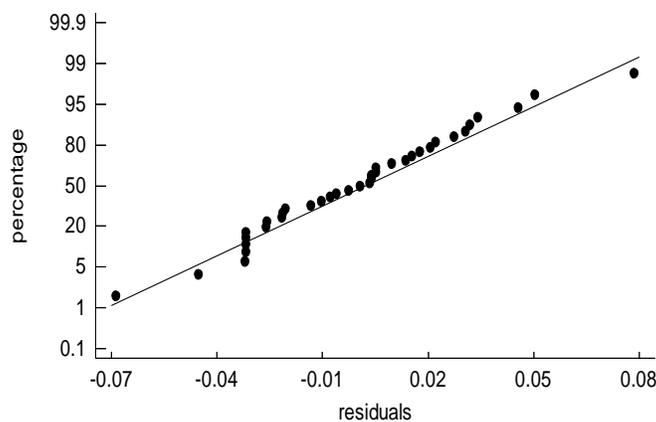


Gráfico 0-22 Probabilidad normal de residuales para \bar{X}_y spherical

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	0.496761	1	0.496761	26.25	0.0000
B:popsize	0.158819	1	0.158819	8.39	0.0081
C:elite	0.0662723	1	0.0662723	3.50	0.0741
D:crossfraction	0.0479696	1	0.0479696	2.53	0.1250
E:maxit	0.256037	1	0.256037	13.53	0.0012
AC	0.219738	1	0.219738	11.61	0.0024
AE	0.0932375	1	0.0932375	4.93	0.0366
BD	0.150712	1	0.150712	7.96	0.0097
BCE	0.213258	1	0.213258	11.27	0.0027
BDE	0.129821	1	0.129821	6.86	0.0153
ABCE	0.0685154	1	0.0685154	3.62	0.0697
ABCDE	0.0860885	1	0.0860885	4.55	0.0438
Total error	0.435318	23	0.0189269		
Total (corr.)	2.42255	35			

Tabla 0-27 ANOVA para S_y spherical

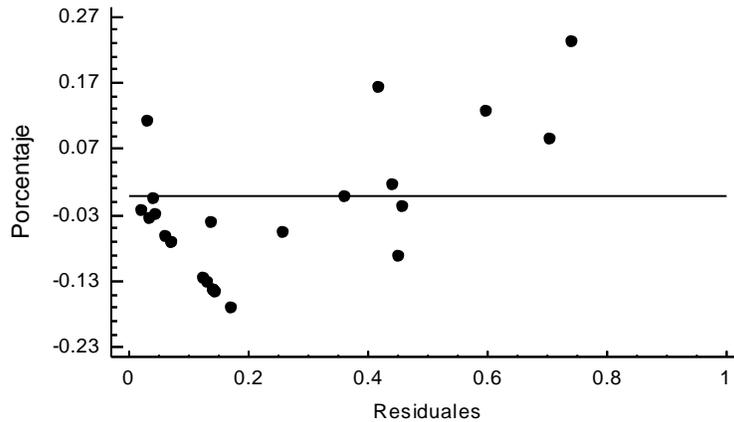


Gráfico 0-23 Residuales Vs predichos para S_y spherical

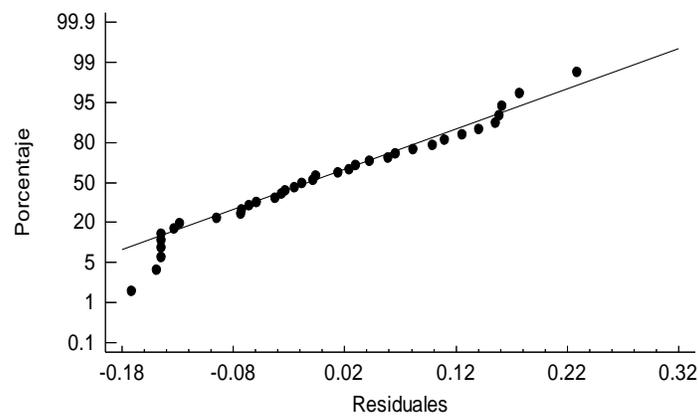


Gráfico 0-24 Probabilidad normal de residuales para S_y spherical

Coefficiente	Estimador
Constante	0.031411
A:dimensión	0.0282293
B:popsiz	-0.0232524
C:elite	-0.000791313
D:crossfraction	0.0199399
E:maxit	-0.0253671
AB	-0.0161443
AD	0.0128318
AE	-0.0185971
BD	-0.0229233
BE	0.0145298
DE	-0.0159364
ABD	-0.0158153
BCE	0.014149
BDE	0.0201674

Tabla 0-29 Coeficientes de regresión para $\bar{X}_{y\ spherical}$

Coefficiente	Estimador
constante	0.144053
A:dimensión	0.124594
B:popsiz	-0.0704492
C:elite	-0.0455083
D:crossfraction	0.0387176
E:maxit	-0.0894491
AC	-0.0828662
AE	-0.0539784
BD	-0.0686275
BCE	0.0816352
BDE	0.0636938
ABCE	0.0462721
ABCDE	-0.0518678

Tabla 0-28 Coeficientes de regresión para $S_{y\ spherical}$

Estadístico	$\bar{X}_{y\ spherical}$	$S_{y\ spherical}$
R ²	83.6345 %	82.0306 %
R ² _{aj}	72.7242 %	72.6552 %
Error estándar de estimación	0.0386043	0.137575
Media del error absoluto	0.0235424	0.0933869
Durbin-Watson	1.60407 (P=0.1213)	2.10899 (P=0.6499)

Tabla 0-30 Estadísticos para los modelos de regresión (Spherical)

4.1.7 Optimización

Dado que el valor de las funciones de prueba es $f(x)=0$, el problema de optimización busca minimizar el valor de la variable de respuesta. Es decir, obtener los parámetros donde la variable de respuesta \bar{X} , sea cercana a la solución conocida (precisión). Y con una desviación estándar de la muestra mínima (exactitud).

Los parámetros óptimos se buscan entre los puntos experimentales del diseño 2^5 , utilizando la función de deseabilidad (sección 2.6.2.1). Los límites utilizados en la ecuación $d_{ij} =$

$$0 \quad Y_{ij} \leq Y_{j^*} \quad Y_{ij} > Y_{j^*} \quad Y_{ij} < Y_{j^*} \quad Y_{ij} \geq Y_{j^*} \quad Y_{ij} < Y_{j^*} \quad Y_{ij} > Y_{j^*} \quad 0$$

$$d_{ij} = \frac{Y_{ij} - Y_{j^*}}{Y_{j^*} - Y_{j^*}} \quad \frac{Y_{j^*} - Y_{ij}}{Y_{j^*} - Y_{j^*}} \quad \frac{Y_{ij} - Y_{j^*}}{Y_{j^*} - Y_{j^*}} \quad \frac{Y_{j^*} - Y_{ij}}{Y_{j^*} - Y_{j^*}} \quad 0$$

2-41, son los valores mínimos y máximos que toman los estimadores de las variables de respuesta (modelos de regresión), dentro de la región experimental. La deseabilidad para cada punto experimental (Tabla 0-67) se calcula a partir de los estimados para cada respuesta.

Variable de respuesta	Mínimo Y_{j^*}	Máximo Y_j^*
\bar{X}_{ackley}	0.4886	2.1078
$\bar{X}_{griewank}$	0.00024445	0.040657
$\bar{X}_{rastrigin}$	1.3433	10.221
$\bar{X}_{rosenbrock}$	-1.8038	30.997
$\bar{X}_{schwefel}$	19.773	543.65
$\bar{X}_{spherical}$	-0.047838	0.2785
S_{ackley}	1.0932	1.5112
$S_{griewank}$	0.0014298	0.10988
$S_{rastrigin}$	2.0027	7.4166
$S_{rosenbrock}$	2.6806	90.425
$S_{schwefel}$	83.966	281.91
$S_{spherical}$	-0.17727	0.74338

Tabla 0-31 límites para el cálculo de la función de deseabilidad (GA)

	Número de variables	Tamaño de población	elites	Fracción de cruza	Máximo de iteraciones	D _g
1	-1	-1	-1	-1	-1	0.75421
2	1	-1	-1	-1	-1	0
3	-1	1	-1	-1	-1	0.75881
4	1	1	-1	-1	-1	0
5	-1	-1	1	-1	-1	0.78855
6	1	-1	1	-1	-1	0.28012
7	-1	1	1	-1	-1	0.8527
8	1	1	1	-1	-1	0.62351
9	-1	-1	-1	1	-1	0.41167
10	1	-1	-1	1	-1	0
11	-1	1	-1	1	-1	0.66551
12	1	1	-1	1	-1	0.48561
13	-1	-1	1	1	-1	0.41799
14	1	-1	1	1	-1	0
15	-1	1	1	1	-1	0
16	1	1	1	1	-1	0.53704
17	-1	-1	-1	-1	1	0.74928
18	1	-1	-1	-1	1	0
19	-1	1	-1	-1	1	0.87806
20	1	1	-1	-1	1	0.5835
21	-1	-1	1	-1	1	0.73247
22	1	-1	1	-1	1	0.3546
23	-1	1	1	-1	1	0.88021
24	1	1	1	-1	1	0.67963
25	-1	-1	-1	1	1	0.48222
26	1	-1	-1	1	1	0
27	-1	1	-1	1	1	0.67738
28	1	1	-1	1	1	0.5218
29	-1	-1	1	1	1	0.52248
30	1	-1	1	1	1	0
31	-1	1	1	1	1	0.59741
32	1	1	1	1	1	0.53102

Tabla 0-32 Deseabilidad para cada punto experimental (GA)

Validación.

Con los niveles óptimos de los factores del experimento de sintonización del GA, se resolvió cada función hasta formar dos grupos de 10 series datos (medias y desviaciones estándar) para nueve y cinco variables, utilizando los límites de la Tabla 0-31 se calcula la deseabilidad, mostrada en la Tabla 0-33.

i	Dga5 (5 dimensiones)	Dga9 (9 dimensiones)
1	0.90612	0.7306
2	0.7934	0.59257
3	0.87233	0.68083
4	0.83147	0.7065
5	0.9254	0.64011
6	0.85823	0.66514
7	0.73554	0.60325
8	0.94262	0.57122
9	0.91618	0.67109
10	0.78633	0.7079

Tabla 0-33 deseabilidades para prueba de hipótesis

El primer paso de la validación es verificar que los datos provienen de una distribución normal. La prueba de Shapiro-Wilk (Tabla 0-34) indica que no se rechaza la hipótesis de que los datos Dga5 y Dga9 provienen de una distribución normal, con un nivel de confianza del 95%. Los gráficos Gráfico 0-25 y Gráfico 0-26 apoyan el supuesto de normalidad.

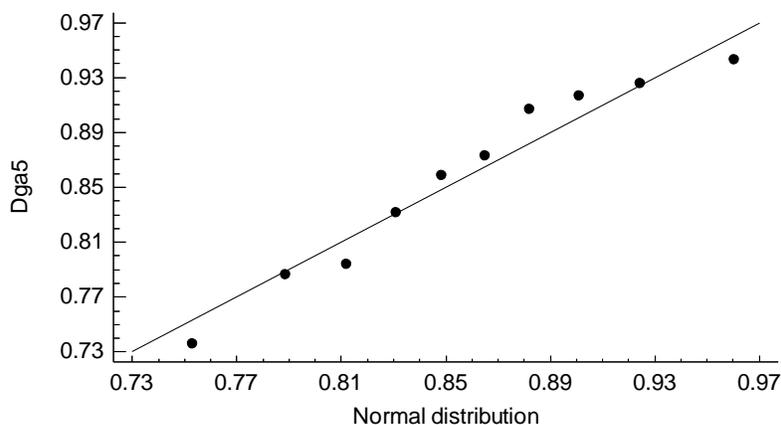


Gráfico 0-25 Cuantil-cuantil para Dag5 Vs distribución normal

Existe evidencia de que los datos pueden representarse por una distribución normal. Esto permite el uso del intervalo de confianza para la media y desviación. Que es el intervalo que contiene a la media poblacional.

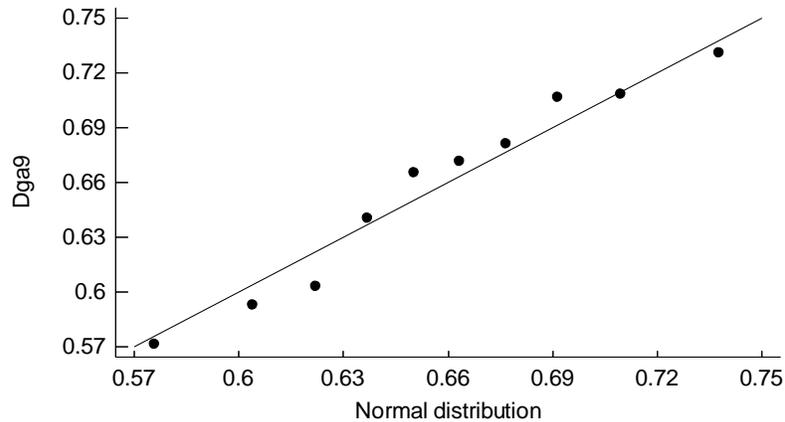


Gráfico 0-26 Cuantil-cuantil para Dag9 Vs distribución normal

	<i>Estadístico</i>	<i>Valor-P</i>
5 variables	0.947503	0.623095
9 variables	0.949174	0.643032

0-34 Prueba de Shapiro-Wilk W

	Media	Desviación estándar	Intervalo de confianza para la media [LIC,LSC]	Intervalo de confianza para la desviación estándar [LIC,LSC]
Dga5	0.856762	0.0688023	[0.807544; 0.90598]	[0.0473247; 0.125606]
Dga9	0.656921	0.0537219	[0.618491; 0.695351]	[0.0369518; 0.098075]

Tabla 0-35 intervalos de confianza del 95% para la media y desviación estándar

La prueba t, se realiza tomando como μ_0 la deseabilidad encontrada en la optimización, 0.88021 para Dga5 y 0.67963 para Dga9.

Para Dga5 se tiene:

$$H_0: \mu_0 = 0.88021$$

$$H_1: \mu_0 \neq 0.88021$$

$$t_{calculada} = -1.07771$$

$$\text{Valor-P} = 0.3092$$

No se rechaza la hipótesis nula con 95% de confianza

Para Dga9:

$$H_0: \mu_0 = 0.67963$$

$$H_1: \mu_0 \neq 0.67963$$

$$t_{calculada} = 1.54716$$

$$\text{Valor-P} = 0.156229$$

No se rechaza la hipótesis nula con 95% de confianza.

Ya sea que se maneje el número de variables en un nivel alto o bajo, los mejores parámetros para el GA son:

Tamaño de la población=50.

Elite=5.

Fracción de cruza=0.7.

Máximo de iteraciones=1000.

4.2 Sintonización de PSO

Para este análisis, se utilizó el valor de la función $f(x)$ para evaluar las variables de respuesta, la unidad experimental es similar, con cien datos por tratamiento

$$\bar{X}_{y \text{ funcion}} = \frac{1}{100} \sum_{i=1}^{100} y_i \quad 0-3$$

$$S_{x \text{ funcion}} = \sqrt{\frac{1}{99} \sum_{i=1}^k (y_i - \bar{X}_{y \text{ funcion}})^2} \quad 0-4$$

Las variables de respuesta son, la media de $f(x)$ para 100 datos - $\bar{X}_{x \text{ funcion}}$ - y la desviación

$S_{x \text{ funcion}}$ - (ecuaciones **$Xy \text{ funcion} = 1100i = 1100yi$**

$Sx \text{ funcion} = 199i = 1kyi - Xy \text{ funcion}^2$ 0-4).

El análisis realizado es como se describió en la sección 4.1, se presenta la tabla ANOVA, los gráficos de residuales contra Predichos y de probabilidad normal, los coeficientes de regresión y los estadísticos de los modelos de regresión ajustados. Seguido de la optimización y validación de los resultados.

En los casos donde los residuales se desviaban de la normalidad, se utilizaron transformaciones $(\ln(Y), \frac{1}{\sqrt{Y}}, \sqrt{Y})$ para normalizar los residuales.

4.2.1 Función de prueba Ackley (ecuación 0-1)

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	778.502	1	778.502	12.56	0.0008
B:floksize	390.135	1	390.135	6.30	0.0149
C:w	7747.39	1	7747.39	125.04	0.0000
D:c1	62.4715	1	62.4715	1.01	0.3195
E:c2	382.555	1	382.555	6.17	0.0159
AC	892.031	1	892.031	14.40	0.0004
BC	393.132	1	393.132	6.35	0.0146
DE	476.6	1	476.6	7.69	0.0074
Total error	3593.63	58	61.9591		
Total (corr.)	14815.1	66			

Tabla 0-36 ANOVA para $\ln(\bar{X}_{x\text{ackley}})$

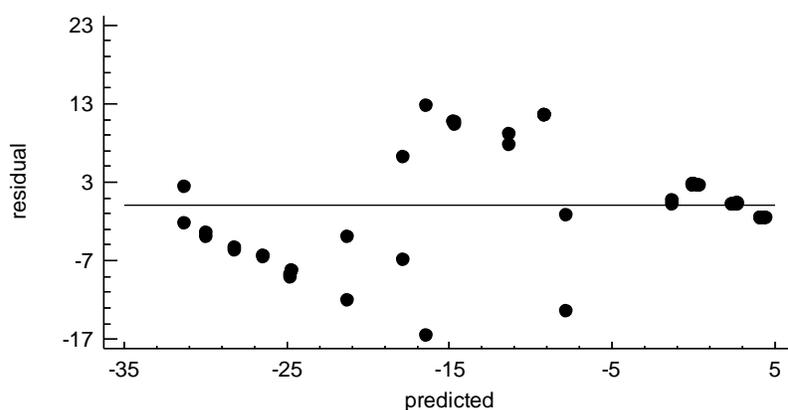


Gráfico 0-27 Residuales Vs predichos para $\ln(\bar{X}_{x\text{ackley}})$

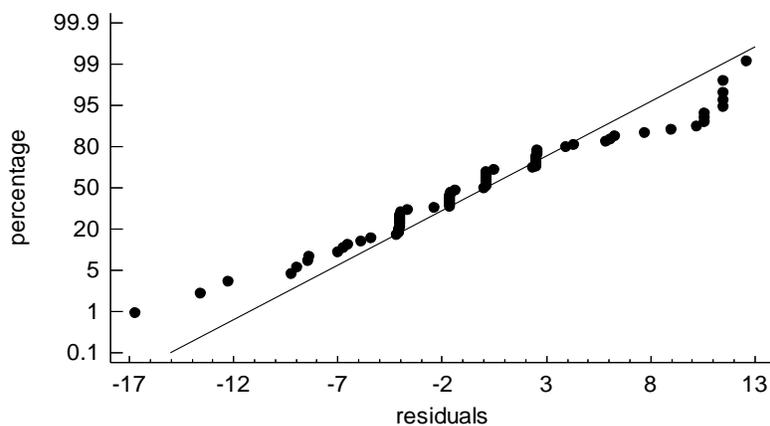


Gráfico 0-28 Probabilidad normal de residuales para $\ln(\bar{X}_{x\text{ackley}})$

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	778.502	1	778.502	12.56	0.0008
B:floksize	390.135	1	390.135	6.30	0.0149
C:w	7747.39	1	7747.39	125.04	0.0000
D:c1	62.4715	1	62.4715	1.01	0.3195
E:c2	382.555	1	382.555	6.17	0.0159
AC	892.031	1	892.031	14.40	0.0004
BC	393.132	1	393.132	6.35	0.0146
DE	476.6	1	476.6	7.69	0.0074
Total error	3593.63	58	61.9591		
Total (corr.)	14815.1	66			

Tabla 0-37 ANOVA para $\ln(S_{yackley})$

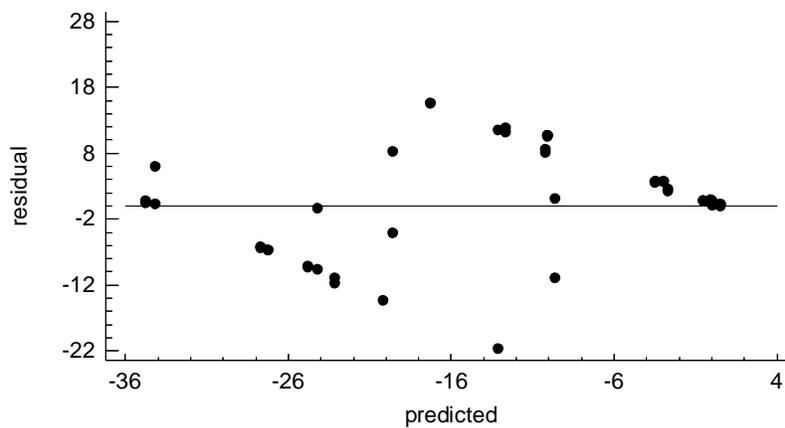


Gráfico 0-29 residuales Vs predichos para $\ln(S_{yackley})$

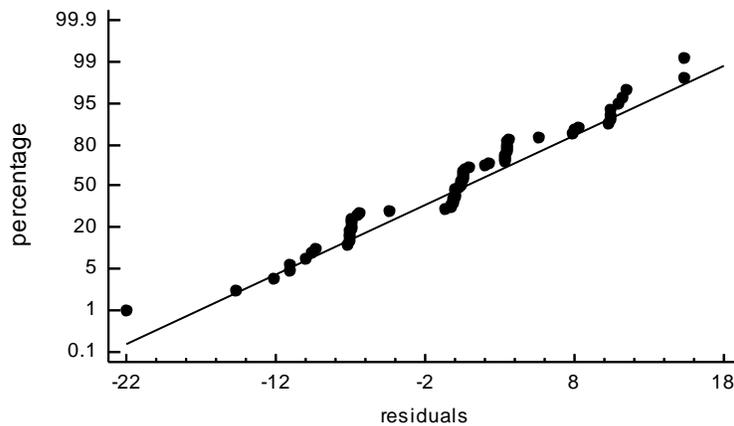


Gráfico 0-30 Probabilidad normal de residuales para $\ln(S_{yackley})$

<i>Coficiente</i>	<i>Estimador</i>
Constante	-10.0343
A:dimensión	3.51924
B:floksize	-2.4913
C:w	11.1019
D:c1	-0.996918
E:c2	-2.46698
AC	-3.76711
BC	2.50085
DE	2.75357

Tabla 0-38 coeficientes de regresión para $\ln(S_{y\ ackley})$

<i>Coficiente</i>	<i>Estimador</i>
Constante	-9.15938
A:dimensión	3.44808
B:floksize	-2.5162
C:w	12.593
D:c1	-1.19975
E:c2	-2.05032
AC	-3.30119
BC	2.49034
DE	2.75473
CDE	-2.75025

Tabla 0-39 coeficientes de regresión para $\ln(\bar{X}_{y\ ackley})$

<i>Estadístico</i>	$\ln(\bar{X}_{y\ ackley})$	$\ln(S_{y\ ackley})$
R ²	83.1006 %	75.7435 %
R ² _{aj}	80.4782 %	72.3978 %
Error estándar de estimación	6.94104	7.87141
Media del error absoluto	4.96493	5.53017
Durbin-Watson	1.53276 (P=0.0223)	1.35994 (P=0.0032)

Tabla 0-40 Estadísticos para los modelos de regresión (Ackley)

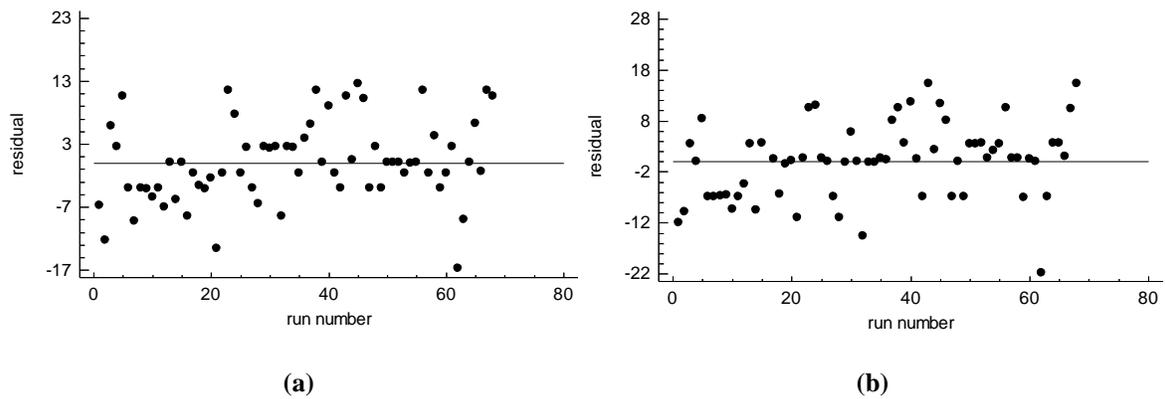


Gráfico 0-31 residuales Vs orden de corrida para $\ln(\bar{X}_{y\ ackley})$ (a) y $\ln(S_{y\ ackley})$ (b)

4.2.2 Función de prueba Griewank (ecuación 0-2)

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	28.704	1	28.704	966.45	0.0000
B:floksize	0.810775	1	0.810775	27.30	0.0000
C:w	608.759	1	608.759	20496.64	0.0000
D:c1	0.00739817	1	0.00739817	0.25	0.6197
E:c2	0.239062	1	0.239062	8.05	0.0064
F:maxit	0.239463	1	0.239463	8.06	0.0063
AC	2.51391	1	2.51391	84.64	0.0000
DE	0.666372	1	0.666372	22.44	0.0000
Total error	1.63352	55	0.0297004		
Total (corr.)	643.574	63			

Tabla 0-41 ANOVA para $\ln(\bar{X}_{y \text{ griewnk}})$

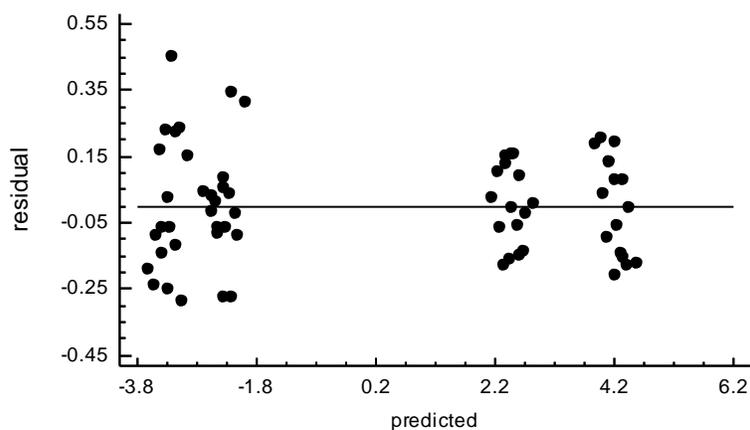


Gráfico 0-32 Residuales Vs predichos para $\ln(\bar{X}_{y \text{ griewnk}})$

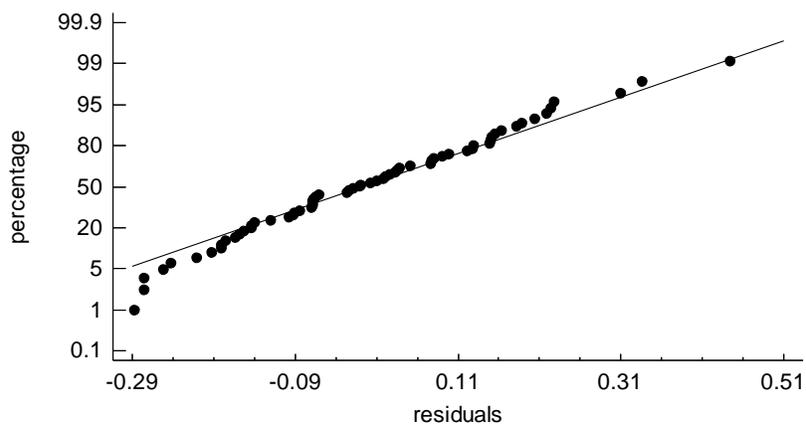


Gráfico 0-33 Probabilidad normal de residuales para $\ln(\bar{X}_{y \text{ griewnk}})$

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	16.5464	1	16.5464	374.77	0.0000
B:floksize	0.75532	1	0.75532	17.11	0.0001
C:w	530.091	1	530.091	12006.42	0.0000
D:c1	0.00630113	1	0.00630113	0.14	0.7070
E:c2	0.156387	1	0.156387	3.54	0.0651
F:maxit	0.24249	1	0.24249	5.49	0.0227
AC	2.38253	1	2.38253	53.96	0.0000
DE	0.669234	1	0.669234	15.16	0.0003
Total error	2.42829	55	0.0441507		
Total (corr.)	553.278	63			

Tabla 0-42 ANOVA para $\ln(S_{y\ griewnk})$

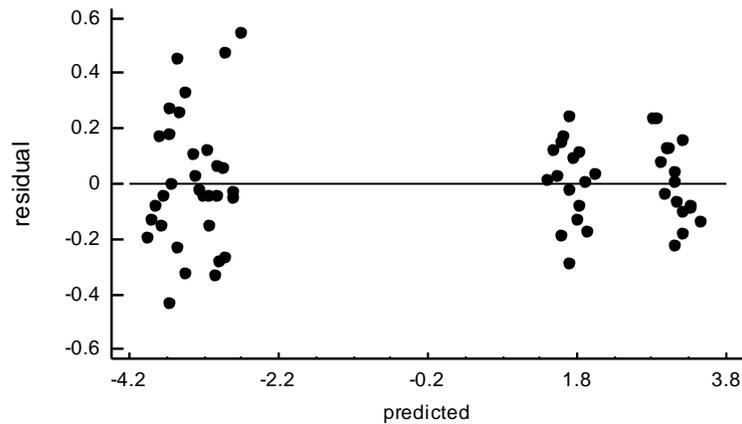


Gráfico 0-34 Residuales Vs predichos para $\ln(S_{y\ griewnk})$

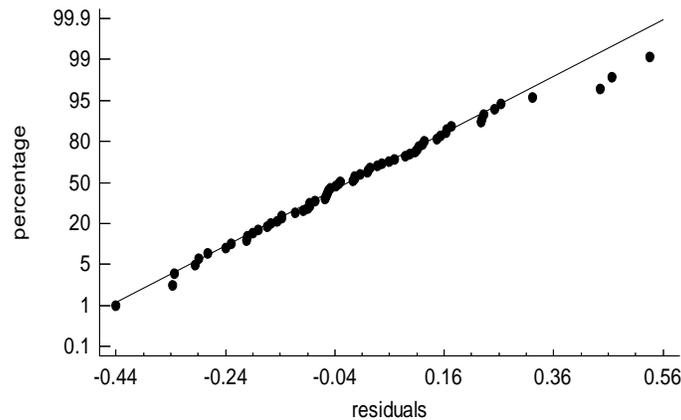


Gráfico 0-35 Probabilidad normal de residuales para $\ln(S_{y\ griewnk})$

<i>Coficiente</i>	<i>Estimador</i>
Constante	-0.440732
A:dimensión	0.508466
B:floksize	-0.108636
C:w	2.87796
D:c1	0.00992246
E:c2	0.0494323
F:maxit	-0.0615541
AC	0.192943
DE	0.102258

Tabla 0-43 Coeficientes de regresión
para $\ln(\bar{X}_{y\ griewnk})$

<i>Coficiente</i>	<i>Estimador</i>
Constante	0.288224
A:dimensión	0.669701
B:floksize	-0.112554
C:w	3.08413
D:c1	0.0107516
E:c2	0.0611175
F:maxit	-0.0611687
AC	0.198192
DE	0.10204

Tabla 0-44 Coeficientes de regresión
para $\ln(S_{y\ griewnk})$

Estadístico	$\ln(\bar{X}_{y\ griewnk})$	$\ln(S_{y\ griewnk})$
R ²	99.7462 %	99.5611 %
R ² _{aj}	99.7093 %	99.4973 %
Error estándar de estimación	0.172338	0.210121
Media del error absoluto	0.130829	0.152176
Durbin-Watson	2.0813 (P=0.6442)	2.06799 (P=0.6243)

Tabla 0-45 Estadísticos para los modelos de regresión (Griewank)

4.2.3 Función de prueba Rastrigin (ecuación 0-3)

Fuente de variación	Suma de cuadrados	G.I.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	15183.2	1	15183.2	5819.66	0.0000
B:floksize	108.171	1	108.171	41.46	0.0000
C:w	45800.8	1	45800.8	17555.26	0.0000
D:c1	0.0271006	1	0.0271006	0.01	0.9192
E:c2	13.0097	1	13.0097	4.99	0.0296
AC	10960.2	1	10960.2	4201.02	0.0000
CD	42.6026	1	42.6026	16.33	0.0002
CE	85.5766	1	85.5766	32.80	0.0000
Total error	143.492	55	2.60895		
Total (corr.)	72337.1	63			

Tabla 0-46 ANOVA para $\bar{X}_{y\text{rastrigin}}$

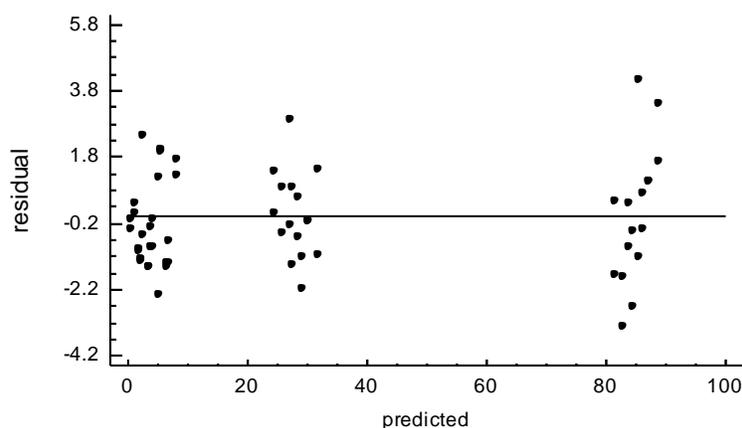


Gráfico 0-36 Residuales Vs predichos para $\bar{X}_{y\text{rastrigin}}$

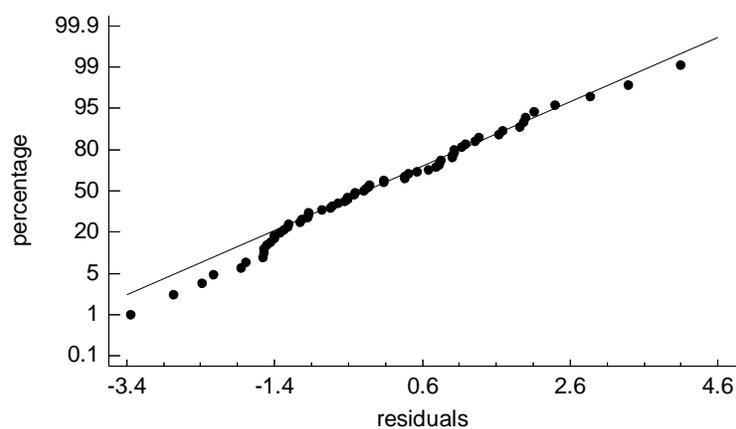


Gráfico 0-37 Probabilidad normal de residuales para $\bar{X}_{y\text{rastrigin}}$

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	201.205	1	201.205	372.96	0.0000
B:floksize	9.10858	1	9.10858	16.88	0.0001
C:w	658.577	1	658.577	1220.77	0.0000
D:c1	1.04195	1	1.04195	1.93	0.1700
AC	19.0655	1	19.0655	35.34	0.0000
CD	9.88745	1	9.88745	18.33	0.0001
Total error	30.7501	57	0.539475		
Total (corr.)	929.635	63			

Tabla 0-47 ANOVA para $S_{y\text{rastirgin}}$

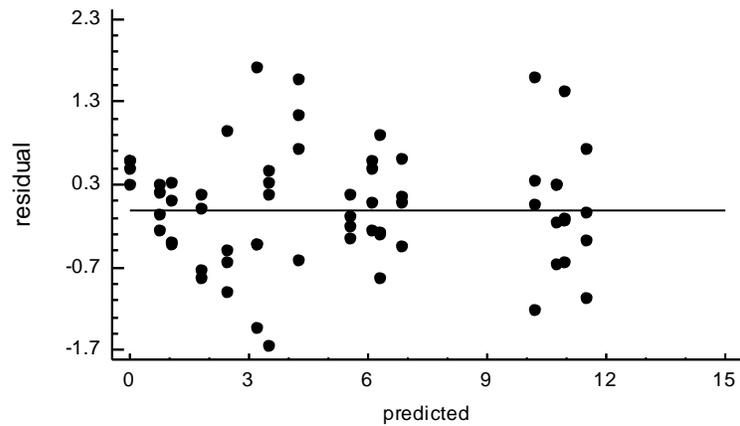


Gráfico 0-38 Residuales Vs predichos para $S_{y\text{rastirgin}}$

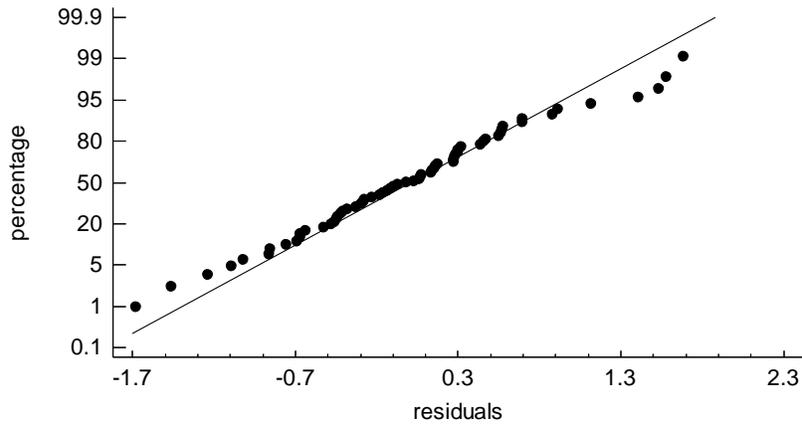


Gráfico 0-39 Probabilidad normal de residuales para $S_{y\text{rastirgin}}$

<i>Coficiente</i>	<i>Estimador</i>
Constante	30.0098
A:dimensión	15.4025
B:floksize	-1.30006
C:w	26.7514
D:c1	-0.0205778
E:c2	0.450861
AC	13.0864
CD	0.815884
CE	1.15635

Tabla 0-49 Coeficientes de regresión para

$\bar{X}_{y\text{rastrigin}}$

<i>Coficiente</i>	<i>Estimador</i>
Constante	5.36835
A:dimensión	1.77309
B:floksize	-0.377255
C:w	3.20784
D:c1	-0.127595
AC	0.5458
CD	0.393054

Tabla 0-48 Coeficientes de regresión para

$S_{y\text{rastrigin}}$

Estadístico	$\bar{X}_{y\text{rastrigin}}$	$S_{y\text{rastrigin}}$
R ²	99.8016 %	96.6922 %
R ² _{aj}	99.7728 %	96.3441 %
Error estándar de estimación	1.61522	0.73449
Media del error absoluto	1.22955	0.5393
Durbin-Watson	1.66902 (P=0.0682)	1.94139 (P=0.3841)

Tabla 0-50 Estadísticos para los modelos de regresión (Rastrigin)

4.2.4 Función de prueba Rosenbrock (ecuación 0-4)

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	266974.	1	266974.	2749.64	0.0000
B:floksize	10813.6	1	10813.6	111.37	0.0000
C:w	138415.	1	138415.	1425.57	0.0000
D:c1	3897.64	1	3897.64	40.14	0.0000
E:c2	29103.6	1	29103.6	299.75	0.0000
AB	4675.79	1	4675.79	48.16	0.0000
AC	36374.1	1	36374.1	374.63	0.0000
AD	2716.99	1	2716.99	27.98	0.0000
AE	20276.6	1	20276.6	208.83	0.0000
BC	1048.42	1	1048.42	10.80	0.0018
CD	4915.37	1	4915.37	50.62	0.0000
CE	4925.14	1	4925.14	50.73	0.0000
ACD	3853.76	1	3853.76	39.69	0.0000
ACE	5271.0	1	5271.0	54.29	0.0000
CDE	572.133	1	572.133	5.89	0.0187
Total error	5048.89	52	97.094		
Total (corr.)	538881.	67			

Tabla 0-51 ANOVA para $\sqrt{\bar{X}_y}$ rosenbrock

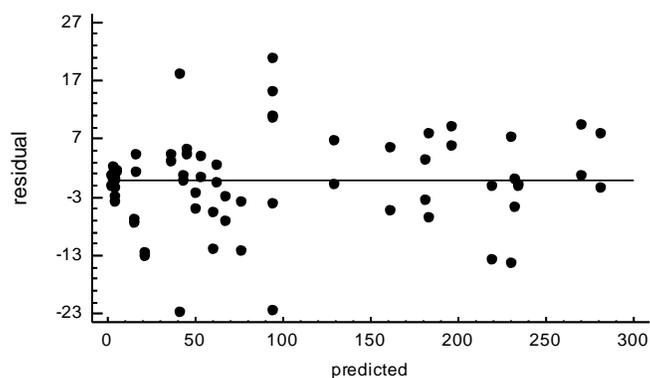


Gráfico 0-40 Residuales Vs predichos para $\sqrt{\bar{X}_y}$ rosenbrock

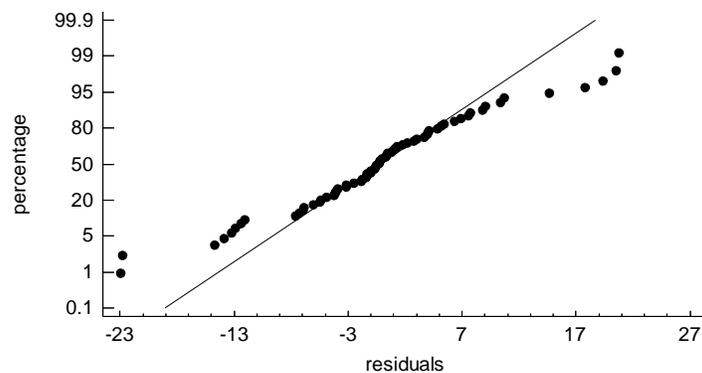


Gráfico 0-41 Probabilidad normal de residuales para $\sqrt{\bar{X}_y}$ rosenbrock

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	256221.	1	256221.	997.18	0.0000
B:floksize	14175.6	1	14175.6	55.17	0.0000
C:w	63997.2	1	63997.2	249.07	0.0000
D:c1	3988.72	1	3988.72	15.52	0.0003
E:c2	31356.3	1	31356.3	122.04	0.0000
F:maxit	4.99527	1	4.99527	0.02	0.8897
AB	4037.0	1	4037.0	15.71	0.0002
AC	7192.35	1	7192.35	27.99	0.0000
AD	2509.62	1	2509.62	9.77	0.0030
AE	19343.0	1	19343.0	75.28	0.0000
BC	2141.86	1	2141.86	8.34	0.0057
CD	4361.27	1	4361.27	16.97	0.0001
CE	7303.39	1	7303.39	28.42	0.0000
ACD	2645.32	1	2645.32	10.30	0.0023
ACE	6985.13	1	6985.13	27.19	0.0000
ADE	1312.96	1	1312.96	5.11	0.0282
CEF	1021.57	1	1021.57	3.98	0.0516
Total error	12847.2	50	256.945		
Total (corr.)	441445.	67			

Tabla 0-52 ANOVA para $\sqrt{S_{y\text{rosenbrock}}}$

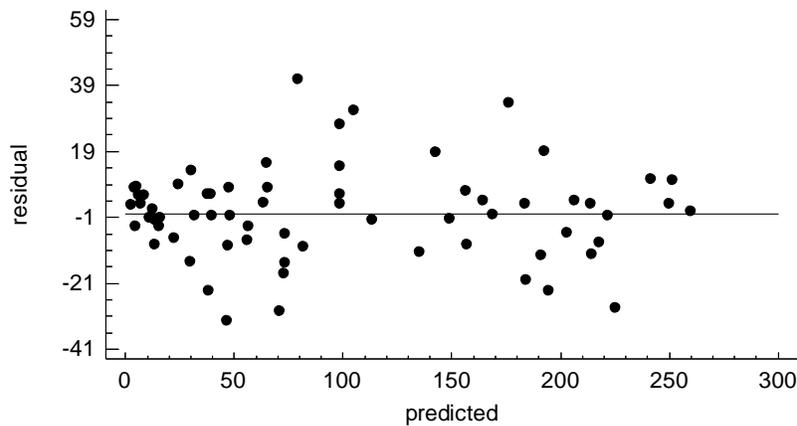


Gráfico 0-42 Residuales Vs predichos para $\sqrt{S_{y\text{rosenbrock}}}$

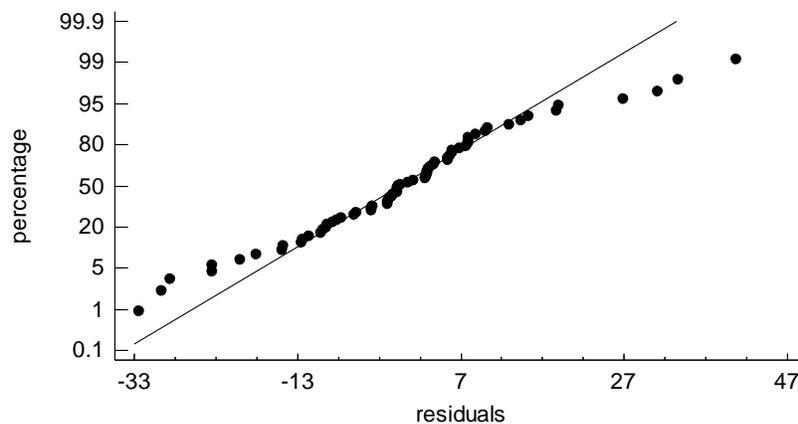


Gráfico 0-43 Probabilidad normal de residuales para $\sqrt{S_{y\text{rosenbrock}}}$

	Estimador
Constante	94.8442
A:dimensión	64.5869
B:floksize	-12.9986
C:w	46.5051
D:c1	7.80389
E:c2	21.3247
AB	-8.54747
AC	23.84
AD	6.51559
AE	17.7995
BC	-4.04741
CD	-8.76371
CE	-8.77242
ACD	-7.75983
ACE	-9.0752
CDE	-2.98991

**Tabla 0-53 coeficientes de
regresión para $\sqrt{\bar{X}_{y\text{ rosenbrock}}}$**

Coficiente	Estimador
constante	98.5416
A:dimensión	63.2729
B:floksize	-14.8827
C:w	31.6221
D:c1	7.89454
E:c2	22.1346
F:maxit	-0.279376
AB	-7.94218
AC	10.601
AD	6.26201
AE	17.3849
BC	-5.78503
CD	-8.25499
CE	-10.6825
ACD	-6.42908
ACE	-10.4471
ADE	-4.52935
CEF	-3.99525

**Tabla 0-54 coeficientes de
regresión para $\sqrt{S_{y\text{ rosenbrock}}}$**

Estadístico	$\sqrt{\bar{X}_{y\text{ rosenbrock}}}$	$\sqrt{S_{y\text{ rosenbrock}}}$
R ²	99.06%	97.09%
R ² _{aj}	98.79%	96.10%
Error estándar de estimación	9.85363	16.0295
Media del error absoluto	6.16876	10.0195
Durbin-Watson	2.23957 (P=0.8504)	1.86745 (P=0.3047)

Tabla 0-55 Estadísticos para los modelos de regresión (Rosenbrock)

4.2.5 Función de prueba Schwefel (ecuación 0-5)

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	5.3218E6	1	5.3218E6	7208.85	0.0000
B:floksize	45638.6	1	45638.6	61.82	0.0000
C:w	1.21615E7	1	1.21615E7	16473.89	0.0000
D:c1	227.714	1	227.714	0.31	0.5811
E:c2	1631.16	1	1631.16	2.21	0.1433
F:maxit	12627.5	1	12627.5	17.11	0.0001
AB	3583.76	1	3583.76	4.85	0.0321
AC	4.74755E6	1	4.74755E6	6430.98	0.0000
BC	19279.3	1	19279.3	26.12	0.0000
CD	18352.1	1	18352.1	24.86	0.0000
CE	34193.2	1	34193.2	46.32	0.0000
CF	14636.2	1	14636.2	19.83	0.0000
Total error	37649.8	51	738.231		
Total (corr.)	2.24187E7	63			

Tabla 0-56 ANOVA para $\bar{X}_{y\ schwefel}$

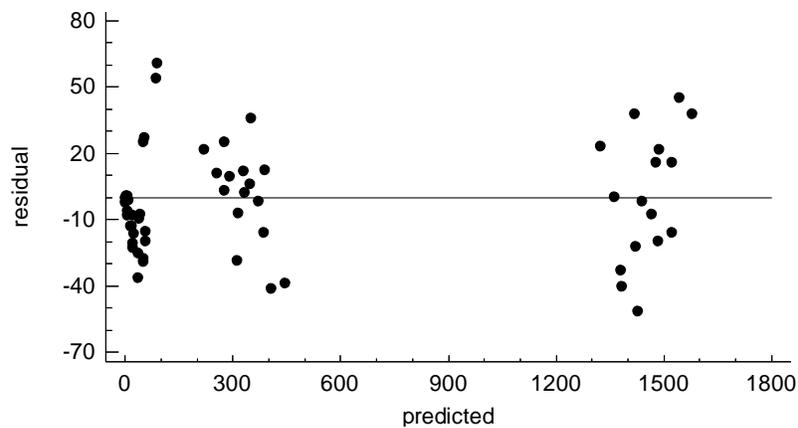


Gráfico 0-44 Residuales Vs predichos para $\bar{X}_{y\ schwefel}$

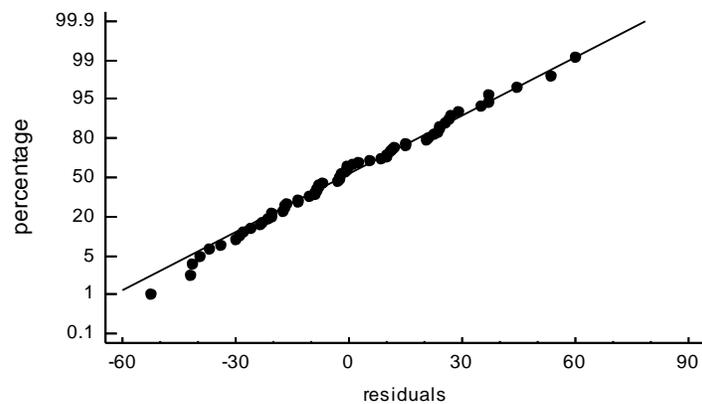


Gráfico 0-45 Probabilidad normal de residuales para $\bar{X}_{y\ schwefel}$

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	107183.	1	107183.	367.94	0.0000
B:floksize	7705.61	1	7705.61	26.45	0.0000
C:w	363687.	1	363687.	1248.46	0.0000
D:c1	5186.45	1	5186.45	17.80	0.0001
E:c2	9712.98	1	9712.98	33.34	0.0000
AC	20031.6	1	20031.6	68.76	0.0000
AE	4208.98	1	4208.98	14.45	0.0004
CD	11995.6	1	11995.6	41.18	0.0000
CE	14041.9	1	14041.9	48.20	0.0000
Total error	15730.6	54	291.308		
Total (corr.)	559483.	63			

Tabla 0-57 ANOVA para $S_{y\ schwefel}$

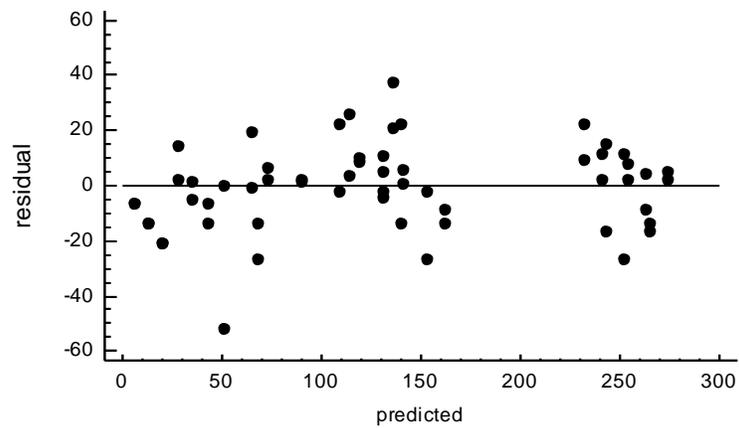


Gráfico 0-46 Residuales Vs predichos para $S_{y\ schwefel}$

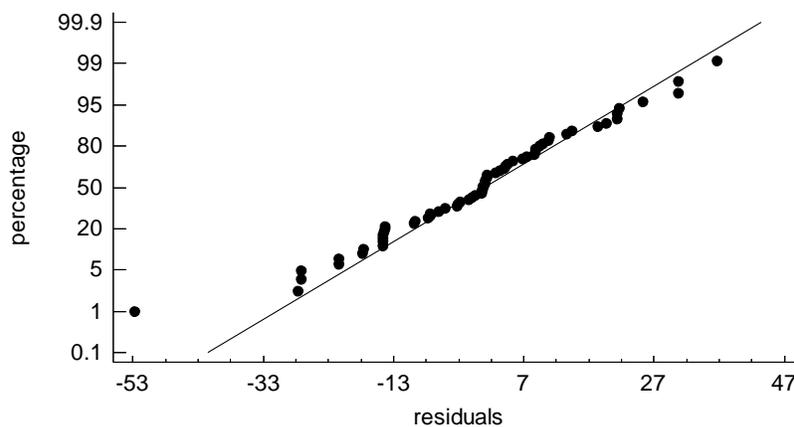


Gráfico 0-47 Probabilidad normal de residuales para $S_{y\ schwefel}$

Coefficiente	Estimador
Constante	458.294
A:dimensión	288.363
B:floksize	-26.704
C:w	435.917
D:c1	1.88627
E:c2	5.04845
F:maxit	-14.0465
AB	-7.48306
AC	272.361
BC	-17.3562
CD	16.9337
CE	23.1143
CF	-15.1225

Tabla 0-59 Coeficientes de regresión para $\bar{X}_{y\ schwefel}$

Coefficiente	Estimador
Constante	120.402
A:dimensión	40.9234
B:floksize	-10.9727
C:w	75.3831
D:c1	-9.00213
E:c2	-12.3193
AC	17.6916
AE	-8.10958
CD	13.6906
CE	14.8123

Tabla 0-58 Coeficientes de regresión para $S_{y\ schwefel}$

Estadístico	$\bar{X}_{y\ schwefel}$	$S_{y\ schwefel}$
R ²	99.8321 %	97.1884 %
R ² _{aj}	99.7925 %	96.7198 %
Error estándar de estimación	27.1704	17.0677
Media del error absoluto	19.5824	11.707
Durbin-Watson	2.03724 (P=0.5093)	2.09328 (P=0.5718)

Tabla 0-60 Estadísticos para los modelos de regresión (Schwefel)

4.2.6 Función de prueba Spherical (ecuación 0-6)

Fuente de variación	Suma de cuadrados	G.I.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	24566.0	1	24566.0	60.67	0.0000
B:floksize	862.166	1	862.166	2.13	0.1505
C:w	396091.	1	396091.	978.23	0.0000
D:c1	20767.1	1	20767.1	51.29	0.0000
E:c2	47642.3	1	47642.3	117.66	0.0000
F:maxit	44020.6	1	44020.6	108.72	0.0000
AC	22335.9	1	22335.9	55.16	0.0000
CD	20701.8	1	20701.8	51.13	0.0000
CE	47181.5	1	47181.5	116.52	0.0000
CF	43988.9	1	43988.9	108.64	0.0000
EF	5589.22	1	5589.22	13.80	0.0005
Total error	21055.2	52	404.907		
Total (corr.)	694801.	63			

Tabla 0-61 ANOVA para $\ln(\bar{X}_{y\ spherical})$

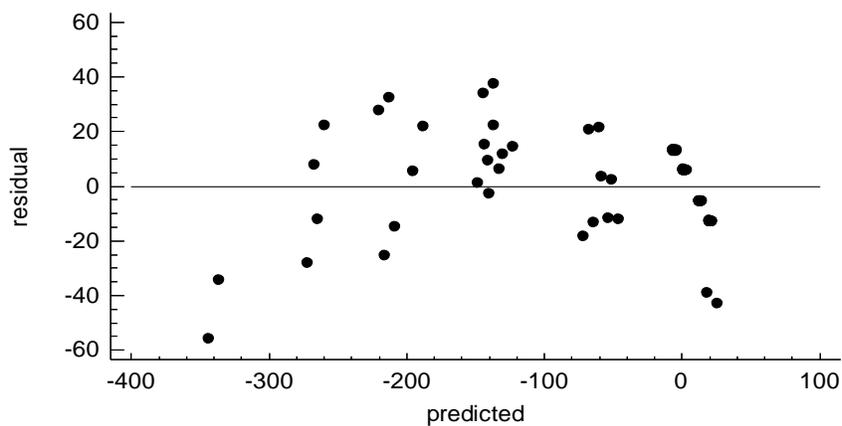


Gráfico 0-48 Residuales Vs predichos para $\ln(\bar{X}_{y\ spherical})$

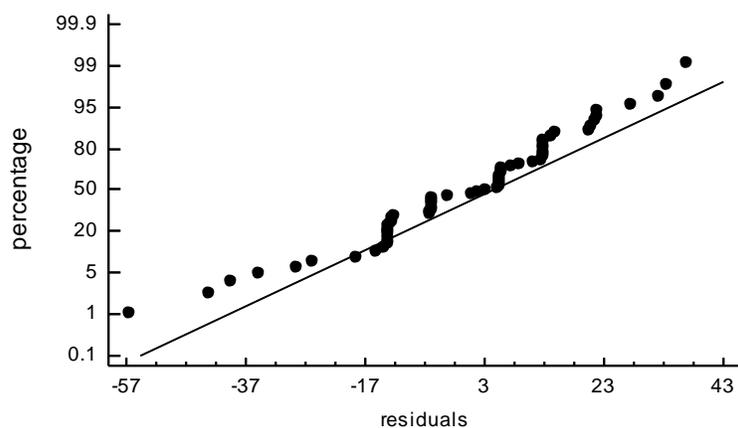


Gráfico 0-49 Probabilidad normal de residuales para $\ln(\bar{X}_{y\ spherical})$

Fuente de variación	Suma de cuadrados	G.I.	Cuadrados medios	Relación-F	Valor-P
A:dimensión	21186.0	1	21186.0	63.09	0.0000
B:floksize	425.997	1	425.997	1.27	0.2653
C:w	365295.	1	365295.	1087.75	0.0000
D:c1	18028.5	1	18028.5	53.68	0.0000
E:c2	42829.2	1	42829.2	127.53	0.0000
F:maxit	39484.3	1	39484.3	117.57	0.0000
AC	19612.0	1	19612.0	58.40	0.0000
CD	17925.7	1	17925.7	53.38	0.0000
CE	42453.6	1	42453.6	126.42	0.0000
CF	39334.2	1	39334.2	117.13	0.0000
EF	4294.18	1	4294.18	12.79	0.0008
Total error	17127.1	51	335.825		
Total (corr.)	571465.	62			

Tabla 0-62 ANOVA para $\ln(S_{y\ spherical})$

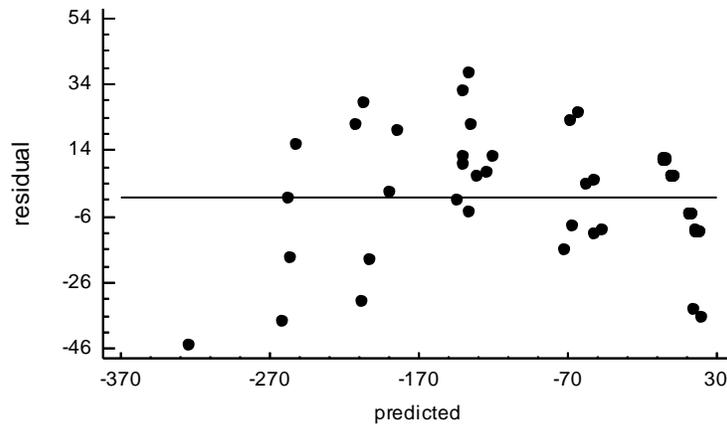


Gráfico 0-50 Residuales Vs predichos para $\ln(S_{y\ spherical})$

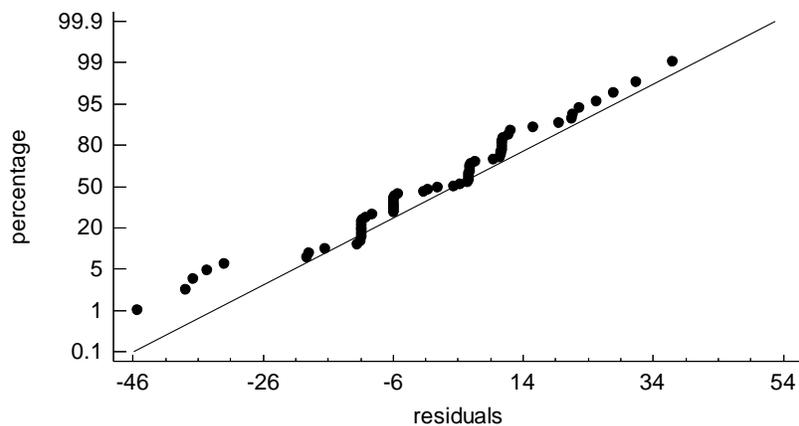


Gráfico 0-51 Probabilidad normal de residuos para $\ln(S_{y\ spherical})$

Coefficiente	Estimador
constante	-70.6345
A:dimensión	19.5919
B:floksize	-3.67033
C:w	78.6697
D:c1	18.0135
E:c2	27.2839
F:maxit	-26.2264
AC	-18.6815
CD	-17.9852
CE	-27.1516
CF	26.2169
EF	9.34514

Tabla 0-64 Coeficientes de regresión para

$\ln(\bar{X}_y \text{ spherical})$

Coefficiente	Estimador
constante	-69.1233
A:dimensión	18.3684
B:floksize	-2.60465
C:w	76.2725
D:c1	16.9444
E:c2	26.1166
F:maxit	-25.076
AC	-17.6729
CD	-16.896
CE	-26.0018
CF	25.0283
EF	8.26964

Tabla 0-63 coeficientes de regresión para

$\ln(S_y \text{ spherical})$

Estadístico	$\ln(\bar{X}_y \text{ spherical})$	$\ln(S_y \text{ spherical})$
R ²	96.9696 %	97.003 %
R ² _{aj}	96.3286 %	96.3565 %
Error estándar de estimación	20.1223	18.3255
Media del error absoluto	14.4826	13.0411
Durbin-Watson	1.9095 (P=0.3222)	1.92236

Tabla

4.2.7 Optimización

Para este caso se también se utiliza la función de deseabilidad, definida en la sección REF _Ref277085702 \r \h * MERGEFORMAT 2.6.2.1. Los límites utilizados en la

$$d_{ij} = \frac{Y_{ij} - Y_{j^*}}{Y_j - Y_{j^*}} \quad Y_{j^*} < Y_{ij} < Y_j \quad Y_{j^*} - Y_{ij} < Y_j - Y_{j^*} \quad Y_{j^*} < Y_{ij} < Y_j \quad 0$$

2-41 se enlistan en la Tabla 0-66. La deseabilidad para cada punto experimental (Tabla 0-67) se calcula a partir de los estimados para cada respuesta.

Variable de respuesta	Mínimo Y_{j^*}	Máximo Y_j
\bar{X}_{ackley}	4.87E-18	954.244569
$\bar{X}_{griewank}$	0.027500706	98.2857252
$\bar{X}_{rastrigin}$	-1.8997108	88.9526772
$\bar{X}_{rosenbrock}$	10.68198099	79433.9152
$\bar{X}_{schwefel}$	-29.67902	1582.62998
$\bar{X}_{spherical}$	6.80E-150	2.3244E+11
S_{ackley}	4.51E-17	1886.51785
$S_{griewank}$	0.019328532	32.1508812
$S_{rastrigin}$	0.035316	11.537794
$S_{rosenbrock}$	7.625738653	67791.9672
$S_{schwefel}$	-30.90035	275.67785
$S_{spherical}$	2.45E-143	1377601802

Tabla 0-66 límites para el cálculo del a función de deseabilidad (PSO)

	Número de variables	Número de partículas	W	C1	C2	Máximo de iteraciones	D
1	-1	-1	-1	-1	-1	-1	0.943
2	1	-1	-1	-1	-1	-1	0.88302
3	-1	1	-1	-1	-1	-1	0.95983
4	1	1	-1	-1	-1	-1	0.90805
5	-1	-1	1	-1	-1	-1	0.50942
6	1	-1	1	-1	-1	-1	0
7	-1	1	1	-1	-1	-1	0
8	1	1	1	-1	-1	-1	0.25543
9	-1	-1	-1	1	-1	-1	0.97054
10	1	-1	-1	1	-1	-1	0.89847
11	-1	1	-1	1	-1	-1	0.98571
12	1	1	-1	1	-1	-1	0.93064
13	-1	-1	1	1	-1	-1	0.7801
14	1	-1	1	1	-1	-1	0
15	-1	1	1	1	-1	-1	0.81756
16	1	1	1	1	-1	-1	0.33065
17	-1	-1	-1	-1	1	-1	0.96008
18	1	-1	-1	-1	1	-1	0.86157
19	-1	1	-1	-1	1	-1	0.97584
20	1	1	-1	-1	1	-1	0.90618
21	-1	-1	1	-1	1	-1	0.77846
22	1	-1	1	-1	1	-1	0
23	-1	1	1	-1	1	-1	0.81571
24	1	1	1	-1	1	-1	0.31701
25	-1	-1	-1	1	1	-1	0.98458
26	1	-1	-1	1	1	-1	0
27	-1	1	-1	1	1	-1	0.99911
28	1	1	-1	1	1	-1	0.88191
29	-1	-1	1	1	1	-1	0.75566
30	1	-1	1	1	1	-1	0
31	-1	1	1	1	1	-1	0.79577
32	1	1	1	1	1	-1	0.25862

Tabla 0-67 Deseabilidad para cada punto experimental (PSO)

	Número de variables	Número de partículas	W	C1	C2	Máximo de iteraciones	D
33	-1	-1	-1	-1	-1	1	0.94313
34	1	-1	-1	-1	-1	1	0.88376
35	-1	1	-1	-1	-1	1	0.95991
36	1	1	-1	-1	-1	1	0.90813
37	-1	-1	1	-1	-1	1	0.51396
38	1	-1	1	-1	-1	1	0
39	-1	1	1	-1	-1	1	0
40	1	1	1	-1	-1	1	0.27062
41	-1	-1	-1	1	-1	1	0.97063
42	1	-1	-1	1	-1	1	0.90089
43	-1	1	-1	1	-1	1	0.98575
44	1	1	-1	1	-1	1	0.93216
45	-1	-1	1	1	-1	1	0.78678
46	1	-1	1	1	-1	1	0
47	-1	1	1	1	-1	1	0.82235
48	1	1	1	1	-1	1	0.34636
49	-1	-1	-1	-1	1	1	0.95976
50	1	-1	-1	-1	1	1	0.85663
51	-1	1	-1	-1	1	1	0.97556
52	1	1	-1	-1	1	1	0.90288
53	-1	-1	1	-1	1	1	0.78407
54	1	-1	1	-1	1	1	0
55	-1	1	1	-1	1	1	0.82205
56	1	1	1	-1	1	1	0.34001
57	-1	-1	-1	1	1	1	0.98397
58	1	-1	-1	1	1	1	0.79932
59	-1	1	-1	1	1	1	0.99853
60	1	1	-1	1	1	1	0.87603
61	-1	-1	1	1	1	1	0.76215
62	1	-1	1	1	1	1	0
63	-1	1	1	1	1	1	0.80294
64	1	1	1	1	1	1	0.28856

Tabla 0-32 Deseabilidad para cada punto experimental (PSO) (continuación)

Validación.

Con los niveles óptimos de los factores del experimento de sintonización del PSO, se resolvió cada función hasta formar dos grupos de 10 series datos (medias y desviaciones estándar) para nueve y cinco variables, utilizando los límites de la Tabla 0-65, se calcula la deseabilidad, mostrada en la Tabla 0-68.

i	Dpso5 (5 dimensiones)	Dpso9 (9 dimensiones)
1	0.97774	0.94329
2	0.97789	0.94351
3	0.97399	0.93776
4	0.97634	0.9408
5	0.97274	0.93227
6	0.97048	0.92433
7	0.97752	0.94584
8	0.97358	0.94637
9	0.96314	0.94101
10	0.97784	0.9419

Tabla 0-68 deseabilidades para prueba de hipótesis

El primer paso es verificar que los datos provienen de una distribución normal. La prueba de Shapiro-Wilk (Tabla 0-69), indica que no se rechaza la hipótesis de que los datos Dpso5 y Dpso9 provienen de una distribución normal, con un nivel de confianza del 95%. Los gráficos Gráfico 0-52 y Gráfico 0-26Gráfico 0-53 apoyan el supuesto de normalidad.

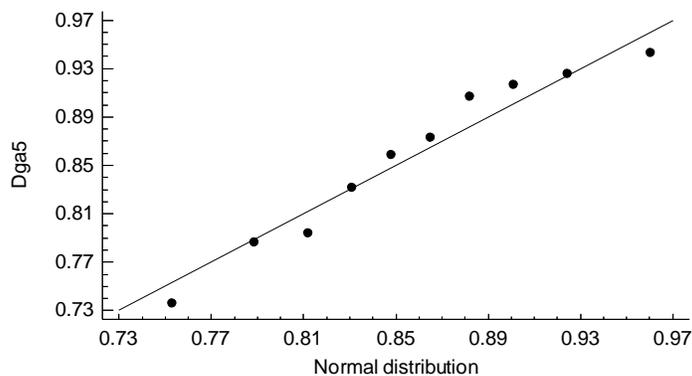


Gráfico 0-52 Cuantil-cuantil para Dpso5 Vs distribución normal

Existe evidencia de que los datos pueden representarse por una distribución normal. Esto permite el uso del intervalo de confianza para la media y desviación. Que es el intervalo que contiene la media real.

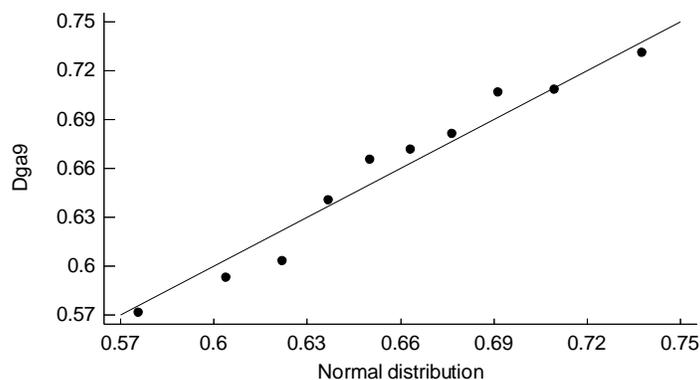


Gráfico 0-53 Cuantil-cuantil para Dpso9 Vs distribución normal

	<i>Estadístico</i>	<i>Valor-P</i>
5 variables	0.810747	0.0195846
9 variables	0.949161	0.642869

Tabla 0-69 Prueba de Shapiro-Wilk W

	media	Desviación estándar	Intervalo de confianza para la media [LIC,LSC]	Intervalo de confianza para la desviación estándar [LIC,LSC]
Dpso5	0.974125	0.00463913	[0.970806; 0.977444]	[0.00319096; 0.00846924]
Dpso9	0.939709	0.00677095	[0.934865; 0.944553]	[0.0046573; 0.0123611]

Tabla 0-70 intervalos de confianza del 95% para la media y desviación estándar

La prueba t, se realiza tomando como μ_0 la deseabilidad encontrada en la optimización, 0.99911 para Dpso5 y 0.93216 para Dpso9. Estos valores caen fuera de los límites de confianza para la media.

Como $0.99911 > LSC$, Para Dpso5 se tiene:

$$H_0: \mu_0 = 0.99911$$

$$H_1: \mu_0 > 0.99911$$

$$t_{calculada} = -17.031$$

$$\text{Valor-P} = 1.0$$

No se rechaza la hipótesis nula con 95% de confianza

Como $0.93216 < LIC$, Para Dga9:

$$H_0: \mu_0 = 0.93216$$

$$H_1: \mu_0 < 0.93216$$

$$t_{calculada} = 3.52569$$

$$\text{Valor-P} = 0.996771$$

No se rechaza la hipótesis nula con 95% de confianza.

Ya sea que se maneje el número de variables en un nivel alto o bajo, los mejores parámetros para el PSO son:

Número de variables	Número de partículas	W	C1	C2	Máximo de iteraciones
5	50	0.4	2	2.5	500
9	50	0.4	2	2	1000

Tabla

4.3 Comparación de algoritmos (GA y PSO)

El

GA					PSO				
1	2	3	4	5	1	2	3	4	5
0.58317	0.30229	0.083442	0.59802	0.84965	0.58318	0.30609	0.083442	0.59802	0.83831
0.58318	0.30609	0.083442	0.59802	0.84965	0.58318	0.30609	0.083442	0.59802	0.83831
0.58318	0.30609	0.083442	0.57557	0.84965	0.58318	0.30609	0.083442	0.59802	0.84965
0.58318	0.30229	0.083442	0.59802	0.84965	0.58318	0.30609	0.083442	0.59802	0.84965
0.58318	0.30609	0.083442	0.59802	0.84965	0.58318	0.30229	0.083442	0.57557	0.84965
0.58317	0.30229	0.083442	0.59802	0.8383	0.58318	0.30609	0.083442	0.57557	0.84965
0.58316	0.30229	0.083442	0.59802	0.84965	0.58318	0.30609	0.083442	0.59802	0.83831
0.58317	0.30609	0.083442	0.59802	0.84965	0.57756	0.30609	0.083442	0.59802	0.84965
0.58318	0.30229	0.083442	0.59802	0.83831	0.58318	0.30229	0.083442	0.59802	0.83831
0.58317	0.30229	0.083442	0.59802	0.84965	0.57756	0.30609	0.083442	0.59802	0.84965

Tabla

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
Model	6.9216	9	0.769067	48427.75	0.0000
Residual	0.00142926	90	0.0000158807		
Total (Corr.)	6.92303	99			

Tabla

Fuente de variación	Suma de cuadrados	G.l.	Cuadrados medios	Relación-F	Valor-P
Algoritmo	0.000040386	1	0.000040386	0.00	0.9947
Caso(Algoritmo)	6.92156	8	0.865195	54480.90	0.0000
Residual	0.00142926	90	0.0000158807		
Total (corr.)	6.92303	99			

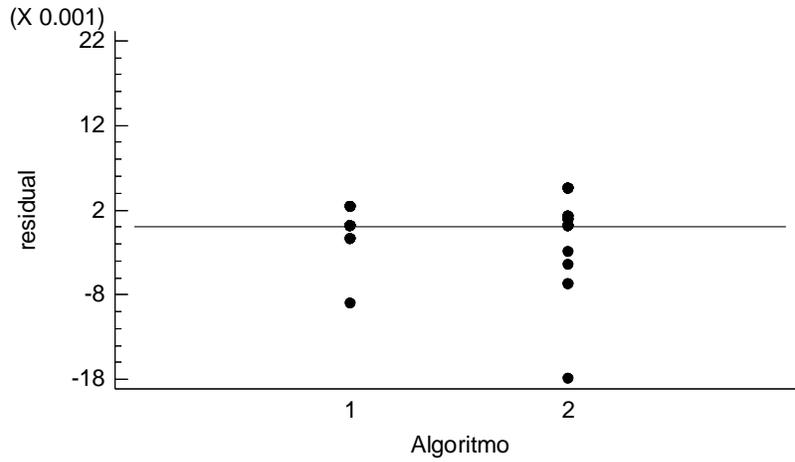
Tabla

Estadístico	D
R^2	99.9794 %
R^2_{aj}	99.9773 %
Error estándar de estimación	0.00398506
Media del error absoluto	0.0021102
Durbin-Watson	2.05788

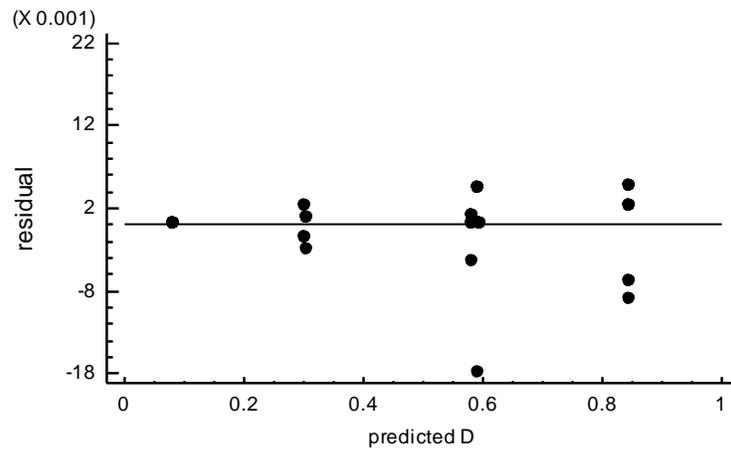
Tabla

<i>Fuente de variación</i>	<i>Suma de cuadrados</i>	<i>G.I.</i>	<i>Cuadrados medios</i>	<i>Componente de varianza</i>	<i>%</i>
TOTAL (CORREJIDO)	6.92303	99			
Algoritmo	0.000040386	1	0.000040386	0.0	0.00
Caso	6.92156	8	0.865195	0.0865179	99.98
ERROR	0.00142926	90	0.0000158807	0.0000158807	0.02

Tabla

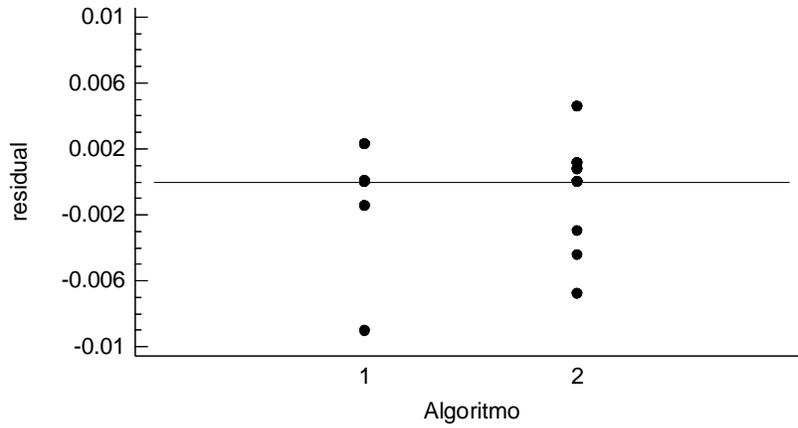


Gráfico



Gráfico

Los



Gráfico

Existe evidencia estadística para concluir que no existe diferencia entre el promedio de las soluciones encontradas por los algoritmos, pero es significativa entre cada caso de estudio. Un análisis del gráfico anterior, revela que la variación dentro del caso es menor para el PSO, dado que la distancia entre puntos es menor para este algoritmo.

CAPÍTULO 5.

CONCLUSIONES

Para este estudio, la sintonización de los algoritmos fue decisiva, y el diseño de experimentos fue una herramienta efectiva para lograrlo. En la etapa de sintonización se pone en evidencia las capacidades de los dos algoritmos. Resulta de particular interés el desempeño del GA y del PSO frente a las funciones de prueba Rosenbrock y Schwefel. Los parámetros seleccionados permiten a los algoritmos encontrar solución con media y desviación estándar mínimas para todas las funciones.

Los resultados obtenidos ponen en evidencia que el GA pierde precisión y exactitud al aumentar el número de variables, es por esto los puntos con deseabilidades mayores a 0.9 se presentan en los tratamientos donde el número de variables está en nivel bajo. En contraposición, los tratamientos con deseabilidad igual a cero, fueron aquellos donde el número de variables estaba en nivel alto.

Para el caso del PSO, no pierde precisión y exactitud al aumentar variables. El parámetro $V_{k+1} = w_k + C1r1LBX - X + C2r2GBX - X$ 2-51 es el que tiene mayor peso sobre la media y desviación estándar, los tratamientos con deseabilidades mayores a 0.85, incluyen a w en nivel bajo, mientras que la mayoría de los tratamientos con deseabilidad igual a cero, incluyen w en nivel alto. El cambio en el nivel del número de variables se compensa con el ajuste de $C2$ y del máximo de iteraciones. Para 5 variables, $C2$ se fija en 1 y el máximo de iteraciones en -1. Para 9 variables, $C2$ disminuye a -1.

Considérese las funciones Rosenbrock y Schwefel. Mientras que el GA resuelve la función Rosenbrock con mejor precisión y exactitud que el PSO, para la función Schwefel ocurre lo contrario (Tabla 5-1).

	PSO				GA			
	5 variables		9 variables		5 variables		9 variables	
	\bar{X}_y	S_y	\bar{X}_y	S_y	\bar{X}_y	S_y	\bar{X}_y	S_y
Rosenbrock	14.2777	100.9807	77.6978	219.1679	0.8281	1.5379	7.2167	19.3136
Schwefel	6.4E-05	0	6.4E-05	1.7E-13	58.6278	101.1740	225.6004	211.5809

Tabla 5-1 Resultados para la funciones Rosenbrock y Schwefel en el tratamiento óptimo

Para el caso del PSO, el tratamiento que presenta menor media y desviación estándar para la función Rosenbrock (pag. 49) es [-1, 1, -1, -1, -1 1] con deseabilidad observada de 0.96. Este valor es menor que la deseabilidad del tratamiento óptimo (0.98). El tratamiento que minimiza la media y desviación estándar para la función Rosenbrock repercute en la precisión y exactitud de las otras funciones.

Para el GA, el tratamiento que presenta menor media y desviación estándar para la función Schwefel (pag. 51) es [-1, 1, -1, -1, -1] con deseabilidad observada de 0.94, menor que la deseabilidad del tratamiento óptimo (0.95). El tratamiento que minimiza la media y desviación estándar para la función Schwefel representa un detrimento en la precisión y exactitud de las otras funciones.

De los párrafos anteriores se deduce que la media y desviación estándar para la función Schwefel son variables antagónicas en la sintonización del GA. De manera complementaria, la media y desviación estándar para la función Rosenbrock son variables antagónicas en la sintonización del PSO.

Es importante señalar que la sintonización fue muy importante en la comparación de los algoritmos. En un experimento previo, se varió el tamaño de población en cada algoritmo, para resolver el caso de Derringer, y la variable de respuesta fue la media de mil corridas.. Los parámetros seleccionados para cada algoritmo (Tabla 5-2), son diferentes a los encontrados en la sintonización, la población inicial fue de 20, 30, 40 y 50 individuos o partículas.

PSO	GA
<i>Iteraciones =100</i>	<i>Generaciones=100</i>
<i>w=0.4</i>	<i>Cruza: diseminada</i>
<i>C1=0.5</i>	<i>Fracción de cruza:=0.8</i>
<i>C2=1.6</i>	<i>Selección :torneo 4</i>
	<i>Elites=0</i>

Tabla 5-2 Parámetros de los algoritmos no sintonizados

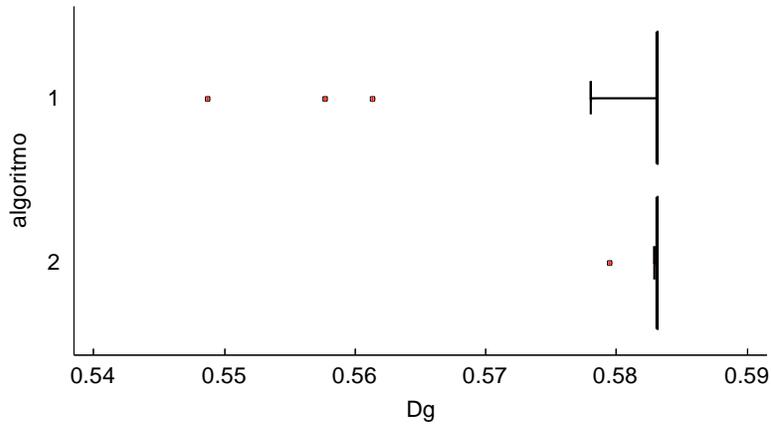


Gráfico 5-1 Diagrama de caja para el experimento del caso Derringer

En el Gráfico 5-1 muestra que el GA (1) presenta un rango amplio de soluciones, y a primera instancia parece que hay diferencia entre los dos algoritmos, pero la prueba estadística demuestra que no hay diferencia entre las medias, debido al traslape de los intervalos de diferencia mínima significativa o LSD (Gráfico 5-2).

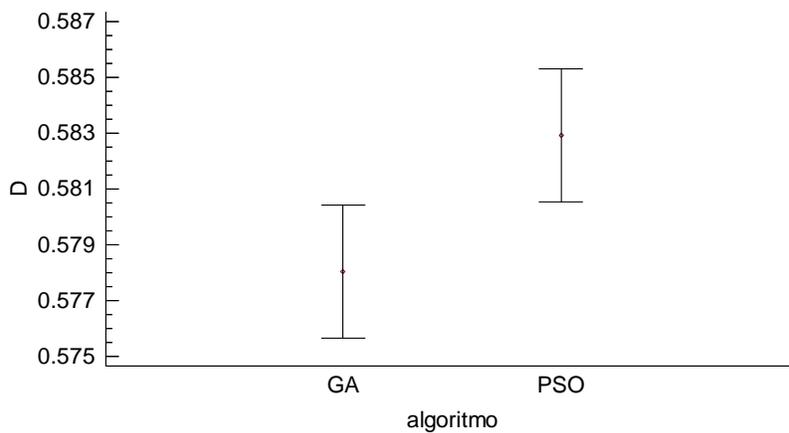


Gráfico 5-2 Medias e intervalo de diferencia mínima significativa (no sintonizado)

El uso de las funciones de benchmarking para sintonizar los algoritmos, resulta efectivo en la reducción del rango del GA, y aumentando el traslape con el PSO (Gráfico 5-3).

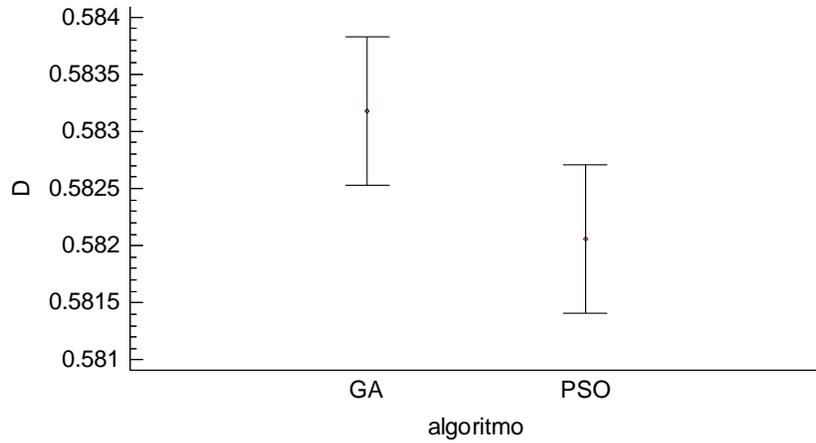


Gráfico 5-3 Medias e intervalo de diferencia mínima significativa (sintonizado)

La afirmación de que el PSO presenta menor variación dentro de cada caso puede apoyarse graficando el valor de la función (D) contra las variables independientes correspondientes. En los casos de Derringer, Del Castillo, Contreras y Wold, se encontraron al menos dos soluciones. El caso de Hendricks presenta una desviación estándar de cero.

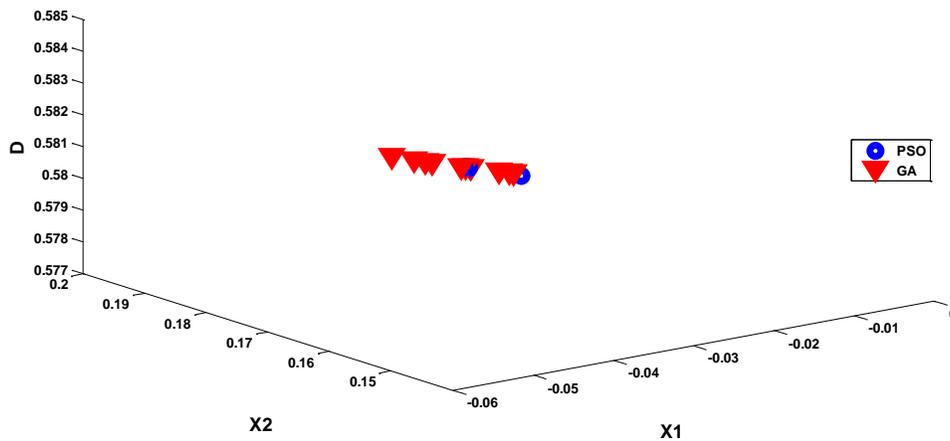


Gráfico 5-4 D Vs X1 Vs X2 para el caso Derringer

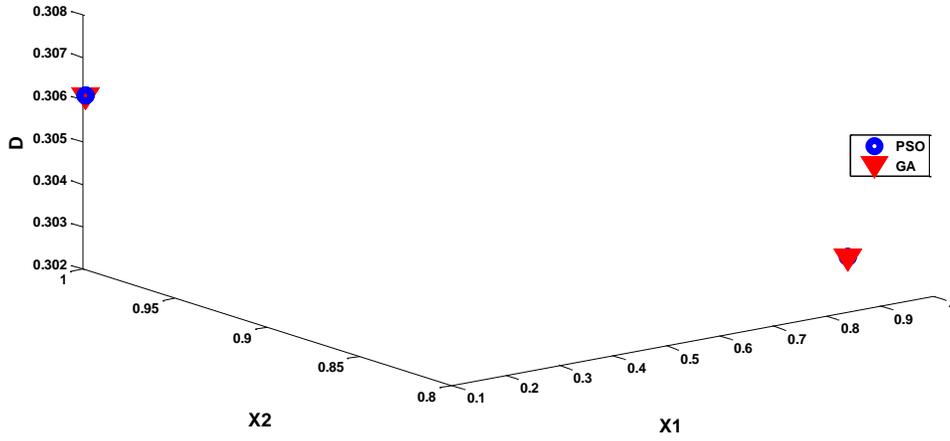


Gráfico 5-5 D Vs X1 Vs X2 para el caso Del Castillo

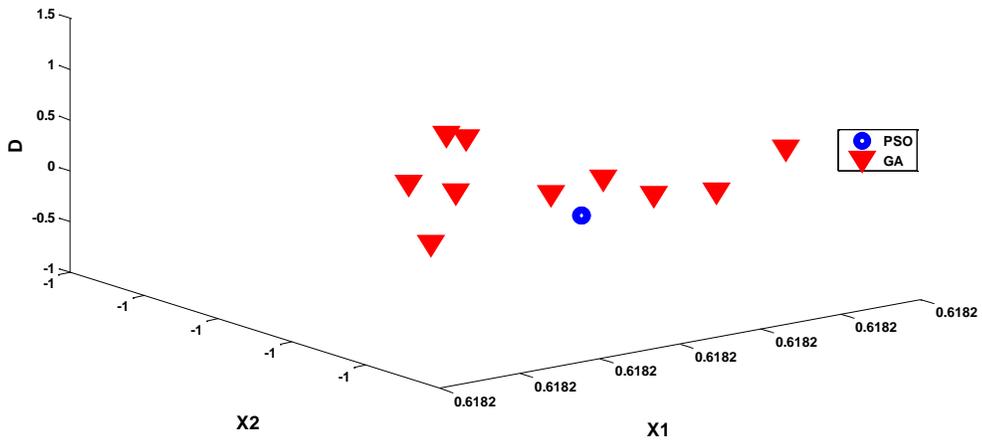


Gráfico 5-6 D Vs X1 Vs X2 para el caso Hendricks

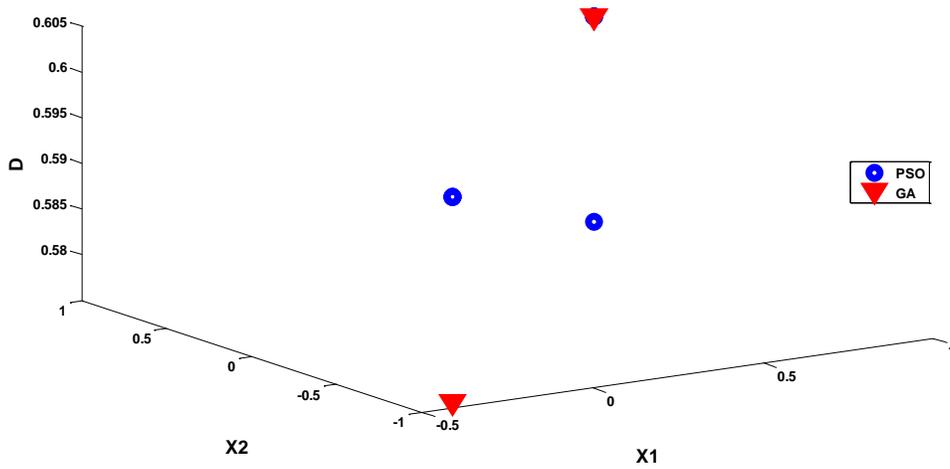


Gráfico 5-7 D Vs X1 Vs X2 para el caso Contreras

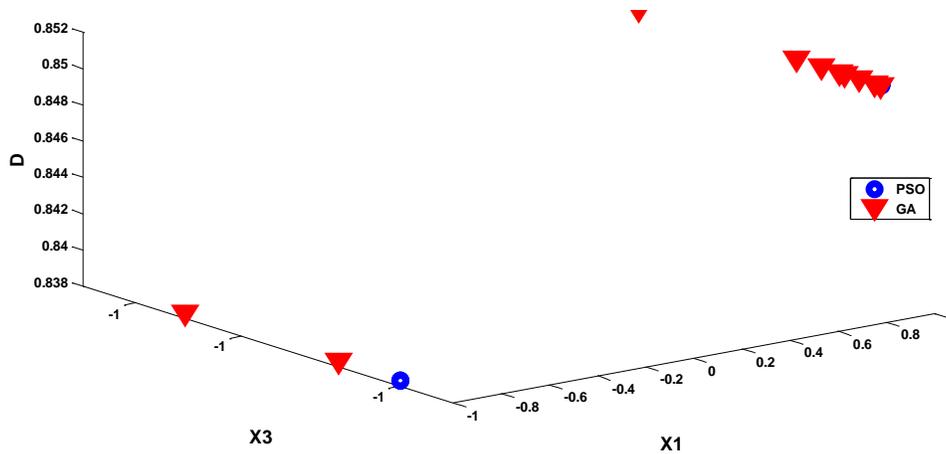


Gráfico 5-8 D Vs X1 Vs X3 para el caso Wold

Los gráficos anteriores presentan un PSO mas preciso que el GA, estos diagramas resultan útiles para evaluar la dispersión de las soluciones, y se verifica la afirmación de que el PSO presenta menor variación dentro de cada caso. El Gráfico 5-7 muestra los dos puntos atípicos del experimento anidado.

Existe evidencia estadística, para afirmar que los dos algoritmos tienen desempeños similares para resolver problemas de optimización de respuesta múltiple. El PSO presenta menor variabilidad dentro de cada caso. Esta ligera ventaja se suma a la facilidad de su programación y tiempo de cómputo menor.

Los resultados de este estudio se probaran en aplicaciones reales, y la metodología de comparación resulta efectiva para comparar otros algoritmos y las variaciones de estos.

BIBLIOGRAFIA

1. **Obermiller, Daniel J.** Multiple Response Optimization using JMP. *Proceedings of SAS Users Group Internacional Conference*. 1997.
2. **Fogliatto, Flavio S.** A Survey of Techniques for Optimizing Multiresponse Experiments. *Proceedings of ENEGEP*. 1998.
3. **Myers, R. H. y Montgomery, D. C.** *Response Surface Methodology: Process and Product Optimization Using Designed Experiments, 2nd ed.* New York : John Wiley & Sons Inc., 2003.
4. **Derringer, G. y Suich, R.** "Simultaneous Optimization of Several Response Variables". *Journal of Quality Technology*. 1980, Vol. 12, 4, págs. 214-219.
5. **Gutierrez Pulido, Humberto y De La Vara Salazar, Roman.** *Análisis y diseño de experimentos*. Mexico D.F : Mc Graw Hill , 2008.
6. **Mehnen J., Trautmann H.** Integration of Expert's Preferences in Pareto Optimization by Desirability. *Proceedings of the 5th CIRP International Seminar on Intelligent Computation in Manufacturing Engineering*. 2006, Vol. 5.
7. **Hassan, R., Cohanin, B. y de Weck, O.** A Comparison of Particle Swarm Optimization and the Genetic Algorithm Austin, TX. In *Proceedings of the 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*. 2005.
8. **Lovbjerg M., Rasmussen T. K. and Krink T.** "Hybrid particle swarm optimizer with breeding and subpopulation". in *Proc. 3rd Genetic Evolutionary Computation Conf. San Francisco, CA*. 2001.
9. **Muñoz Zavala A. E., Arturo Hernández Aguirre, Villa Diharce E. R.** Particle Evolutionary Swarm for Continuous Optimization Problems (PESCOP). *COMCEV, IV Congreso Mexicano de Computación Evolutiva*. 2008, págs. 27-32.
10. **Shi, X. H., y otros.** An improved GA and a novel PSO-GA-based hybrid algorithm. *Information Processing Letters*. 2005, 93, págs. 255–261.
11. **Panduro, M. A., y otros.** "A comparison of genetic algorithms, particle swarm optimization and the differential evolution method for the design of scannable circular antenna arrays". *Progress in electromagnetics Research B*. 2009, Vol. 13, págs. 171-186.

12. **Valdez, F., Melin, P. y Castillo, O.** An improved evolutionary method with fuzzy logic for combining Particle Swarm Optimization and Genetic Algorithms. *Applied Soft Computing Journal*. (2010), doi:10.1016/j.asoc.2010.10.010.
13. **Sabata, S. L., Ali, L. y K., Udatab S.** Integrated Learning Particle Swarm Optimizer for global optimization. *Applied Soft Computing Journal*. 2011, Vol. 11, 1, págs. 574–584.
14. **Montgomery, Douglas C.** *Design and analysis of experiments 5th edition*. New york : John Wiley & Sons Inc., 2001.
15. **Vose, Michael D.** *The simple genetic algorithm: foundations and theory*. Cambridge, MA : MIT Press, 1999.
16. **Kennedy, J. and Eberhart, R.** “Particle Swarm Optimization”. *Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia* . 1995, págs. 1942-1945.
17. **Del CAstillo, E., Montgomery, D.C. y McCarville, D.R.** Modified Desirability Functions for Multiple Response Optimization. *Journal of Quality Technology*. Julio de 1996, Vol. 28, No.3.
18. **M, Contreras L. y B., Bruins H. B. and Trout J. R.** Heating Characteristics of Beef Stew in MRE Pouches as Influenced by Critical Process and Product Parameters. *Working Paper 100, Center for Advanced Food Technology, Rutgers University*. 1995.
19. **Hendriks, M. M. W. B., de Boer, J. H., Smilde A.K. and Doornbos, D. A.** Multicriteria Decision Making. *Chemometrics and Intelligent Laboratory Systems*. 1992, Vol. 16, 3.
20. **Wold, S., Carlson, R. y Skagerberg, B.** Statistical Optimization a a Means to Reduce Risks in industrial Prosses. *The Enviromental Professional*. 1989, Vol. 11, págs. 127-131.
21. **Ng, Szu Hui.** A Bayesian Model-Averaging Approach for Multiple-Response Optimization. *Journal of Quality Technology*. enero de 2010, Vol. 42, 1, págs. 52-68.
22. **Ortiz, F., JR, Simpson, J.R., Pignatiello,J.J.,JR, Heredia Langner, A.** A Genetic Algorithm Approach to Multiple-Response optimization. *Journal of Quality Technology* . 2004, Vol. 36, 4.
23. **Goldberg, D.** *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA : Addison-Wesley, 1989.

24. **Freund John E., Miller Irwin, Miller Mariless.** *Estadística Matemática con Aplicaciones sexta edición.* Mexico : Pearson Educación, 2000.
25. **Montgomery, Douglas C. y Runger, George C.** *Applied Statistics and Probability for Engineers Third Edition.* New York, NY : John Willey & Sons Inc., 2003.