



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Sistemas Inteligentes
y Aplicaciones Numéricas en Ingeniería

Estudio sobre la Codificación en el Diseño Óptimo de Estructuras Metálicas



Máster en Sistemas Inteligentes y Aplicaciones
Numéricas en Ingeniería

Trabajo Fin de Máster

Autor: Noelia Díaz Padilla

Tutores: David Juan Greiner Sánchez
José María Emperador Alzola

Las Palmas de Gran Canaria, diciembre 2012

A mis padres, por haberme enseñado todas las cosas importantes que sé.

AGRADECIMIENTOS

A mis tutores, David Greiner y José María Emperador, por los conocimientos transmitidos en un campo tan nuevo como interesante para mí, y por el intenso apoyo y la disponibilidad que me han brindado estos meses.

Al Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería (SIANI), cuyo equipamiento de cálculo ha sido clave para el desarrollo de este trabajo, y especialmente a Juan Ignacio, por la ayuda prestada para poner a mi disposición estos equipos.

A mi familia, por su cariño y apoyo, así como por esforzarse en liberarme de todo cuanto han podido para que me centrara en este trabajo. A mis amigos también por su apoyo y por entender mis ausencias.

Tabla de contenido

1.	PRESENTACIÓN DEL PROBLEMA. OPTIMIZACIÓN ESTRUCTURAL.....	2
1.1.	Problema propuesto.....	3
2.	ALGORITMOS GENÉTICOS en la resolución de problemas estructurales.....	5
2.1.	Conceptos.....	6
2.2.	Codificación del cromosoma en el problema estructural.....	8
2.3.	Definición de la función objetivo.....	10
3.	CODIFICACIÓN DEL CROMOSOMA. CÓDIGOS BINARIOS	12
3.1.	Códigos binarios	12
3.1.1.	Códigos ponderados.....	13
3.1.2.	Códigos no ponderados.....	13
3.2.	Aplicación de diferentes códigos binarios al problema presentado	14
3.3.	Creación y representación de códigos binarios diferentes	16
3.4.	Estrategia empleada para la generación de codificaciones	16
4.	FASE EXPERIMENTAL	22
4.1.	Problema empleado. Variables seleccionadas.....	26
4.2.	Tratamiento de los resultados	27
4.3.	Resultados obtenidos.....	28
4.4.	Observación de resultados. Gráficos.....	29
5.	ANÁLISIS ESTADÍSTICO DE LOS RESULTADOS OBTENIDOS	45
5.1.	Tests de hipótesis. Teoría e implementación.....	47
5.2.	Resultados obtenidos.....	49
5.2.1.	Test de normalidad Lillieforts	50
5.2.2.	Tests de significancia Kruskal- Wallis y multicomparación	51
6.	CONCLUSIONES Y LÍNEAS FUTURAS.....	87
7.	BIBLIOGRAFÍA Y REFERENCIAS	90

1. PRESENTACIÓN DEL PROBLEMA. OPTIMIZACIÓN ESTRUCTURAL

La **optimización de pórticos metálicos** de nudos rígidos en régimen elástico destaca en importancia por el gran número de construcciones que responden a esta tipología estructural, tanto en ingeniería industrial, civil o en arquitectura. La dificultad que entraña la resolución de un problema compuesto por multitud de variables, así como de parámetros de entrada, implica que los métodos tradicionales de cálculos limiten la capacidad de la técnica a explorar nuevas y óptimas soluciones. En este sentido, los **algoritmos evolutivos** suponen una técnica ampliamente aceptada hoy en día en multitud de campos de aplicación (no sólo en los anteriormente descritos): la evaluación de funciones extremadamente complejas y con gran cantidad de variables en juego no permitía con anterioridad la búsqueda de soluciones óptimas.

El Instituto Universitario SIANI de la ULPGC trabaja desde hace años en el diseño óptimo de estructuras, como una de sus líneas de investigación, en la que se enmarca la actividad investigadora de los tutores de este TFM. En este sentido, se pueden citar diversas publicaciones de relevancia que desde el año 2000 sitúan al SIANI en el campo del diseño óptimo de estructuras mediante el empleo de computación evolutiva. Algunas referencias son [1], [2], [3], [4] Y [5].

Los algoritmos genéticos son optimizadores globales, robustos y con capacidad de búsqueda de soluciones en pos del mínimo global del problema (función objetivo). Estos algoritmos aseguran una exploración amplia de soluciones que evita en muchas ocasiones que la solución no sea un mínimo local y explore el campo de soluciones hacia otros posiblemente mejores.

Para ello, codifican en un cromosoma toda la información variable del problema (con codificación binaria), que sirve como transportador de la información a lo largo de la función objetivo. El resultado final que se obtenga de la función objetivo para un cromosoma o individuo concreto sirve como retroalimentación para generar un nuevo individuo, con la intención de que éste mejore el resultado obtenido respecto al anterior. El empleo de la codificación binaria para representar en el cromosoma cada una de las variables en juego permite desligar la naturaleza de la variable del problema, generando una herramienta de cálculo basada en una cadena de compuesta de ceros y unos independiente de la tipología de problema, variables que intervienen, soluciones posibles a tales variables, etc. De forma sucinta, se puede decir que se asigna un valor

decimal a cada una de las posibles soluciones a cada variable, para que en el cromosoma esa solución se vea representada por el número binario correspondiente al valor decimal.

Sin embargo, el empleo de un código digital no homogéneo como el binario puede generar puntos críticos de cálculo, en los cuales dos soluciones reales consecutivas no presentan tal cercanía desde el punto de vista de la similitud genética. Concretamente, y centrando el ejemplo en el problema que se desarrolla en este trabajo, si en una barra concreta el perfil representado por el número decimal 7 (en binario de 4 bits sería 0111) se acerca a cumplir con todas las restricciones impuestas para la barra pero no llega a cumplir alguna de ellas por muy poco, lo lógico sería que el cromosoma mute a uno ligeramente diferente para intentar acercarse a una solución cercana que sí cumpla. Sin embargo, el perfil inmediatamente superior se encuentra representado por el decimal 8 (en binario, 1000), por lo que será complicado que el algoritmo genético entienda que estaba muy cerca de la solución óptima, pero para alcanzarla necesita un cambio completo del gen. Este es el principio en el cual se sustenta el trabajo aquí planteado, y es estudiar y comparar diferentes codificaciones en cálculos estructurales con estrategia discreta, con el fin de comprobar la influencia que tiene la codificación empleada en el resultado óptimo alcanzado, así como en el tiempo empleado para llegar al valor óptimo.

La *hipótesis de partida* es que efectivamente existe una influencia del código en este sentido, si bien se quiere cuantificar la misma para diferentes codificaciones y también buscar posibles parámetros representativos de la codificación que se relacionen con el mejor o peor comportamiento de la misma en el algoritmo genético de optimización.

3

1.1. Problema propuesto

En el *gráfico 1* se muestra la estructura planteada como caso de estudio en el presente trabajo, que ya fue propuesto previamente por D. Greiner, JM. Emperador y G. Winter en [3] como caso de estudio para optimización de estructuras mono y multiobjetivo. En el gráfico se incluye un esquema de la estructura con la numeración de las barras, nodos así como las cargas puntuales (expresadas en toneladas). En todas las vigas existe una carga uniformemente repartida de 4.076 T/m.

La luz de todas las vigas es de 5,6m y la altura de pilares de 2,80 m. Se consideran perfiles HEB para los pilares e IPE para las vigas, siendo las tensiones admisibles de 2,0 T/cm² y 2,2 T/cm² respectivamente. Se tiene en cuenta asimismo una restricción de desplazamiento en el punto medio de todas las vigas de 1/300=1,86 cm. La densidad y módulo de elasticidad E son las del acero: 7,85 T/m³ y 2100 T/cm².

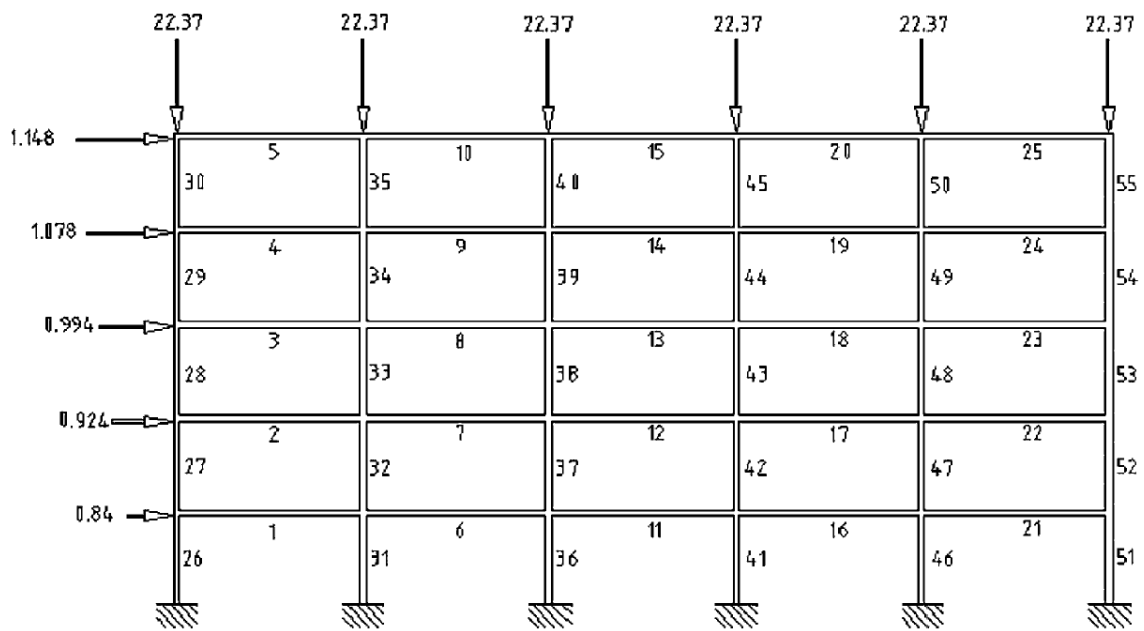


Gráfico 1. Estructura metálica de nudos rígidos objeto de estudio.

Esta estructura ha sido calculada mediante el empleo de un programa desarrollado por el Instituto Universitario SIANI, que tal y como ha sido mencionado con anterioridad, cuenta con una línea de investigación específica para el diseño óptimo de estructuras. El programa empleado para implementar la optimización mediante métodos evolutivos, y que será expuesto en los siguientes apartados, también ha sido diseñado por el propio instituto.

2. ALGORITMOS GENÉTICOS en la resolución de problemas estructurales

Los algoritmos genéticos se basan en el proceso evolutivo de la naturaleza, del que recogen los conceptos generales para, a través de un proceso evolutivo, partir de una solución inicial y “evolucionar” hasta la solución que cumpla con los objetivos propuestos. La evolución humana se puede representar como la reproducción de seres que, a partir de la recombinación y mutación de su *código genético o ADN*, evolucionan en base a la selección natural, donde priman las formas que mejor se adaptan al medio natural y fracasan en la evolución aquellas que no se adaptan correctamente al medio.

Para la implementación de los algoritmos genéticos en optimización es necesario codificar la información concerniente al problema en un *cromosoma*, de tal manera que el algoritmo use este cromosoma para extraer la información necesaria para resolver la función objetivo. El resultado obtenido de la función objetivo o función aptitud y su cercanía al resultado óptimo que se busca se empleará para retroalimentar la información sobre el cromosoma empleado y promocionarlo hacia la siguiente generación de tal manera que “mute” hacia otro que pueda proporcionar un mejor resultado de la función objetivo. Sin embargo estos algoritmos no trabajan con un único cromosoma, sino con un conjunto de individuos, de mayor o menor tamaño, llamado *población*.

El esquema general de un proceso genético se resume a continuación:

```
Crear la población inicial (P);
Evaluar mediante la función objetivo la población inicial (P);
Mientras (no condición de terminación)
{
    P1=seleccionar probabilísticamente (P);
    P2=cruzar probabilísticamente(P1);
    P3=mutar probabilísticamente(P2);
P4=crear nueva población (P, P3);
P=P4;
Evaluación de P;
Verificar condición de terminación;
}
```


2.1. Conceptos

Los operadores empleados en cada una de estas etapas son: operador de selección, cruce y mutación. A continuación se resumen algunos conceptos relacionados con la ejecución de algoritmos evolutivos, entre ellos los operadores anteriormente nombrados:

Cromosoma: El cromosoma contiene la información variable del problema, que coincide por tanto con las incógnitas que desean ser resueltas aportando una solución óptima del mismo. El resto de información no variable (parámetros) está reflejada en la función objetivo, que "captura" del cromosoma cada trozo correspondiente a la información de cada variable y rellena los valores en la función objetivo para que pueda ser evaluada. Es necesario que el cromosoma tenga la información organizada, de tal modo que el algoritmo sepa qué trozo del cromosoma pertenece al valor de cada variable. También es importante que cuando el cromosoma varíe en el proceso evolutivo, siempre se produzcan nuevos individuos coherentes con el problema, y que puedan ser correctamente evaluados.

La codificación del tradicionalmente empleada en algoritmos evolutivos es la binaria, siendo esta la de menor cardinalidad, por lo que permite que la búsqueda y evolución de la información a través de los criterios de evolución sea independiente a la naturaleza de la variable, lo cual permite que puedan ser aplicados en multitud de problemas, incluso cuando en un mismo problema intervienen variables de distinta naturaleza.

Selección: la operación de selección consiste en elegir en base a un criterio establecido aquellos individuos de la población anterior que mejor se hayan adaptado al medio, es decir, que mejor respuesta tengan de la función aptitud, para que sirvan como población intermedia para generar la nueva población filial. Para ello, primero hay que asignar las probabilidades individuales de selección y después aplicarlas. Existen diferentes procedimientos para ello: de manera proporcional al valor de la función objetivo, ordenadamente en función del valor de la función objetivo, según procedimiento de selección escalada, selección uniforme (donde no se tiene en cuenta el valor de la función aptitud)... Cuando ya han sido asignadas las probabilidades, entonces hay que escoger los individuos en función de ellas. Existen también diferentes procedimientos para ello, algunos son. Uno de los procedimientos más comúnmente empleados es el procedimiento de ruleta (De Jong, 1975), que mediante la generación de números aleatorios y buscando aquel individuo con probabilidad lo más cercana posible al número aleatorio. La selección universal estocástica tiene un procedimiento similar, ya que a partir de un solo número aleatorio se generan el resto de números que buscar en las probabilidades ya asignadas). Otro

procedimiento ampliamente empleado es la selección por torneo (Brindle, 1981), donde primero se escoge aleatoriamente una cantidad de individuos (población de torneo) y "enfrentando" a dos individuos al torneo, se escoge aquel de los dos que tenga mejor valor de la función aptitud.

Cruce: el operador de cruce se emplea para intercambiar información entre individuos de la población, creando de ese modo nuevos individuos a partir de los seleccionados. Este operador se aplica sobre los individuos que han sido seleccionados, lo cual implica que el cruce de información se hace entre individuos que poseen un valor óptimo de función aptitud. Existen diferentes procedimientos de cruce, como el cruce un punto, el cruce dos puntos o el uniforme. En el primero se establece un punto del cromosoma en el cual los dos individuos de la población intermedia escogidos son partidos y cruzados entre sí, generando dos nuevos individuos que llevan parte del cromosoma de cada uno de sus progenitores. El cruce dos puntos funciona exactamente igual, sólo que este cruce se realiza en dos puntos. Finalmente, el cruce uniforme establece el cruce de los cromosomas en función de los espacios existentes entre genes, y mediante una probabilidad de cruce uniforme.

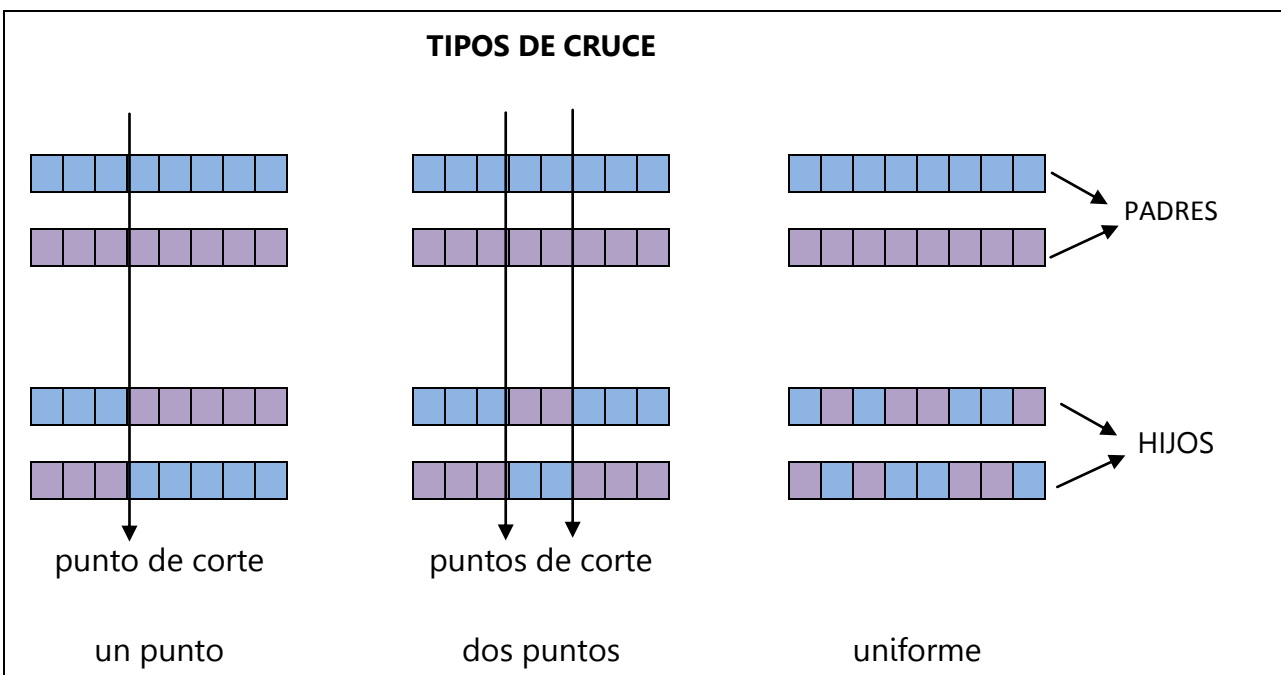


Gráfico 2. Ejemplos gráficos de tipos de cruce.

Los tipos de cruce un punto y uniforme son los más distintos por su modo de actuar. Un análisis pormenorizado de su modo de actuación hace ver que el cruce 1 punto favorece la preservación y propagación de bloques de construcción en el cromosoma de corta longitud, mientras que es propenso a destruir los de larga longitud. El cruce uniforme es más disruptivo

con el cromosoma, por lo que son útiles en poblaciones pequeñas para generar importantes transformaciones del cromosoma que permitan una mayor exploración de posibles soluciones.

Mutación: este operador también tiene como fin aumentar el espacio de búsqueda, realizando leves modificaciones en los descendientes generados tras el cruce de progenitores, y en el cual se invierten unos pocos genes (cambiar 0 por 1 y viceversa). Teniendo en cuenta que se trata de una mutación inducida sobre los genes que no parte de otro individuo que aporte ese trozo de información, probabilidades de mutación altas pueden desvirtuar la población filial generada, produciendo campos de búsqueda excesivamente exploratorios.

La definición de los parámetros anteriores guiará la evolución del resultado del algoritmo hacia espacios de búsqueda diferentes, o limitará este entre unos pocos espacios limitados por poblaciones pequeñas o poco distribuidas. Este problema, que se conoce como el equilibrio entre la exploración y la explotación, debe ser evaluado en el proceso de ejecución, con el fin de encontrar dicho equilibrio [8]. Los conceptos que expresan el nivel de exploración de un algoritmo, o la explotación de los mismos individuos o individuos muy similares se conocen como *diversidad poblacional* y *presión de selección*. Pequeñas poblaciones y porcentajes de mutación muy pequeños pueden llevar a una alta presión de selección, o lo que es lo mismo, a la limitación del campo de exploración a un foco concreto del espacio de búsqueda. Sin embargo, grandes valores de mutación y poblaciones podrían generar un alto coste computacional invertido en exploración de grandes campos. El equilibrio debe establecerse en función de las características de cada problema.

Definición de función objetivo: la función objetivo es aquella que engloba la información necesaria para que el algoritmo genético efectúe la búsqueda del óptimo (individuo que aporta una solución óptima de la función objetivo). Esta función mide la adaptación del individuo al entorno del problema, lo que también se conoce como la aptitud del individuo en la función.

2.2. Codificación del cromosoma en el problema estructural

En primer lugar se define a continuación toda la información que constituye y define el problema descrito y optimizado en el presente trabajo [3]:

- Coordenadas nodales de la estructura
- Conectividades entre barras

- Características geométricas de la sección de la barra (área, momento de inercia, módulo de rigidez, radio de giro, área del alma, etc.).
- Características de material/es de las barras.
- Vínculos de apoyo o asentamientos de ciertos nodos.
- Casos de carga de la estructura
- Hipótesis de carga de la estructura
- Peso de la estructura
- Número de perfiles diferentes utilizados
- Límites máximos tensionales, de esbeltez o desplazamiento de ciertos puntos estructurales admisibles.

La conjunción de todos los parámetros anteriores proporciona una estructura concreta, cuya solución puede o no ser válida (en caso de que en algún punto de la misma no se respeten los valores máximos admisibles para alguna de las restricciones establecidas sobre la estructura según normativa) y óptima (si, en caso de cumplir con las restricciones, la estructura está sobredimensionada y el coste de la misma supera el mínimo óptimo existente pero no localizado). Si se observa el listado de parámetros anterior se puede concluir que de manera general, el diseño de una estructura metálica de nudos rígidos implica que todos los parámetros excepto las características geométricas de las barras serán parámetros de definición del problema (barras, sus nudos, conectividades y apoyos, cargas e hipótesis de carga de la misma...). Es por ello que en el problema que aquí se resuelve, la optimización de los perfiles escogidos para cada una de las barras es el objetivo de la misma, para lo cual se plantea una función aptitud que, bajo el cumplimiento de todas las restricciones, reporte el peso mínimo de la estructura metálica diseñada.

La codificación genética de cada uno de los individuos o soluciones candidatas contendrá la información necesaria para la definición de las características geométricas citadas de cada una de las barras, que es aquella información variable que distingue una solución de otra. Concretamente, para la optimización discreta del problema empleando perfiles reales contemplados según el Código Técnico de la Edificación (CTE), se hace necesario describir las características de cada uno de los perfiles comerciales que pretenden emplearse para que puedan ser seleccionados. De esta manera, el cromosoma contendría la información necesaria para referenciar qué perfil se asigna a cada una de las barras que conforman la estructura, y es el algoritmo de cálculo de la estructura el que se encarga de cargar desde un fichero las características geométricas que describen cada perfil para calcular la función objetivo. Por tanto, en el caso que se calcula en este problema, se establece una estructura formada por 55 barras

donde se establecen un total de 16 perfiles comerciales como las posibles soluciones. Por tanto, y conociendo que para representar cada una de las 16 barras mediante un código binario por 4 bits, el tamaño total del cromosoma será de:

$$4\text{bits/perfil} \times 55 \text{ barras} = 220 \text{ bits/estructura.}$$

2.3. Definición de la función objetivo

El problema estructural planteado se basa en la optimización del peso total de la función (monoobjetivo), restringido a partir de condiciones tensionales de la estructura. La minimización del peso estructural responde a la premisa de minimizar el coste de adquisición de materia prima de la estructura diseñada, puesto que en el caso de estructura metálica, este coste está en correlación directa con el peso total de la misma. Esta minimización de peso tiene como límite aquellas condiciones que hagan cumplir al pórtico diseñado con su función: servir de soporte a las cargas objeto de diseño sin colapsar o deformarse excesivamente [3].

La minimización del peso restringido constituye un problema condicionado, ya que está sometido a una serie de restricciones que dependen de la hipótesis de carga que sobre la estructura se plantea. Las restricciones consideradas de manera genética son las siguientes:

Tensión de comparación menor a la tensión límite (generalmente es de 2600 kgp/cm² en estructuras de acero, por ser el límite elástico):

$$\sigma_{co} - \sigma_{lim} \leq 0$$

Esbeltez lambda a compresión menor a 200 (límite impuesto por normativa):

$$\lambda - \lambda_{lim} \leq 0$$

Desplazamientos límites (en cada uno de los tres grados de libertad posibles) determinados en ciertos puntos, nudos o eventualmente centros de vanos en vigas:

$$u_{co} - u_{lim} \leq 0$$

Con todo ello la función objetivo peso restringido resulta ser la siguiente:

$$F_{objetivo} = \left[\sum_{i=1}^{Nbarras} A_i \cdot \rho_i \cdot l_i \right] \cdot \left[1 + k \cdot \sum_{j=1}^{Nviolac} (viol_j - 1) \right]$$

Dónde:

A_i : área de sección de barra i .

ρ_i : densidad de barra i .

l_i : longitud de barra i .

k : constante que regula la equivalencia entre restricción y peso, y que para la relación adoptada es de 0,6.

$Viol_j$: para cada una de las restricciones violadas, es la relación entre el valor que supone violación de la restricción (sea, tensión, desplazamiento o esbeltez) y su límite de referencia.

El método escogido para **condicionar** el resultado de la función objetivo según los resultados obtenidos para cada una de las restricciones del problema es el método de la *función penalty* [9]. Este método transforma un problema sometido a restricciones en otro sin ella, incorporándolas como un costo a la función objetivo cuando una de las restricciones es violada (o varias). De esta manera, se establece para cada una de las restricciones como será penalizada sobre el peso total de la estructura, de tal manera que cuanto más se viole la restricción mayor sea la penalización, y que el alto valor final de la función aptitud sirva para retroalimentar el problema y generar un cromosoma algo más alejado del actualmente ensayado. Si por el contrario se descartase un cromosoma por violar una restricción sin cuantificar sobre la función aptitud cómo de grave es tal violación, se estaría desechando la posibilidad de obtener información útil sobre cómo de apta o no apta se la solución sobre la función objetivo.

Ya que se ha planteado un algoritmo genético de minimización, no es necesario transformar la función objetivo (función de adaptación) para que el objetivo de la función sea la minimización y no la maximización.

3. CODIFICACIÓN DEL CROMOSOMA. CÓDIGOS BINARIOS

En un ambiente de sistemas digitales se denomina codificación a la asignación de un significado a una configuración de bits.

Al modelar problemas es usual encontrar variables que pueden tomar múltiples valores, se denomina codificación al proceso de convertir esas variables en códigos más simples, y en el caso de los algoritmos evolutivos, a códigos más consecuentes con el proceso de evolución natural. Generalmente son los códigos binarios (de base 2, {0,1}), los más usualmente empleados, si bien existen soluciones que requieren de otro tipo de codificaciones para poder ser correctamente resueltos.

3.1. Códigos binarios

12

Los códigos binarios se formulan a partir de una base de tamaño 2 {0,1}, que, mediante la combinación de estos valores genera una representación equivalente a una palabra en el código original o de partida al que "sustituye". En el caso concreto de la representación de números decimales mediante códigos binarios, cabe destacar que la capacidad de creación que tiene un código tiene una relación de 2^n , donde n es el número de bits empleados. Es decir, que para un código de 4 bits, se posible crear hasta un total de **$2^4=16$ códigos representativos diferentes**.

Un código binario es continuo, si las combinaciones correspondientes a números decimales consecutivos son adyacentes, es decir, aquellas que varía solo en un bit en códigos consecutivos. Un código binario es cíclico cuando además la última combinación es adyacente a la primera. El código de Johnson es de tipo binario, cíclico continuo y progresivo. Que sea progresivo limita que el código de Johnson vea reducida su capacidad de generar códigos diferentes, de tal manera que sólo puede representar un total de **$2 \cdot n$** códigos. En el caso concreto de n=4 bits, el código de Johnson solo genera 8 códigos, frente a los 16 que puede generar uno no progresivo (sería necesario emplear un código de 8 bits para poder representar 16 valores diferentes mediante codificación tipo Johnson). Tal es el caso del código Gray, que es binario, cíclico y continuo, pero no progresivo. Se hablará en detalle sobre esta codificación.

3.1.1. Códigos ponderados

Se denomina código ponderado a uno en el que a cada posición de la secuencia binaria se le asigna un peso o ponderación. El código ponderado más empleado es el **BCD**, que representa a los dígitos decimales por secuencias de bits en sistema binario[9]. BCD es el acrónimo de **Binary Coded Decimal**, es decir, decimal codificado en binario. En la Tabla 1 se puede ver la equivalencia entre los 16 primeros números decimales y su equivalencia en binario (sobre 4 bits).

La posición de más a la izquierda es la más significativa, y tiene peso 8. El bit menos significativo, el de la derecha, tiene ponderación 1. Como las ponderaciones son potencias de 2, se puede comprobar la equivalencia entre el código binario ponderado y el decimal mediante el siguiente cálculo:

Para $n=7$:

$$0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 4 + 2 + 1 = 7$$

3.1.2. Códigos no ponderados

Los códigos no ponderados se basan en la ordenación de una serie de números binarios en función de algún criterio, sin que en este tenga que existir una asignación de pesos o ponderaciones para cada posición del bit. Uno de los códigos no ponderados más habituales es el código Gray, que además es de tipo reflejado [11]. La característica peculiar que destaca de una codificación de Gray es que entre dos números consecutivos binarios sólo existe un bit de diferencia (adyacente), de tal manera que la transición de un número al siguiente es bastante homogénea. Esto implica que en codificaciones Gray, los códigos de elementos del dominio que están cerca (*fenotipo*) también están cerca en su código (*genotipo*). Una de las aplicaciones más comunes de los códigos Gray es en el control de sistemas mecánicos, donde simplifica enormemente la ejecución de una tarea de manera progresiva (por ejemplo, el movimiento de un engranaje).

Los códigos reflejados presentan una inversión del código a partir de la mitad de su generación, es decir, que para escribir el segundo 50% del código, se emplea el primer 50% invertido (cambiando unos por ceros y viceversa) y reflejado (la posición 0 invertida en la n , la posición 1 invertida en la $n-1$...):

Gray	Binario Pesado	Nº decimal
0000	0000	0
0001	0001	1
0011	0010	2
0010	0011	3
0110	0100	4
0111	0101	5
0101	0110	6
0100	0111	7
1100	1000	8
1101	1001	9
1111	1010	10
1110	1011	11
1010	1100	12
1011	1101	13
1001	1110	14
1000	1111	15

Tabla 1. Comparación entre código Gray y Binario pesado para los 16 primeros números enteros.

Tal y como puede verse en [4], el empleo de diferentes códigos deriva en la asignación de una representación binaria diferente para cada una de los valores discretos que puede tomar cada una de las variables implicadas.

Estos códigos no ponderan el valor de cada posición mediante una función de exponencial o de otro tipo, de tal manera que no existe una función matemática que permita calcular a qué número decimal corresponde un código binario. Esto puede suponer, según el uso que se pretenda hacer del código, un problema, ya que si la variable representada necesita de un genotipo muy largo, el "traductor" de un código a otro debe contener almacenado toda la cadena y buscarla para traducirla, no puede emplear una función matemática que la convierta. Sin embargo, las aplicaciones generales para las que se aplican este tipo de códigos se basan precisamente en la simplicidad, y en la necesidad de unos pocos códigos que representen un número acotado de variables, por lo que el número de bits es más que manejable.

3.2. Aplicación de diferentes códigos binarios al problema presentado

Tal y como puede verse en la *tabla 1*, para un mismo orden de números decimales (del 0 al 15) se establecen diferentes códigos binarios para representar un mismo valor decimal,

representando en cada caso las reglas del código en cuestión. Teniendo en cuenta que el objetivo de este trabajo es comprobar si el empleo de diferentes códigos binarios, con diferentes estructuras, influyen en el resultado final obtenido en el proceso de optimización de una estructura, así como en la evolución del algoritmo genético de cálculo a lo largo de las evaluaciones, en este apartado se pretende establecer unas bases o principios que sustenten el procedimiento experimental que se ha seguido, así como las justificaciones que han llevado al mismo.

Variables a representar en el cromosoma

En el problema planteado, el cromosoma debe albergar, de manera ordenada, el perfil asignado en cada individuo a cada una de las 55 barras que contiene la estructura. Tal y como se calculó con anterioridad, esto corresponde a un total de 220 bits/estructura, teniendo en cuenta que para cada barra existirían 16 perfiles posibles (y que por tanto cada tipo de perfil quedaría perfectamente definido e identificado con un código de 4 bits).

Como aclaración cabe destacar que, aunque existan tanto perfiles tipo IPE como HEB, las primeras sólo son válidas para las vigas y las segundas, para los pilares, con lo cual cada barra sólo tiene 16 soluciones posibles y como cada barra está perfectamente identificada con su "trozo" de cromosoma, se descarta toda posibilidad de fallo por el hecho de que dos perfiles tengan igual codificación, ya que en ningún caso serán competencia en una misma barra.

Homogeneidad en la representación de soluciones

Tal y como se ha venido expresando a lo largo de los anteriores apartados, uno de los inconvenientes de usar el código binario natural (y más ampliamente extendido) es que en algunos puntos presenta dos códigos consecutivos muy divergentes, en ocasiones incluso opuestos, por lo que la correlación natural existente entre números decimales consecutivos no queda fielmente reflejada en su código genético representativo. Analizando este hecho desde el visor de la evolución en la naturaleza, que funciona por imitación de aquello que mejor se adapta y desechando lo que no, resulta fácil pensar que estos puntos puedan resultar críticos a la hora de optimizar un problema: el algoritmo evolutivo tenderá a crear nuevos individuos con cromosomas muy parecidos a aquellos que han resultado buenos, y sin embargo no se estará dando cuenta que la solución óptima y consecutiva está en un código muy divergente. En resumen, cabe pensar que codificaciones más cercanas a la evolución lógica y progresiva (de transición lenta o suave) puedan reportar resultados mejores.

Medidores de la variabilidad en la codificación

Se establecen en este sentido posibles hipótesis para evaluar el comportamiento de una codificación. En este trabajo se ha optado por caracterizar cada codificación a partir de:

- N° de bits diferentes entre códigos sucesivos.
- Suma total de bits diferentes en la codificación, valor medio y varianza.
- Cantidad de variaciones totales de código (diferencias de 4 bits existentes).

3.3. Creación y representación de códigos binarios diferentes

El estudio de la combinatoria permite saber que, para un total de 16 números diferentes, existe un total de 16! (factorial de 16) posibilidades diferentes de combinación. Esto supone una cantidad total de 20.922.789.888.000 combinaciones posibles (casi 21 billones). Por lo que si una vez se crean N combinaciones diferentes de los 16 números decimales comprendidos entre el 0 y el 15, y posteriormente se transforman a código binario natural, se obtendrían N códigos diferentes para representar a las mismas 16 variables.

Tal y como se puede ver en la tabla 1, el código Gray consiste en asignar un código binario según unas reglas para cada uno de los números comprendidos entre el 0 y el 15, ordenados según el orden decimal lógico. Sin embargo, no existen reglas para generar todos y cada uno de los 21 billones de códigos posibles que pueden obtenerse. Es por ello que la estrategia que se ha planteado en el presente trabajo para encontrar diferentes codificaciones para emplearlas en la optimización propuesta sea la siguiente:

16

3.4. Estrategia empleada para la generación de codificaciones

Se pretende dar representación mediante diferentes códigos binarios a un total de 16 perfiles metálicos que están ordenados naturalmente según el aumento de su área, lo que es sinónimo de robustez, y también de aumento de peso. Se desea que este orden creciente naturalmente establecido, se vea representado por diferentes códigos binarios de distinta naturaleza. Para ello se sabe que cada perfil debe ir identificado por un número decimal diferente, y que es necesario emplear un lenguaje de transformación entre cada número decimal y el código binario en cuestión. Ya que este sólo existe para unos pocos códigos binarios, lo que se plantea es que la

asignación de números decimales a cada tipo de perfil no tenga que ser estrictamente creciente, sino que esta asignación sea un mero enlace entre el perfil comercial y su código binario correspondiente. Esto permite que se emplee un único lenguaje de transformación entre el número decimal y el binario, que sea el más extendido y matematizable: el binario natural.

Empleando esta estrategia, se pueden generar multitud de códigos binarios mediante la reordenación aleatoria de los 16 números decimales, y convirtiéndolos posteriormente a decimal. A modo de ejemplo se muestra a continuación una comparación entre dos códigos de Gray, expresados de diferente manera:

Código binario de tipo Gray	Traducción mediante código Gray	Traducción mediante código binario natural
0000	0	0
0001	1	1
0011	2	3
0010	3	2
0110	4	6
0111	5	7
0101	6	5
0100	7	4
1100	8	12
1101	9	13
1111	10	15
1110	11	14
1010	12	10
1011	13	11
1001	14	9
1000	15	8

Tabla 2. Diferentes traducciones de código binario a decimal.

El objetivo del presente trabajo, y en general de la investigación en pos de mejorar las técnicas de optimización evolutiva, se fundamentan en que la representación binaria de la variable sea congruente y óptima con el resultado obtenido, y trasladado a el problema estructural que aquí se plantea, la relación entre la representación binaria de cada perfil, por lo que el número decimal asignado a cada perfil funciona como mero transmisor o conector entre ambos lenguajes. Este es el fundamento que se ha establecido para realizar el trabajo presentado.

Se ha optado por realizar primero un muestreo sobre todas las ordenaciones posibles (para lo cual se ha empleado un procedimiento de generación de ordenaciones aleatorio), y de la

evaluación de los resultados obtenido se han seleccionado gran parte de las ordenaciones finalmente empleadas. Otras, sin embargo, han sido escogidas por su especial relevancia, independientemente de que fuesen o no generadas en el muestreo.

```

int numeros []={0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15};
for (i=0; i<16; i++)
{
a=rand()%n;
codigo_int[i]=numeros[a];
for (k=a; k<n-1; k++)
{
numeros[k]=numeros[k+1];
}
n-=1;
}
    
```

Código en ANSI C para la generación de ordenaciones aleatorias.

El muestreo realizado fue de un total de 20.000.000 de ordenaciones generadas, lo que corresponde aproximadamente con una millonésima parte del total de posibilidades existentes. La distribución de las ordenaciones respecto a la suma total de bits diferentes existentes en el código se grafican en el siguiente diagrama de barras:

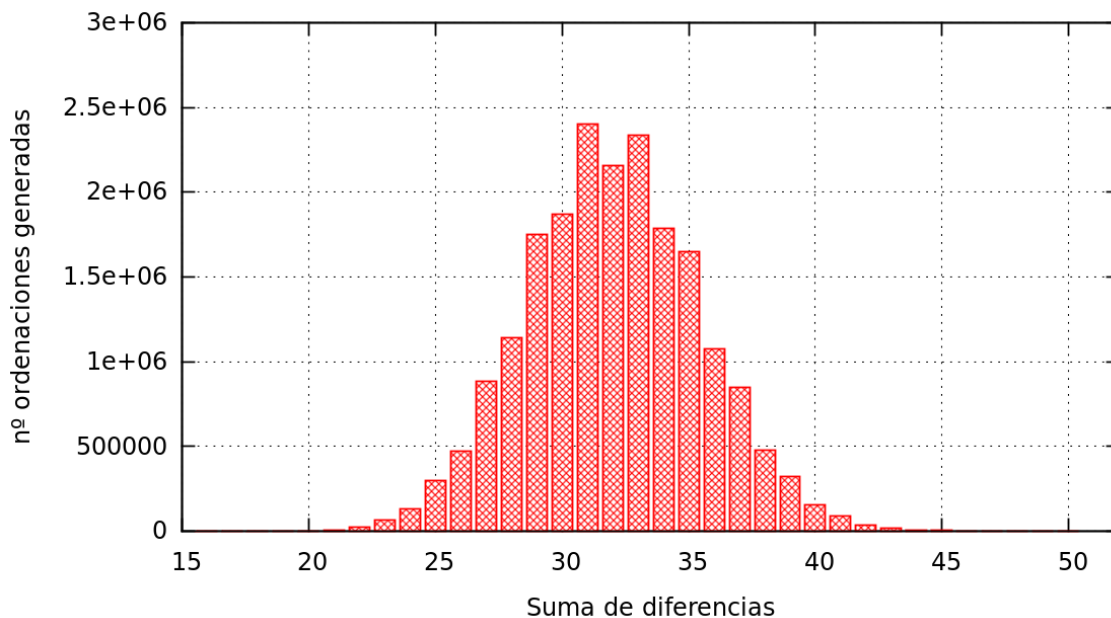


Gráfico 3. Diagrama de barras representando la distribución de los 20.000.000 de ordenaciones generados en función de la suma total de bits diferentes encontrados.

En la tabla 3 además se tabulan los datos representados en al gráfica anterior, así como los valores de varianza mínima y máxima encontrados para cada suma total de bits diferentes:

Suma	Cantidad ordenaciones calculadas	Var. min-Var. máx
16	3	0,25
17	18	0,34-0,50
18	90	0,40-0,54
19	499	0,44-0,77
20	1978	0,47-0,87
21	7886	0,49-0,88
22	21076	0,50-1,02
23	63850	0,50-1,09
24	128862	0,50-1,08
25	298292	0,47-1,14
26	474079	0,44-1,18
27	884366	0,40-1,33
28	1142700	0,34-1,31
29	1748322	0,25-1,34
30	1870579	0,37-1,41
31	2399617	0,25-1,39
32	2158419	0,50-1,36
33	2335619	0,40-1,33
34	1786698	0,44-1,39
35	1647188	0,47-1,40
36	1075855	0,49-1,40
37	847975	0,50-1,45
38	475243	0,50-1,41
39	322039	0,49-1,40
40	155023	0,47-1,35
41	89100	0,44-1,34
42	36705	0,40-1,28
43	18011	0,34-1,26
44	6317	0,25-1,18
45	2534	0,37-1,21
46	721	0,44-1,18
47	263	0,50-1,02
48	55	0,40-0,91
49	15	0,57-0,93
50	3	0,7

Tabla 3. Resumen de los resultados obtenidos para una generación de 20.000.000 de ordenaciones.

Cabe destacar que:

- Se obtuvieron ordenaciones de hasta 35 sumas totales diferentes (desde 16 hasta 50). A su vez se analizó la varianza presentada en cada caso, lo cual ascendió a un total de 390 grupos diferentes de ordenaciones que compartían igual suma y varianza en las diferencias de bits entre números consecutivos.
- No se obtuvo una sola ordenación de tipo Gray, en la cual la suma total de bits se de 15 (1 bit distinto en cada dos códigos consecutivos). Sí se obtuvieron 3 ordenaciones de suma total 16.
- En otro muestreo realizado para testear el correcto funcionamiento del programa (con una muestra menor), se llegó a obtener una ordenación con una suma total de 51. Fue la más desfavorable obtenida desde este punto de vista.
- Tan sólo se obtuvo 1 ordenación con suma total igual a 50.
- Destaca la baja presencia de ordenaciones con altas varianzas (entre 1 y 1,4 bits de varianza, dependiendo de la suma total).

20

Finalmente, se resume a continuación brevemente algunos criterios seguidos para seleccionar hasta un total de 20 ordenaciones (o codificaciones) que emplear en la fase experimental de este trabajo. No obstante en el apartado de Desarrollo experimental se encuentran las 20 ordenaciones perfectamente caracterizadas:

- Código binario natural. Es un estándar en este tipo de problemas por lo que debe ser comparado.
- 3 códigos Gray diferentes: no sólo se quiere contrastar si el Gray reporta mejores resultados que otras codificaciones, sino además si existe diferencia significativa entre varias codificaciones Gray diferentes.
- 4 codificaciones con alto número de variaciones: se seleccionan varias codificaciones de las obtenidas en el muestreo aleatorio realizado, con suma total igual a 46, 49 y 50. También se selecciona la que durante toda la realización del trabajo ha sido la ordenación con mayor suma de bits distintos, cuyo valor es de 51.
- 3 codificaciones que han sido denominadas "Antigray". En este caso se partió de un código Gray y mediante la inversión de códigos alternativos (cambiar ceros por unos y

viceversa en uno de cada dos códigos), se generó otro que presentase una variación alta pero uniforme de bits. Concretamente presentan una variación de 3 bits de media, una suma total de 45 con varianza 0.

- 9 ordenaciones más, divididas de 3 en 3, para varios valores medios que cubriesen un espectro más amplio de sumas totales de bits distintos. Se escogieron 3 ordenaciones con suma 27, otras 3 con suma 33 y otras 3 con suma 39. En cada caso además se siguió otro criterio establecido en base a la varianza: 1 del grupo de menor varianza, otra del grupo de mayor varianza, y la tercera del grupo más cercano al valor medio entre la varianza mayor y la menor.

Una explicación pormenorizada de la fase experimental llevada a cabo, así como de las ordenaciones empleadas y de los resultados obtenidos pueden encontrarse en los siguientes apartados.

4. FASE EXPERIMENTAL

Tal y como se ha planteado con anterioridad, el objetivo de este trabajo se sustenta en la hipótesis de que el empleo de diferentes codificaciones numéricas para representar las posibles soluciones de las incógnitas del problema influye en el resultado final del mismo. En estudios previos ya se determinó que el empleo de una codificación tipo Gray presentaba, para el cálculo de una estructura metálica de nudos rígidos, mejores resultados en la función objetivo que si se empleaba la codificación binaria, tal y como puede verse en [4].

A partir de los cálculos expuestos en el apartado anterior sobre la generación de codificaciones aleatorias, se han seleccionado 20 codificaciones diferentes para su empleo en el cálculo del problema elegido. Estas codificaciones, que inicialmente se crean y almacenan como números decimales (entre el 0 y el 15), son posteriormente transformadas a código binario, y es sobre esta codificación sobre la que se tienen en cuenta los siguientes cálculos para su caracterización:

- Cálculo de bits diferentes entre números consecutivo.
- Suma total de diferencias existentes entre los 16 números que conforman el código, así como la media.
- Desviación estándar del número de bits diferentes en cada ordenación.
- Número de variaciones de 4 bits existentes en la codificación.

Cabe recordar en este punto que, dado que el código empleado tiene un total de 16 números, para representar a las 16 tipos de perfiles comerciales que se han incluido en el estudio (16 perfiles tipo IPE y otros 16 tipo HEB), la representación binaria de los números decimales enteros comprendidos entre el 0 y el 15 pueden ser representados con tan sólo 4 bits.

A continuación se muestra la subrutina empleada para la caracterización de cada codificación generada:

```

/*Cálculo de diferencia entre números consecutivos, suma, media y
varianza en una codificación*/
/*codigo_int[j] contiene la codificación en números enteros
generada*/
for (j=0; j<15; j++) /*Cálculo de diferencias totales en la
codificación*/
{
    diferencias[j]=(codigo_int[j]^codigo_int[j+1]);
    C= diferencias[j];
    dif=0;
    for (i=0; i<4; i++)
    {
        dif+=(C&1);
        cuenta+=(C&1);
        C>>=1;
    }
    vector_cuenta[j]=dif;
}
media=cuenta/15;
float suma_desviacion, desviacion;
suma_desviacion=0;
desviacion=0;

for (i=0; i<15; i++)
{
    suma_desviacion+=(vector_cuenta[i]-
media)*(vector_cuenta[i]-media);
}
desviacion=sqrt(suma_desviacion/15); /*desviación típica
poblacional*/

```

Código en ANSI C para caracterizar las ordenaciones según cantidad de bits diferentes en números consecutivos.

Las codificaciones elegidas se muestran en la siguiente tabla. La nomenclatura empleada para diferenciarlas se basa en la suma total de diferencias encontradas en los bits del código, (SXX), una letra que identifica si es la primera (A), segunda (B) o tercera (C) codificación que se ha escogido con un mismo valor de suma, y finalmente las primeras cifras significativas de la desviación estándar calculada para la suma total de diferencias (Y,YY). Finalmente, y para facilitar el estudio, estas codificaciones se han clasificado siguiendo un orden creciente de suma en primer lugar, y de varianza como segundo criterio. De esta forma se facilita la visualización de los resultados ya sean tabulados o graficados.

En la siguiente tabla se muestra cada una de las ordenaciones empleadas, descrita por:

- Número decimal, código binario correspondiente (a partir del código binario natural).
- Bits diferentes entre códigos consecutivos, suma total, y varianza.
- Número de variaciones de 4 bits detectadas

S15A0			S15B0			S15C0			S26A0,92			S27A0,40			S27B0,83			S27C1,33			S33A0,40			S33B0,91			S33C1,38		
0	0000		12	1100		8	1000		0	0000		3	0011		9	1001		9	1001		4	0100		6	0110		4	0100	
1	0001	1	13	1101	1	0	0000	1	1	0001	1	9	1001	2	10	1010	2	1	0001	1	1	0001	2	2	0010	1	0	0000	1
3	0011	1	15	1111	1	1	0001	1	2	0010	2	12	1100	2	11	1011	1	0	0000	1	7	0111	2	9	1001	3	15	1111	4
2	0010	1	14	1110	1	3	0011	1	3	0011	1	5	0101	2	15	1111	1	15	1111	4	13	1101	2	4	0100	3	13	1101	1
6	0110	1	10	1010	1	2	0010	1	4	0100	3	13	1101	1	12	1100	2	13	1101	1	0	0000	3	3	0011	3	2	0010	4
7	0111	1	11	1011	1	6	0110	1	5	0101	1	11	1011	2	8	1000	1	2	0010	4	5	0101	2	5	0101	2	9	1001	3
5	0101	1	9	1001	1	7	0111	1	6	0110	2	1	0001	2	0	0000	1	6	0110	1	6	0110	1	14	1110	3	8	1000	1
4	0100	1	8	1000	1	5	0101	1	7	0111	1	8	1000	2	7	0111	3	7	0111	1	15	1111	2	13	1101	2	12	1100	1
12	1100	1	0	0000	1	4	0100	1	8	1000	4	4	0100	2	2	0010	2	5	0101	1	12	1100	2	10	1010	3	3	0011	4
13	1101	1	1	0001	1	12	1100	1	9	1001	1	14	1110	2	4	0100	2	4	0100	1	11	1011	3	8	1000	1	11	1011	1
15	1111	1	3	0011	1	13	1101	1	10	1010	2	7	0111	2	13	1101	2	11	1011	4	2	0010	2	0	0000	1	7	0111	2
14	1110	1	2	0010	1	15	1111	1	11	1011	1	15	1111	1	5	0101	1	3	0011	1	8	1000	2	11	1011	3	6	0110	1
10	1010	1	6	0110	1	14	1110	1	12	1100	3	10	1010	2	3	0011	2	12	1100	4	14	1110	2	15	1111	1	14	1110	1
11	1011	1	7	0111	1	10	1010	1	13	1101	1	6	0110	2	6	0110	2	14	1110	1	9	1001	3	7	0111	1	1	0001	4
9	1001	1	5	0101	1	11	1011	1	14	1110	2	0	0000	2	14	1110	1	10	1010	1	3	0011	2	12	1100	3	5	0101	1
8	1000	1	4	0100	1	9	1001	1	15	1111	1	2	0010	1	1	0001	4	8	1000	1	10	1010	3	1	0001	3	10	1010	4
suma	15		suma	15		suma	15		suma	26		suma	27		suma	27		suma	27		suma	33,00		suma	33		suma	33	
var.	0		var.	0		var.	0		var.	0,92		var.	0,4		var.	0,83		var.	1,33		var.	0,40		var.	0,91		var.	1	
c.4	0		c.4	0		c.4	0		c.4	1		c.4	0		c.4	1		c.4	4		c.4	0		c.4	0		c.4	5	

Tabla 4. Codificaciones empleadas en este trabajo caracterizadas.

Columna1: n° decimal, Columna2: n° binario, Columna3: bits diferentes con el n° decimal anterior.

S39A0,49			S39B0,95			S39C1,40			S45A0			S45B0			S45C0			S46A0,92			S49A0,57			S50A0,47			S51A0,61		
3	0011		9	1001		7	0111		4	0100		10	1010		6	0110		13	1101		2	0010		14	1110		15	1111	
4	0100	3	13	1101	1	8	1000	4	9	1001	3	4	0100	3	8	1000	3	12	1100	1	12	1100	3	1	0001	4	0	0000	4
10	1010	3	7	0111	2	9	1001	1	7	0111	3	9	1001	3	5	0101	3	2	0010	3	7	0111	3	12	1100	3	14	1110	3
0	0000	2	12	1100	3	6	0110	4	0	0000	3	7	0111	3	11	1011	3	5	0101	3	8	1000	4	7	0111	3	1	0001	4
13	1101	3	6	0110	2	5	0101	2	14	1110	3	0	0000	3	12	1100	3	10	1010	4	15	1111	3	10	1010	3	13	1101	3
6	0110	3	8	1000	3	10	1010	4	3	0011	3	14	1110	3	2	0010	3	1	0001	3	0	0000	4	5	0101	4	2	0010	4
12	1100	2	4	0100	2	11	1011	1	13	1101	3	3	0011	3	15	1111	3	14	1110	4	14	1110	3	8	1000	3	5	0101	3
11	1011	3	3	0011	3	4	0100	4	6	0110	3	13	1101	3	1	0001	3	7	0111	2	1	0001	4	15	1111	3	10	1010	4
1	0001	2	15	1111	2	12	1100	1	8	1000	3	6	0110	3	10	1010	3	8	1000	4	6	0110	3	0	0000	4	4	0100	3
15	1111	3	0	0000	4	14	1110	1	5	0101	3	8	1000	3	4	0100	3	3	0011	3	9	1001	4	13	1101	3	11	1011	3
2	0010	3	2	0010	1	1	0001	4	11	1011	3	5	0101	3	9	1001	3	6	0110	2	4	0100	3	3	0011	3	6	0110	3
8	1000	2	5	0101	3	3	0011	1	12	1100	3	11	1011	3	7	0111	3	9	1001	4	3	0011	3	4	0100	3	9	1001	4
5	0101	3	10	1010	4	2	0010	1	2	0010	3	12	1100	3	0	0000	3	15	1111	2	13	1101	3	11	1011	4	7	0111	3
9	1001	2	1	0001	3	13	1101	4	15	1111	3	2	0010	3	14	1110	3	0	0000	4	11	1011	2	6	0110	3	12	1100	3
14	1110	3	14	1110	4	0	0000	3	1	0001	3	15	1111	3	3	0011	3	11	1011	3	5	0101	3	9	1001	4	3	0011	4
7	0111	2	11	1011	2	15	1111	4	10	1010	3	1	0001	3	13	1101	3	4	0100	4	10	1010	4	2	0010	3	8	1000	3
suma	39		suma	39		suma	39		suma	45		suma	45		suma	45		suma	46		suma	49		suma	50		suma	51	
var.	0,49		var.	0,95		var.	1,40		var.	0		var.	0		var.	0,92		var.	0,57		var.	0,47		var.	0,61		var.	0,61	
c. 4	0		c. 4	3		c. 4	7		c. 4	0		c. 4	0		c. 4	0		c. 4	6		c. 4	5		c. 4	5		c. 4	6	

Tabla 4 (bis). Codificaciones empleadas en este trabajo caracterizadas.

Columna1: nº decimal, Columna2: nº binario, Columna3: bits diferentes con el nº decimal anterior.

4.1. Problema empleado. Variables seleccionadas.

El problema finalmente seleccionado para su resolución como caso de estudio ya fue inicialmente descrito en el apartado de Presentación del Problema. Se trata de una estructura metálica de nudos rígidos compuesta por un total de 55 barras, con cargas puntuales, así como cargas uniformemente distribuidas. En el gráfico 1 puede visualizarse la estructura, con la numeración de barras y nudos empleada en el programa de cálculo de la misma. La estrategia elegida es de tipo real (discreta), con lo cual los posibles perfiles que pueden emplearse en cada una de las barras son perfiles comerciales, y los valores de los parámetros empleados para caracterizar cada barra, y por tanto para calcular la estructura, son discretos y no continuos. La función objetivo planteada es la de peso restringido como solución del problema.

Respecto al algoritmo de cálculo evolutivo empleado, se trata de un algoritmo genético generacional monoobjetivo desarrollado en el Instituto Universitario SIANI. Los parámetros de cálculo que han sido variados en el estudio son: tipo de cruce del gen (cruce de tipo UNIFORME, o cruce en 1 PUNTO), tamaño de la población empleado (50 y 150 individuos), y probabilidad de mutación del gen 0,4% y 0,8%. Estos tres parámetros, probados cada uno con dos valores diferentes, genera un total de:

2 tipos de cruce x 2 tamaños de población x 2 porcentajes de mutación = 8 casos de ensayo.

E 1	E 2	E 3	E 4	E 5	E 6	E 7	E 8
MUTACIÓN 0,4%				MUTACIÓN 0,8%			
CRUCE 1 PUNTO		CRUCE UNIFORME		CRUCE 1 PUNTO		CRUCE UNIFORME	
T. POB 50	T. POB 150	T. POB 50	T. POB 150	T. POB 50	T. POB 150	T. POB 50	T. POB 150

Tabla 5. Resumen explicativo de los ensayos realizados.

Cada una de las 20 codificaciones elegidas ha sido testada en los 8 casos de ensayo, lo cual ha generado finalmente un total de 160 ensayos realizados. Además, para la correcta comparación de los resultados obtenidos, se han repetido los ensayos un total de 100 veces, de modo que la muestra sea lo suficientemente grande para poder realizar estimaciones estadísticas representativas.

4.2. Tratamiento de los resultados

Dado el altísimo volumen de datos que se han generado durante el proceso de cálculo, se han realizado tres procedimientos diferentes para extraer los datos necesarios para la interpretación de los resultados. Nótese que para cada ejecución se obtiene un archivo con la población final (popfin.XX.txt), así como un archivo que posee los resultados obtenidos a lo largo de las evaluaciones ejecutadas. Cabe destacar que se ha limitado el total de evaluaciones a 1.000.000 (tanto en el caso de población de 50 individuos como en el de 150). Es por ello que el archivo final que contiene los resultados (resmej.XX.txt), contiene un total de 2001 líneas (imprime una línea con los resultados cada 50 generaciones, con un total de 10,000 generaciones). Esto genera un total de 1.000.050 evaluaciones para poblaciones de 50 individuos y 1.000.150 evaluaciones para 150 individuos. Cada archivo almacena en cada una de sus 2001 líneas los valores correspondientes a:

nº generaciones, nº evaluaciones, peso, valor de las restricciones, nombre de los perfiles calculados para cada una de las 55 barras que forman la estructura para la generación indicada.

1. De una parte se ha realizado una estadística de las 100 reproducciones realizadas en cada caso, mediante el cual se ha generado en cada caso un archivo conteniendo la siguiente información:

nº generaciones, nº evaluaciones, peso medio restringido, varianza del peso medio restringido, mejor valor obtenido del peso medio restringido en las 100 generaciones, y el nombre de los perfiles comerciales calculados para el mejor valor en cada una de las 55 barras de la estructura, ordenadas según el número identificativo de cada barra.

Estos archivos de estadística han sido empleados para graficar la evolución del valor de la función objetivo (peso restringido) con el avance del cálculo (número de evaluaciones realizadas).

2. En segundo lugar se han extraído y tabulado los datos estadísticos (es decir, extraídos de los archivos generados en el punto 1) correspondientes a 100.000 y 1.000.000 de evaluaciones. Concretamente sólo se han extraído los datos correspondientes a : nº evaluaciones, peso medio restringido, varianza del peso medio restringido, mejor valor. Y en segundo lugar se extraen los datos correspondientes a dos valores referencia establecidos de la función objetivo: 10.900 kg y 10.200 kg (se conoce previamente que el

peso óptimo de la función es de 10.127,1 kg [4], y para estos valores de peso restringido se extrae y tabula el número de evaluaciones que se ha necesitado en cada caso para alcanzar dicho valor (se ha extraído el número de evaluaciones para el primer número inmediatamente inferior que el valor referencia).

3. Finalmente, se han extractado todos los resultados obtenidos de peso restringido para 100.000 y 1.000.000 de evaluaciones, en cada uno de los 100 casos resueltos para una misma codificación y ensayo. Estos datos han sido empleados para el análisis estadístico de los mismos y que se presentarán en el siguiente apartado.

4.3. Resultados obtenidos

En las siguientes páginas pueden encontrarse:

- a. Tabla resumen de los resultados obtenidos en los diferentes ensayos llevados a cabo con las 20 ordenaciones. Corresponde a las estadísticas realizadas sobre las 100 repeticiones de cada ensayo, por lo que el contenido de la tabla es el siguiente:

- Media, desviación y mejor resultados para 100.000 evaluaciones.
- Media, desviación y mejor resultados para 1.000.000 evaluaciones.
- Nº de evaluaciones necesarias para alcanzar un valor de peso restringido de 10.900 (valor medio). Para medirlo se ha tabulado el nº de evaluaciones correspondiente al valor inmediatamente inferior a 10.900.
- Nº de evaluaciones necesarias para alcanzar un valor de peso restringido de 10.200 (valor medio). Para medirlo se ha tabulado el nº de evaluaciones correspondiente al valor inmediatamente inferior a 10.200
-

En estos dos últimos casos, cuando no se ha alcanzado el valor medio deseado, se encuentra en la tabla el valor 0 tanto en el nº de evaluaciones como en el valor de peso medio restringido.

- b. Gráficos:

- Gráfico 4. Comparación de MEDIAS de los resultados obtenidos de las codificaciones

S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,4%.

- Gráfico 5. Comparación de MEDIAS de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,8%.
- Gráfico 6. Comparación de MEJOR VALOR de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,4%.
- Gráfico 7. Comparación de MEJOR VALOR de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,8%.
- Gráfico 8. Comparación de DESVIACIÓN de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,4%.
- Gráfico 9. Comparación de DESVIACIÓN de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,8%.
- Gráfico 10. Comparación de MEDIA de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,4%.
- Gráfico 11. Comparación de MEDIA de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,8%.
- Gráfico 12. Comparación de MEJOR VALOR de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,4%.
- Gráfico 13. Comparación de MEJOR VALOR de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,8%.
- Gráfico 14. Comparación de DESVIACIÓN de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,4%.
- Gráfico 15. Comparación de DESVIACIÓN de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,8%.

4.4. Observación de resultados. Gráficos

En este apartado se resumen algunas observaciones que se han realizado sobre los datos tabulados, así como en las gráficas que se han realizado:

- En todos los ensayos realizados, sólo las codificaciones tipo Gray (S15A0, S15B0 y S15C0) han alcanzado valores medios de peso restringido de la estructura inferiores a 10.200 kg. Por tanto, son las únicas codificaciones que han sido capaces de llegar al final del cálculo con un valor medio cercano al óptimo (10127,1 kg).
- Las varianzas para los casos resueltos con codificaciones tipo Gray son considerablemente inferiores que en el resto de casos.
- De manera general, a 100.000 evaluaciones es la codificación S39A0,49 la que presenta un valor medio más alto (progresión del algoritmo más lenta). Sin embargo, es la codificación S27C1,33 la que peor resultado final presenta, es decir, la que tiene mayor valor medio tras 1.000.000 de evaluaciones.
- Aparentemente, las poblaciones mayores (150 individuos) facilitan una evolución más rápida hacia el óptimo que las poblaciones pequeñas (50 individuos).
- En los gráficos 5 y 6 de medias se puede observar como los resultados se distribuyen en la gráfica formando 3 grandes bandas que agrupan los resultados para una misma codificación, encontrándose luego pequeñas variaciones experimentadas por el empleo de diferentes parámetros en el proceso de cálculo. Sin embargo, se deduce que lo más significativo es el empleo de una codificación diferente. Este efecto puede verse también en las gráficas de mejor valor y desviación.
- Se observa que el comportamiento de la codificación con mayor suma de diferencias (S51A0,61) es mejor que para la ordenación (S27C1,33). Parece tener por tanto mayor influencia la varianza en el número de bits, o en todo caso la cantidad de variaciones de 4 bits, más que la suma total de bits diferentes encontradas en la codificación.

E. 1PT P50 M4	100.000 EVALUACIONES			1.000.000 EVALUACIONES			MEDIA <10.900		MEDIA <10.200	
	MEDIA	DESV	MEJ.RES.	MEDIA	DESV	MEJ.RES.	EVAL.	MEDIA	EVAL.	MEDIA
S15A0	10252,44	160,00	10127,10	10174,47	91,55	10127,10	10060	10885,19	513560	10199,89
S15B0	10230,81	126,29	10127,10	10175,47	90,16	10127,10	10060	10840,20	216560	10199,96
S15C0	10236,27	134,47	10127,10	10180,22	85,46	10127,10	9560	10880,21	420060	10199,82
S26A0,92	11556,14	324,51	10817,30	10824,10	228,41	10247,30	647560	10899,99	N.A.	N.A.
S27A0,40	11392,39	371,55	10648,00	10430,34	274,56	10127,10	230060	10898,97	N.A.	N.A.
S27B0,83	11331,25	344,93	10680,10	10507,76	233,51	10127,10	236560	10898,82	N.A.	N.A.
S27C1,33	11607,80	254,82	11065,80	11511,70	232,14	11062,10	N.A.	N.A.	N.A.	N.A.
S33A0,40	12151,64	9853,82	10730,60	10443,10	197,16	10148,80	175560	10898,40	N.A.	N.A.
S33B0,91	11661,27	292,04	11078,40	11046,50	205,40	10650,20	N.A.	N.A.	N.A.	N.A.
S33C1,38	10937,38	302,17	10234,00	10624,36	215,25	10234,00	128560	10898,63	N.A.	N.A.
S39A0,49	11900,55	309,32	11201,90	11068,97	254,66	10589,90	N.A.	N.A.	N.A.	N.A.
S39B0,95	11083,73	230,83	10589,20	10765,83	165,93	10470,60	326060	10899,50	N.A.	N.A.
S39C1,40	11564,52	348,17	10863,60	10864,95	272,46	10408,60	802560	10899,66	N.A.	N.A.
S45A0	11680,35	315,07	10946,70	10836,13	236,82	10489,70	637560	10899,79	N.A.	N.A.
S45B0	11796,81	302,78	11287,80	10998,31	231,86	10622,30	N.A.	N.A.	N.A.	N.A.
S45C0	11771,20	265,20	11123,40	10980,75	182,59	10634,60	N.A.	N.A.	N.A.	N.A.
S46A0,92	11386,36	230,94	10858,10	10931,78	165,78	10573,90	N.A.	N.A.	N.A.	N.A.
S50A0,57	11123,79	176,59	10797,30	10895,93	156,47	10604,00	922560	10899,98	N.A.	N.A.
S49A0,47	10979,39	243,00	10596,50	10808,22	183,49	10467,50	241060	10899,82	N.A.	N.A.
S51A0,61	11142,69	260,47	10723,40	10843,23	155,24	10557,00	474060	10899,97	N.A.	N.A.

Tabla 6. Resultados obtenidos para las 20 ordenaciones en el ensayo con cruce 1 punto, población de 50 individuos y mutación 0.4%.

E. 1PT P150 M4	100.000 EVAL.			1.000.000 EVAL			MEDIA <10.900		MEDIA <10.200	
	MEDIA	DESV	MEJ.RES.	MEDIA	DESV	MEJ.RES.	EVAL.	MEDIA	EVAL.	MEDIA
S15A0	10243,52	161,43	10127,10	10178,00	118,15	10127,10	17660	10898,67	222660	10199,97
S15B0	10230,91	114,13	10127,10	10165,06	76,93	10127,10	17660	10891,55	190160	10199,96
S15C0	10219,73	104,06	10127,10	10164,22	66,64	10127,10	17160	10855,46	146660	10199,91
S26A0,92	11458,29	319,77	10742,10	10780,55	222,53	10334,10	490160	10899,17	N.A.	N.A.
S27A0,40	11282,37	334,38	10711,30	10353,62	183,27	10127,10	197660	10899,07	N.A.	N.A.
S27B0,83	11238,68	271,03	10755,90	10435,02	197,31	10131,70	218160	10899,53	N.A.	N.A.
S27C1,33	11492,78	251,42	10892,00	11393,00	245,74	10802,70	N.A.	N.A.	N.A.	N.A.
S33A0,40	11202,51	232,62	10750,40	10412,18	199,01	10142,50	187660	10899,92	N.A.	N.A.
S33B0,91	11568,82	308,83	10935,90	11006,38	222,50	10552,80	N.A.	N.A.	N.A.	N.A.
S33C1,38	10864,23	284,28	10351,20	10634,14	227,47	10317,50	81660	10899,23	N.A.	N.A.
S39A0,49	11851,15	262,11	11290,10	10997,07	221,03	10456,60	N.A.	N.A.	N.A.	N.A.
S39B0,95	11023,41	174,12	10719,00	10725,09	162,74	10345,30	228160	10899,34	N.A.	N.A.
S39C1,40	11471,95	317,74	10801,90	10717,17	229,81	10258,30	406160	10899,74	N.A.	N.A.
S45A0	11592,29	304,90	10913,50	10785,75	201,62	10470,20	558660	10899,75	N.A.	N.A.
S45B0	11689,91	260,76	11184,70	10929,25	174,79	10610,20	N.A.	N.A.	N.A.	N.A.
S45C0	11729,60	244,77	11184,30	10974,08	210,35	10599,70	N.A.	N.A.	N.A.	N.A.
S46A0,92	11316,97	203,58	10915,40	10899,52	152,91	10592,60	997160	10899,97	N.A.	N.A.
S50A0,57	11084,91	155,68	10761,00	10879,90	151,59	10572,00	694660	10899,75	N.A.	N.A.
S49A0,47	10918,67	150,27	10628,20	10777,42	127,24	10465,60	119660	10899,97	N.A.	N.A.
S51A0,61	11111,96	210,39	10610,10	10794,65	136,83	10531,20	295160	10899,78	N.A.	N.A.

Tabla 7. Resultados obtenidos para las 20 ordenaciones en el ensayo con cruce 1 punto, población de 150 individuos y mutación 0.4%.

E. 1PT P50 M8	100.000 EVAL.			1.000.000 EVAL			MEDIA <10.900		MEDIA <10.200	
	MEDIA	DESV	MEJ.RES.	MEDIA	DESV	MEJ.RES.	EVAL	MEDIA	EVAL	MEDIA
S15A0	10227,28	127,08	10127,10	10179,28	108,23	10127,10	10060	10894,03	243560	10199,82
S15B0	10223,70	111,43	10127,10	10159,26	67,06	10127,10	10060	10874,26	161060	10199,80
S15C0	10234,69	121,50	10127,10	10173,14	95,40	10127,10	10060	10866,82	294060	10199,98
S26A0,92	11238,31	280,66	10761,00	10705,52	240,22	10297,20	313060	10899,23	N.A.	N.A.
S27A0,40	11009,35	345,87	10333,50	10354,25	216,78	10127,10	121060	10899,19	N.A.	N.A.
S27B0,83	10985,32	231,57	10506,20	10375,18	165,33	10127,10	119560	10898,41	N.A.	N.A.
S27C1,33	11506,43	243,12	10977,50	11353,46	255,44	10833,30	N.A.	N.A.	N.A.	N.A.
S33A0,40	10976,98	274,70	10552,80	10413,84	224,84	10127,10	118560	10899,50	N.A.	N.A.
S33B0,91	11447,09	235,50	10846,30	11006,76	213,44	10582,10	N.A.	N.A.	N.A.	N.A.
S33C1,38	10795,59	267,61	10299,80	10592,02	218,28	10250,40	50560	10899,91	N.A.	N.A.
S39A0,49	11639,80	305,26	10958,80	10945,36	234,11	10404,50	N.A.	N.A.	N.A.	N.A.
S39B0,95	11030,01	208,11	10592,80	10732,56	189,93	10324,90	215560	10899,44	N.A.	N.A.
S39C1,40	11296,75	335,00	10623,40	10750,20	214,57	10302,70	339560	10899,99	N.A.	N.A.
S45A0	11453,35	300,33	10856,00	10770,00	205,13	10368,00	431060	10899,86	N.A.	N.A.
S45B0	11544,80	295,21	10952,20	10917,93	202,80	10536,90	N.A.	N.A.	N.A.	N.A.
S45C0	11535,22	275,22	10917,50	10907,58	208,10	10519,40	N.A.	N.A.	N.A.	N.A.
S46A0,92	11185,31	215,46	10741,40	10885,55	176,87	10515,70	815060	10899,95	N.A.	N.A.
S50A0,57	11049,90	158,04	10726,90	10837,41	141,20	10574,00	441060	10900,00	N.A.	N.A.
S49A0,47	10941,74	184,40	10641,00	10771,24	156,23	10526,70	156060	10899,96	N.A.	N.A.
S51A0,61	11031,27	223,02	10647,80	10790,63	167,80	10520,40	241060	10899,14	N.A.	N.A.

Tabla 8. Resultados obtenidos para las 20 ordenaciones en el ensayo con cruce 1 punto, población de 50 individuos y mutación 0.8%.

E. 1PT P150 M8	100.000 EVAL.			1.000.000 EVAL			MEDIA <10.900		MEDIA <10.200	
	MEDIA	DESV	MEJ.RES.	MEDIA	DESV	MEJ.RES.	EVAL.	MEDIA	EVAL.	MEDIA
S15A0	10209,77	98,60	10127,10	10152,95	57,07	10127,10	18660	10874,66	119160	10199,72
S15B0	10205,85	100,61	10127,10	10150,82	61,26	10127,10	18660	10859,93	109660	10199,69
S15C0	10225,51	116,89	10127,10	10160,08	68,23	10127,10	18160	10891,43	145660	10199,90
S26A0,92	11083,43	279,86	10444,70	10596,89	207,49	10276,10	162160	10898,99	N.A.	N.A.
S27A0,40	11002,40	278,63	10540,90	10285,61	144,83	10127,10	118160	10898,32	N.A.	N.A.
S27B0,83	11063,45	236,59	10566,10	10358,88	184,79	10127,10	132160	10897,51	N.A.	N.A.
S27C1,33	11383,28	221,38	10694,10	11211,10	211,80	10601,20	N.A.	N.A.	N.A.	N.A.
S33A0,40	10998,09	206,54	10640,50	10342,89	151,02	10151,90	126160	10899,78	N.A.	N.A.
S33B0,91	11364,09	261,54	10876,40	10922,36	212,45	10530,60	N.A.	N.A.	N.A.	N.A.
S33C1,38	10769,22	242,05	10292,80	10559,26	194,98	10211,00	55660	10898,40	N.A.	N.A.
S39A0,49	11643,49	280,54	10927,10	10907,31	232,40	10306,60	N.A.	N.A.	N.A.	N.A.
S39B0,95	10960,15	221,04	10606,50	10677,14	179,26	10304,40	145660	10899,90	N.A.	N.A.
S39C1,40	11185,70	271,47	10457,00	10676,27	213,92	10239,60	235160	10899,34	N.A.	N.A.
S45A0	11358,82	272,96	10830,40	10703,18	176,28	10383,10	308660	10899,56	N.A.	N.A.
S45B0	11498,49	232,77	10974,40	10866,09	187,95	10476,60	750660	10899,93	N.A.	N.A.
S45C0	11497,33	264,04	11020,10	10833,55	174,65	10531,50	580570	10899,80	N.A.	N.A.
S46A0,92	11149,20	183,26	10742,90	10846,79	162,65	10392,40	500160	10899,84	N.A.	N.A.
S50A0,57	11015,83	153,82	10738,40	10804,88	135,39	10573,70	260660	10899,73	N.A.	N.A.
S49A0,47	10893,01	172,81	10592,20	10734,16	145,53	10481,80	97160	10899,91	N.A.	N.A.
S51A0,61	11013,57	203,34	10604,90	10776,63	157,95	10457,40	193660	10899,90	N.A.	N.A.

Tabla 9. Resultados obtenidos para las 20 ordenaciones en el ensayo con cruce 1 punto, población de 150 individuos y mutación 0.8%.

E. UNI P50 M4	100.000 EVAL.			1.000.000 EVAL			MEDIA <10.900		MEDIA <10.200	
	MEDIA	DESV	MEJ.RES.	MEDIA	DESV	MEJ.RES.	EVAL.	MEDIA	EVAL.	MEDIA
S15A0	10213,09	113,53	10127,10	10169,29	90,84	10127,10	7060	10827,11	133060	10199,62
S15B0	10224,54	112,20	10127,10	10171,36	86,76	10127,10	7060	10860,60	227560	10199,97
S15C0	10236,52	145,50	10127,10	10178,99	97,55	10127,10	7060	10822,92	202560	10199,88
S26A0,92	11673,97	311,54	10944,30	10932,79	243,66	10538,80	N.A.	N.A.	N.A.	N.A.
S27A0,40	11635,14	402,56	10971,30	10514,12	268,11	10165,90	334560	10899,68	N.A.	N.A.
S27B0,83	11440,55	256,20	10889,60	10552,38	223,36	10134,80	297060	10899,98	N.A.	N.A.
S27C1,33	11701,74	279,60	10964,30	11608,72	260,50	10954,50	N.A.	N.A.	N.A.	N.A.
S33A0,40	11515,69	264,90	10845,60	10670,17	220,36	10293,00	494060	10899,21	N.A.	N.A.
S33B0,91	11800,73	341,46	11144,70	11121,70	243,76	10544,70	N.A.	N.A.	N.A.	N.A.
S33C1,38	10996,58	330,68	10397,00	10671,87	262,35	10284,90	168560	10899,79	N.A.	N.A.
S39A0,49	12216,76	346,72	11614,90	11353,16	297,11	10706,50	N.A.	N.A.	N.A.	N.A.
S39B0,95	11163,68	243,64	10694,80	10867,99	216,47	10477,00	743060	10899,70	N.A.	N.A.
S39C1,40	11726,78	350,41	10832,40	10928,24	320,16	10356,00	N.A.	N.A.	N.A.	N.A.
S45A0	11930,24	371,57	11200,10	11080,50	253,62	10554,60	N.A.	N.A.	N.A.	N.A.
S45B0	12026,11	320,65	11351,40	11303,69	273,24	10775,30	N.A.	N.A.	N.A.	N.A.
S45C0	12072,58	333,38	11432,30	11330,51	291,86	10676,40	N.A.	N.A.	N.A.	N.A.
S46A0,92	11500,05	265,56	11040,30	10999,15	175,66	10665,40	N.A.	N.A.	N.A.	N.A.
S50A0,57	11180,44	191,26	10735,60	10876,69	140,38	10531,30	753560	10899,46	N.A.	N.A.
S49A0,47	10987,89	236,85	10610,80	10822,45	153,60	10499,60	343060	10899,96	N.A.	N.A.
S51A0,61	11266,02	239,78	10760,10	10876,70	170,94	10478,30	775060	10899,82	N.A.	N.A.

Tabla 10. Resultados obtenidos para las 20 ordenaciones en el ensayo con cruce uniforme, población de 50 individuos y mutación 0.4%.

E. UNI P150 M4	100.000 EVAL.			1.000.000 EVAL			MEDIA <10.900		MEDIA <10.200	
	MEDIA	DESV	MEJ.RES.	MEDIA	DESV	MEJ.RES.	EVAL.	MEDIA	EVAL.	MEDIA
S15A0	10174,91	84,36	10127,10	10145,79	49,50	10127,10	11160	10848,88	62160	10199,92
S15B0	10198,67	96,32	10127,10	10153,91	57,77	10127,10	11660	10838,45	99660	10198,71
S15C0	10194,23	90,25	10127,10	10166,02	73,51	10127,10	10160	10876,76	79660	10199,77
S26A0,92	11642,84	252,15	10960,80	10846,66	224,08	10420,50	779160	10899,86	N.A.	N.A.
S27A0,40	11558,70	344,27	10551,10	10544,95	226,60	10160,50	379160	10899,58	N.A.	N.A.
S27B0,83	11417,99	309,28	10703,80	10549,69	204,40	10160,50	302160	10898,62	N.A.	N.A.
S27C1,33	11744,73	198,32	11195,30	11645,99	209,35	11015,30	N.A.	N.A.	N.A.	N.A.
S33A0,40	11570,84	224,86	11079,00	10768,97	236,29	10304,40	701660	10899,93	N.A.	N.A.
S33B0,91	11903,63	330,59	11186,00	11156,23	236,98	10733,10	N.A.	N.A.	N.A.	N.A.
S33C1,38	10925,27	325,43	10300,30	10623,22	237,24	10244,40	113160	10899,86	N.A.	N.A.
S39A0,49	12301,88	338,72	11468,90	11417,49	312,27	10773,30	N.A.	N.A.	N.A.	N.A.
S39B0,95	11015,42	183,66	10647,30	10762,00	172,10	10436,10	333660	10899,98	N.A.	N.A.
S39C1,40	11686,44	310,09	10884,70	10875,69	265,71	10290,60	845660	10899,71	N.A.	N.A.
S45A0	11801,31	320,11	11070,40	11109,39	269,22	10573,50	N.A.	N.A.	N.A.	N.A.
S45B0	12020,64	314,45	11291,20	11350,01	280,69	10764,70	N.A.	N.A.	N.A.	N.A.
S45C0	11934,87	244,81	11239,50	11346,61	233,26	10840,50	N.A.	N.A.	N.A.	N.A.
S46A0,92	11533,92	223,19	10959,90	11012,85	159,81	10692,80	N.A.	N.A.	N.A.	N.A.
S50A0,57	11211,22	186,06	10795,30	10871,26	134,10	10509,50	735660	10899,51	N.A.	N.A.
S49A0,47	10938,95	190,39	10530,20	10792,14	157,21	10446,40	140160	10899,53	N.A.	N.A.
S51A0,61	11205,53	242,42	10706,00	10849,51	164,36	10559,00	506660	10899,43	N.A.	N.A.

Tabla 11. Resultados obtenidos para las 20 ordenaciones en el ensayo con cruce uniforme, población de 150 individuos y mutación 0.4%.

E. UNI P50 M8	100.000 EVAL.			1.000.000 EVAL			MEDIA <10.900		MEDIA <10.200	
	MEDIA	DESV	MES.RES.	MEDIA	DESV	MES.REJ.	EVAL.	MEDIA	EVAL.	MEDIA
S15A0	10192,72	86,37	10127,10	10153,07	57,33	10127,10	7560	10863,46	91060	10199,47
S15B0	10213,28	108,92	10127,10	10156,57	71,27	10127,10	7560	10860,29	132060	10199,95
S15C0	10205,81	102,97	10127,10	10166,04	71,42	10127,10	7060	10850,32	111060	10199,61
S26A0,92	11399,99	317,39	10801,00	10772,97	226,28	10295,70	508560	10899,78	N.A.	N.A.
S27A0,40	11316,36	359,80	10631,30	10436,58	231,64	10132,80	201060	10899,50	N.A.	N.A.
S27B0,83	11146,19	303,60	10584,50	10441,25	207,78	10131,70	168560	10899,29	N.A.	N.A.
S27C1,33	11617,60	223,98	11074,80	11444,70	251,77	10973,30	N.A.	N.A.	N.A.	N.A.
S33A0,40	11285,03	257,58	10807,90	10496,95	192,37	10152,10	266060	10899,29	N.A.	N.A.
S33B0,91	11541,66	282,23	10939,70	11062,00	208,58	10573,00	N.A.	N.A.	N.A.	N.A.
S33C1,38	10881,83	353,75	10391,90	10643,74	285,42	10223,80	93060	10899,90	N.A.	N.A.
S39A0,49	11981,91	296,19	11183,40	11186,35	309,89	10557,90	N.A.	N.A.	N.A.	N.A.
S39B0,95	11073,76	210,28	10623,80	10741,28	181,98	10391,50	323060	10899,92	N.A.	N.A.
S39C1,40	11390,40	346,35	10591,10	10773,21	267,22	10279,00	434060	10899,93	N.A.	N.A.
S45A0	11691,89	363,91	10715,50	10940,11	224,84	10474,60	N.A.	N.A.	N.A.	N.A.
S45B0	11835,00	318,36	10951,60	11099,57	253,97	10564,50	N.A.	N.A.	N.A.	N.A.
S45C0	11819,88	292,77	11178,80	11103,65	232,16	10652,40	N.A.	N.A.	N.A.	N.A.
S46A0,92	11334,56	231,49	10793,10	10948,10	170,66	10539,00	N.A.	N.A.	N.A.	N.A.
S50A0,57	11084,54	174,57	10673,30	10833,59	138,33	10483,60	401560	10899,95	N.A.	N.A.
S49A0,47	10917,29	155,46	10533,50	10780,70	127,17	10484,50	118560	10899,20	N.A.	N.A.
S51A0,61	11049,63	216,10	10581,80	10815,52	162,66	10489,70	278060	10899,90	N.A.	N.A.

Tabla 12. Resultados obtenidos para las 20 ordenaciones en el ensayo con cruce uniforme, población de 50 individuos y mutación 0.8%.

E. UNI P150 M8	100.000 EVAL.			1.000.000 EVAL			MEDIA <10.900		MEDIA <10.200	
	MEDIA	DESV	MEJ.RES.	MEDIA	DESV	MEJ.RES.	EVAL.	MEDIA	EVAL.	MEDIA
S15A0	10167,47	75,31	10127,10	10132,91	24,23	10127,10	13160	10873,57	58660	10199,71
S15B0	10190,53	92,18	10127,10	10144,98	60,50	10127,10	13160	10899,72	82660	10199,44
S15C0	10170,19	80,21	10127,10	10140,04	42,68	10127,10	12660	10835,65	55660	10199,98
S26A0,92	11310,74	290,20	10791,30	10699,22	206,18	10350,40	303160	10899,61	N.A.	N.A.
S27A0,40	11461,01	284,25	10885,30	10476,46	251,12	10132,80	250160	10898,81	N.A.	N.A.
S27B0,83	11198,87	268,03	10547,30	10386,39	171,47	10127,10	164660	10899,24	N.A.	N.A.
S27C1,33	11712,13	216,65	11270,30	11538,67	239,13	10980,30	N.A.	N.A.	N.A.	N.A.
S33A0,40	11450,92	220,19	10981,90	10568,05	211,78	10249,70	359660	10899,97	N.A.	N.A.
S33B0,91	11648,74	270,98	11120,10	11085,80	205,18	10569,10	N.A.	N.A.	N.A.	N.A.
S33C1,38	10750,49	219,74	10374,20	10546,48	174,40	10187,30	51160	10899,45	N.A.	N.A.
S39A0,49	12113,47	289,09	11606,00	11266,47	259,74	10717,90	N.A.	N.A.	N.A.	N.A.
S39B0,95	10951,44	202,69	10544,00	10655,29	159,14	10287,30	142160	10899,77	N.A.	N.A.
S39C1,40	11451,60	285,26	10906,60	10776,56	223,36	10348,00	426660	10899,55	N.A.	N.A.
S45A0	11601,98	292,73	10934,60	10894,68	212,60	10488,30	985160	10899,71	N.A.	N.A.
S45B0	11867,51	282,84	11253,90	11140,09	220,91	10654,80	N.A.	N.A.	N.A.	N.A.
S45C0	11899,41	275,68	11333,30	11137,66	208,09	10730,90	N.A.	N.A.	N.A.	N.A.
S46A0,92	11398,25	219,70	10971,10	10934,19	147,22	10598,80	N.A.	N.A.	N.A.	N.A.
S50A0,57	11123,17	186,11	10692,60	10845,41	159,91	10453,90	450660	10899,75	N.A.	N.A.
S49A0,47	10903,07	137,07	10542,50	10749,30	123,61	10354,50	103160	10899,65	N.A.	N.A.
S51A0,61	11119,58	201,03	10591,30	10828,01	155,24	10490,40	341160	10899,95	N.A.	N.A.

Tabla 13. Resultados obtenidos para las 20 ordenaciones en el ensayo con cruce uniforme, población de 150 individuos y mutación 0.8%.

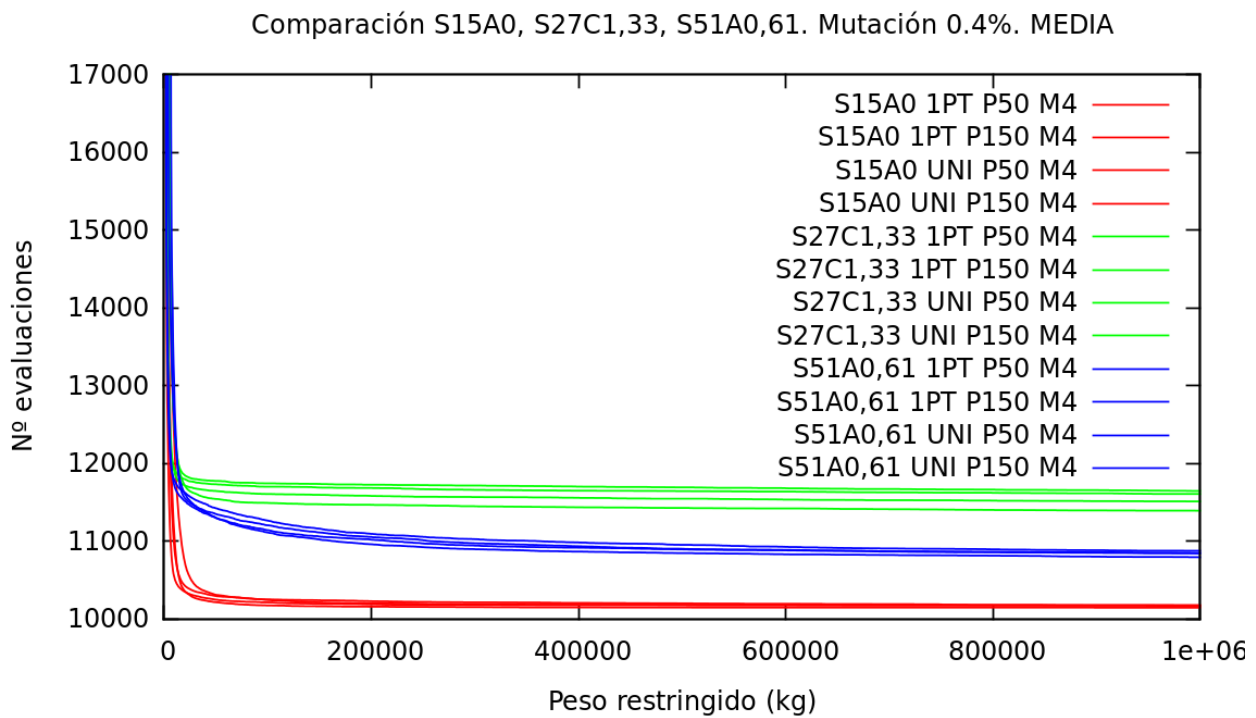


Gráfico 4. Comparación de MEDIAS de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,4%.

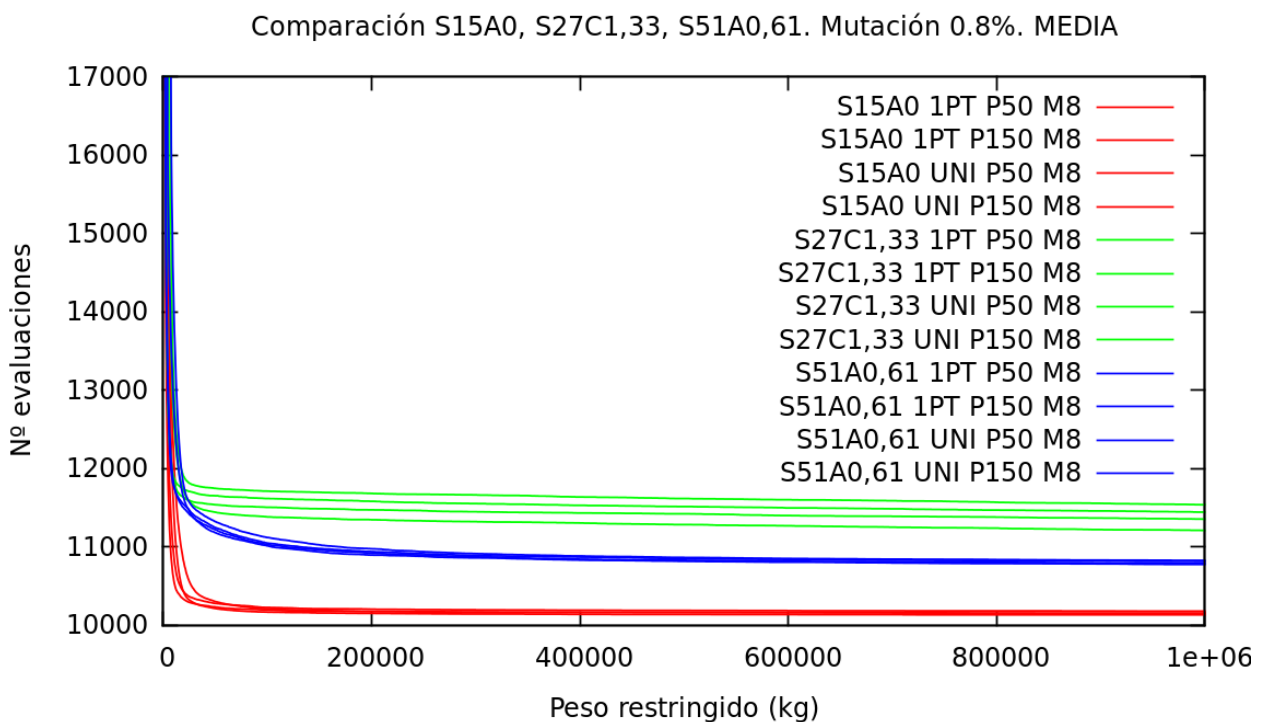


Gráfico 5. Comparación de MEDIAS de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,8%.

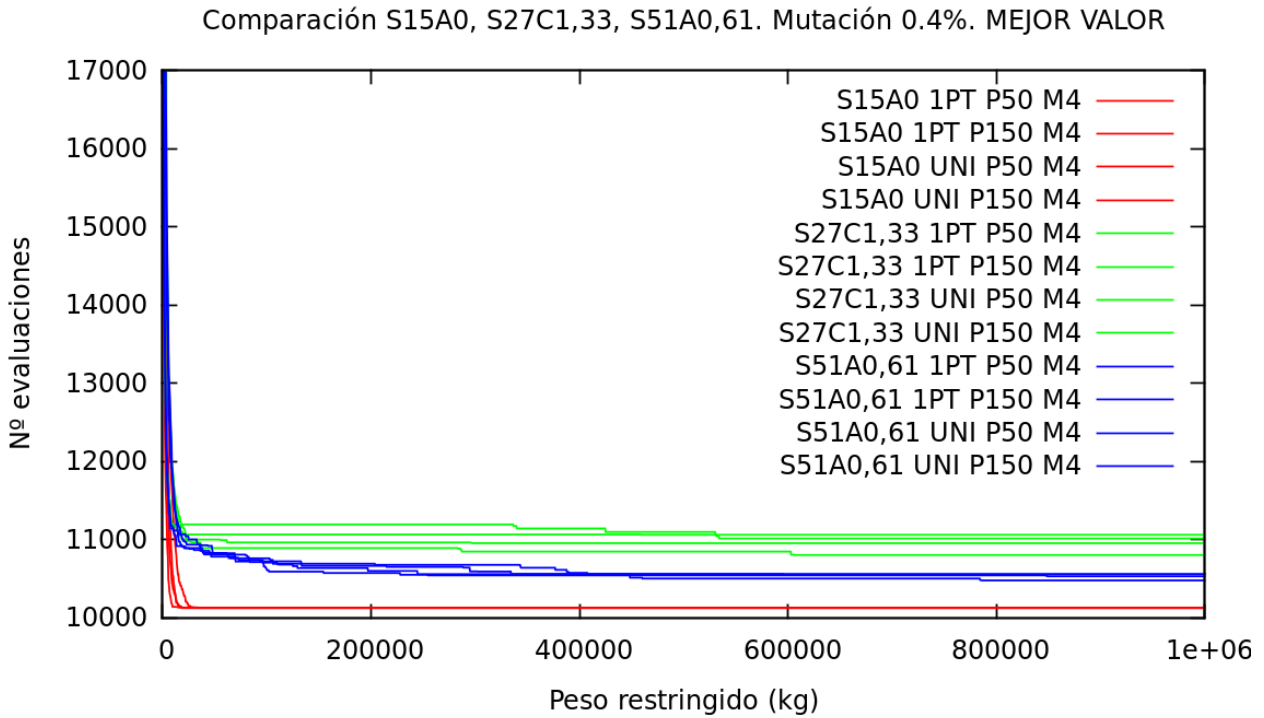


Gráfico 6. Comparación de MEJOR VALOR de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,4%.

40

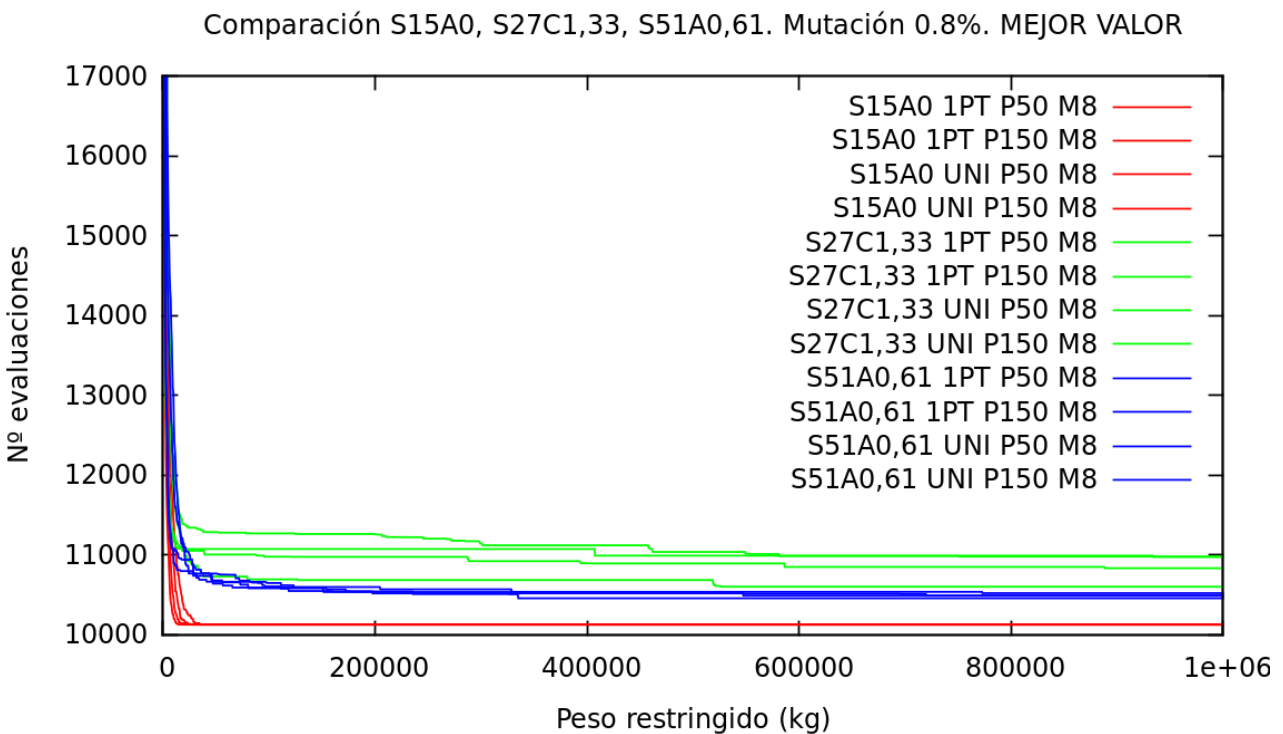


Gráfico 7. Comparación de MEJOR VALOR de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,8%.

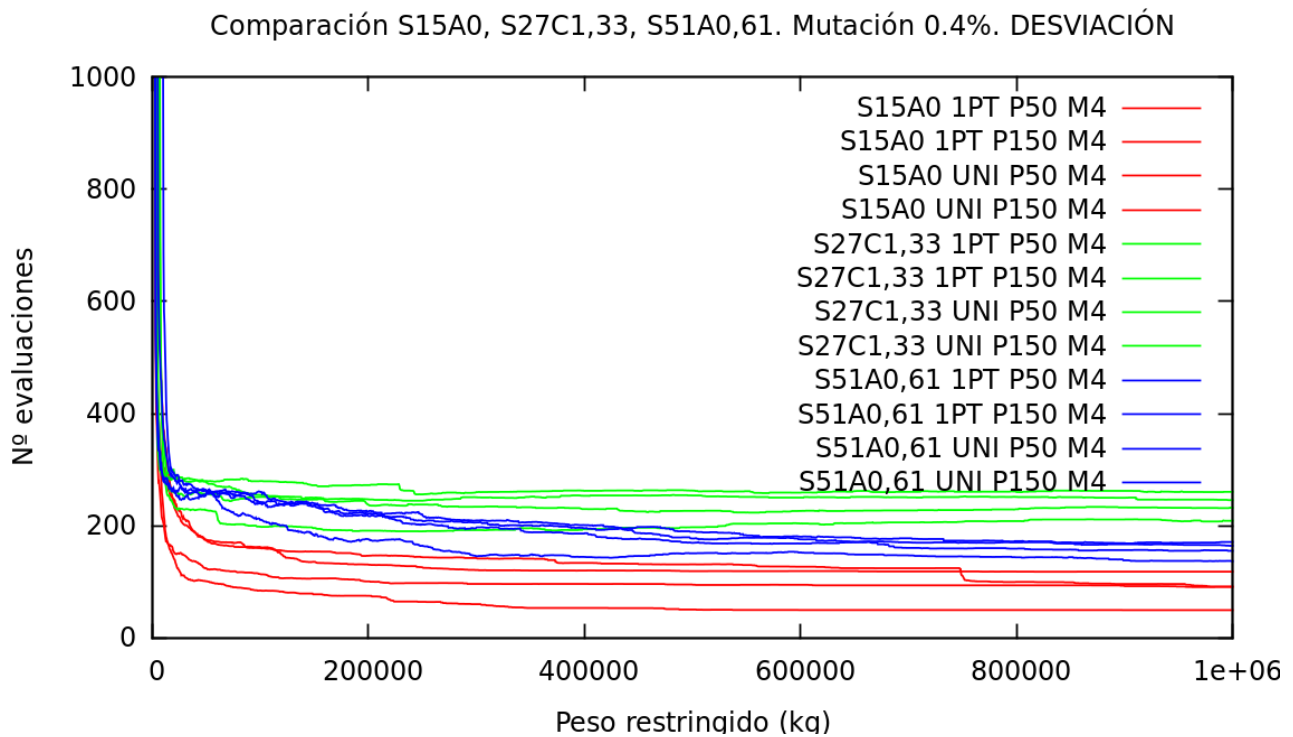


Gráfico 8. Comparación de DESVIACIÓN de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,4%.

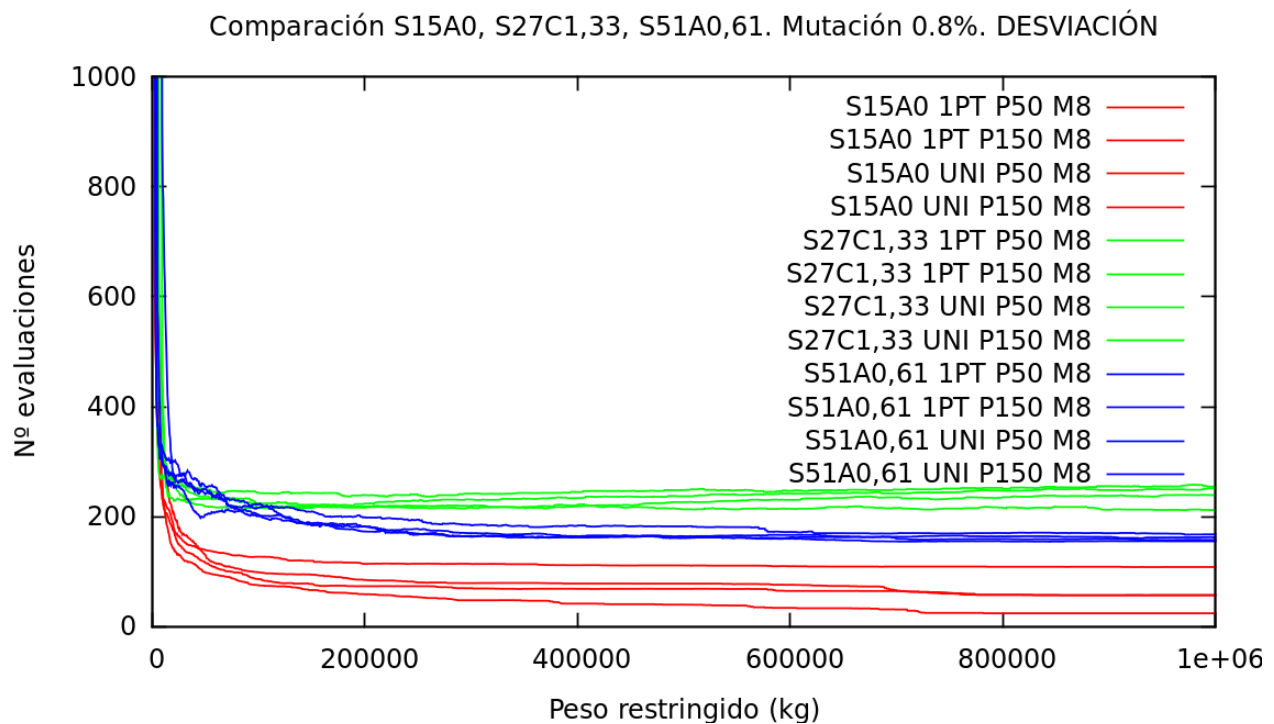


Gráfico 9. Comparación de DESVIACIÓN de los resultados obtenidos de las codificaciones S15A0, S27C1,33 y S51A0,61. Todos los ensayos con mutación 0,8%.

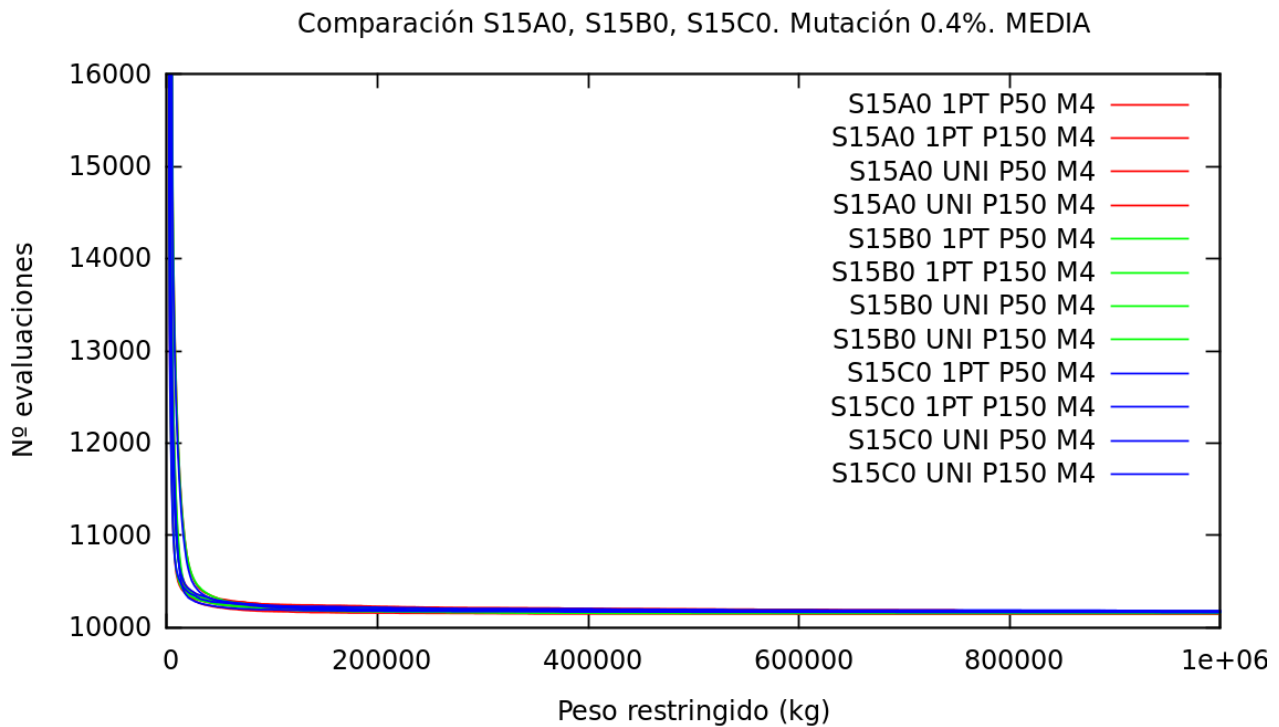


Gráfico 10. Comparación de MEDIA de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,4%.

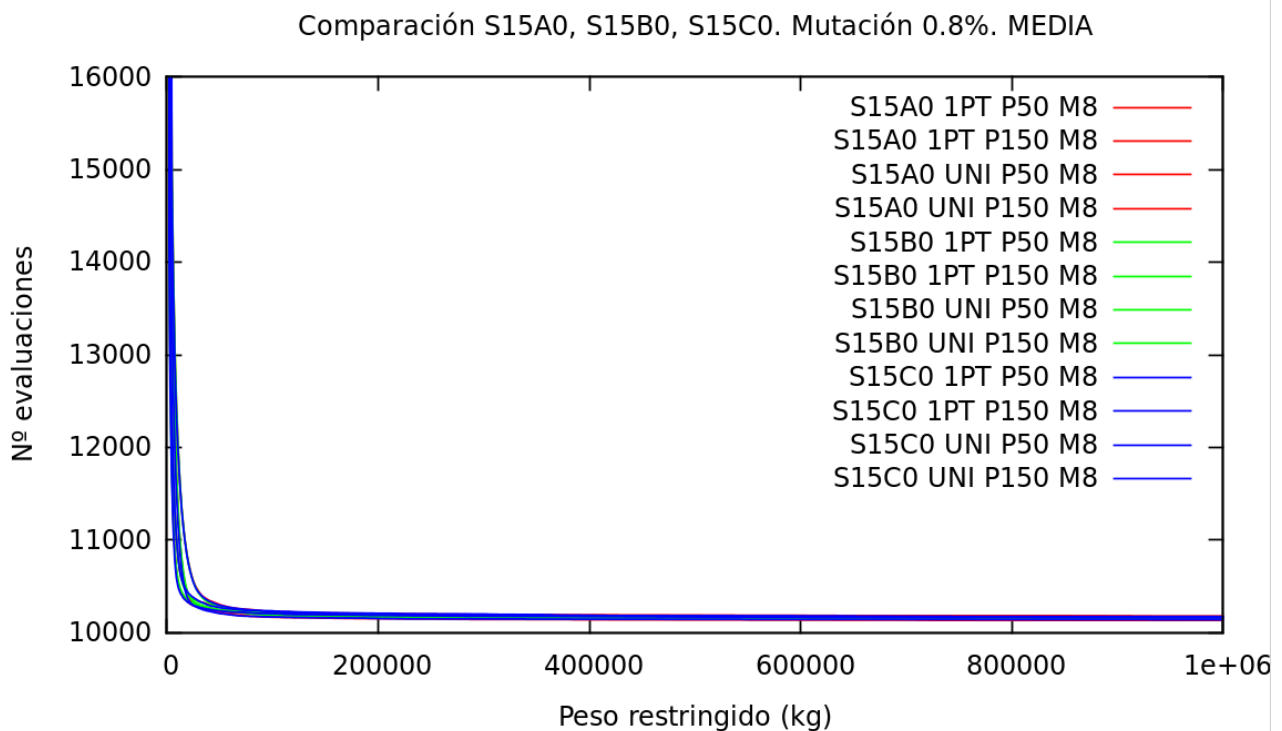


Gráfico 11. Comparación de MEDIA de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,8%

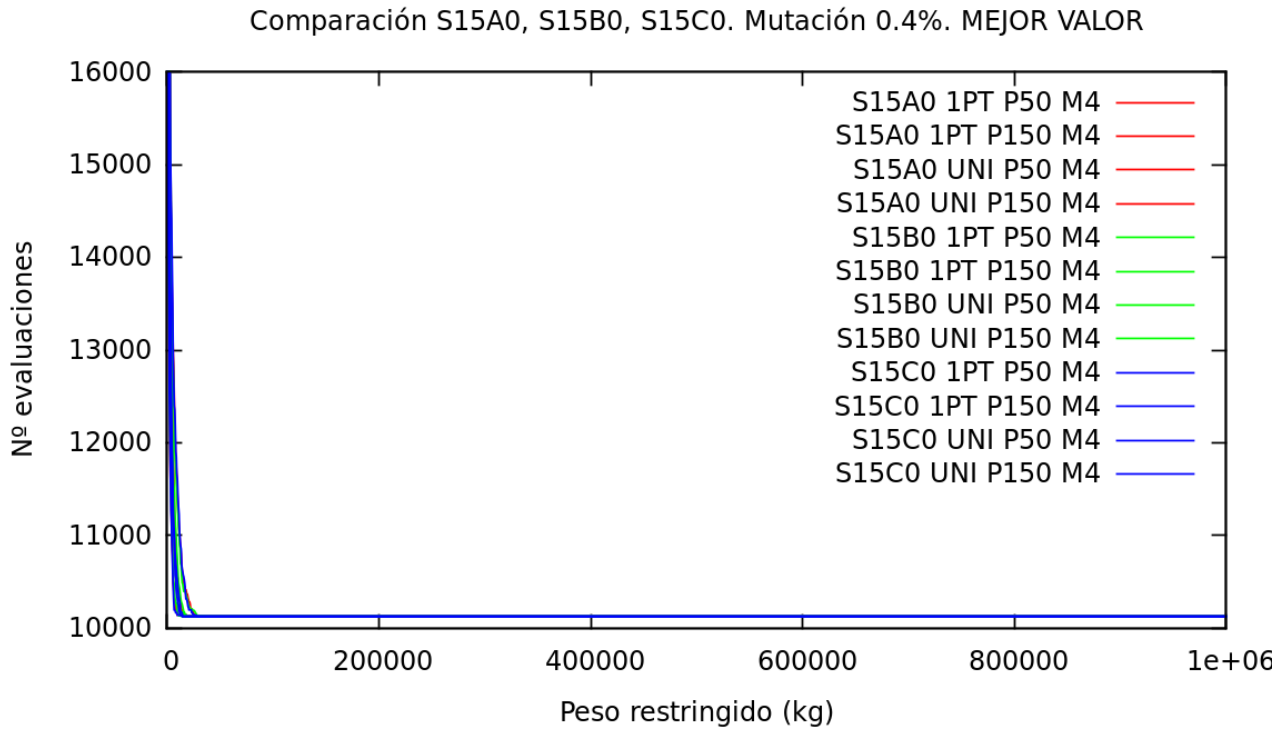


Gráfico 12. Comparación de MEJOR VALOR de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,4%.

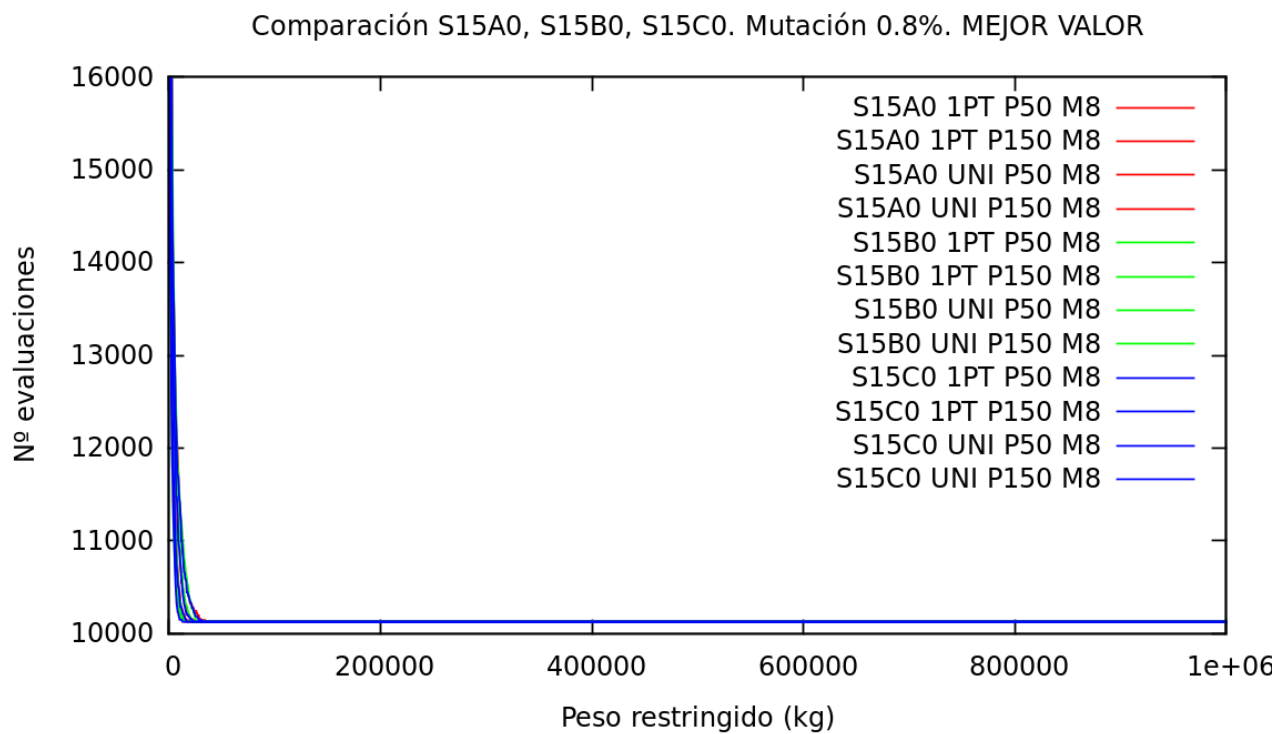


Gráfico 13. Comparación de MEJOR VALOR de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,8%.

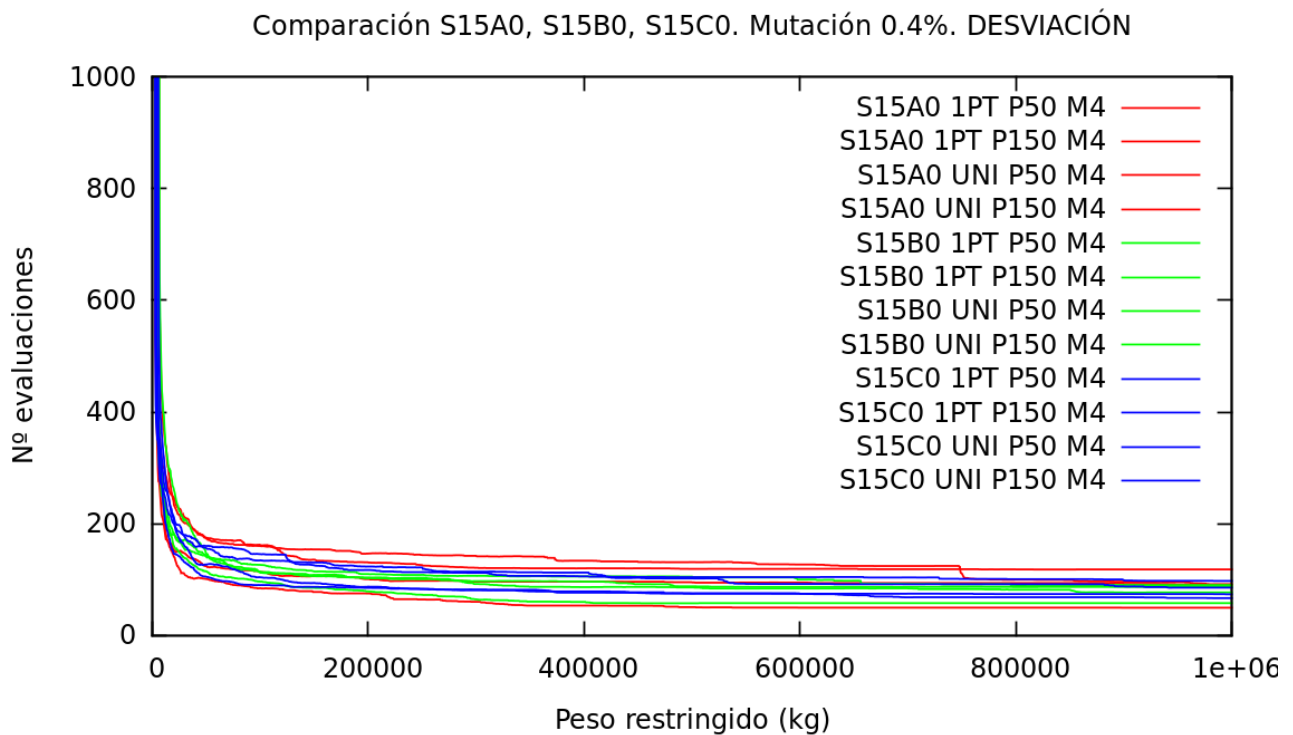


Gráfico 14. Comparación de DESVIACIÓN de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,4%.

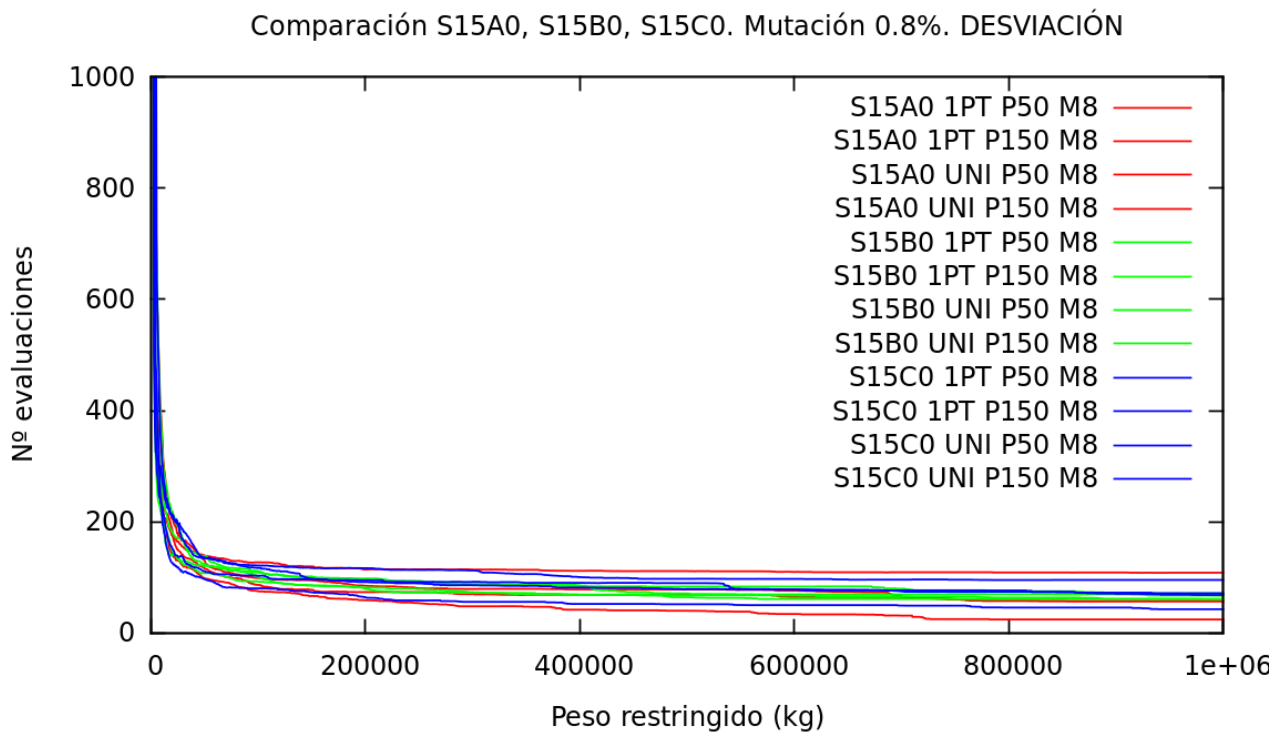


Gráfico 15. Comparación de DESVIACIÓN de los resultados obtenidos de las codificaciones S15A0, S15B0 y S15C0. Todos los ensayos con mutación 0,8%.

5. ANÁLISIS ESTADÍSTICO DE LOS RESULTADOS OBTENIDOS

El objetivo de este análisis estadístico es comprobar las hipótesis establecidas en el apartado anterior a partir de la visualización gráficos de resultados: los resultados obtenidos mediante el empleo de codificaciones tipo Gray son significativamente mejores que en el caso de los obtenidos mediante el resto de codificaciones, incluido con la ordenación binaria.

Se ha establecido por tanto un procedimiento de análisis estadístico que permita realizar tal contrastación. En primer lugar hay que tener en cuenta los parámetros que se han ensayado en la fase experimental, cómo se han agrupado los cálculos y qué parámetros se van a tener en cuenta en el análisis estadístico. Concretamente se han probado un total de 20 codificaciones diferentes (que es el factor en estudio), y estas codificaciones se han probado en un total de 8 ensayos diferentes, mediante la combinación de dos niveles (valores diferentes) en un total de 3 parámetros. En este análisis sólo se valora la influencia de la ordenación empleada en el resultado final de la función objetivo (peso restringido de la estructura en kilogramos), y el resto se han empleado para crear diferentes escenarios de cálculo. Por tanto, el número total de grupos de datos que se tiene es:

$$\begin{aligned} & \mathbf{20 \text{ codificaciones} \times 2 \text{ tipos de cruce} \times 2 \text{ tamaños de población} \times 2 \% \text{ mutación}} \\ & \mathbf{= 160 \text{ grupos de datos}} \end{aligned}$$

Además, y para estudiar la evolución del algoritmo hacia el resultado final de la optimización, se ha decidido extraer y analizar el valor de peso restringido que para cada caso se calculó en la evaluación 100.000 (o lo que es lo mismo, tras el primer 10% de las evaluaciones implementadas). Por tanto, finalmente se obtiene un total de 320 grupos de datos (o muestras), ya que para cada ensayo se han analizado los datos de función objetivo obtenido tanto a 100.000 evaluaciones como a 1.000.000 (valor final del cálculo).

En base a esta distribución, y teniendo en cuenta el objetivo principal de este análisis estadístico (encontrar significancia en los resultados obtenidos respecto al factor de estudio de la variación de ordenación), se plantea por tanto una reagrupamiento de los ensayos realizados, de tal manera que se comparen los resultados obtenidos para diferentes codificaciones dentro de un

mismo ensayo. Se considera un mismo ensayo aquel que tiene el mismo tipo de cruce, tamaño de población y porcentaje de mutación.

Se tienen por tanto 8 ensayos con 20 grupos de datos en cada uno. Se pretende comprobar, para cada uno de los ensayos, que mediante el empleo de las codificaciones de tipo Gray (estándar reflejada) se obtienen resultados significativamente mejores que con las otras codificaciones. A continuación se resumen los ensayos realizados, la nomenclatura empleada en el análisis estadístico y su relación con las leyendas empleadas en las gráficas del apartado anterior:

ENSAYO	CRUCE	TAMAÑO	MUTACIÓN
1PT.P50.M4	1 PUNTO	50	0,4 %
1PT.P150.M4	1 PUNTO	150	0,4 %
1PT.P50.M8	1 PUNTO	50	0,8 %
1PT.P150.M8	1 PUNTO	150	0,8 %
UNI.P50.M4	UNIFORME	50	0,4 %
UNI.P150.M4	UNIFORME	150	0,4 %
UNI.P50.M8	UNIFORME	50	0,8 %
UNI.P150.M8	UNIFORME	150	0,8 %

Tabla 14. Resumen con la nomenclatura de los ensayos y su significado.

El orden que se ha establecido en el análisis estadístico para clasificar las 20 codificaciones es la siguiente:

GRUPO	ORDENACIÓN	GRUPO	ORDENACIÓN
1	S15A0	11	S39A0,49
2	S15B0	12	S39B0,95
3	S15C0	13	S39C1,40
4	S26A0,92	14	S45A0
5	S27A0,40	15	S45B0
6	S27B0,83	16	S45C0
7	S27C1,33	17	S46A0,92
8	S33A0,40	18	S50A0,57
9	S33B0,91	19	S49A0,47
10	S33C1,38	20	S51A0,61

Tabla 15. Tabla resumen con la nomenclatura empleada en el test de hipótesis y la ordenación correspondiente que representa.

Cabe destacar que el procedimiento de análisis estadístico aquí propuesto responde a recomendaciones específicas para computación evolutiva, ya que se han tomado como referencia diferentes tutoriales del Congreso Internacional de Computación Genética y Evolutiva (GECCO), en su edición 2012, [6], [7].

5.1. Tests de hipótesis. Teoría e implementación.

Existen diferentes tests de hipótesis que permiten contrastar grupos de datos con el fin de determinar si no se encuentra que uno de los niveles empleados sea significativamente diferente al resto (hipótesis nula, $h=0$), o si se rechaza la hipótesis nula, con lo que se deduce que existe al menos un nivel significativo respecto del resto con los que ha sido comparado. Cuando se dispone de más de 2 grupos que contrastar (en este caso son 20), se puede emplear un test para multigrupos. En el caso de que los resultados en cada grupo estuviesen distribuidos siguiendo una distribución normal, podría emplearse el test ANOVA1 (ANOVA de 1 factor, de tipo paramétrico), y que corresponde con el test t-Student para multigrupos. Sin embargo, para ello hay que comprobar previamente que todos los grupos de datos de un mismo ensayo están distribuidos normalmente. De no ser así, no podrán emplearse ningún test de tipo paramétrico y tendrá que emplearse un test alternativo no paramétrico (no presuponen nada sobre el comportamiento de la distribución de los datos), como es el caso del test de Kruskal-Wallis (correspondiente al ANOVA1 pero no paramétrico).

47

Test de Lillieforts realiza un test donde la hipótesis nula es que los datos contenidos en el vector provienen de una distribución de tipo normal, frente a la alternativa de que no procede de una distribución normal (cuando se rechaza la hipótesis nula). Por tanto, representando h como la respuesta a la validación o el rechazo de dicha hipótesis, se obtendrá un valor $h=0$ cuando la hipótesis nula se acepte, y $h=1$ cuando se rechace. En este caso se rechaza la hipótesis nula para un nivel de significancia inferior al 5%.

Se trata de un test de normalidad bilateral donde no es necesario especificar los parámetros que caracterizan a la distribución. Es por tanto una modificación del test de Kolmogorov-Smirnov, donde sí que es necesario especificar completamente la distribución. De ahí que en este caso se emplee el test de Lillieforts, ya que es más simple de implementar.

Tal y como se verá a continuación, en todos los ensayos se detectaron grupos de datos que no cumplían con una distribución normal, por lo que en todos los casos fue necesario emplear un test de hipótesis no paramétrico para el estudio de la significancia de las medias. En este sentido, el test de Kruskal-Wallis realiza un test para comparar dos o más grupos de datos.

Cada vector de la matriz analizada debe representar a un grupo independiente de datos, y contener datos mutuamente dependientes entre sí. La función compara la mediana de los datos de cada vector de la matriz, y devuelve el p-valor donde la hipótesis nula ($h=0$) implica que todos los datos pertenecen a la misma población, o a poblaciones con la misma distribución. Cuando el p-valor se acerca a 0 el test duda de la hipótesis nula y sugiere que al menos un conjunto de datos es significativamente diferente del resto. El valor de significancia establecido habitualmente para rechazar o no la hipótesis nula se sitúa entre 0,05 y 0,01 (en este caso se ha empleado 0,01). Se trata por tanto de la versión no paramétrica del clásico test de ANOVA para 1 parámetro para dos o más grupos (que además es la extensión del test t-Student, diseñado sólo para la comparación de dos grupos): el test ANOVA asume (para la hipótesis nula) que todos los datos provienen de la misma distribución, que además es la distribución normal. De ahí que no pueda emplearse cuando existen datos que no cumplen esta distribución. Del test de Kruskal-Wallis se extrae una tabla con el ranking de cada valor, y además un diagrama de cajas que muestra el valor de la mediana, sus intervalos de confianza y los valores más alejados no contenidos en el intervalo de confianza. Será este gráfico el que se valorará a continuación.

Finalmente, para una correcta contrastación de los resultados, se realiza una multicomparación de medias (o medianas en este caso). Para ello, el test emplea los resultados extraídos a partir del test de Kruskal- Wallis (o el que se haya aplicado indistintamente). Este test de multicomparación es un test simple que permite de manera visual detectar cual o cuales son las medias realmente significativas respecto del resto con las que se compara. Como resultado se obtiene tanto una tabla de datos con las multicomparaciones de cada pareja de muestras, como un gráfico donde aparece representada la mediana y su intervalo de confianza. Se considera que una muestra es significativa si su intervalo de confianza no se superpone con ningún otro de los comparados.

En base a los conceptos descritos anteriormente, el procedimiento de análisis estadístico que se ha diseñado es el siguiente:

1. Comprobar si todos los niveles de un mismo ensayo se distribuyen normalmente.
2. Emplear el test de hipótesis correspondiente (ANOVA1 o Kruskal-Wallis).
3. Realizar un análisis multicomparativo para determinar qué nivel es significativo respecto del resto.
4. En caso de ser necesario, se repetirán los puntos 2 y 3 para aquellos niveles que hayan resultado significativos, para determinar si entre ellos existen diferencias.

Se ha empleado el software MATLAB® R2010a para la realización de los cálculos

estadísticos, siendo hoy en día una de las herramientas más potentes de cálculo en el ámbito de la ciencia y la ingeniería. A modo de resumen se muestran a continuación los comandos empleados para emplear los tres métodos de análisis empleados:

Test de Lilliefors: `[h,p]=lillietest(vector_a_estudiar)`

Dónde **h** vale 0 ó 1 según se acepte la hipótesis nula (0) o se rechace, y **p** el valor de significancia obtenido, que es inferior a 0.01 cuando la hipótesis se rechaza, y superior si se acepta.

Para simplificar, se han tabulado solamente los valores de h obtenidos, si bien los valores de p han sido comprobados en cada caso.

Test de Kruskal-Wallis: `[Q,TBL,STATS]=kruskalwallis(matriz_a_estudiar)`

Dónde **Q** el valor de significancia obtenido, que es inferior a 0,05 cuando la hipótesis se rechaza, y superior si se acepta. TBL genera una tabla tipo ANOVA, que muestra el "ranking" de cada valor X del vector y STATS es una estructura de datos resultantes del test preparada específicamente para que pueda ser empleada como datos de entrada en una multicomparación de medianas, ya que el propio test de Kruskal-Wallis sólo es capaz de decir si existe o no al menos un nivel significativo, pero no de qué nivel se trata.

49

Además, muestra también como salida un gráfico de cajas, que junto al valor Q, son los datos mostrados en el apartado de resultados.

Multicomparación de medias o medianas: `multcompare[STATS]`

Este test emplea los resultados listados en STATS para realizar la multicomparación de medias o medianas (según el test de hipótesis que se haya empleado y la naturaleza de los grupos de datos. Genera como salida una tabla multicomparativa de cada mediana con el resto, así como un gráfico (que se muestra a continuación).

5.2. Resultados obtenidos

En este apartado se resumen los resultados obtenidos tras seguir el procedimiento estadístico que se ha diseñado.

5.2.1. Test de normalidad Lillieforts

Tras la realización del test de normalidad a todos las muestras (320 en total) se concluye que en todos los grupos de ensayo existe al menos una muestra que no sigue distribución de tipo normal, por lo que se hace necesario el empleo de tests de tipo no paramétrico para comparar las muestras. En la siguiente tabla se resumen los valores de h obtenidos en cada caso:

Ordenación	MUTACIÓN 0,4%								MUTACIÓN 0,4%							
	1PTO				UNI				1PTO				UNI			
	P50		P150		P50		P150		P50		P150		P50		P150	
	100	1mill	100	1mill	100	1mill	100	1mill	100	1mill	100	1mill	100	1mill	100	1mill
S15A0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
S15B0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
S15C0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
S26A0,92	0	0	0	0	0	1	0	1	1	1	0	1	0	0	1	1
S27A0,40	0	1	1	1	0	1	0	0	1	1	1	1	0	1	0	1
S27B0,83	0	1	0	1	0	1	0	1	0	0	0	1	0	1	0	1
S27C1,33	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
S33A0,40	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1
S33B0,91	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0
S33C1,38	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	0
S39A0,49	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
S39B0,95	1	0	0	1	1	0	0	0	1	0	0	0	0	0	0	1
S39C1,40	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
S45A0	0	1	0	1	0	1	0	0	0	1	1	0	0	0	0	1
S45B0	1	1	0	1	0	0	0	0	1	0	0	1	1	0	0	0
S45C0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0
S46A0,92	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0
S50A0,57	1	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0
S49A0,47	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1
S51A0,61	1	1	0	0	0	1	0	0	0	0	0	0	0	1	0	1

Tabla 16. Resultados obtenidos en el test de normalidad de Lillieforts.

Para evaluar los resultados obtenidos en los diferentes tests de hipótesis multigrupos de Kruskal-Wallis, se mostrará en primer lugar todos los gráficos extraídos para el primer ensayo (Mutación 0,4%, población de 50 individuos, cruce de tipo 1 punto). En el resto se mostrarán sólo los gráficos de cajas y en algún caso concreto si fuere necesario el gráfico de multicomparación.

Cabe destacar que en todos los casos se ha obtenido significancia en el test de Kruskal-Wallis, y principalmente centrado en las tres codificaciones de tipo Gray respecto al resto. Es por ello

que en todos los casos ha sido necesario repetir el test para comparar sólo las muestras para ensayos con codificaciones de Gray, y en algún caso concreto una ordenación más que inicialmente parecía significativa.

5.2.2. Tests de significancia Kruskal- Wallis y multicomparación

Resultados: ensayo 1PTO.P50.M4

Para 100.000 evaluaciones

Se muestran a continuación los gráficos y datos más significativos extraídos del ensayo. En primer lugar cabe destacar el valor de Q que representa la respuesta del test Kruskal-Wallis a la hipótesis nula, y que concretamente es de $7,78 \cdot 10^{-291}$. Como el valor de significancia establecido es de 0,05, la hipótesis nula se rechaza, y se intuye por tanto que existe algún nivel significativo respecto del resto. Para ello se observa el gráfico de cajas, con el fin de evaluar las posibles muestras significativas. En el mismo puede comprobarse como la mediana de los tres primeros grupos (que corresponde con las codificaciones de Gray S15A0, S15B0 y S15C0) es bastante inferior que el resto. También se observa que no existe un patrón de variación uniforme con los valores de suma de variaciones (15 en los tres primeros casos y 51 en el último). De hecho, los últimos grupos con mayor cantidad de bits diferentes en la ordenación presentan resultados mejores que otras codificaciones. Destaca particularmente la muestra correspondiente a la ordenación S39A0,49, por presentar la mayor mediana entre las calculadas. Por otra parte, y como podrá observarse a lo largo del resto de gráficos que van a evaluarse, el intervalo de confianza de las tres primeras codificaciones es mucho menos que en el resto, lo cual concuerda con que en ningún caso las muestras obtenidas a partir de codificaciones tipo Gray hayan cumplido con una distribución normal, sino que los datos se encuentran muy concentrados en torno al valor medio.

Finalmente se realiza la multicomparación de las medianas (gráfico 17), y se comprueba que las tres primeras muestras efectivamente presentan significancia respecto del resto (no se solapan las líneas que representan a los intervalos de confianza), mientras que aparentemente no se aprecian diferencias entre las tres primeras muestras entre sí. Para ello se vuelve a realizar un test de Kruskal-Wallis y una multicomparación sólo para las tres muestras aparentemente significativas. En este caso el valor de Q obtenido fue de 0.8528, con lo cual se acepta la hipótesis nula de que todos los datos pertenecen a poblaciones con igual distribución.

18 muestran el gráfico de cajas y de multicomparación para las muestras correspondientes a las codificaciones S15A0, S15B0 y S15C0 para el ensayo que aquí se evalúa (1PTO.P50.M4). Se comprueba efectivamente que ambas medianas representan a la gran parte de los resultados obtenidos, y sólo unos pocos datos escapan de la mediana y del intervalo de confianza. Además, en la multicomparación se observa cómo no existe entre las 3 muestras ninguna con significancia respecto del resto. Por tanto se puede concluir que para este ensayo, las tres codificaciones de tipo Gray han presentado resultados de peso restringido significativamente mejores que el resto de codificaciones ensayadas, sin mostrar significancia entre ellas, es decir, presentando un comportamiento similar.

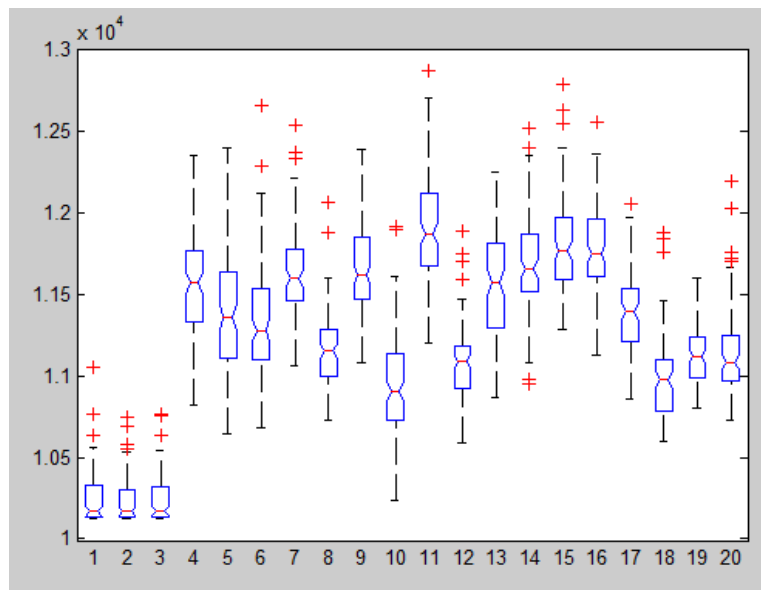


Gráfico 16. Gráfico de cajas para los resultados a 100.000 evaluaciones del ensayo 1PT.P50.M4 para las 20 codificaciones ensayadas.

52

Kruskal-Wallis ANOVA Table					
Source	SS	df	MS	Chi-sq	Prob>Chi-sq
Columns	4.75002e+008	19	25000091	1424.31	7.78002e-291
Error	1.91658e+008	1980	96796.9		
Total	6.6666e+008	1999			

Tabla 17. Tabla Kruskal-Wallis ANOVA con los resultados extraídos del test.

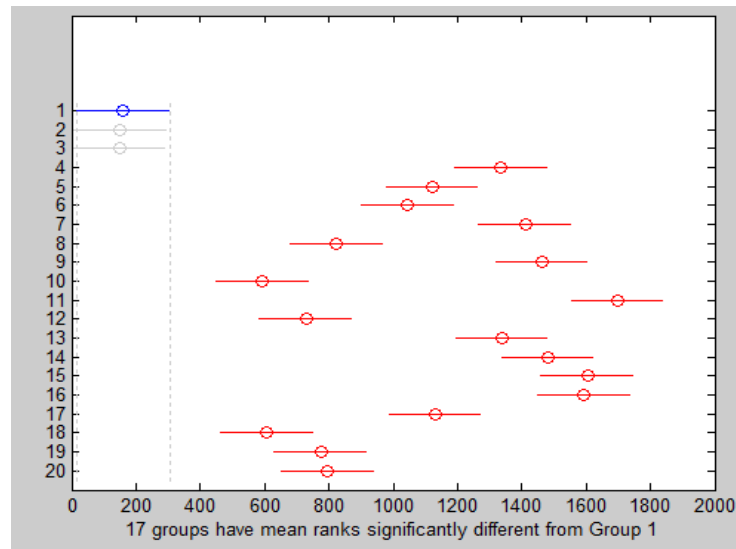


Gráfico 17. Multicomparación de las 20 codificaciones para 100.000 evaluaciones en el ensayo 1PT.P50.M4.

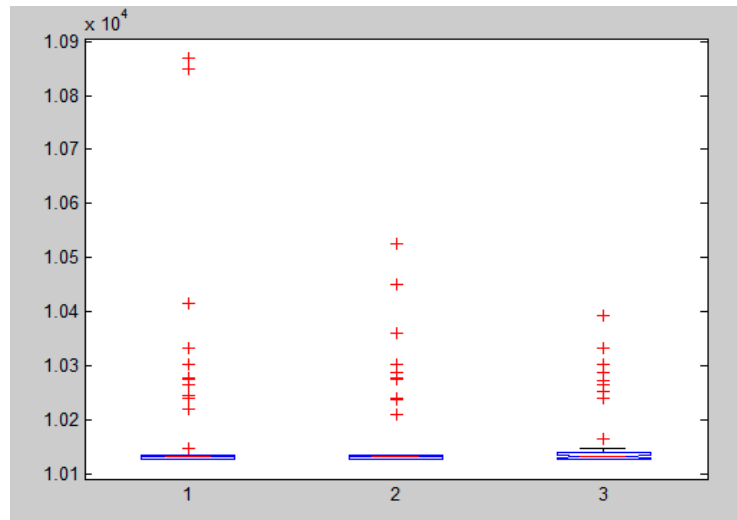


Gráfico 18. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 100.00 evaluaciones en el ensayo 1PT.P50.M4.

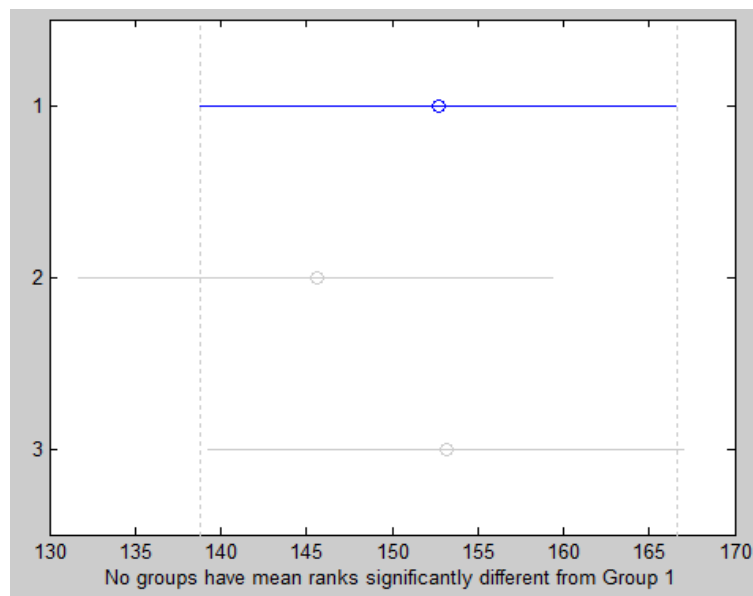


Gráfico 19. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 100.00 evaluaciones en el ensayo 1PT.P50.M4

Para 1.000.000 evaluaciones

En este caso se analizan los resultados obtenidos para la última evaluación realizada en el ensayo con cruce 1 punto, población de 50 individuos y mutación de 0.4%. Al igual que en el caso anterior se muestran primero los gráficos y tablas para el test de Kruskal-Wallis para las 20 codificaciones ensayadas y a continuación para las 3 muestras que resultaron significativas.

A modo de resumen, las codificaciones tipo Gray (grupos 1, 2 y 3 en los gráficos), son significativamente mejores que el resto (mediana del peso restringido inferior al resto), y no se aprecian diferencias significativas entre las tres codificaciones tipo Gray. Además, cabe destacar que el peor comportamiento lo presenta la codificación S27C1,38. Este comportamiento se comprueba tanto en el gráfico de cajas (gráfico 20) como en el de multicomparación (gráfico 21). Además, y teniendo en cuenta que los grupos de datos están ordenados de menor a mayor suma de diferencias en la codificación, se observa como el grupo 4 (correspondiente a la codificación binaria, S26A0,92), presenta resultados bastante peores que codificaciones con mayor suma de diferencias (como los grupos 5 y 6 que corresponden a las codificaciones S27A0,40 y S27A1,33).

Finalmente, y evaluando las gráficas en función de las varianzas que presenta cada codificación en función de las diferencias de bits, se observa un comportamiento irregular, ya que la codificación S27C1,38 (grupo 7 en los gráficos) presenta la segunda mayor varianza de todas las codificaciones, mientras que la codificación con mayor varianza es la S39C1,40 (grupo 13) se encuentra dentro del los valores medios del resto de codificaciones, es decir, no presenta un comportamiento destacablemente peor que el resto.

Con respecto a los tres primeros grupos de datos (codificaciones tipo Gray) se repiten los resultados obtenidos en el apartado anterior, y es que no presentan significancia entre sí (Q=0,6309).

Kruskal-Wallis ANOVA Table					
Source	SS	df	MS	Chi-sq	Prob>Chi-sq
Columns	4.6884e+008	19	24675809	1406.06	6.40328e-287
Error	1.97713e+008	1980	99854.9		
Total	6.66553e+008	1999			

Tabla 18. Tabla Kruskal-Wallis ANOVA con los resultados extraídos del test.

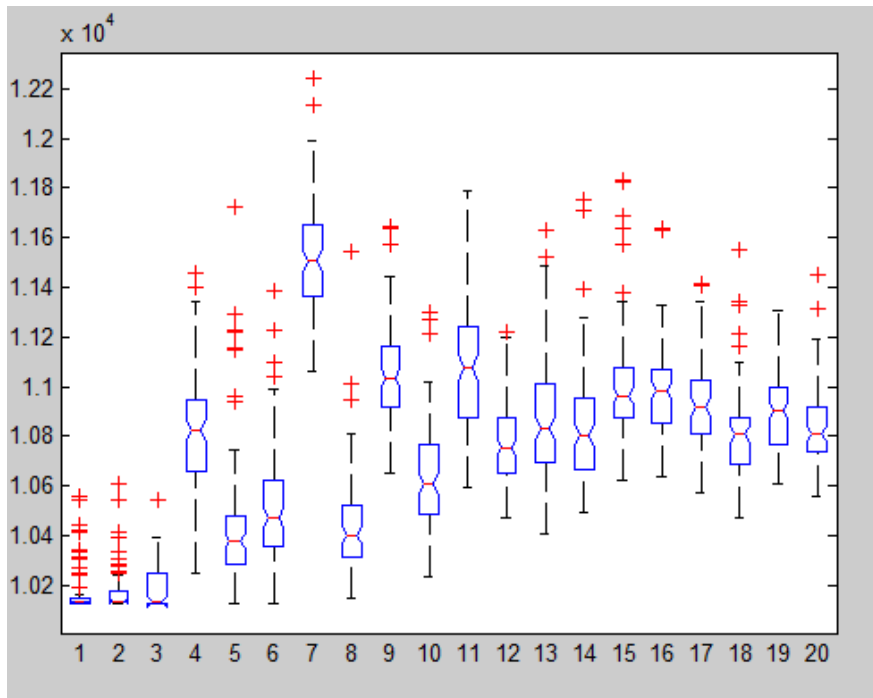


Gráfico 20. Gráfico de cajas para los resultados a 1.000.000 evaluaciones del ensayo 1PT.P50.M4 para las 20 codificaciones ensayadas.

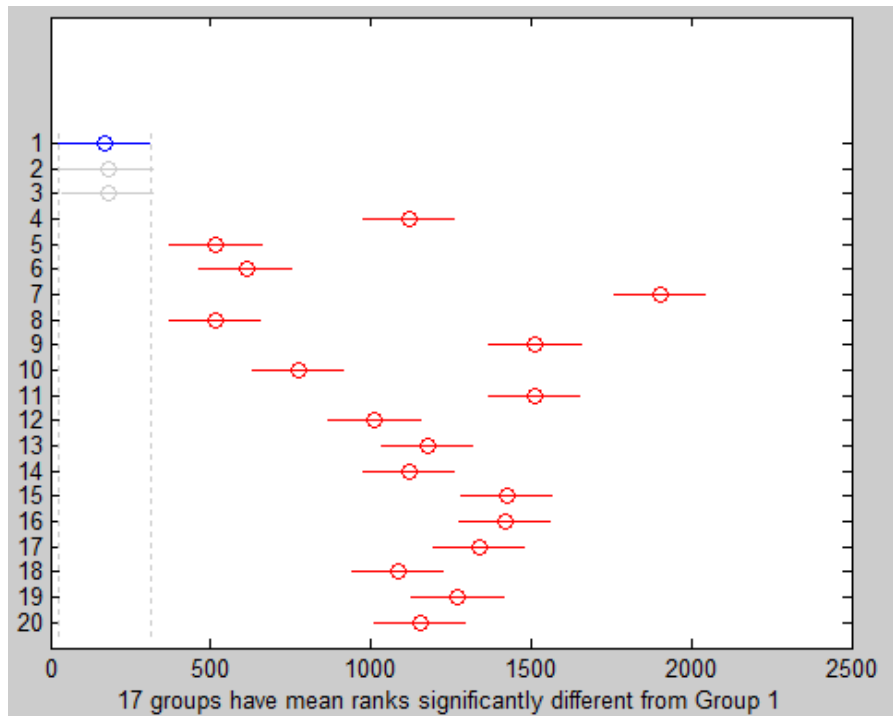


Gráfico 21. Multicomparación de las 20 codificaciones para 1.000.000 evaluaciones en el ensayo 1PT.P50.M4.

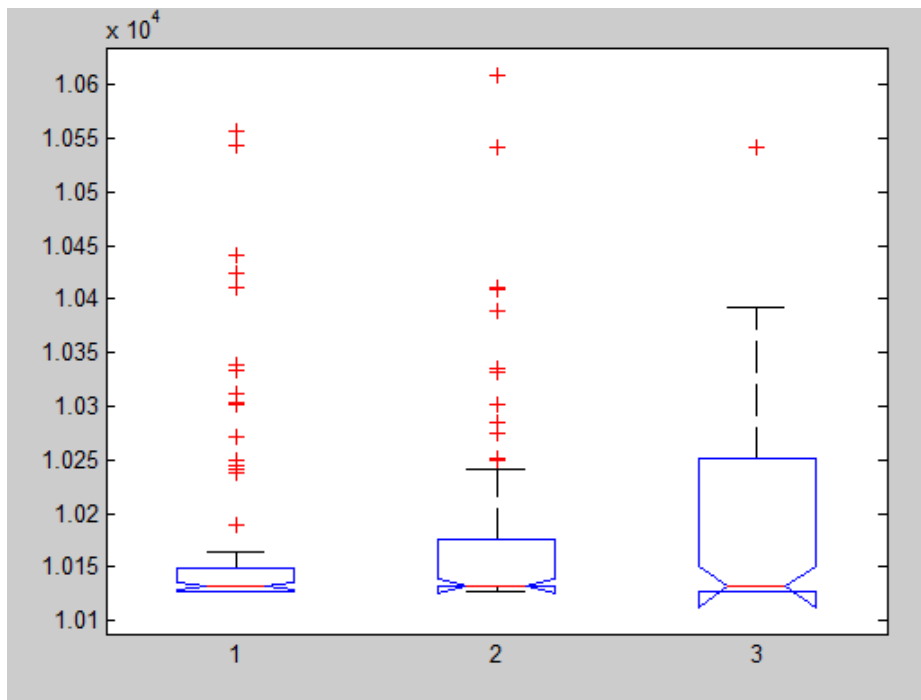


Gráfico 22. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo 1PT.P50.M4.

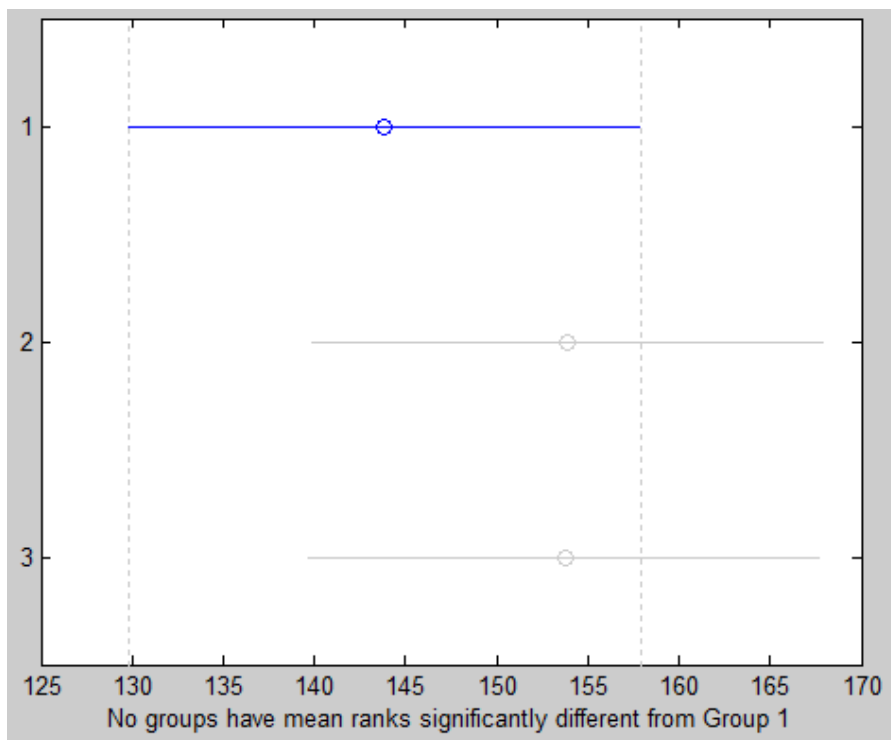


Gráfico 23. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo 1PT.P50.M4.

Resultados: ensayo 1PTO.P150.M4

Con el fin de facilitar la revisión de los datos obtenidos y sus gráficas representativas, a continuación se presentarán sólo los gráficos más significativos para cada grupo de resultados, así como el valor de Q. Previamente se puede encontrar un breve párrafo resumen con algunas observaciones respecto cada grupo de gráficos.

Para 100.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $6,12 \cdot 10^{-298}$). Los peores resultados se encuentran para el grupo 11, correspondiente a la codificación S39A0,49. No se observa tendencia alguna relacionada con el valor de suma de diferencia de bits ni con la varianza, aparentemente.

No existe significancia entre las tres ordenaciones tipo Gray entre sí (valor de Q de 0,9643).

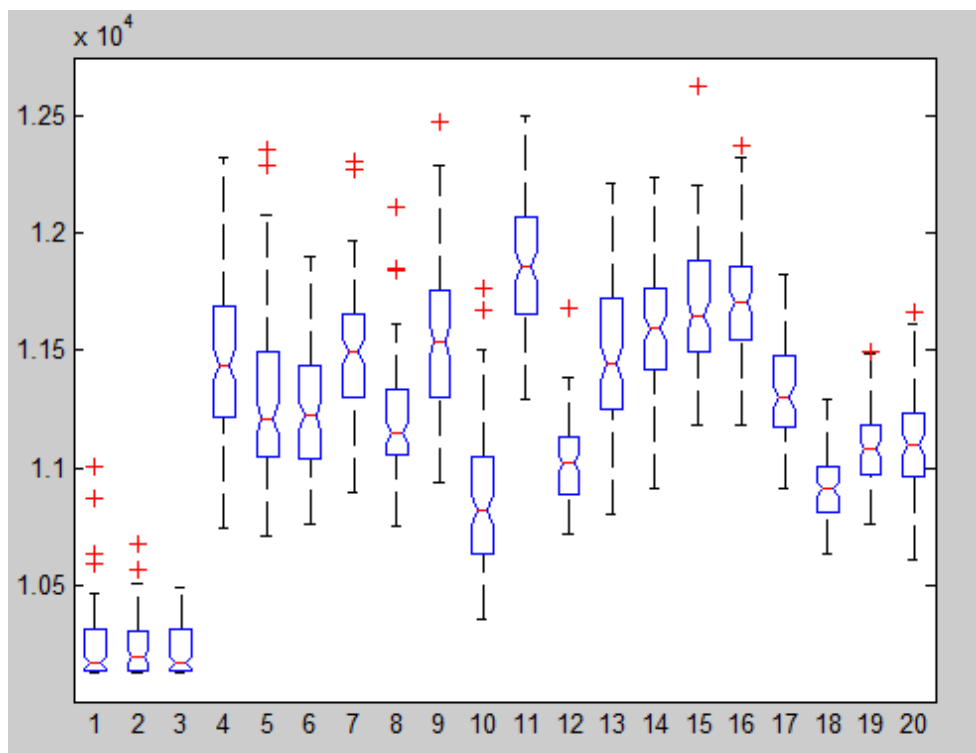


Gráfico 24. Gráfico de cajas para los resultados a 100.000 evaluaciones del ensayo 1PT.P150.M4 para las 20 codificaciones ensayadas.

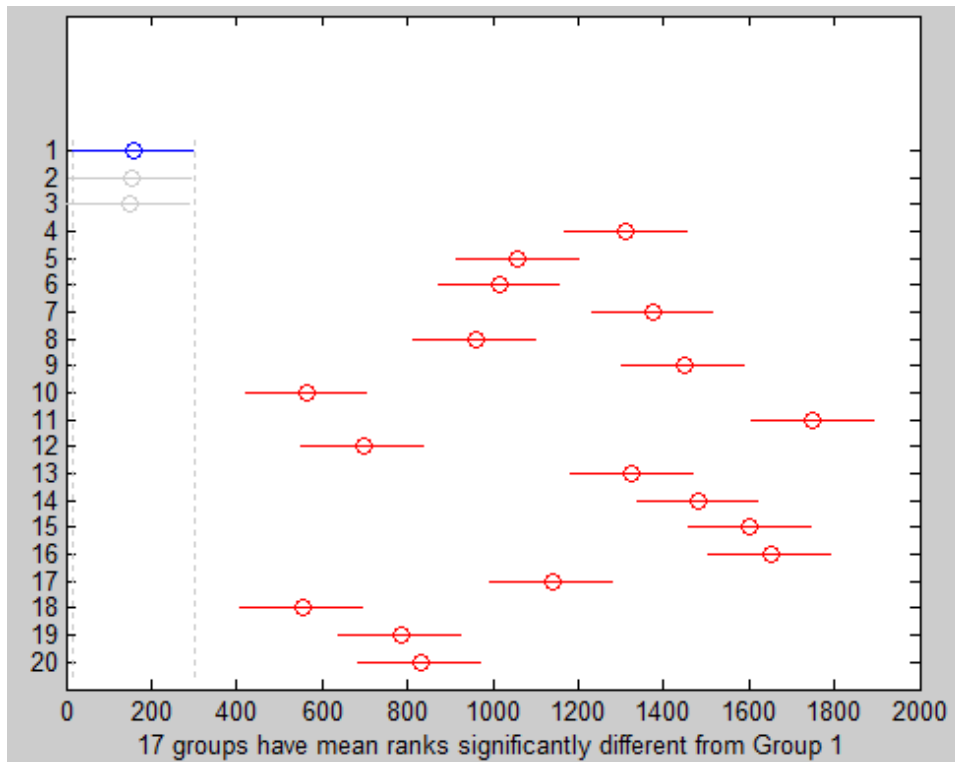


Gráfico 25. Multicomparación de las 20 codificaciones para 100.000 evaluaciones en el ensayo 1PT.P150.M4.

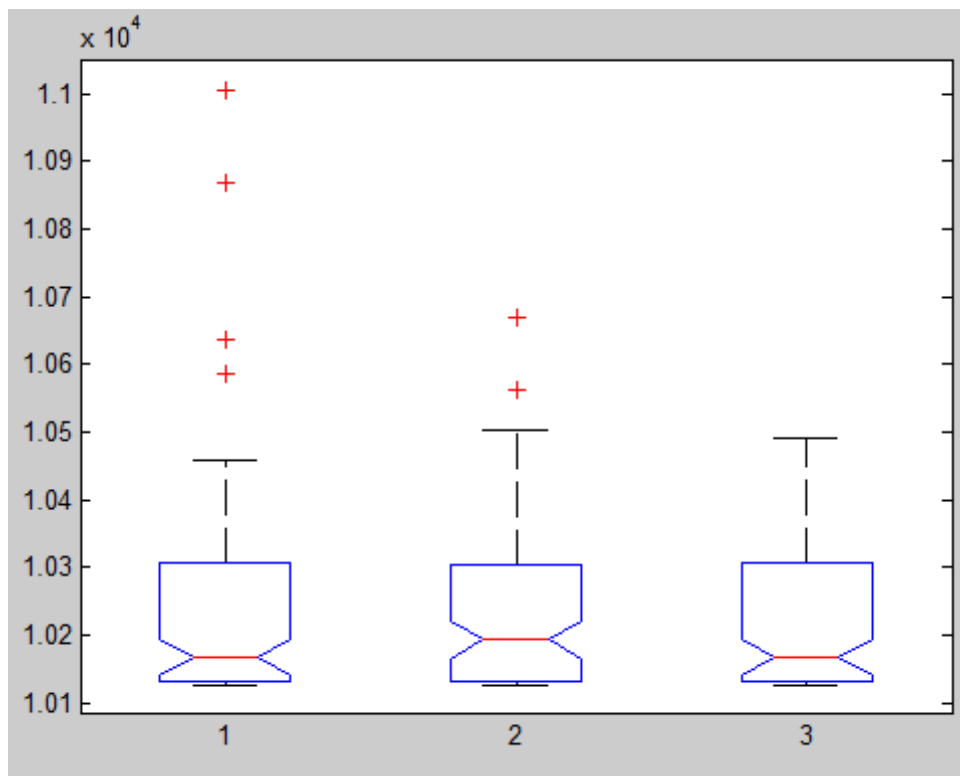


Gráfico 26. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo 1PT.P150.M4.

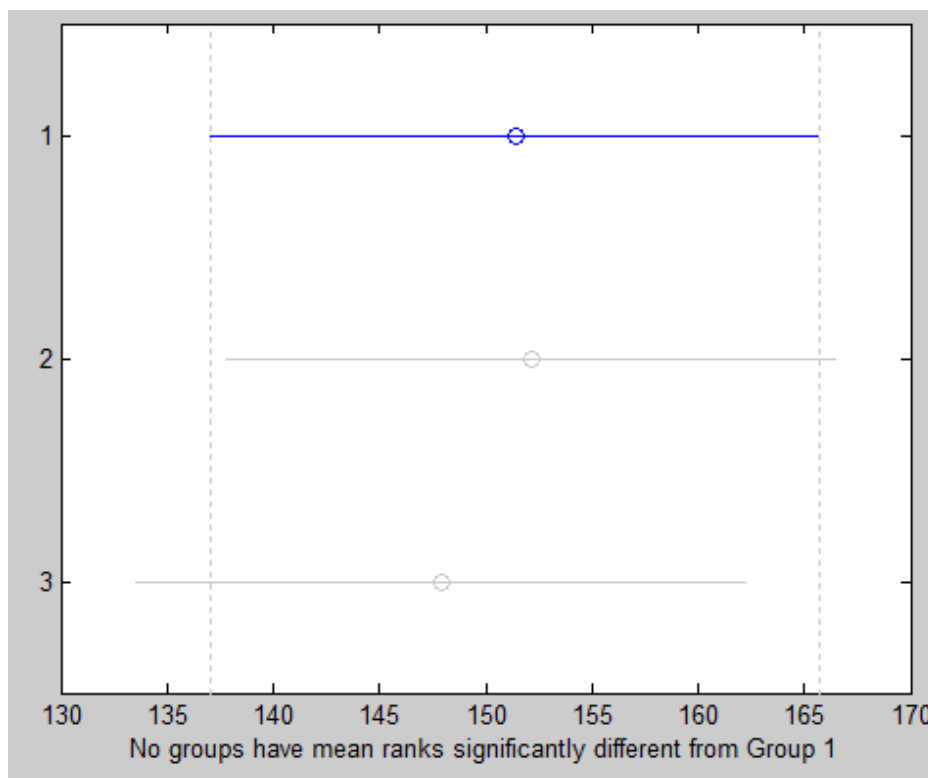


Gráfico 27. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo 1PT.P150.M4.

Para 1.000.000 evaluaciones

59

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $6,98 \cdot 10^{-295}$). Los peores resultados se encuentran para el grupo 7, correspondiente a la codificación **S27A1,33**. No se observa tendencia alguna relacionada con el valor de suma de diferencia de bits ni con la varianza, aparentemente.

Se realiza de nuevo el test para los grupos 1, 2, 3 y 5 (codificaciones S15A0, S15B0, S15C0 y S27A0,40). El valor de Q indica que sí hay significancia ($Q=8,35 \cdot 10^{-49}$) y mediante la gráfica de multicomparación se descarta la posible significancia de los resultados correspondientes a la codificación del grupo 5 (S27A0,40).

Destaca en este caso que, al igual que en el ensayo anterior (1PTO.P50.M4), para 100.000 ordenaciones es el grupo 11 (S39C1,40) el que peores resultados presenta, mientras que al final de las evaluaciones (1.000.000), el valor medio de peso restringido para esta ordenación se encuentra dentro del valor medio para la gran parte de ordenaciones empleadas, mientras que es el grupo 7 (S27C1,33) la que presenta una mediana muy superior al resto.

Finalmente, es destacable la forma que presenta el gráfico de comparación entre los dos ensayos vistos hasta ahora, que tienen forma similar en ambos casos para 1.000.000 de evaluaciones (comparar gráfico 21 y 29).

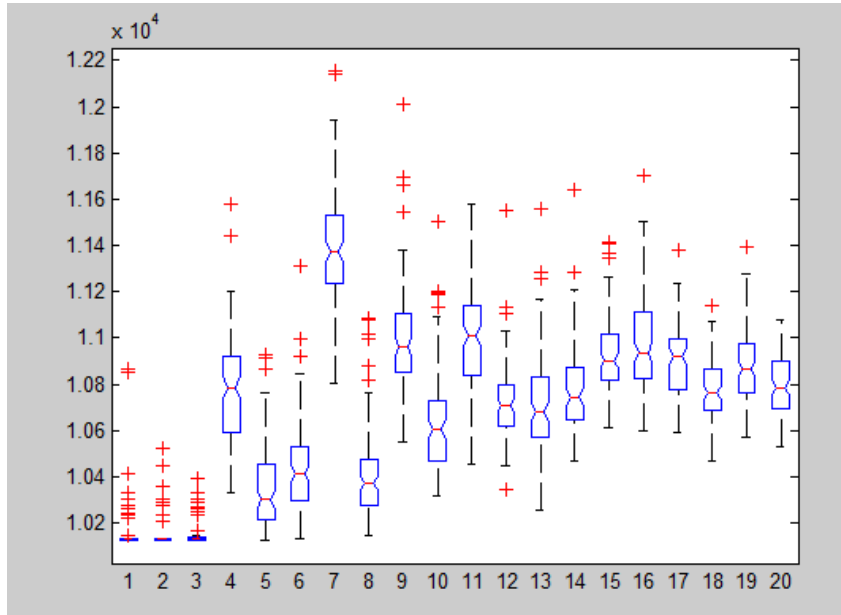


Gráfico 28. Gráfico de cajas para los resultados a 1.000.000 evaluaciones del ensayo 1PT.P150.M4 para las 20 codificaciones ensayadas.

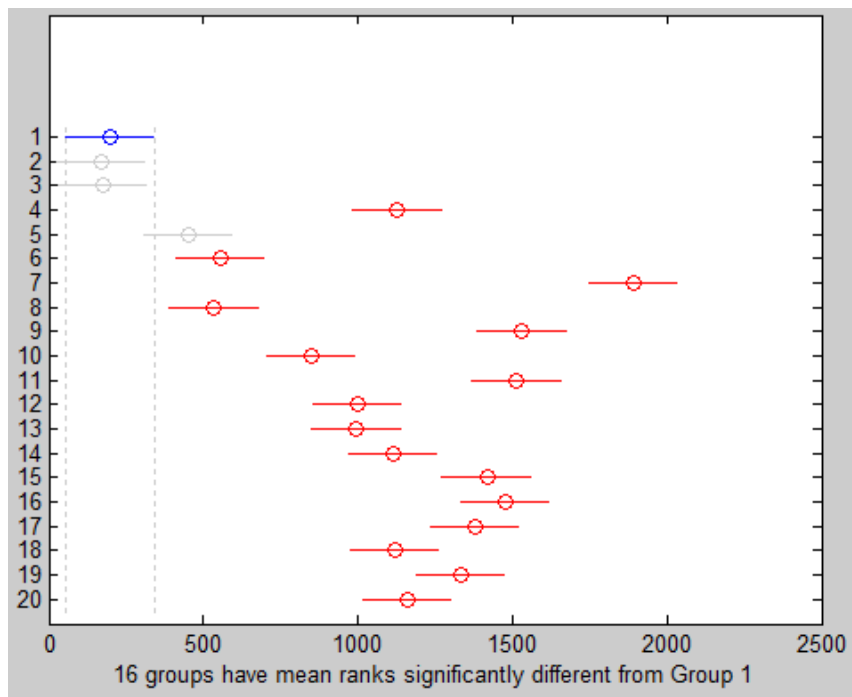


Gráfico 29. Multicomparación de las 20 codificaciones para 1.000.000 evaluaciones en el ensayo 1PT.P150.M4.

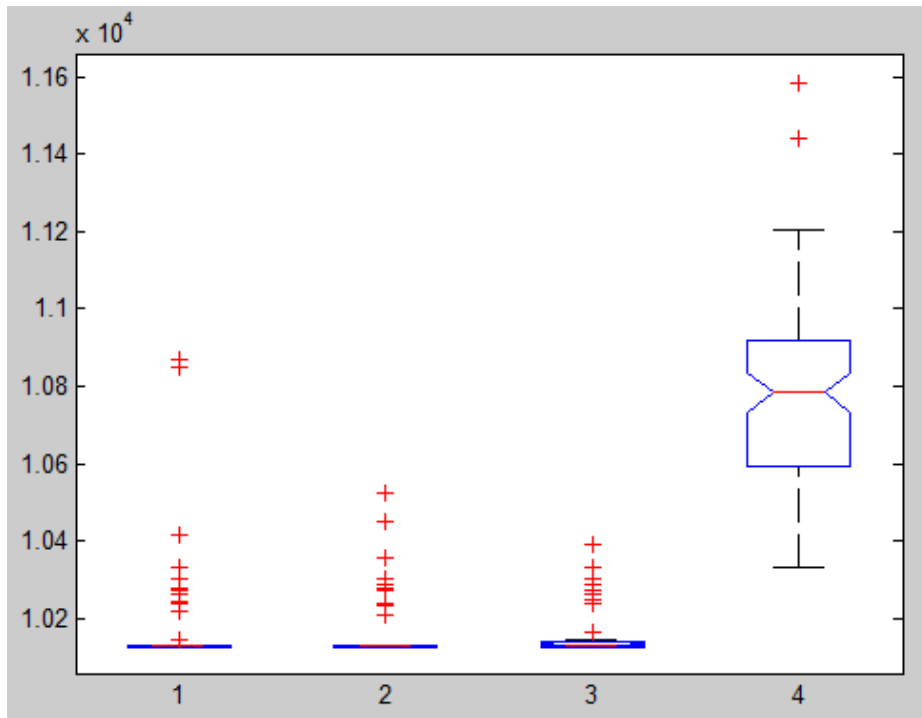


Gráfico 30. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo 1PT.P150.M4.

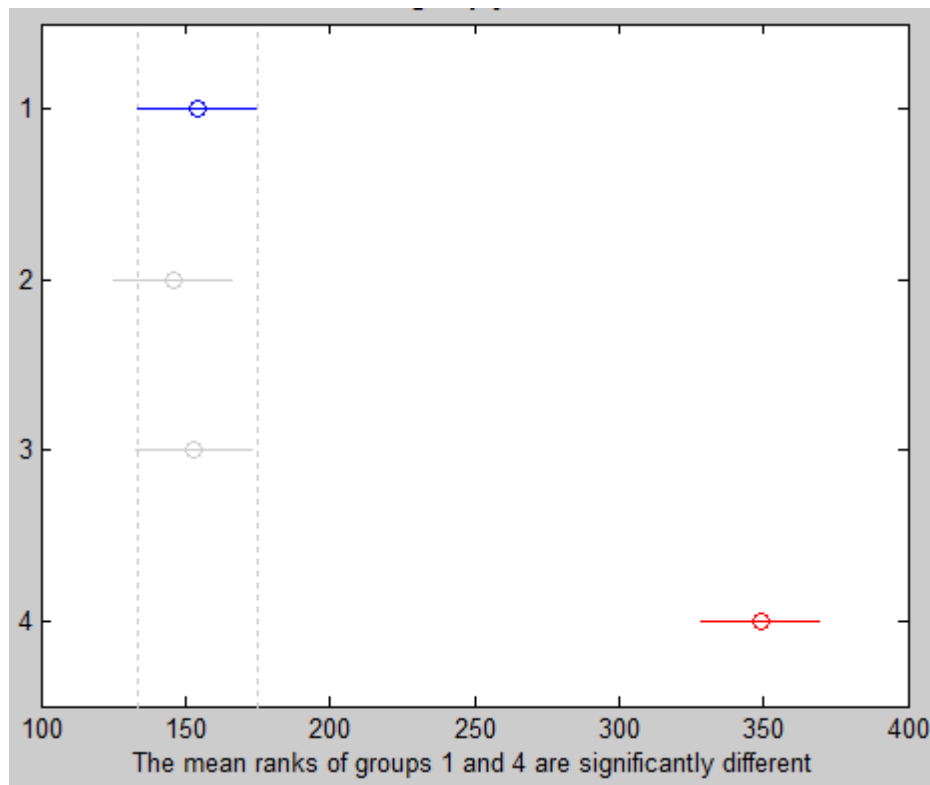


Gráfico 31. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo 1PT.P150.M4.

Resultados: ensayo UNI.P50.M4

Para 100.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $3,20 \cdot 10^{-310}$). Los **peores resultados** se encuentran para el grupo 11, correspondiente a la **codificación S39A0,49**.

No existe significancia entre las tres ordenaciones tipo Gray entre sí (valor de Q de 0,8053).

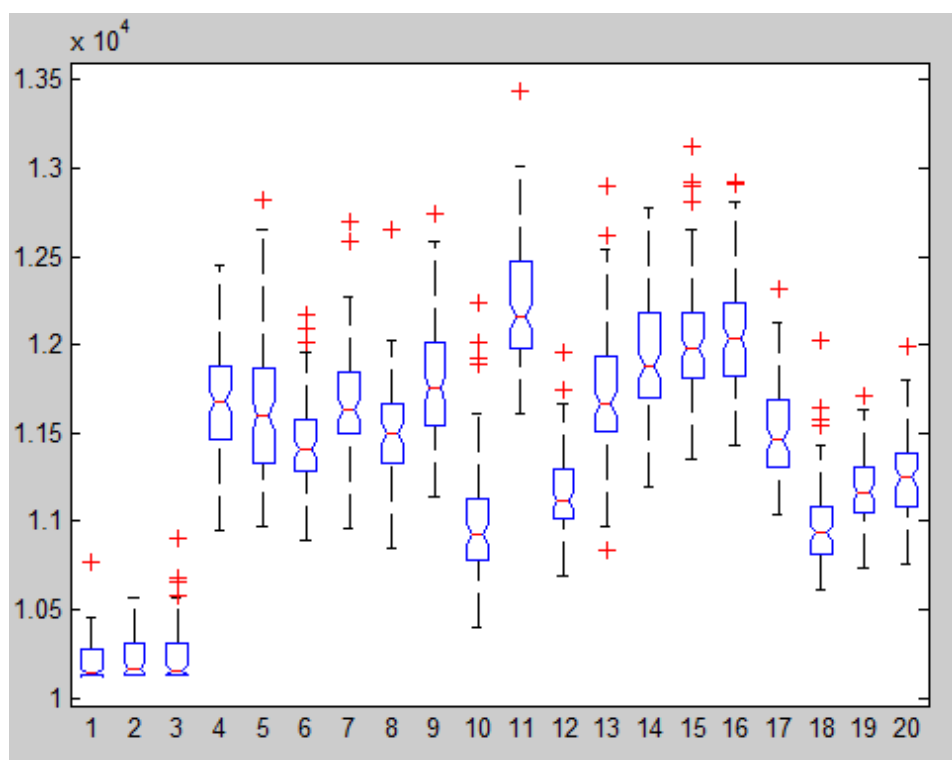


Gráfico 32. Gráfico de cajas para los resultados a 100.000 evaluaciones del ensayo UNI.P50.M4 para las 20 codificaciones ensayadas.

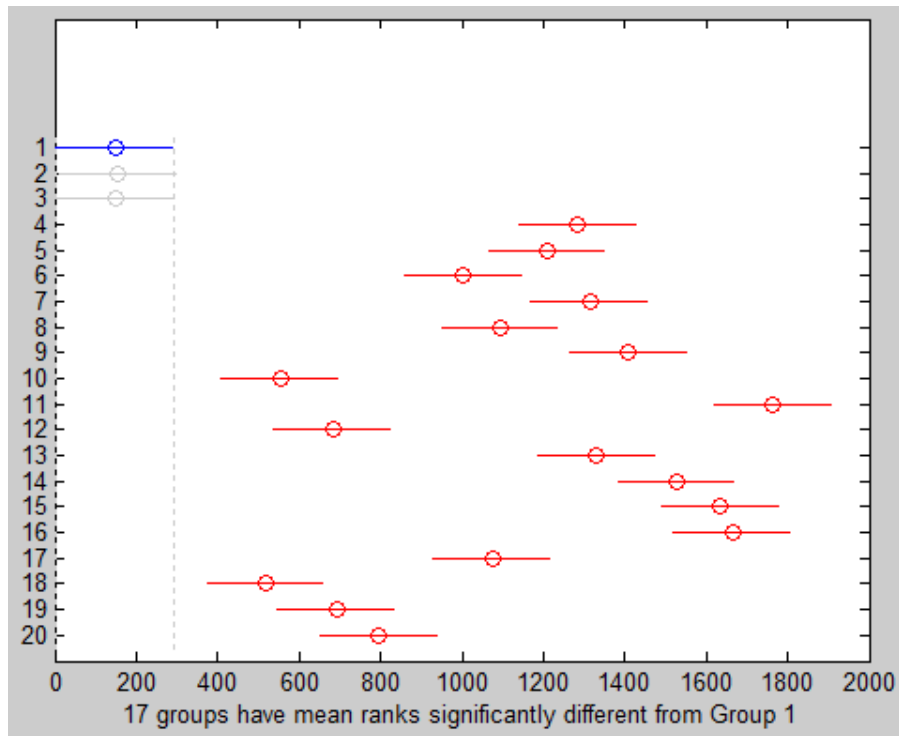


Gráfico 33. Multicomparación de las 20 codificaciones para 100.000 evaluaciones en el ensayo UNI.P50.M4.

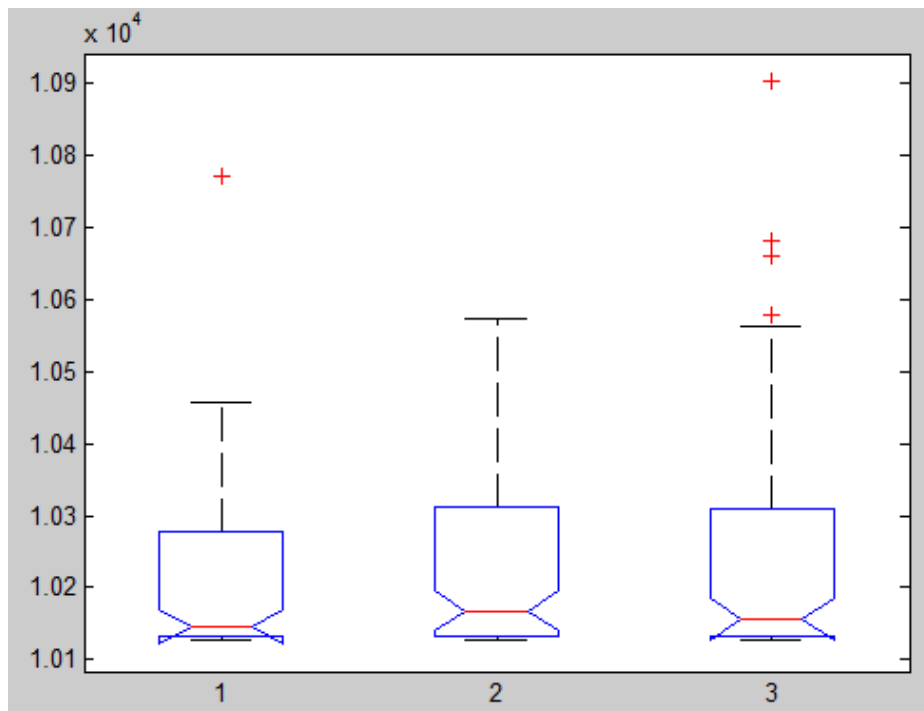


Gráfico 34. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo UNI.P50.M4.

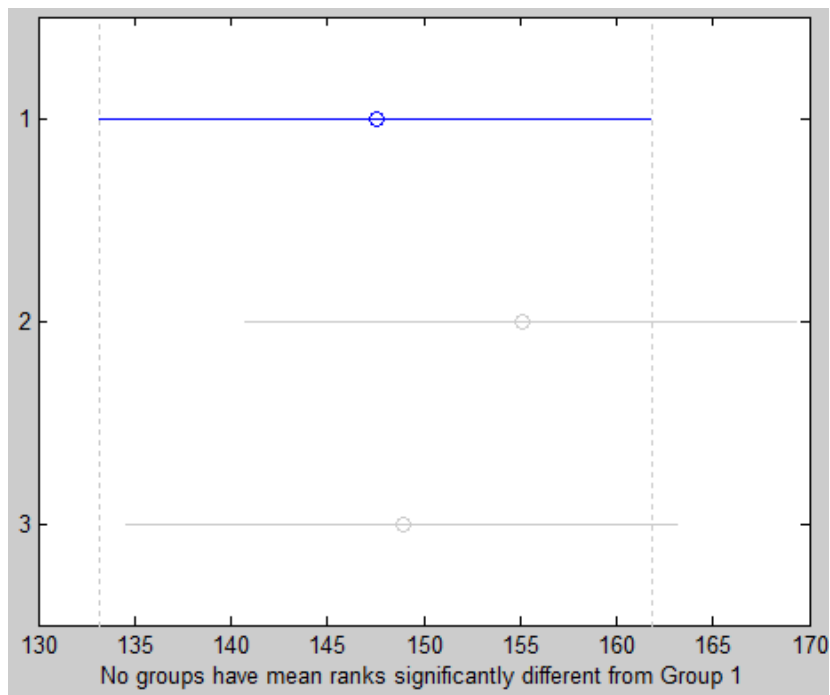


Gráfico 35. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo UNI.P50.M4.

Para 1.000.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $3,70 \cdot 10^{-304}$). Los peores resultados se encuentran para el grupo 7, correspondiente a la codificación **S27A1,33**. No existe significancia entre las tres ordenaciones tipo Gray entre sí (valor de Q de 0,78).

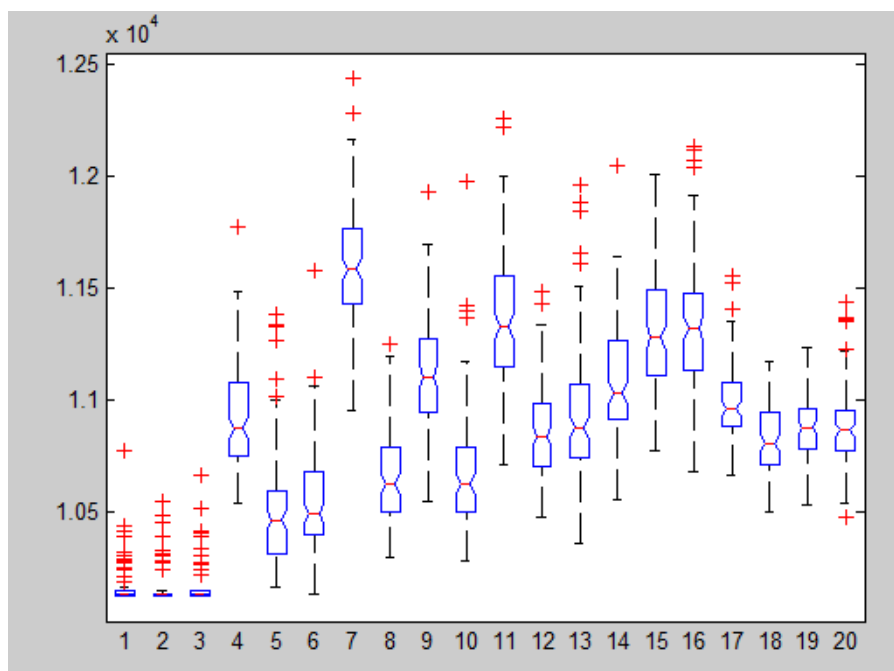


Gráfico 36. Gráfico de cajas para los resultados a 1.000.000 evaluaciones del ensayo UNI.P50.M4 para las 20 codificaciones ensayadas.

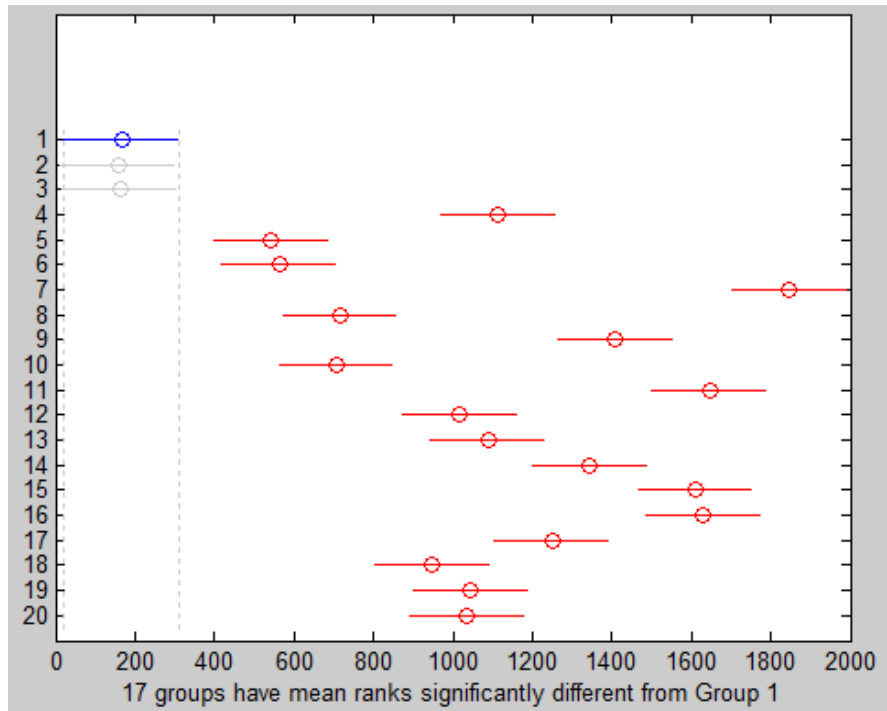


Gráfico 37. Multicomparación de las 20 codificaciones para 1.000.000 evaluaciones en el ensayo UNI.P50.M4.

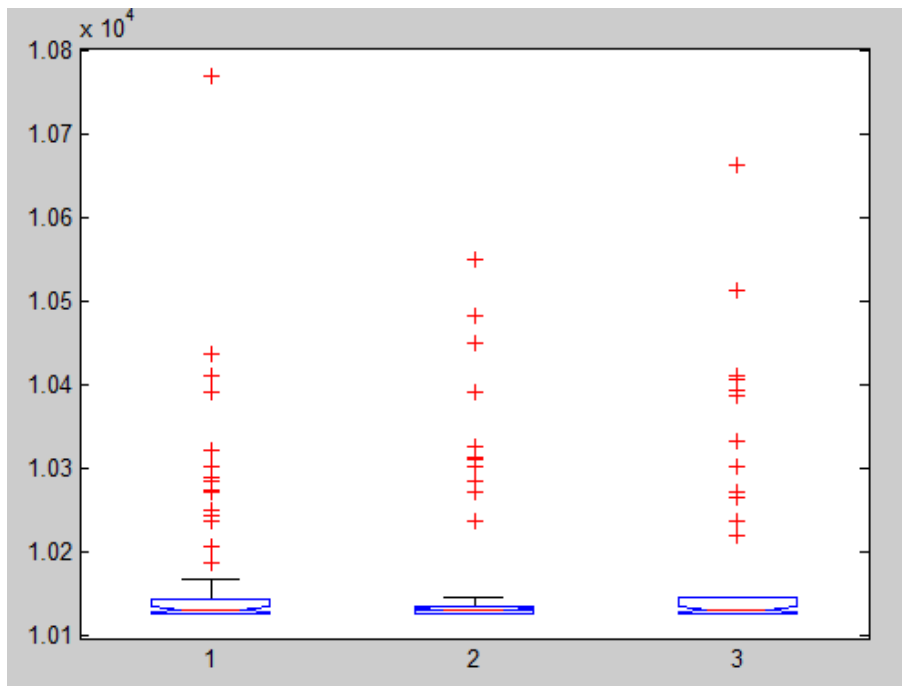


Gráfico 38. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo UNI.P50.M4.

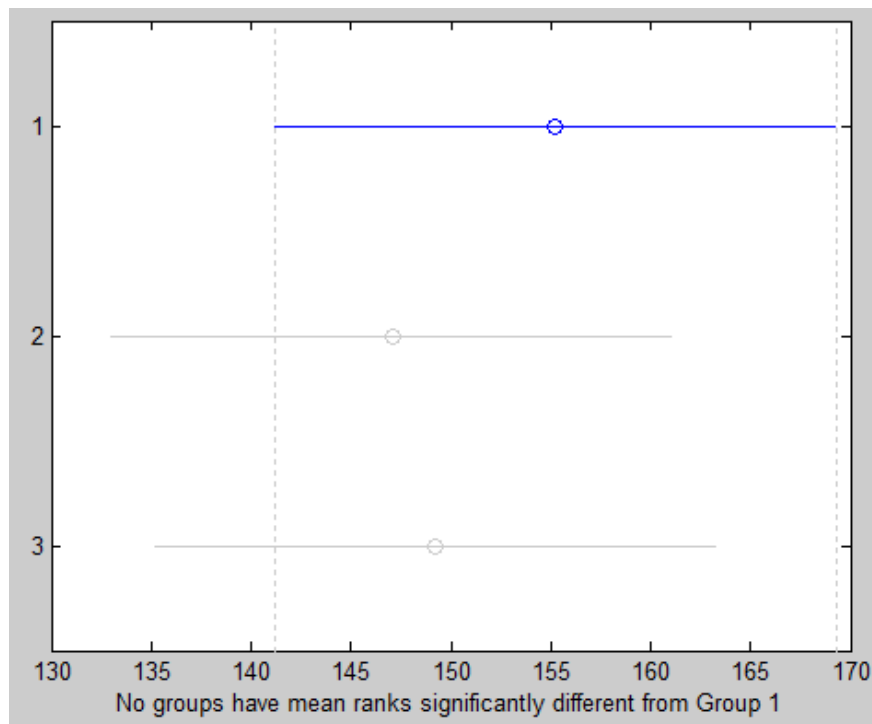


Gráfico 39. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo UNI.P50.M4.

Resultados: ensayo UNI.P150.M4

Para 100.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de 0). Los **peores resultados** se encuentran para el grupo 11, correspondiente a la **codificación S39A0,49**.

No existe significancia entre las tres ordenaciones tipo Gray entre sí (valor de Q de 0,1474).

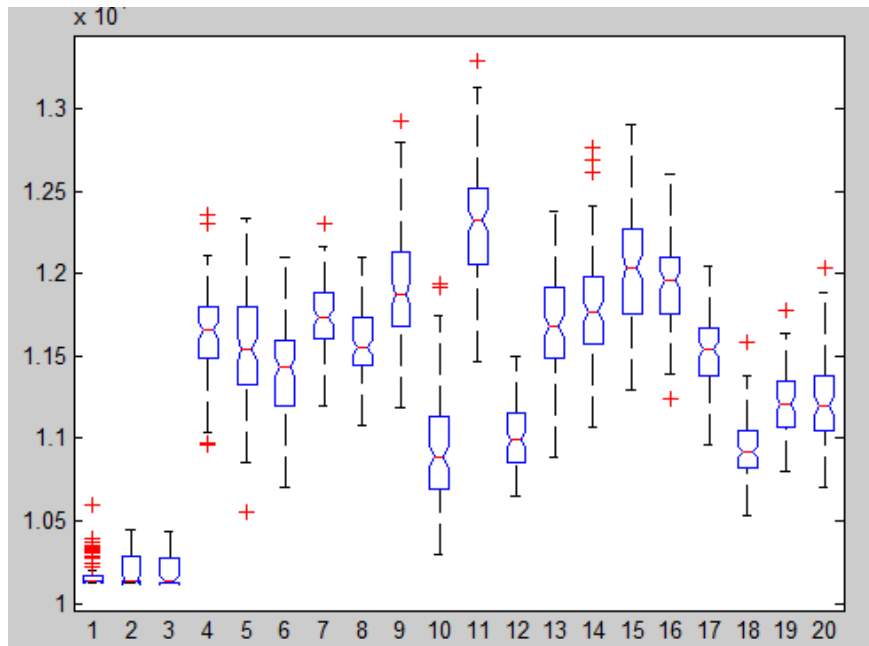


Gráfico 40. Gráfico de cajas para los resultados a 100.000 evaluaciones del ensayo UNI.P150.M4 para las 20 codificaciones ensayadas.

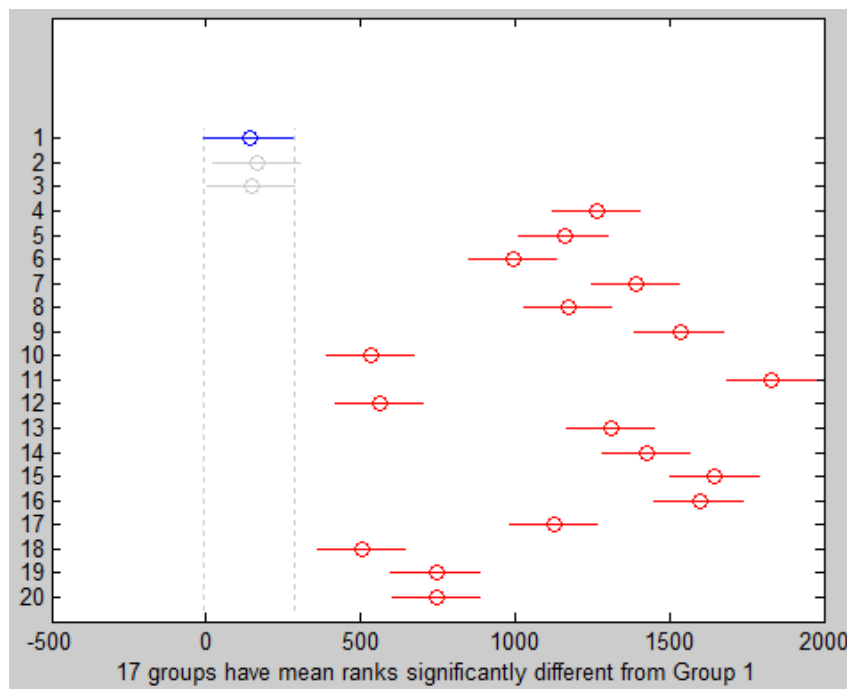


Gráfico 41. Multicomparación de las 20 codificaciones para 100.000 evaluaciones en el ensayo UNI.P150.M4.

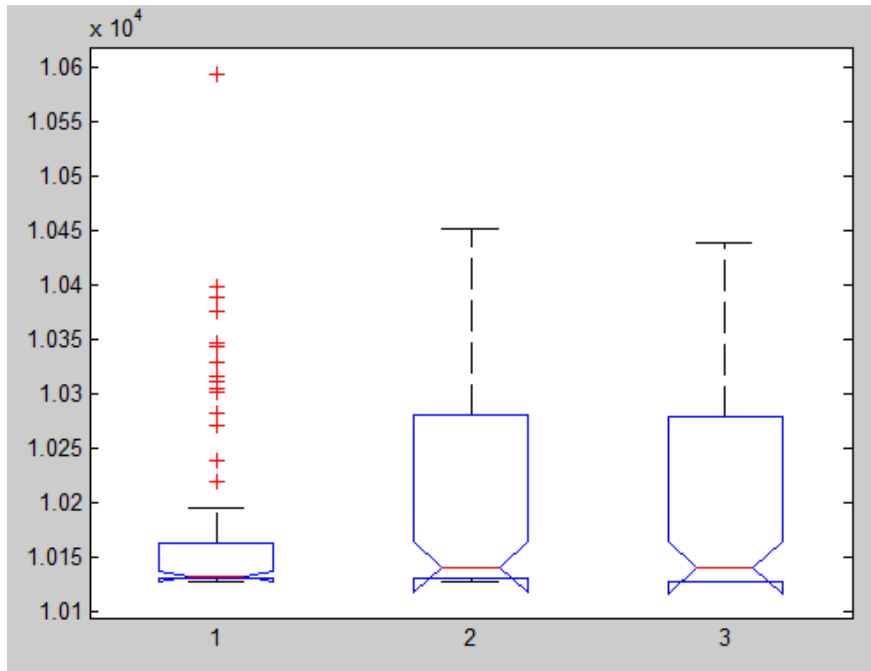


Gráfico 42. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo UNI.P150.M4.

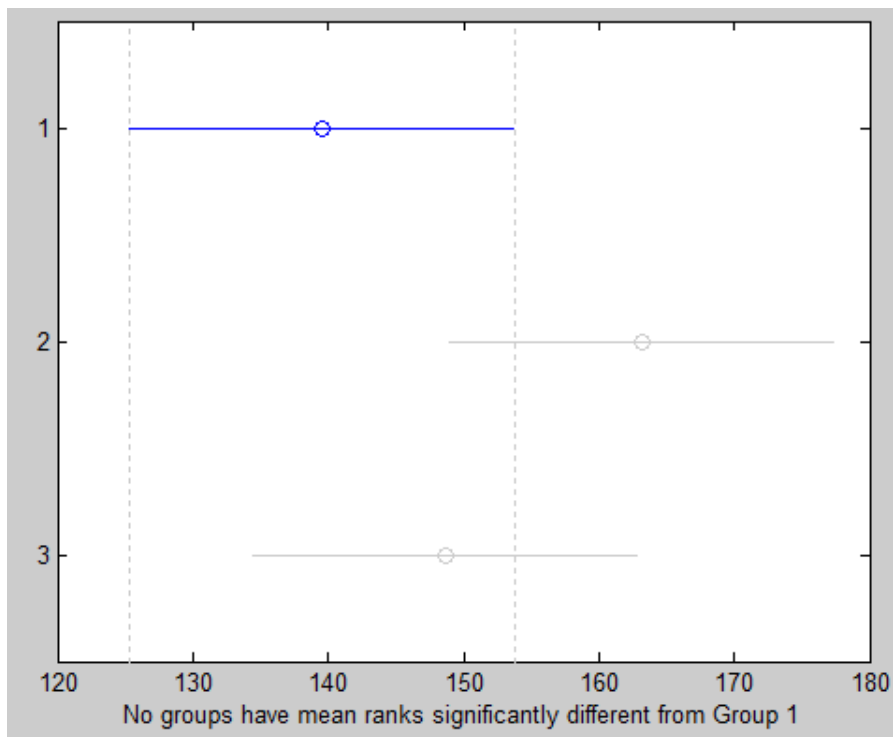


Gráfico 43. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo UNI.P150.M4.

Para 1.000.000 evaluaciones

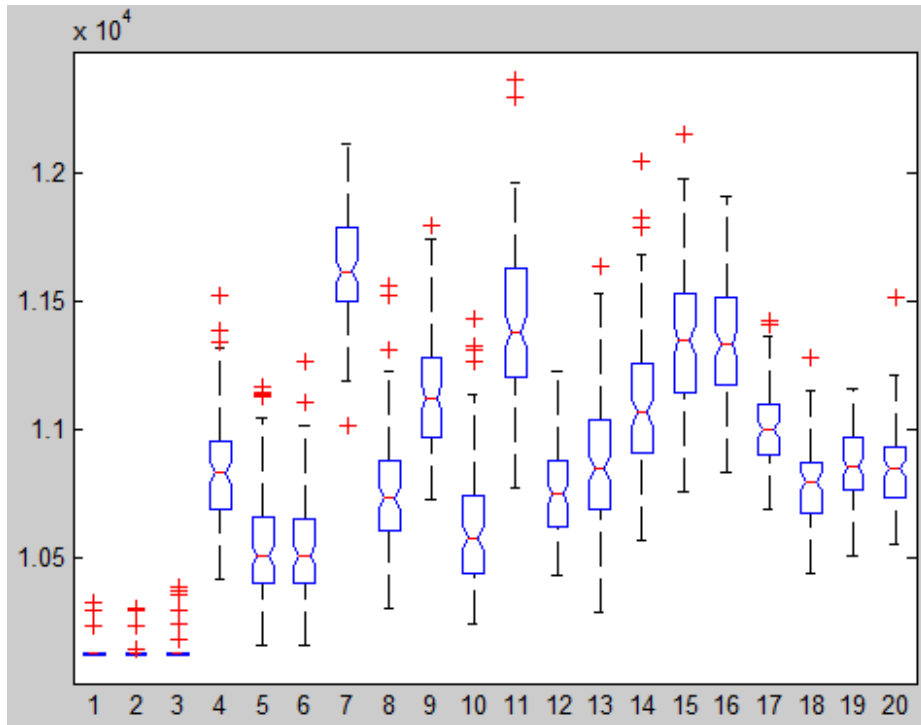


Gráfico 44. Gráfico de cajas para los resultados a 1.000.000 evaluaciones del ensayo UNI.P150.M4 para las 20 codificaciones ensayadas.

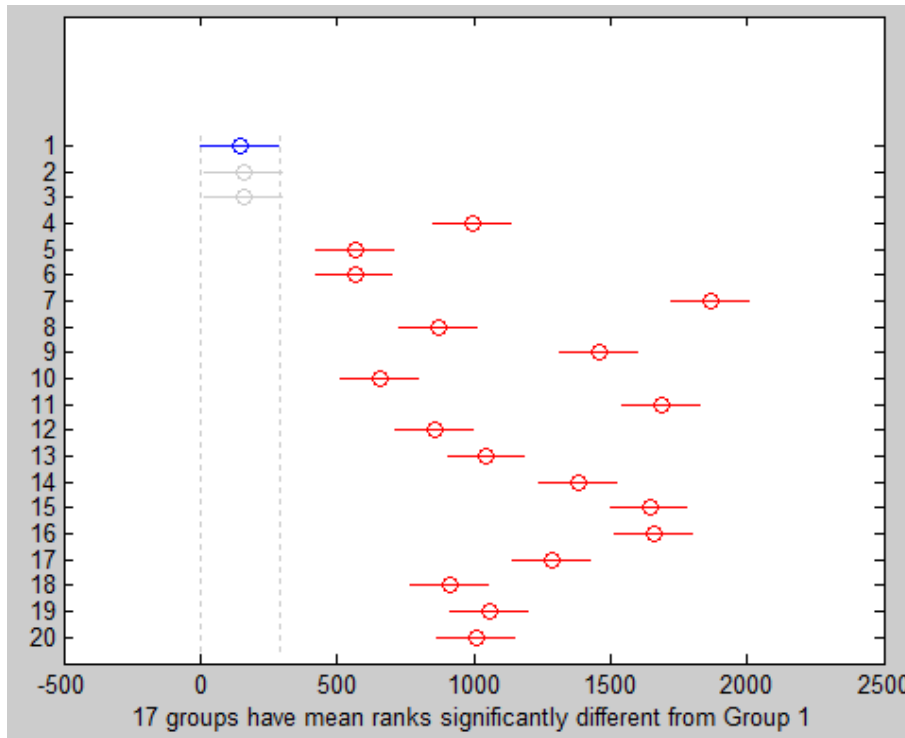


Gráfico 45. Multicomparación de las 20 codificaciones para 1.000.000 evaluaciones en el ensayo UNI.P150.M4.

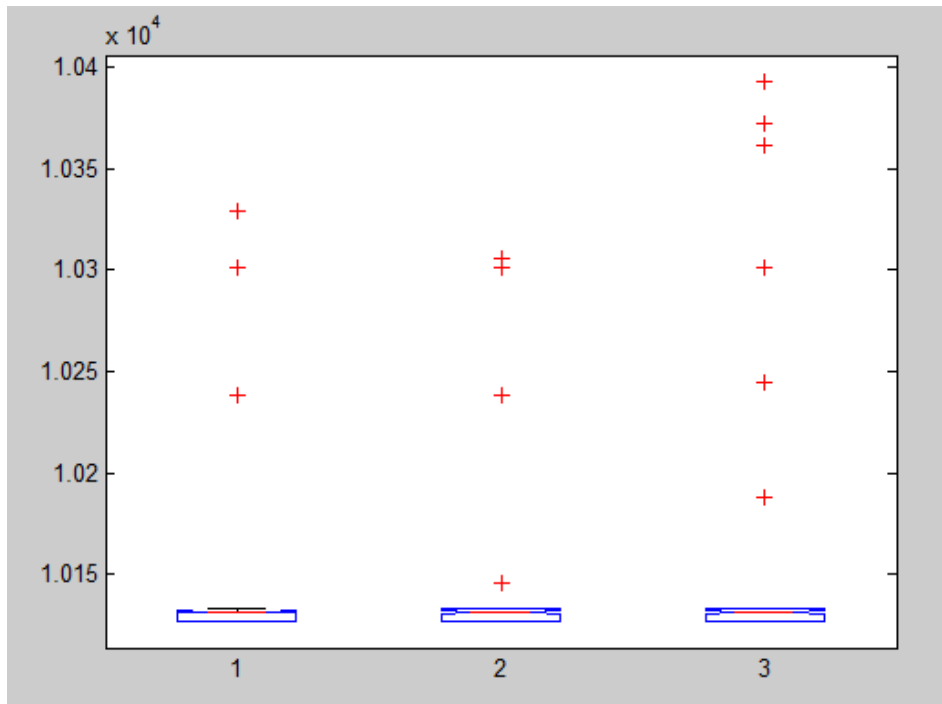


Gráfico 46. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo UNI.P150.M4.

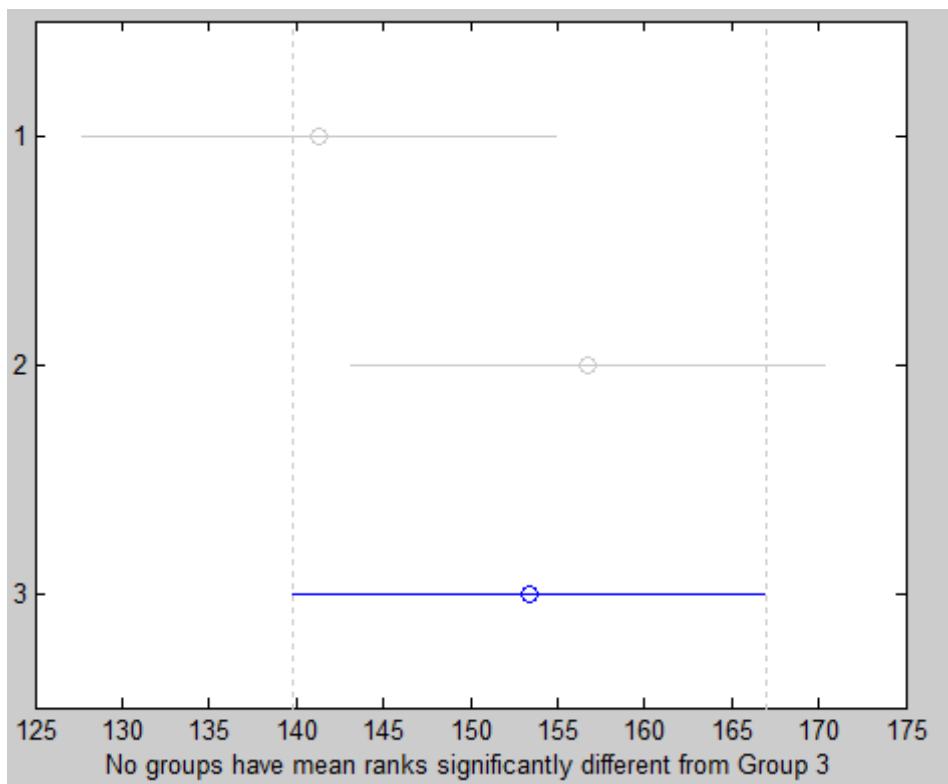


Gráfico 47. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo UNI.P150.M4.

Resultados: ensayo 1PT.P50.M8

Para 100.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $2,34 \cdot 10^{-283}$). Los peores resultados se encuentran para el grupo 11, correspondiente a la codificación S39A0,49.

No existe significancia entre las tres ordenaciones tipo Gray entre sí (valor de Q de 0,4933).

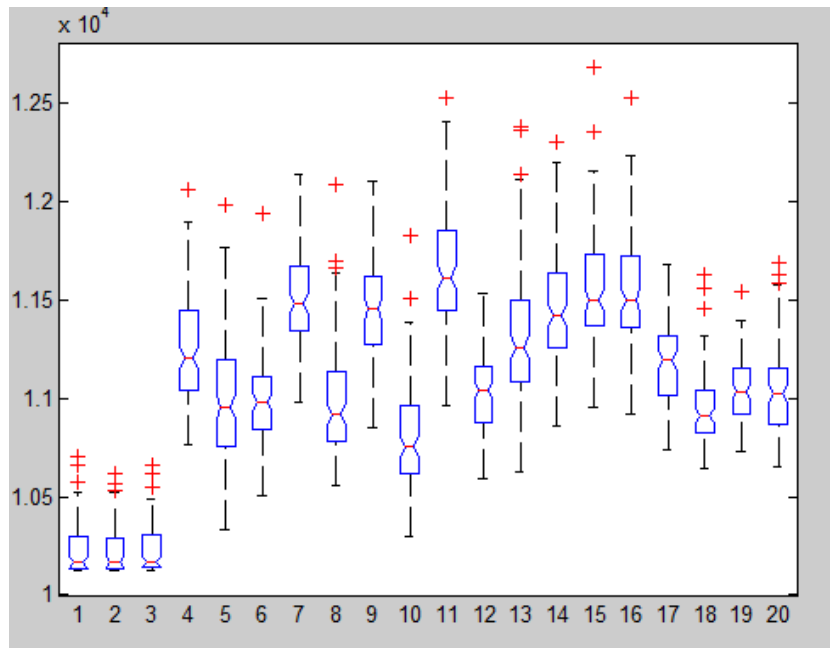


Gráfico 48. Gráfico de cajas para los resultados a 100.000 evaluaciones del ensayo 1PT.50.M8 para las 20 codificaciones ensayadas.

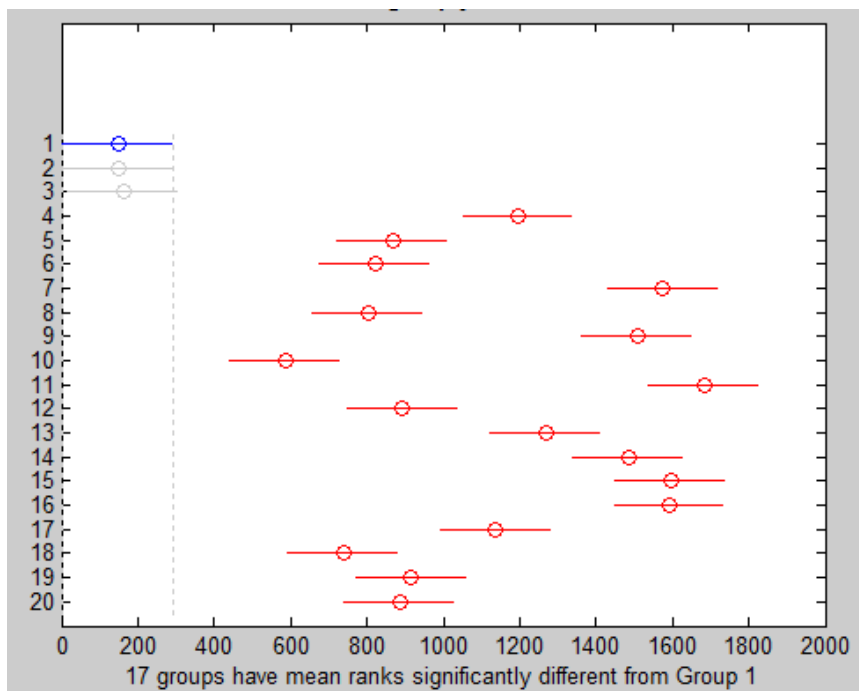


Gráfico 49. Multicomparación de las 20 codificaciones para 100.000 evaluaciones en el ensayo 1PT.P50.M8.

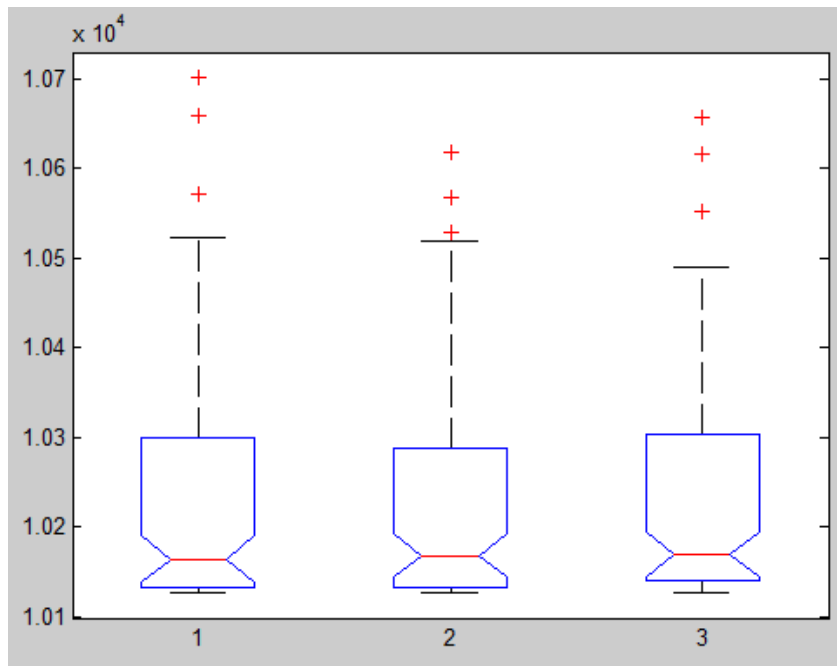


Gráfico 50. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo 1PT.P50.M8.

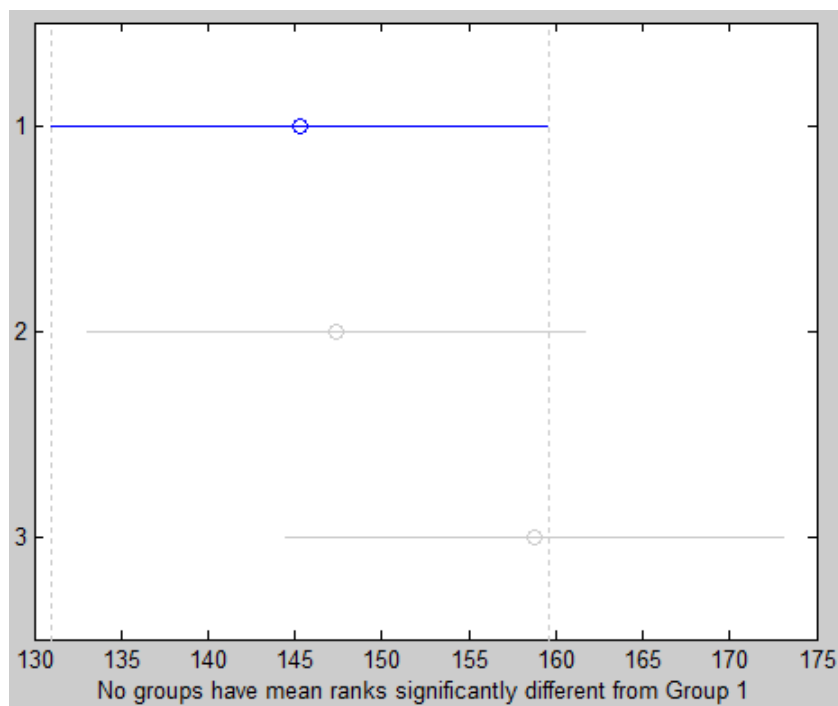


Gráfico 51. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo 1PT.P50.M8.

Para 1.000.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $8,70 \cdot 10^{-287}$). Los **peores resultados** se encuentran para el grupo 7, correspondiente a la **codificación S27A1,33**.

No existe significancia entre las tres ordenaciones tipo Gray entre sí (valor de Q de 0,7525).

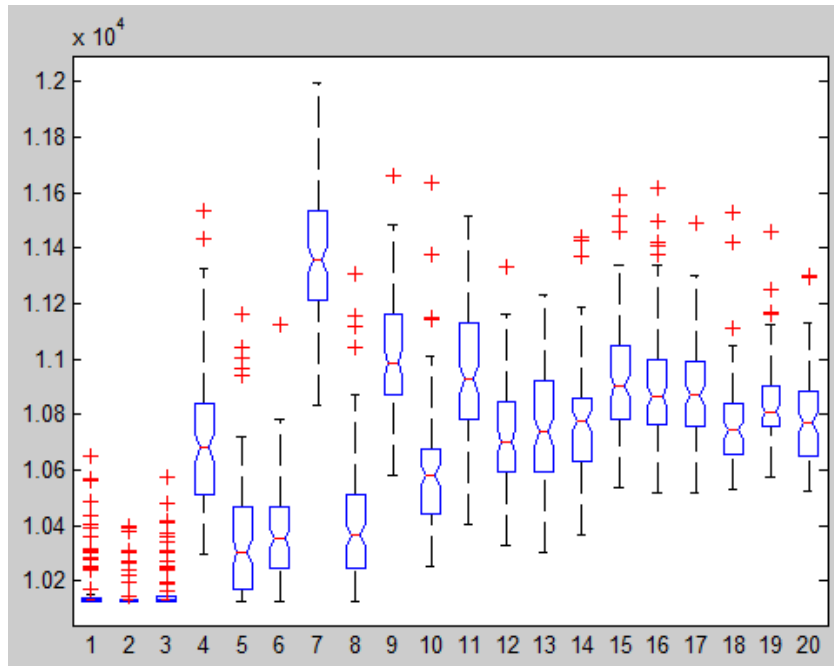


Gráfico 52. Gráfico de cajas para los resultados a 1.000.000 evaluaciones del ensayo 1PT.50.M8 para las 20 codificaciones ensayadas.

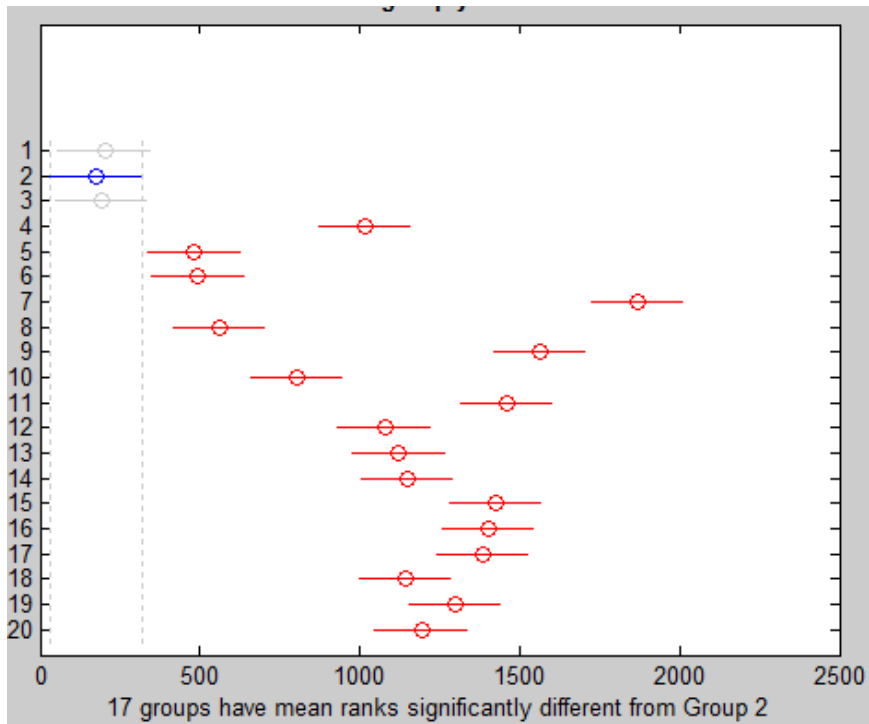


Gráfico 53. Multicomparación de las 20 codificaciones para 1.000.000 evaluaciones en el ensayo 1PT.P50.M8.

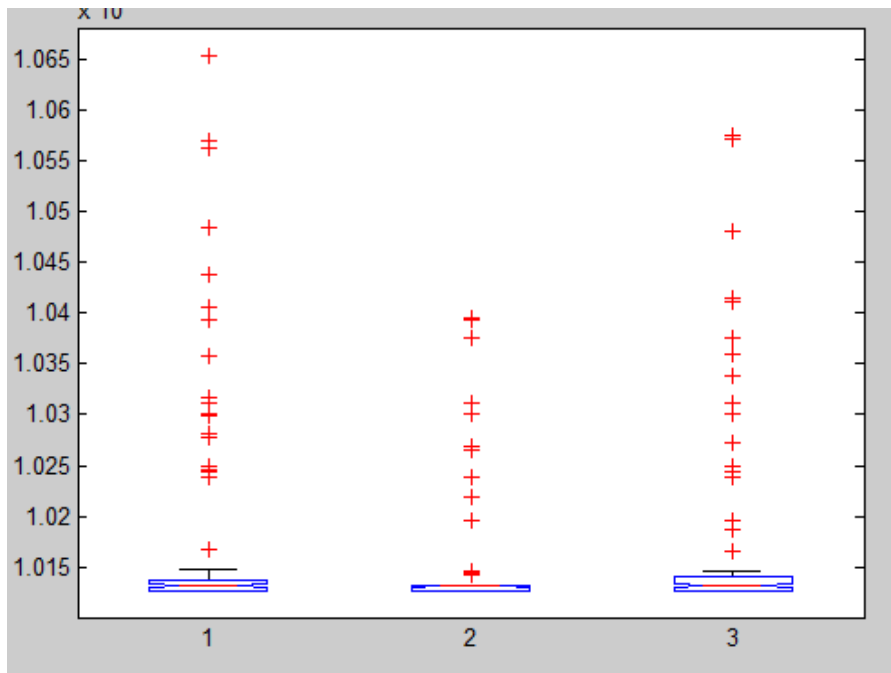


Gráfico 54. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo 1PT.P50.M8.

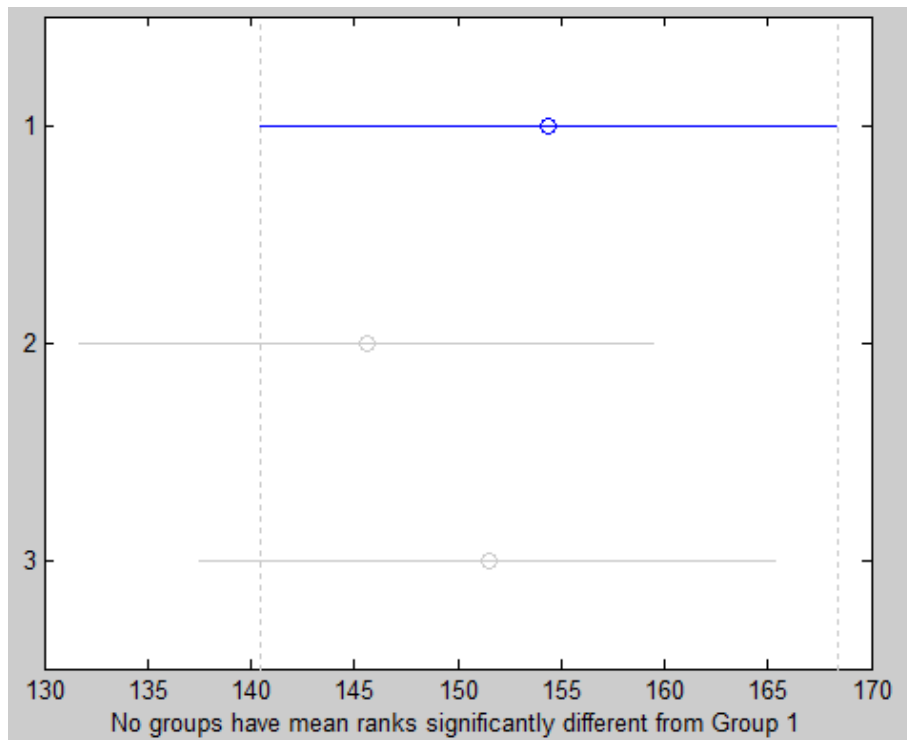


Gráfico 55. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo 1PT.P50.M8.

Resultados: ensayo 1PT.P150.M8

Para 100.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $4,8 \cdot 10^{-283}$). Los **peores resultados** se encuentran para el grupo 11, correspondiente a la **codificación S39A0,49**.

No existe significancia entre las tres ordenaciones tipo Gray entre sí (valor de Q de 0,4808).

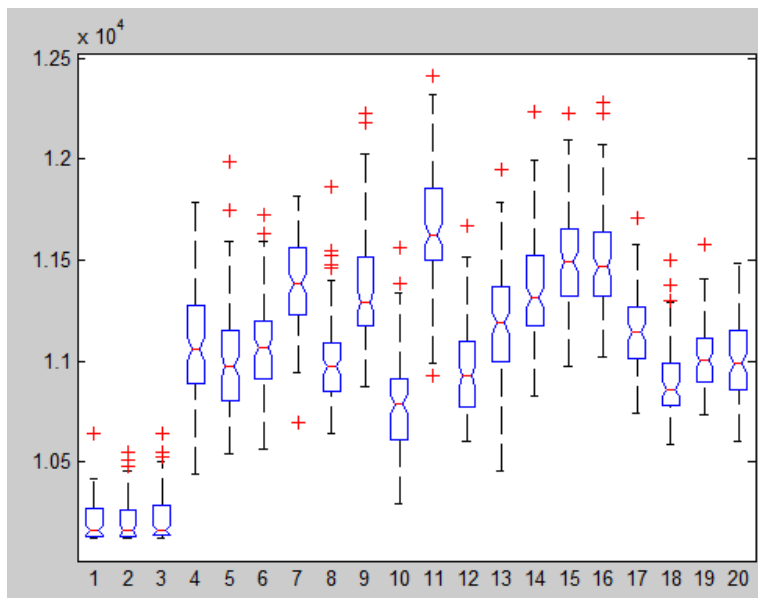


Gráfico 56. Gráfico de cajas para los resultados a 100.000 evaluaciones del ensayo 1PT.150.M8 para las 20 codificaciones ensayadas.

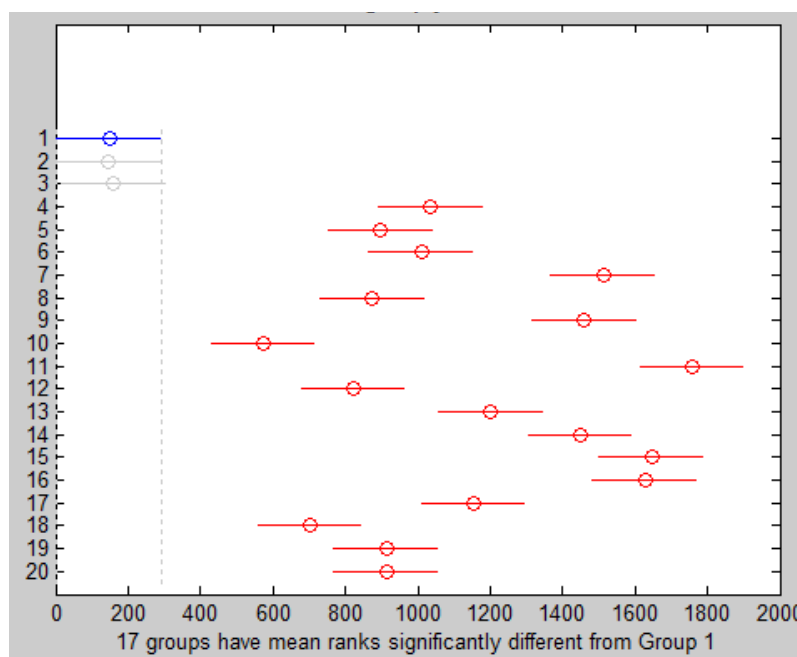


Gráfico 57. Multicomparación de las 20 codificaciones para 100.000 evaluaciones en el ensayo 1PT.P150.M8.

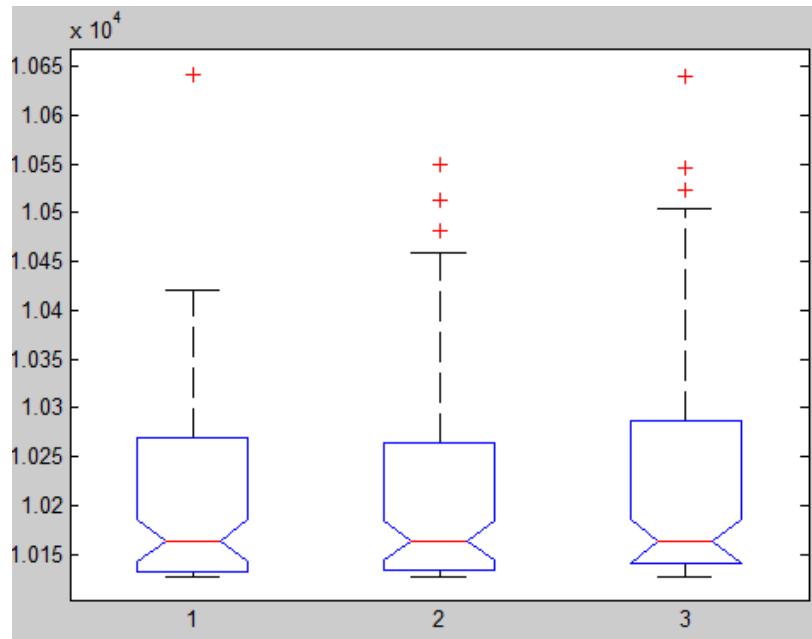


Gráfico 58. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo 1PT.P150.M8.

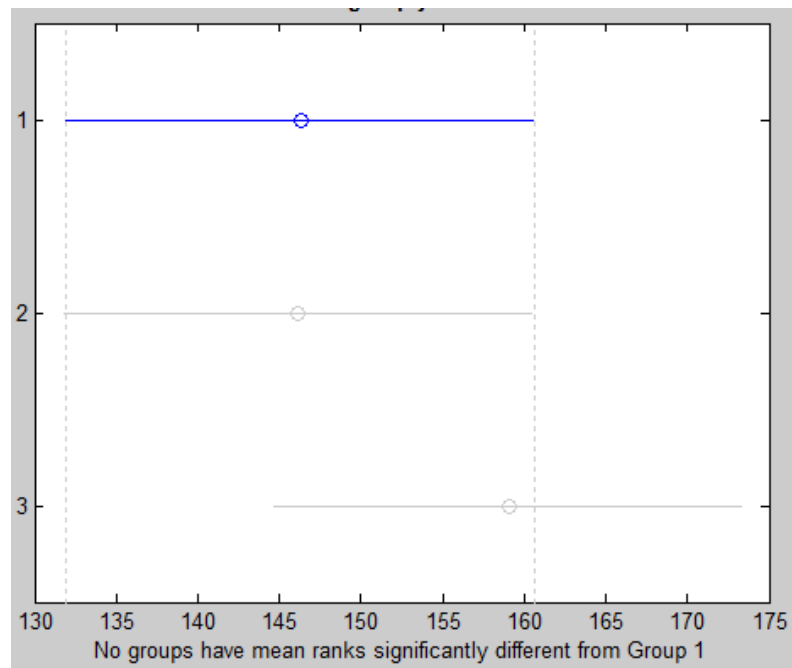


Gráfico 59. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo 1PT.P150.M8.

Para 1.000.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $5,31 \cdot 10^{-299}$). Los **peores resultados** se encuentran para el grupo 7, correspondiente a la **codificación S27A1,33**.

No existe significancia entre las tres ordenaciones tipo Gray entre sí, si bien en este caso se ha incluido en el segundo estudio también la codificación S27A0,40 (grupo 5), comprobándose que es distinta de las otras 3 ($Q=1,02 \cdot 10^{-29}$).

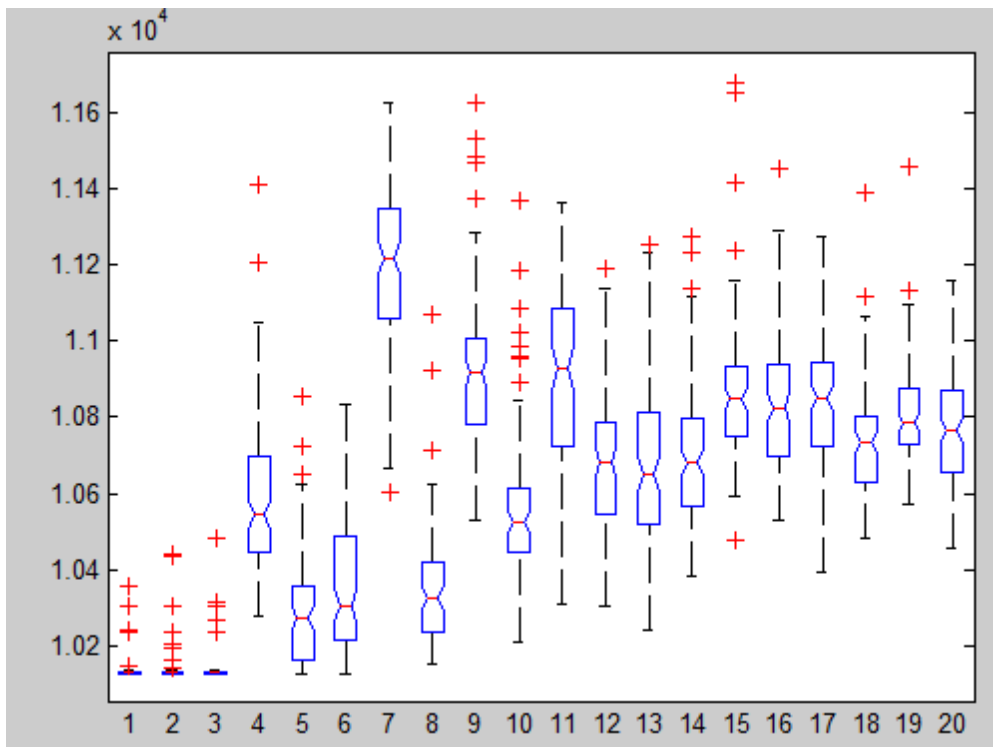


Gráfico 60. Gráfico de cajas para los resultados a 1.000.000 evaluaciones del ensayo 1PT.150.M8 para las 20 codificaciones ensayadas.

77

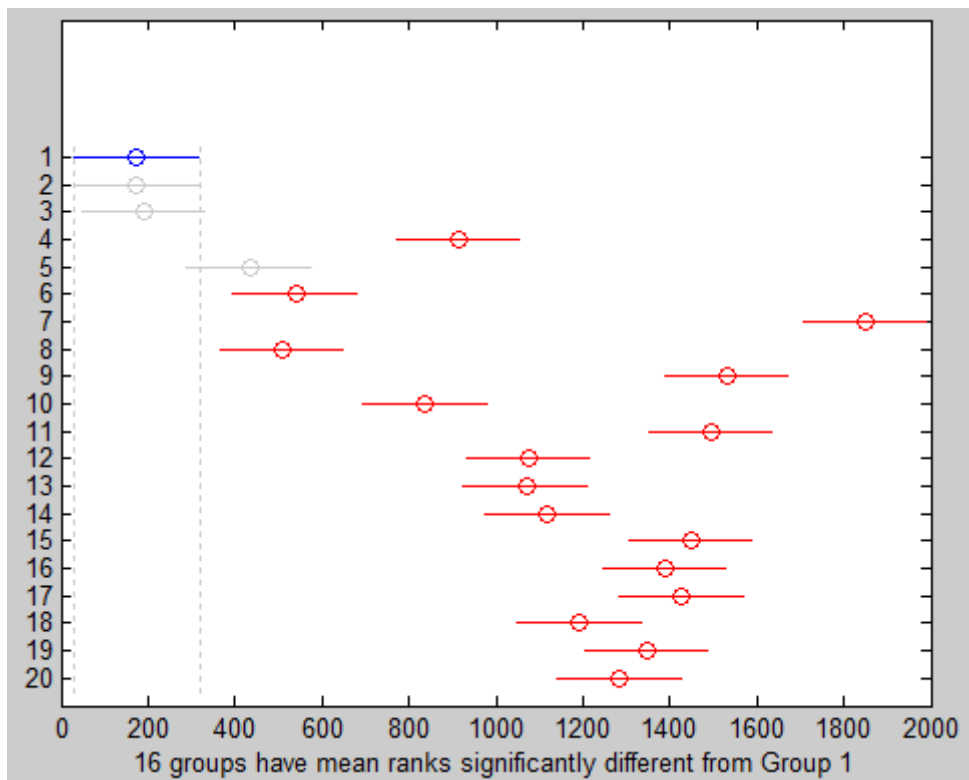


Gráfico 61. Multicomparación de las 20 codificaciones para 1.000.000 evaluaciones en el ensayo 1PT.P150.M8.

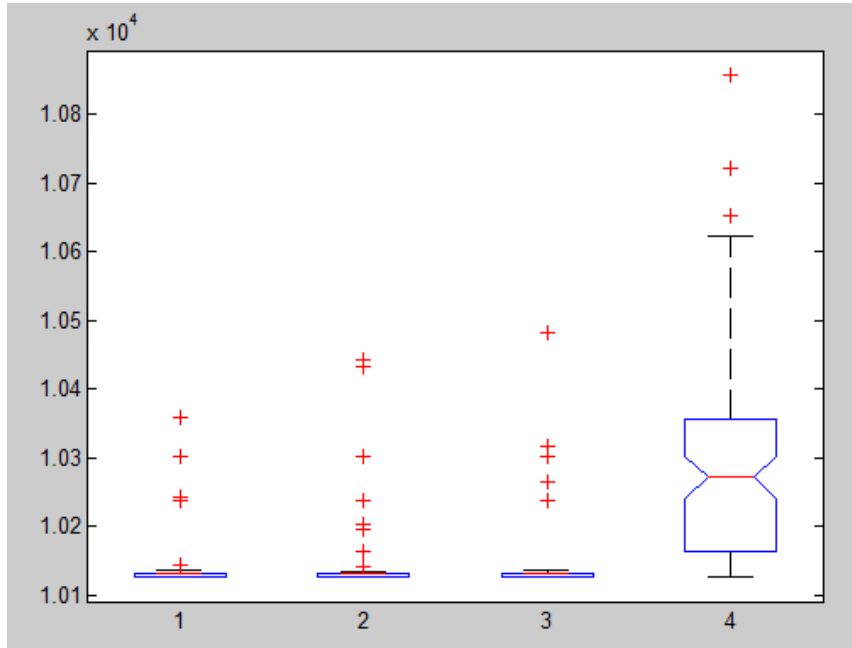


Gráfico 62. Gráfico de cajas de las codificaciones S15A0, S15B0, S15C0 y S27A0,40 para 1.000.000 evaluaciones en el ensayo 1PT.P150.M8.

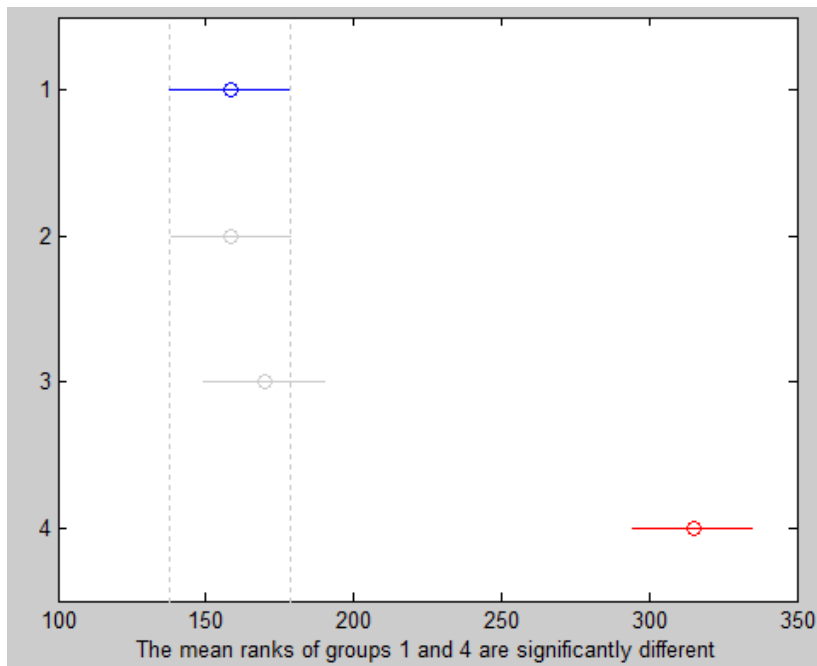


Gráfico 63. Multicomparación de las codificaciones S15A0, S15B0, S15C0 y S27A0,40 para 1.000.000 evaluaciones en el ensayo 1PT.P150.M8.

Resultados: ensayo UNI.P50.M8

Para 100.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $3,9 \cdot 10^{-303}$). Los **peores resultados** se encuentran para el grupo 11, correspondiente a la **codificación S39A0,49**.

No existe significancia entre las tres ordenaciones tipo Gray entre sí (valor de Q de 0,3025).

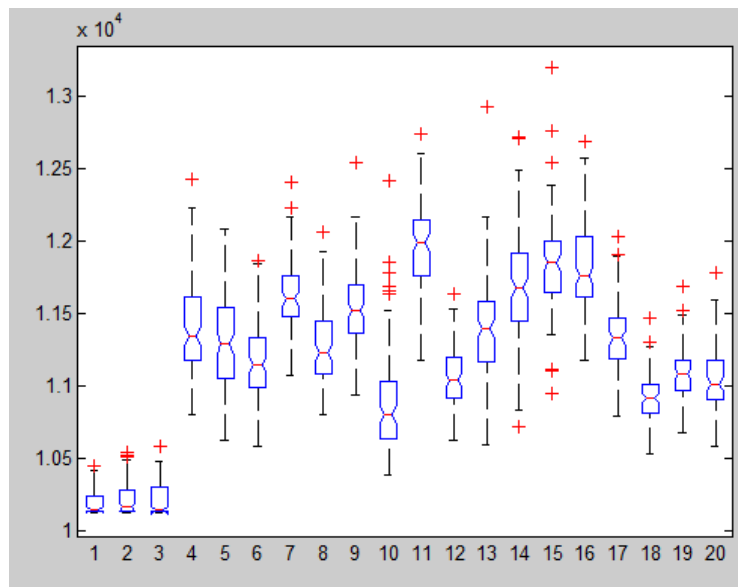


Gráfico 64. Gráfico de cajas para los resultados a 100.000 evaluaciones del ensayo UNI.P50.M8 para las 20 codificaciones ensayadas.

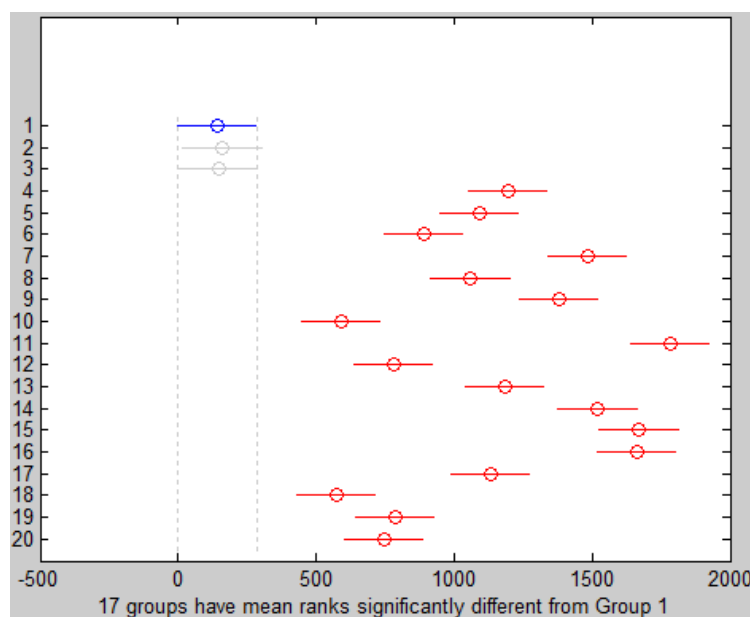


Gráfico 65. Multicomparación de las 20 codificaciones para 100.000 evaluaciones en el ensayo UNI.P50.M8.

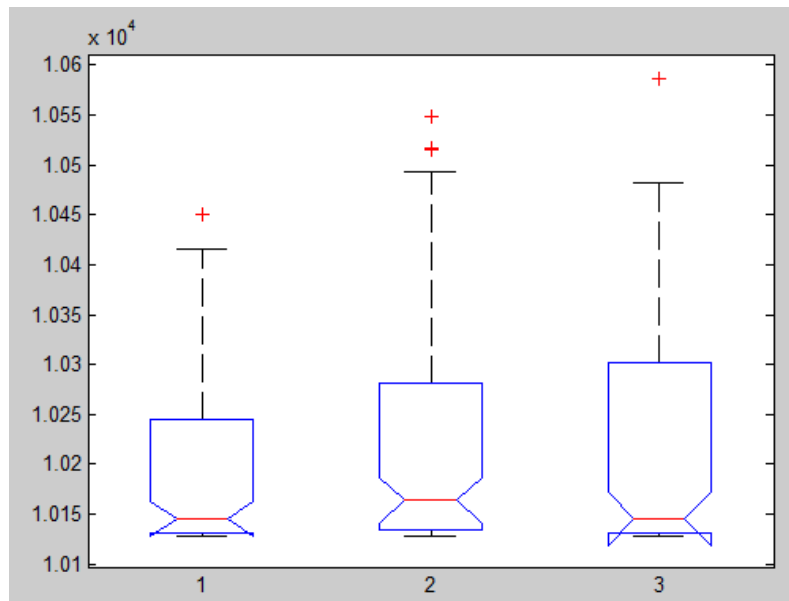


Gráfico 66. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo UNI.P50.M8.

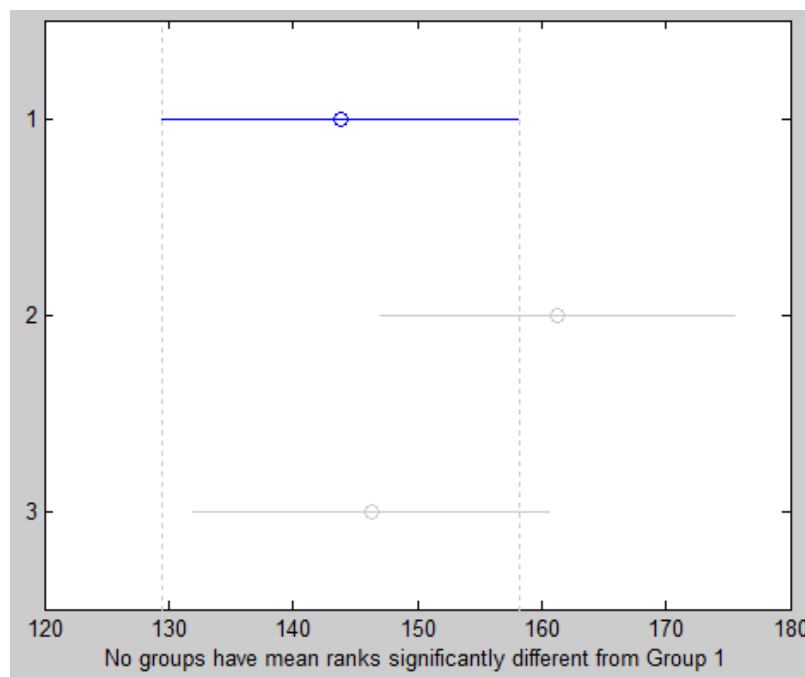


Gráfico 67. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo UNI.P50.M8.

Para 1.000.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de $6,02 \cdot 10^{-304}$). Los **peores resultados** se encuentran para el grupo 7, correspondiente a la **codificación S27A1,33**.

No existe significancia entre las tres ordenaciones tipo Gray entre sí ($Q=0,7008$).

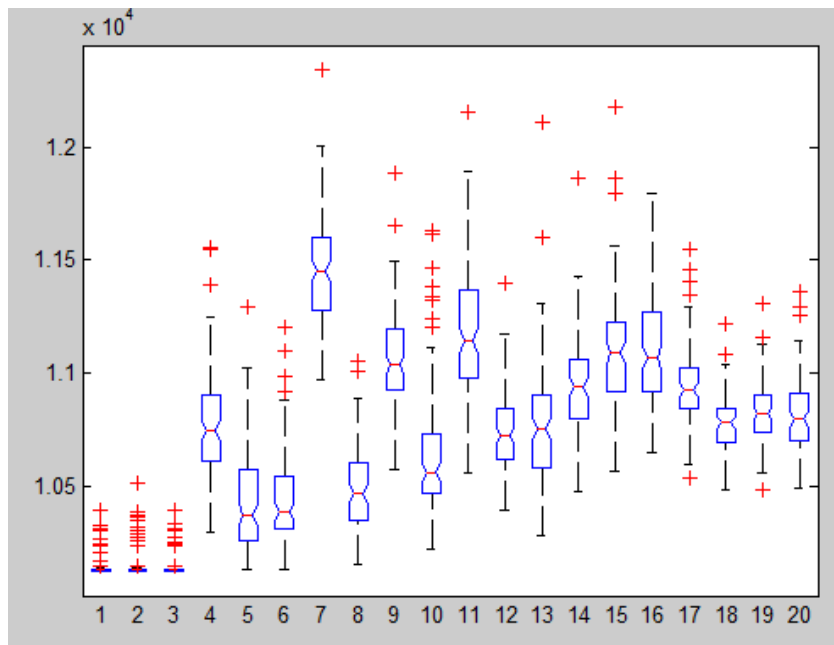


Gráfico 68. Gráfico de cajas para los resultados a 1.000.000 evaluaciones del ensayo UNI.P50.M8 para las 20 codificaciones ensayadas.

81

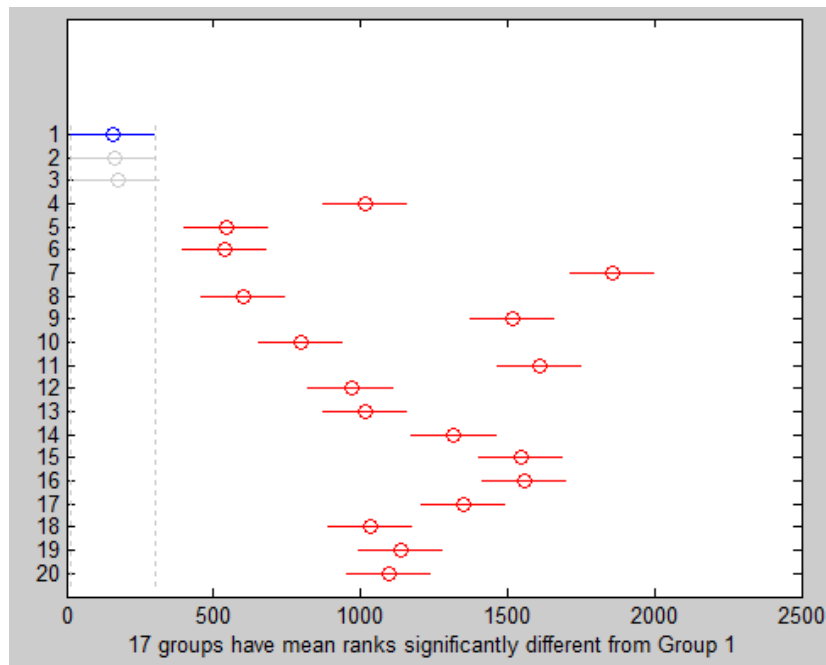


Gráfico 69. Multicomparación de las 20 codificaciones para 1.000.000 evaluaciones en el ensayo UNI.P50.M8.

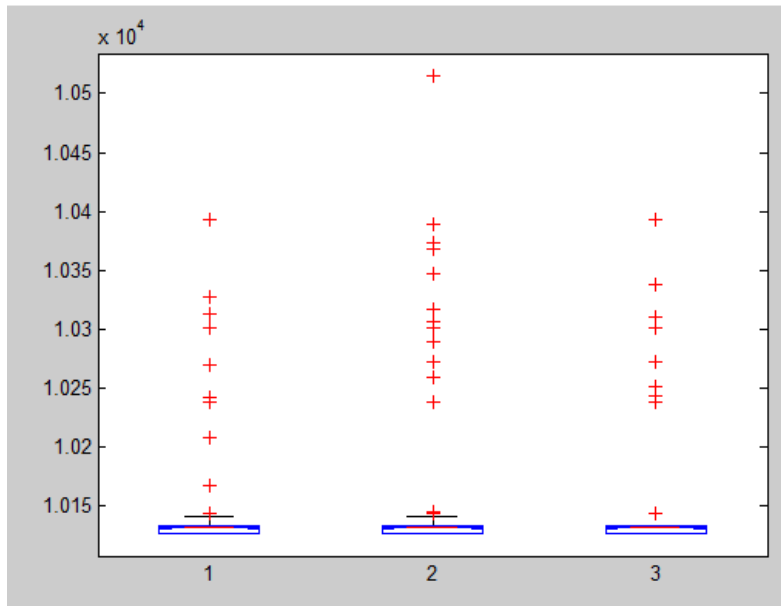


Gráfico 70. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo UNI.P50.M8.

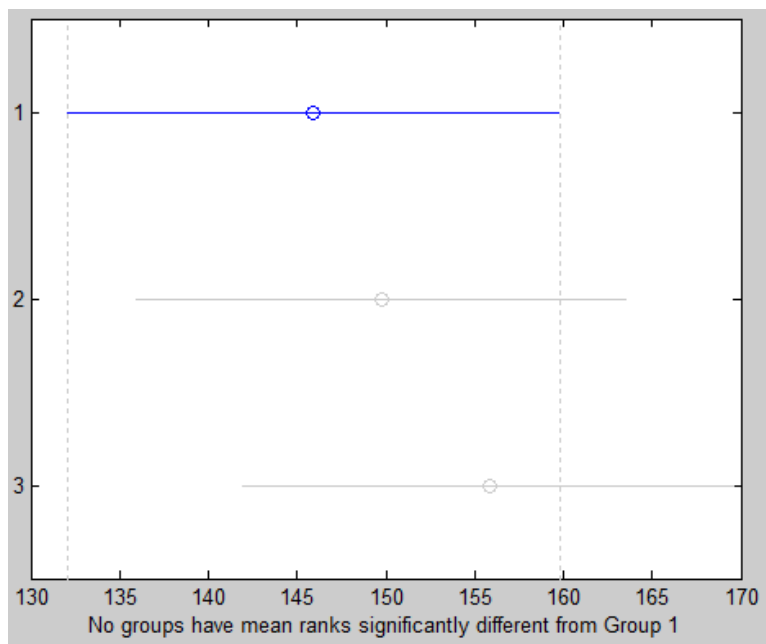


Gráfico 71. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo UNI.P50.M8.

Resultados: ensayo UNI.P150.M8

Para 100.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de 0). Los **peores resultados** se encuentran para el grupo 11, correspondiente a la **codificación S39A0,49**.

No existe significancia entre las tres ordenaciones tipo Gray entre sí (valor de Q de 0,0392).

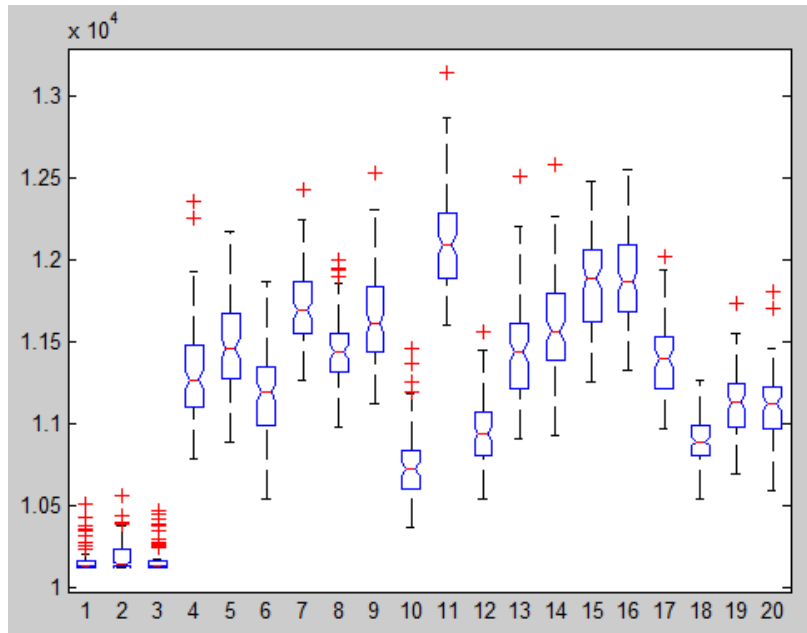


Gráfico 72. Gráfico de cajas para los resultados a 100.000 evaluaciones del ensayo UNI.P150.M8 para las 20 codificaciones ensayadas.

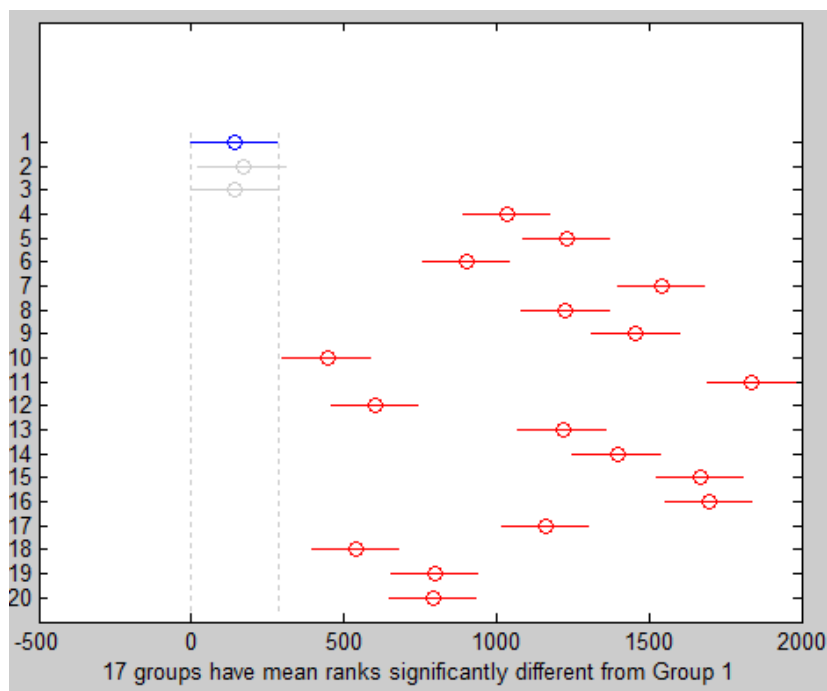


Gráfico X72 Multicomparación de las 20 codificaciones para 100.000 evaluaciones en el ensayo UNI.P150.M8.

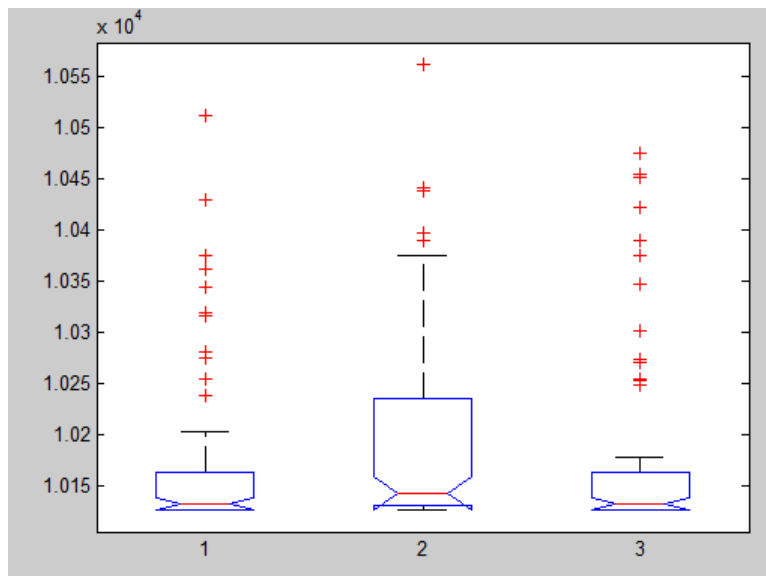


Gráfico 74. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo UNI.P150.M8.

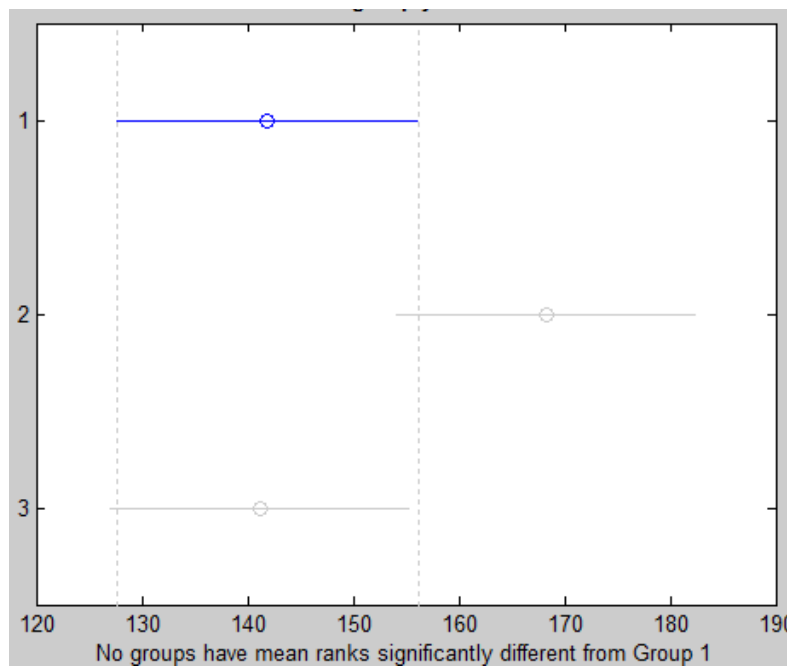


Gráfico 75. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 100.000 evaluaciones en el ensayo UNI.P150.M8.

Para 1.000.000 evaluaciones

Las muestras de datos correspondientes a las codificaciones ensayadas S15A0, S15B0 y S15C0 presentan medianas significativamente inferiores que el resto de las ensayadas (valor de Q para el test de Kruskal-Wallis para las 20 codificaciones de 0). Los **peores resultados** se encuentran para el grupo 7, correspondiente a la **codificación S27A1,33**.

No existe significancia entre las tres ordenaciones tipo Gray entre sí (Q=0,4879).

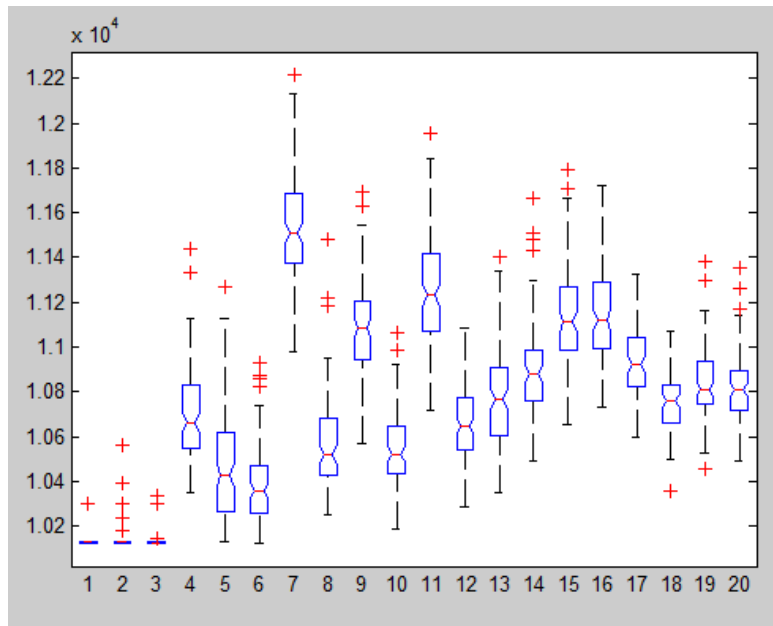


Gráfico 76. Gráfico de cajas para los resultados a 1.000.000 evaluaciones del ensayo UNI.P150.M8 para las 20 codificaciones ensayadas.

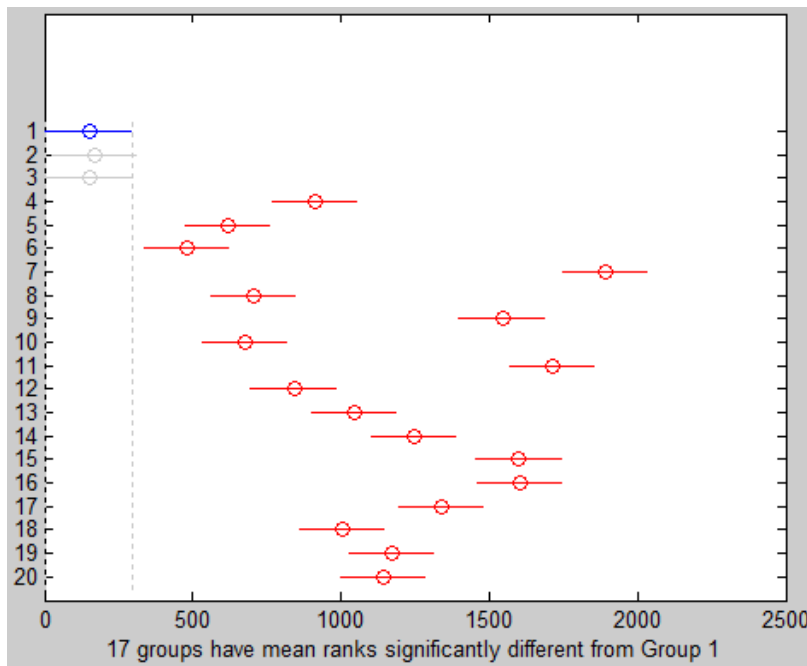


Gráfico 77. Multicomparación de las 20 codificaciones para 1.000.000 evaluaciones en el ensayo UNI.P150.M8.

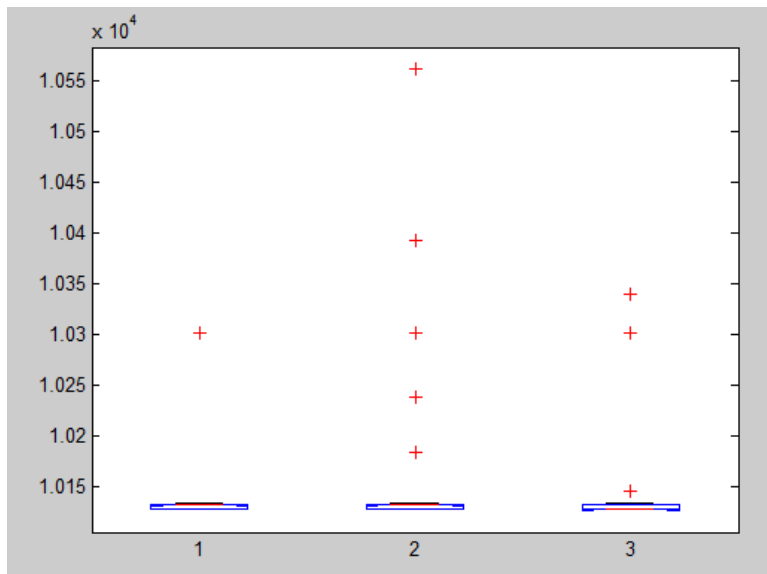


Gráfico 78. Gráfico de cajas de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo UNI.P150.M8.

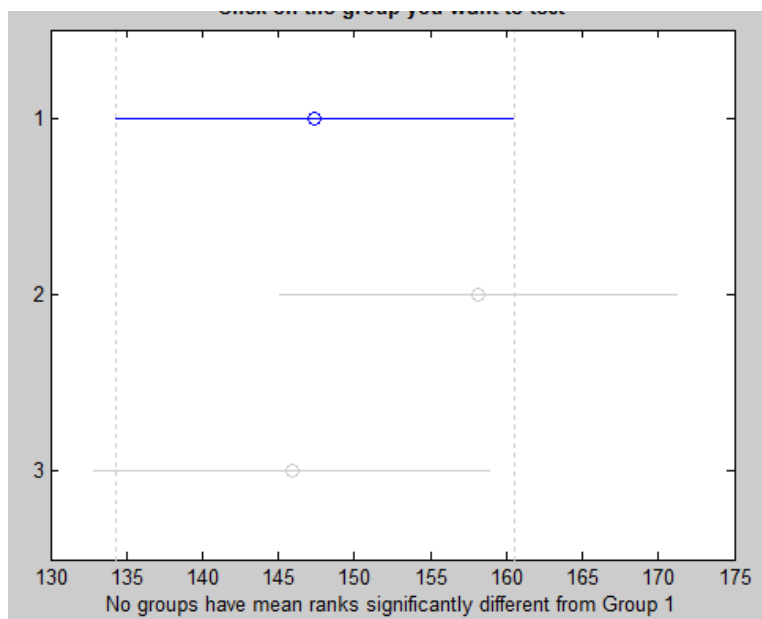


Gráfico 79. Multicomparación de las codificaciones S15A0, S15B0 y S15C0 para 1.000.000 evaluaciones en el ensayo UNI.P150.M8.

6. CONCLUSIONES Y LÍNEAS FUTURAS

El objetivo principal planteado para este trabajo era el de estudiar la influencia que tiene el empleo de diferentes codificaciones binarias al diseño óptimo de estructuras, empleando para ello algoritmos genéticos. Por tanto, se ha seguido una metodología de trabajo basada primeramente en el estudio de diferentes codificaciones, la variedad y características de las mismas. Seguidamente se han planteado unas hipótesis sobre posibles factores de influencia en el comportamiento de las mismas en el problema planteado, todas ellas relacionadas con la variación de bits entre números consecutivos de la codificación, siendo este el factor que a priori se considera como más influyente en el mejor o peor comportamiento de las codificaciones en este tipo de aplicaciones. Finalmente, y mediante un caso test, se han establecido diferentes escenarios de optimización donde un grupo de codificaciones de diferente tipología han sido ensayadas, con el fin de evaluar su comportamiento a partir de los resultados revertidos en cada ensayo. Los resultados obtenidos han sido contrastados mediante gráficas, y también con el empleo de un análisis estadístico que refrende las conclusiones extraídas.

87

A lo largo del trabajo realizado se han extraído las siguientes conclusiones:

- Se han obtenido resultados significativamente mejores mediante el empleo de codificaciones tipo Gray, siendo estos casos los únicos en los que se ha llegado a obtener un resultado medio inferior a 10.200 kg (el óptimo es de 10.127,1kg). Con otras codificaciones sí se ha alcanzado el valor óptimo pero no como valor medio, sino como mejor valor alcanzado en alguna de las 100 repeticiones realizadas.
- No se han detectado diferencias en el comportamiento ni resultados obtenidos entre las tres codificaciones tipo Gray empleadas, presentando las tres un comportamiento casi idéntico en los 8 ensayos realizados.
- De las codificaciones ensayadas, la que peor comportamiento ha presentado (resultado medio a 1.000.000 de evaluaciones) ha sido la denominada S27C1,33, que es una codificación con una suma total de 27 bits diferentes y una varianza de 1,33. Cabe destacar que esta codificación presenta 12 bits diferentes más que las tipo Gray

(15), pero 24 bits menos que la codificación con suma más alta ensayada (S51A0,61). Tiene además la segunda varianza más alta (sólo se ve superada por la S39C1,40). Sin embargo, incluso esta codificación presenta mejores resultados que la anteriormente citada. Se hace necesario a raíz de los resultados de este TFM, por tanto profundizar en el análisis de qué parámetro o parámetros influyen en la correcta obtención del óptimo para el problema ensayado.

- La codificación con mayor cantidad de variaciones de 4 bits que se ha ensayado (S39C1,40, 7 diferencias de 4 detectadas), presenta resultados de los más altos, si bien no es de los peores, por lo que no parece que el hecho de presentar una gran cantidad de variaciones de 4 bits (los más disruptivos con el código) sea el único parámetro influyente.
- De los resultados obtenidos en el análisis estadístico puede parecer que el comportamiento del resto de codificaciones es aparentemente errático, ya que no se ha conseguido establecer un patrón de comportamiento relacionado con ninguna de las hipótesis planteadas: suma total de bits diferentes, varianza de este promedio, cantidad de variaciones de 4 bits contenidas en el código. Sin embargo, sí existe una similitud en el comportamiento de todas las codificaciones, es decir, que no se ha detectado un patrón de comportamiento, pero este debe existir, ya que en los gráficos de multicomparación puede apreciarse todos presentan un grafo similar (cómparese por separado los obtenidos para 100.000 evaluaciones y para 1.000.000). Concretamente, en el caso de los resultados obtenidos para 1.000.000 de evaluaciones puede apreciarse como para todos los ensayos, cada mediana se encuentra situada en la misma parte del gráfico. Además, el hecho de que siempre haya sido la misma codificación la que presenta peor comportamiento a las 100.000 evaluaciones (S39A0,49) y otra diferente quien presenta los peores resultados en todos los casos a 1.000.000 (S27C1,33) apoya esta conclusión.
- Existen codificaciones (no sólo las tipo Gray) que presentan mejores resultados que la binaria natural, como son la S27A0,40 y la S27B0,83. Ambas tienen mayor suma total de bits distintos superior en 1 punto, pero varianzas menores.
- Si bien no ha podido obtenerse una conclusión clara sobre los factores que influyen en el peor comportamiento de las codificaciones, sí que se ha podido contrastar que las codificaciones que aquellas codificaciones **Gray**, que presenten un patrón de evolución suave, con pequeñas modificaciones entre sí, producen de manera general mejores resultados que el resto. Además, y dado que se han empleado 3

codificaciones de este tipo diferentes, se puede concluir que no afecta la posición de los códigos con mayor cantidad de unos o aquellos con más cantidad de ceros para unas variables u otras. Es decir, si analizamos la manera en la que cada uno de los tres códigos avanza de número a número, desde un punto de vista de una codificación ponderada, no parece existir un patrón de relación entre el peso del código y la variable a la que representa. Por tanto, el algoritmo genético no parece entender sobre el peso del código, y no le afecta que las variables de mayor área estén indistintamente representadas por codificaciones con gran cantidad de unos en las variables más robustas y viceversa. En conclusión, la codificación Gray ha resultado ser capaz de obtener mejores resultados con significancia estadística frente al resto de codificaciones, y con mayor influencia que el resto de parámetros del algoritmo evolutivo (tamaño de población, cruce o probabilidad de mutación).

LÍNEAS FUTURAS DE TRABAJO

Gracias a la realización de este trabajo se han planteado diferentes líneas futuras de investigación que sirvan como continuación al trabajo realizado, con el firme objetivo de aportar luz y ampliar horizontes en el empleo de los algoritmos genéticos para el cálculo óptimo no sólo en el área del diseño de estructuras, sino en el amplísimo campo de aplicaciones que éstos presentan. Algunos de los trabajos que se han detectado como posibles para continuar con esta línea de investigación son:

- Estudio pormenorizado de la influencia de los parámetros de optimización y su posible sinergia con el empleo de codificaciones óptimas (tipo de selección, cruce, mutación, etc.).
- Abordaje del problema desde otros puntos de vista diferentes que puedan llevar a la búsqueda de un patrón de comportamiento de las codificaciones.
- Estudio avanzado de la influencia de la codificación tipo Gray (generación de nuevas codificaciones Gray y su comparación).
- Aplicación a otros casos tests para estudiar la generalidad de las conclusiones aquí extraídas.

7. BIBLIOGRAFÍA Y REFERENCIAS

- [1]: D. Greiner, J.M. Emperador, G. Winter, "*Multiobjective Optimisation of Bar Structures by Pareto-GA*", in European Congress on Computational Methods in Applied Sciences and Engineering (2000), CIMNE.
- [2]: D. Greiner, G. Winter, J.M. Emperador, "Optimising Frame Structures by different strategies of GA", *Finite Elem. in Anays. and Design*, Elsevier, 37 (5) pp.381-402, (2001).
- [3]: D. Greiner, JM. Emperador, G. Winter, "*Single and multiobjective frame optimization by evolutionary algorithms and the auto-adaptive rebirth operator*". *Computer Methods in Applied Mechanics and Engineering*, 2004, 193: 3711–3743.
- [4]: D. Greiner, G. Winter, JM. Emperador, B. Galván, "*Gray Coding in Evolutionary Multicriteria Optimization: Application in Frame Structural Optimum Design*". In: CA. Coello Coello, A. Hernández Aguirre, E. Zitzler (eds) *Evolutionary Multi-Criterion Optimization*, Lecture Notes in Computer Science, 2005, 3410: 576–591.
- [5]: D. Greiner, B. Galván, JM. Emperador, M. Méndez, G. Winter, "*Introducing Reference Point using g-dominance in Optimum Design considering Uncertainties: An Application in Structural Engineering*". In: R. Takahashi et al (eds), *Evolutionary Multi-Criterion Optimization*, Lecture Notes in Computer Science, 2011, 6576: 389–403.
- [6]: M. Wineberg. *Statistical Analysis for Evolutionary Computation: Indroduction*. Genetic and Evolutionary Computation Conference (Tutorial). Julio 2012, Philadelphia, USA.
- [7]: T. Bartz-Beielstein, M. Preuss, M. Zaefferer. *Statistical Analysis of Optimization Algorithms with R*. Genetic and Evolutionary Computation Conference (Tutorial). Julio 2012, Philadelphia, USA.
- [8]: J. H. Holland. *Algoritmos Genéticos*. En "Investigación y Ciencia", septiembre 1992, pp 38-45.
- [9]: D. E. Goldberg , M. P. Santani. *Engineering optimization via genetic algorithm*. Proc Ninth Conf. on Electronic Computation. ASCE, New York, NY (1986), pp 471-482.
- [10]: E. Mandado Pérez, Y. Mandado Rodríguez. *Sistemas Electrónicos Digitales*, 9ª edición. Ed. Marcombo, España (2007).
- [11]: C. H. Roth, Jr. *Fundamentos de diseño lógico*, 5ª edición. Ed. Thomson, España (2005).