

# Diseño de Experimentos para la Evaluación de un Algoritmo Genético de Programación de Producción

Alexander A. CORREA ESPINAL

Profesor Escuela de Ingeniería de la Organización, Universidad Nacional de Colombia  
Medellín, Código postal 05004, Colombia, [alcorrea@unal.edu.co](mailto:alcorrea@unal.edu.co)

Elkin RODRIGUEZ VELASQUEZ

Profesor Escuela de Ingeniería de la Organización, Universidad Nacional de Colombia  
Medellín, Código postal 05004, Colombia, [erodrigu@unal.edu.co](mailto:erodrigu@unal.edu.co)

Jhoan S. CADAVID JARAMILLO

Ingeniero Industrial, Universidad Nacional de Colombia  
Medellín, Código postal 05004, Colombia, [jscadav0@unal.edu.co](mailto:jscadav0@unal.edu.co)

## RESUMEN

En este estudio corresponde a la continuación del trabajo de investigación presentado en el CИСCI 2009, donde se muestra la evaluación previa de un algoritmo genético (GA) de programación de producción, mediante el diseño de experimentos, sobre cuatro elementos que componen un GA como son: el % de Mutación, % de Cruce, tamaño de Población y el número de generaciones. El cual no presentaba un resultado contundente y por lo que se recomendaba mejorar el diseño de experimentos usado en fases posteriores. Para el nuevo estudio propuesto en este trabajo, se utilizan diseños factoriales  $2^k$  y  $2^k$  con puntos centrales, que permiten detectar la significancia estadística de efectos principales y algunas interacciones, donde se encuentra que la siguiente combinación de factores o parámetros del algoritmo genético desarrollado: porcentaje de mutación, en 10%; porcentaje de cruce, en un 25%; número de generaciones, en 650; y el tamaño de la población inicial en 3000, permiten obtener el menor valor del *Makespan* o el tiempo de terminación de todos los trabajos en todas las operaciones. Lo que tiene como consecuencia en el aumento de la productividad del proceso de producción programado con este algoritmo.

**Palabras claves:** Algoritmo Genético, Diseño de Experimentos, *Flexible Job shop*, Programación de producción, *Makespan*. Diseños factoriales  $2^k$ .

## 1. INTRODUCCIÓN

Este trabajo es el complemento de la investigación presentada en el CИСCI 2009, titulada: "Evaluación de la metodología de algoritmos genético con diseño experimental para la programación de plantas con configuración tipo *job shop*." [1], en este trabajo se mostró un GA (Algoritmo Genético) desarrollado para el problema FJSP, con el objetivo de disminuir el *Makespan* a partir de evaluar cuatro parámetros del algoritmo, donde se encontraron algunas dificultades en relación a la variabilidad del diseño, que deberían ser superadas. A partir de estas dificultades se determinó realizar una serie de acciones en la planeación y ejecución del diseño experimental como el aumento del número de replicas y cambio en los niveles de los factores, que permitió superarlas y definir los niveles de los factores o parámetros del algoritmo que permitan obtener el menor *Makespan*.

La estructura de presentación del presente artículo es la siguiente: El estado del arte se describe en la sección 2; La metodología utilizada para determinar los niveles de los

factores, el número de replicas y el modelo de diseño experimental usado, se presentan en la sección 3; Los resultados de la experimentación se exhiben en la sección 4; Y en la sección 5, se presentan las conclusiones y trabajos futuros.

## 2. ESTADO DEL ARTE

En el artículo de Correa *et al* (2008) [8], se presenta el estado del arte sobre la aplicación del algoritmo genético en la programación de producción. Para complementar dicho trabajo se hace un barrido de la aplicación del diseño de experimentos en la evaluación de algoritmos heurísticos en la programación de producción. En esta revisión bibliográfica se destaca que los algoritmos heurísticos han sido explorados utilizando métodos estadísticos de comparación simple para evaluar diferentes métodos y son pocos los estudios encontrados en donde sean abordados mediante diseños experimentales, entre estos pocos estudios, se halla el de Gutiérrez y Mejía (2004) [3], quienes estudian la influencia de los parámetros de un GA respecto a las variables de interés en su estudio, que son el *Makespan* y la suma de las tardanzas: Para lo cual analizan los parámetros de probabilidad de mutación, probabilidad de cruce, tamaño de la población y número de generaciones a 7, 8, 7 y 7 niveles, respectivamente. El estudio se realizó sobre un conjunto de problemas generados aleatoriamente, cada uno de los cuales era compuesto de 100 trabajos a ser programados en 5 y 10 máquinas, los tiempos de procesamiento para los trabajos fueron generados aleatoriamente de una distribución uniforme [0-100].

Los autores Gutiérrez y Mejía (2004) [3], toman los valores de los niveles en los cuales los factores muestran cambios y tendencias que favorecen la disminución del *Makespan* y que no implican un aumento considerable del tiempo computacional requerido para la secuenciación, con ello, la elección de mejor probabilidad de mutación y probabilidad de cruce, es de 20% para cada uno. Para el caso del tamaño de población, igual a 80. El número de generaciones, un nivel igual a 2000. Estos resultados a los que llegaron los autores son muy cercanos a los resultados obtenidos en el trabajo realizado por Franca *et al.* (1994) [4].

Otra investigación donde se estudia un planteamiento diferente a los GA muestran a investigadores como Figlali *et al.* (2007)[5], quienes desarrollan un algoritmo basado en agentes colonia de hormigas para el problema del *Job Shop scheduling* e intentan minimizar el *makespan*, donde los parámetros del algoritmo fueron calibrados mediante el uso de un diseño

experimental factorial completo [6], analizando las interacciones presentadas por los factores evaluados.

### 3. METODOLOGÍA

La evaluación del algoritmo genético de secuenciación de tareas en un ambiente productivo *flexible job shop*, se hizo a través de diseños de experimentos factoriales  $2^k$  y  $2^k$  con puntos centrales, dado el carácter exploratorio del mismo. Para evaluar la influencia de los factores que intervienen en el proceso de búsqueda del algoritmo, en la disminución del tiempo de secuenciación de las tareas (trabajos), tomando como la variable respuesta el *Makespan*.

Con el objetivo de reducir el valor del *Makespan*, se realizó un plan experimental con base en un diseño factorial  $2^k$ . Así, se diseñó una etapa de exploración, que mostrara alguna influencia por parte de los factores y/o interacciones entre estos, y conducir la investigación hacia una nueva fase que prometiera más información sobre el comportamiento de los factores, como se presenta en Correa *et al.* (2009) [1]. El experimento se ejecutó en las salas de cómputo de la Escuela de Ingeniería de la Organización de la Facultad de Minas de la Universidad Nacional de Colombia, sede Medellín. El algoritmo genético desarrollado se hizo en Macros de Excel, el diseño factorial se diseñó y ejecutó con ayuda del software estadístico Minitab 15 [7].

Para llevar a cabo las experimentaciones, se ajustaba el escenario que comprendía los tiempos de operación de cada máquina para cada trabajo, el número de trabajos a secuenciar, y un nivel por cada factor estudiado como lo establece los lineamientos propuestos por Montgomery (2005) [2] y Box *et al.* (2005) [6]; una vez ajustado el escenario, se realizaron de forma aleatoria las corridas en los equipos de cómputo, de acuerdo al plan experimental, definido en los diseños factoriales. Una vez obtenidos todos los resultados del experimento, se analizaron en el *software* estadístico Minitab 15 [7].

### 4. RESULTADOS EXPERIMENTALES

La investigación sobre la influencia de los factores contempló una fase inicial de experimentación, en la cual se reportó el uso de un diseño experimental  $2^4$ , con 3 replicas, los factores y niveles de este diseño se muestran en la tabla 1.

**Tabla 1. Factores y niveles analizados en fase de exploración**

Factores	Niveles	
	Bajo(-)	Alto(+)
Porcentaje de mutación	10	30
Porcentaje de cruce	10	40
Tamaño de la población inicial	80	240
Numero de generaciones	100	300

En esta fase inicial de experimentación, se concluyó que era necesario afinar el modelo, con el fin de determinar la significancia de los factores e interacciones analizadas, pues la variabilidad obtenida con la experimentación propuesta, no permitía obtener resultados confiables, para lo cual fue necesario definir fases de experimentación adicionales, que permitieran obtener resultados más prometedores, ver el trabajo de Correa *et al.* (2009) [1].

A partir de los resultados obtenidos en la fase inicial de experimentación, se decidió realizar ajustes en las fases de experimentación posteriores, con el fin mejorar la calidad del modelo y así detectar la significancia de factores e interacciones, para cual se definió un diseño factorial  $2^k$  con puntos centrales de tres replicas, además, se cambio el valor de los niveles de los factores: tamaño de población inicial y numero de generaciones, tal como se presentan en la tabla 2.

**Tabla 2. Factores y niveles analizados en fase final**

Factores	Niveles		
	Bajo	Punto central	Alto
% Mutación	10	20	30
% Cruce	10	25	40
Tamaño población inicial	1000	3000	5000
#Generaciones	300	650	1000

El experimento propuesto, consistió de 58 corridas o pruebas experimentales. En este experimento se utilizaron las ejecuciones realizadas en dos experimentaciones anteriores, en las cuales se encontró una alta variabilidad en la variable respuesta (*Makespan*) asociada a los escenarios evaluados.

Para disminuir esta variabilidad se analizaron todas las corridas experimentales del diseño seleccionado, encontrando el valor de la respuesta más alejado del promedio del escenario o corrida y reemplazándolo por el promedio de las corridas correspondientes a la misma combinación de niveles. Por ejemplo, para la corrida experimental que se muestra en la tabla 3, el proceso de suavizado fue el siguiente:

**Tabla 3. Corrida experimental**

% Mutación	% Cruce	Tamaño Población	Generaciones
10	10	1000	300

Se promedian las corridas con la misma combinación de niveles, como se presenta en la tabla 4, de donde se obtiene la media del *Makespan* para la corrida o escenario:

**Tabla 4. Análisis del *Makespan* del escenario**

% Mutación	% Cruce	Tamaño Población	Generaciones	Makespan
10	10	1000	300	14
10	10	1000	300	13
10	10	1000	300	20

Media del *Makespan* para el escenario = 15,66

El valor de *Makespan* igual a 20 unidades, lo consideramos como un punto atípico, debido a la alta diferencia respecto a la media, caso que no es igual para los dos registros de *Makespan* obtenidos en el mismo escenario. Sin considerar valores decimales, la media del escenario es un *Makespan* igual a 15 unidades, que es un valor factible. De esta forma, se reemplaza el valor más alejado de la media por el valor promedio obtenido, como se muestra en los siguientes resultados para el escenario analizado en la tabla 5.

**Tabla 5. Reemplazo de valor alejado de la media de respuesta del escenario**

% Mutación	% Cruce	Tamaño Población	Generaciones	Makespan
10	10	1000	300	14
10	10	1000	300	13
10	10	1000	300	15

Media del *Makespan* para el escenario = 14

Todos los escenarios se sometieron al mismo análisis y los que presentaron una alta variabilidad debido a un punto atípico, se ajustaron utilizando el procedimiento mostrado en el escenario anterior.

Con los ajustes realizados a algunos escenarios, se tomaron los resultados asociados al *Makespan* y se llevaron al análisis experimental.

#### 4.1 Análisis experimental

La tabla 6, muestra la configuración de Diseño Experimental analizado.

**Tabla 6. Configuración del Diseño Experimental**

<b>Full Factorial Design</b>	
Factors:4	Base Design:4; 16
Runs:58	Replicates:3
Blocks:	Center pts (total):10

**Tabla 8. Efectos y coeficientes estimados para el modelo**

<b>Factorial Fit: MAKESPAN versus % Mutación; % Cruce; ...</b>						
Estimated Effects and Coefficients for MAKESPAN (coded units)						
Term	Effect	Coef	SE Coef	T	P	
<b>Constant</b>		<b>15,875</b>	0,3715	42,74	<b>0,000</b>	
% Mutación	0,000	0,000	0,3715	0,00	1,000	
% Cruce	0,500	0,250	0,3715	0,67	0,505	
Tamaño Población	0,333	0,167	0,3715	0,45	0,656	
<b>Generaciones</b>	-4,250	<b>-2,125</b>	0,3715	-5,72	<b>0,000</b>	
% Mutación*% Cruce	0,417	0,208	0,3715	0,56	0,578	
% Mutación*Tamaño Población	-1,250	-0,625	0,3715	-1,68	0,100	
% Mutación*Generaciones	0,833	0,417	0,3715	1,12	0,269	
<b>% Cruce*Tamaño Población</b>	-3,417	<b>-1,708</b>	0,3715	-4,60	<b>0,000</b>	
% Cruce*Generaciones	-1,167	-0,583	0,3715	-1,57	0,124	
Tamaño Población*Generaciones	-0,333	-0,167	0,3715	-0,45	0,656	
% Mutación*% Cruce*Tamaño Población	-1,333	-0,667	0,3715	-1,79	0,080	
% Mutación*% Cruce*Generaciones	-0,750	-0,375	0,3715	-1,01	0,319	
% Mutación*Tamaño Población*Generaciones	0,917	0,458	0,3715	1,23	0,224	
% Cruce*Tamaño Población*Generaciones	3,583	1,792	0,3715	4,82	0,000	
% Mutación*% Cruce*Tamaño Población*Generaciones	0,167	0,083	0,3715	0,22	0,824	
Ct Pt		-2,975	0,8946	-3,33	0,002	
S = 2,57363      PRESS = 644,316						
R-Sq = 71,27%      R-Sq(pred) = 31,85% <b>R-Sq(adj) = 60,07%</b>						

Los datos son ingresados al Minitab, y este arroja los siguientes resultados, ver tablas 7 y 8, en las que se presenta el ANOVA y los efectos correspondientes.

**Tabla 7. ANOVA del modelo**

Analysis of Variance for MAKESPAN (coded units)				
Source	DF	Seq SS	F	P
<b>Main Effects</b>	<b>4</b>	<b>221,083</b>	<b>8,34</b>	<b>0,000</b>
<b>2-Way Interactions</b>	<b>6</b>	<b>186,917</b>	<b>4,70</b>	<b>0,001</b>
<b>3-Way Interactions</b>	<b>4</b>	<b>192,250</b>	<b>7,26</b>	<b>0,000</b>
4-Way Interactions	1	0,333	0,05	0,824
Curvature	1	73,247	11,06	0,002
Residual Error	41	271,567	271,567	6,6236
Pure Error	41	271,567	271,567	6,6236
Total	57	945,397		

De la tabla 7, se observa que los efectos principales y algunas interacciones son significativos.

El valor R-Sq(adj) obtenido en los experimentos anteriores se encontraba en promedio inferior al 30%, para este experimento se incrementó hasta un 60,07% que es un valor muy bueno para las investigaciones de exploración, este R-Sq(adj) hallado indica que cerca del 60% de la variación de los datos es explicada por el modelo ANOVA. Con este valor se puede generar un modelo lo suficientemente representativo a partir de las interacciones y efectos principales significativos.

De la tabla 8, se muestra que el efecto principal asociado al factor, número de generaciones y la interacción doble entre % de Cruce y Tamaño de población inicial, son significativos.

Con base en la tabla 8, el modelo resultante tiene la forma:

$Y$ : *Makespan*

$$Y = 15,875 - 2,125(\text{Generaciones}) - 1,708(\% \text{Cruce} * \text{TamañoPoblación}) + \varepsilon_{ijk}$$

Para ratificar los resultados de la tabla 8, se presenta la grafica de medias del factor generaciones, ver figura 1 y la grafica de la interacción porcentaje de cruce y tamaño de la población inicial, ver figura 2.

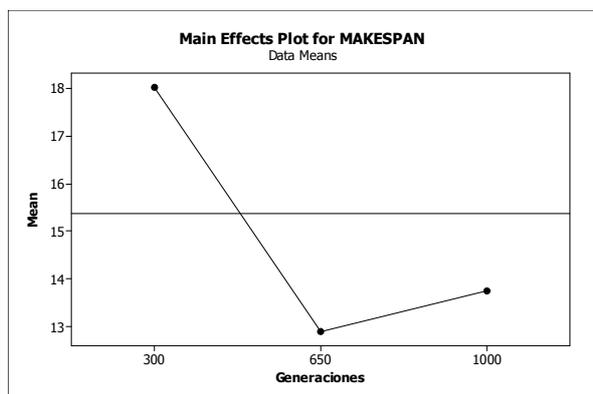


Figura 1. Efecto principal Generaciones

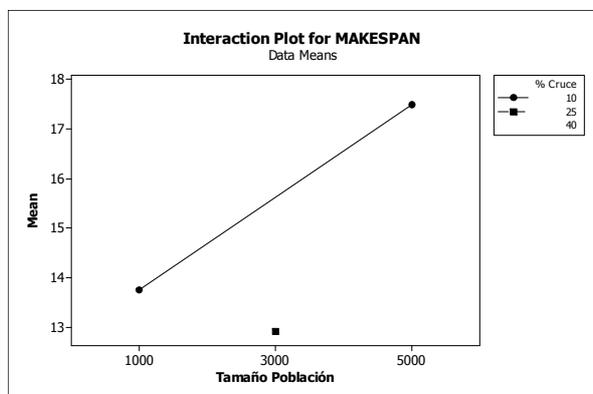


Figura 2. Interacción doble entre % de Cruce y Tamaño de población

Del análisis de las figuras 1 y 2, se puede evidenciar que los niveles de los factores que minimizan la variable respuesta son: % porcentaje de cruce, en un 25%, Numero de generaciones en 650 y el tamaño de la población inicial en 3000 y dado que el factor porcentaje de mutación no es significativo se toma el valor del nivel medio, es decir un % de 20.

Para garantizar validez del modelo estadístico y los análisis realizados, se evaluaron los residuales, con el fin de identificar el cumplimiento de los supuestos del modelo, como son: normalidad, varianza constante e independencia, para evidenciar esto, se presenta la figura 3.

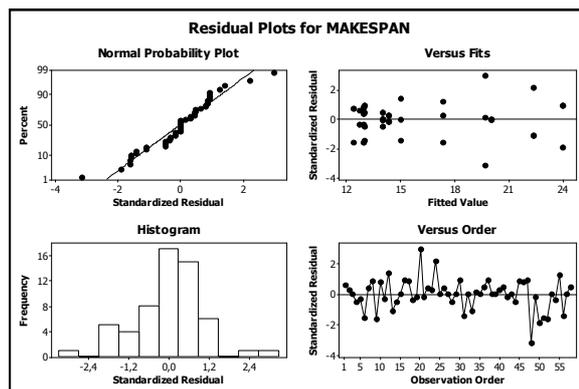


Figura 3. Validación de los supuestos del modelo

De acuerdo a la figura 3, los supuestos de normalidad, varianza constante e independencia se cumplen para los residuales, por lo tanto las conclusiones que se tomen a partir del modelo son validas estadísticamente.

## 5. CONCLUSIONES

Se ha mostrado que mediante el tratamiento de la alta variabilidad fue posible evidenciar significancia de algunos de los efectos principales e interacciones de los factores, que en experimentaciones previas no fue posible identificar.

Al cambiar los niveles de los factores tamaño de población inicial y numero de generaciones, fue posible mejorar el ajuste del modelo.

En resumen, podríamos afirmar que si usamos la siguiente combinación de factores como parámetros del algoritmo genético desarrollado: porcentaje de mutación en 20%, porcentaje de cruce en un 25%, Numero de generaciones en 650 y el tamaño de la población inicial en 3000, obtendríamos en promedio 13 unidades, el menor valor del *Makespan*, que es lo más eficiente a la hora de realizar la programación de producción con el algoritmo desarrollado.

## 6. AGRADECIMIENTOS

Los autores quieren presentar los más sinceros agradecimientos a la Dirección de Investigación de la Universidad Nacional de Colombia, Sede Medellín (DIME) por la financiación del proyecto de investigación No.20201007088, del cual este es uno de sus resultados. Adicionalmente, queremos agradecer el apoyo prestado por la ingeniera industrial, Isabel Londoño Restrepo en la realización de este trabajo.

## 7. REFERENCIAS

- [1]. A. Correa., E. Rodríguez y M. Londoño., “Evaluación de la metodología de algoritmos genéticos con diseño experimental para la programación de plantas con configuración tipo job shop”, **Octava conferencia iberoamericana en Sistemas, Cibernética e Informática**, Orlando, Florida, EE.UU. 2009.
- [2]. D. Montgomery, **Design and Analysis of Experiments**, Wiley.2005.
- [3]. E. Gutiérrez y G. Mejía, “Evaluación de algoritmos genéticos para el problema de máquinas en paralelo con

tiempos de alistamiento dependientes de la secuencia y restricciones en las fechas de entrega” **III Congreso colombiano y I conferencia andina internacional de investigación de operaciones**, Reporte de investigación, Universidad de los Andes, 2004.

[4]. P.M. Franca, M. Gendreau, G. Laporte, F. Muller. “A composed heuristic for the identical parallel machine scheduling problem with minimum makespan objective”, **Computers and Operations Research**. 22(2), 210-305,1994.

[5]. N. Figlali, C. Özcalel,O. y Engin. “Investigation of ant system parameter interactions by using design of experiments for Job Shop Scheduling problems”; **Computers & Industrial Engineering**, 1-22, 2007.

[6].G.E.P Box W.G., Hunter & J.S. Hunter. “**Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building**”; 2005.

[7]. MINITAB 15, **Meet Minitab 15. Minitab** .2008.

[8]. A. Correa, E. Rodríguez y M.I. Londoño. “Secuenciación de operaciones para configuraciones de planta tipo flexible Job Shop: Estado del arte”. **Revista Avances en sistemas e Informática**, Vol. 5, 151-161, 2008.