

# Análisis de rendimiento de algoritmos de optimización mono-objetivo frente a un conjunto de resultados obtenidos de un mismo problema

Erick Salvador Alvarez Valencia, CIMAT A.C.

## I. INTRODUCCIÓN

En el presente trabajo se mostrará un análisis detallado sobre los resultados generados por varios métodos que se basan en técnicas de optimización mono-objetivo ante diferentes instancias del problema conocido como 2DPP. Esto anterior con el fin de detectar varios aspectos importantes que se pueden captar propiamente de estos métodos, tales como: la diversidad poblacional a lo largo del tiempo, la evolución de la calidad a lo largo del tiempo, comparativas de normalidad, varianza y media para determinar qué método es mejor, entre otras. Además se pretende responder a dos preguntas importantes formuladas a manera de hipótesis las cuales son: El supuesto de que el comportamiento relativo de los métodos depende del tamaño de la instancia y el supuesto de que los diferentes métodos manejan la diversidad de forma diferente y esto influye a largo plazo en los resultados. Estos dos supuestos serán respondidos aplicando las pruebas mencionadas anteriormente. Finalmente se darán algunas conclusiones con respecto a los métodos analizados en este reporte.

mds  
Febrero, 2018

## II. MOTIVACIÓN

En el área de optimización estocástica es muy común que se propongan diferentes métodos para atacar un mismo problema, obteniendo así diversa variedad de información respecto a dichos métodos. La cuestión de interés al final del día es el cómo medir qué tan bueno fue un método en comparación a otro si los dos usan distintas estrategias para atacar el problema.

Lo que se mostrará en este reporte son los resultados de analizar 10 métodos que fueron usados para resolver el problema conocido como 2DPP (Two Dimensional Packing Problem), el cual consiste en: Tratar de idear nuevas formas de evolucionar con un cromosoma 2D. El objetivo de este problema es empaquetar mejor una cuadrícula de modo que se maximice la suma de los puntajes puntuales para cada par de números adyacentes. El truco es que una vez que recolectas los puntos para un par en particular, ¡no puedes juntar esos puntos por segunda vez en la misma cuadrícula. Finalmente lo que nos interesa de este problema es que consiste en un problema de maximización, por lo cual fueron dados un conjunto de resultados obtenidos con tres tipos

instancias, pequeña, mediana y grande, las cuales fueron tratadas con 10 métodos distintos: CHC, CLR\_1, CLR\_5, COMB, DETCR, Multi, MultiDynamic, RW, Sawtooth y SPOBCR.

De la misma forma, al realizar el análisis de los métodos descritos anteriormente se quiere responder a dos cuestiones que fueron creadas a manera de hipótesis:

1. Se cree que el comportamiento relativo de los métodos puede depender del tamaño de la instancia.
2. Se cree que diferentes métodos manejan la diversidad de forma drásticamente diferente y que esto puede influir en los resultados a corto y largo plazo.

Para la hipótesis no. 2 se usarán solo los datos de la instancia llamada: *GECCO-contest* la cual es una instancia en el grupo de las grandes, esto ya que se requiere analizar la evolución de la diversidad poblacional y sólo dicha instancia contiene la información suficiente para determinar la diversidad.

## III. EVOLUCIÓN DE LA DIVERSIDAD

Primeramente se generará el análisis de diversidad poblacional mencionado en la parte anterior ya que es un buen punto de partida, y además nos servirá a lo largo del análisis. Para comenzar, el interés por medir la diversidad poblacional se genera ya que esta puede darnos información sobre lo que está pasando en el proceso de búsqueda a lo largo de tiempo. Si la diversidad en una población baja, esto puede indicar que la mayoría de los individuos de la misma están convergiendo hacia cierto espacio de búsqueda o visto de otra forma, el optimizador encontró algún mínimo local. Idealmente la variabilidad de una población debería ir descendiendo a lo largo del tiempo, el hecho de graficar este concepto en relación al tiempo nos indica cómo es que el optimizador está realizando la búsqueda, por ejemplo, si la variabilidad desciende muy rápido en corto tiempo entonces podemos concluir que el algoritmo es muy elitista y busca siempre poblaciones que den mejores resultados, lo cual no siempre es bueno ya que se quedará con el primer óptimo local que encuentre y los resultados pueden no ser siempre muy buenos. Por el contrario, si la variabilidad se mantiene alta al rededor del tiempo, puede indicar que el algoritmo sigue recorriendo gran parte del espacio de búsqueda y aún no encuentra un buen óptimo local.

Una buena forma de medir la variabilidad poblacional (y es la que se usará en este caso) es con el concepto de entropía, ya

que esta puede medir la cantidad de variabilidad que aporta la población. A continuación se muestra la fórmula usada generalmente para calcular la entropía poblacional en un cierto tiempo  $t$ :

$$H(n) = \frac{1}{n} \sum_{i=1}^n H_i(t) \quad (1)$$

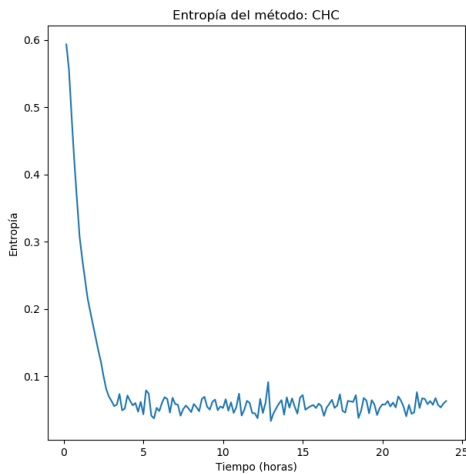
Donde  $H_i(t)$  lo definimos como:

$$H_i(t) = - \sum_{j=1}^{v_i} P_{ij} \log_{v_i}(P_{ij}) \quad (2)$$

La entropía de una población en un tiempo  $t$  la calculamos promediando las entropías generadas por cada variable (gen) que contengan los individuos de la población.

En la instancia llamada *GECCO-contest* nos dan 10 corridas por cada método implementado, para calcular la evolución de la entropía poblacional de cada método al rededor de todas sus corridas, se calcula la entropía poblacional por cada tiempo contenido en cada corrida, esto nos generará un vector de entropías por cada corrida, ahora, como cada corrida contiene la información poblacional tomada en la misma discretización del tiempo, entonces podemos promediar estos vectores para encontrar una sola línea de evolución de variabilidad.

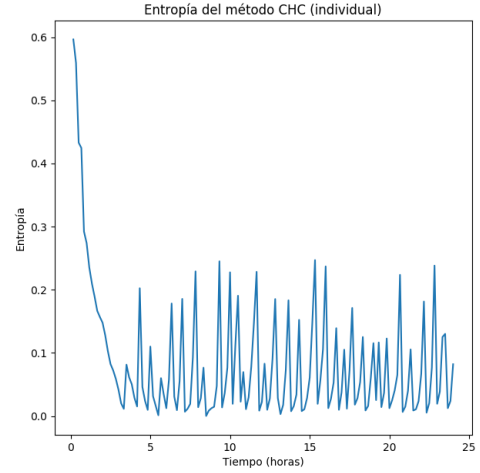
Utilizando el método descrito anteriormente se calculó la evolución de la variabilidad media de todos los métodos con respecto al tiempo. A continuación se mostrarán las gráficas generadas de dichas entropías con sus respectivos análisis. Para poder visualizar estas gráficas de la mejor manera hay que aclarar que el tiempo de las mismas se indica en horas y el valor de la entropía siempre está entre 0 y 1.



(a) Figura 1. Evolución de la entropía poblacional del método CHC.

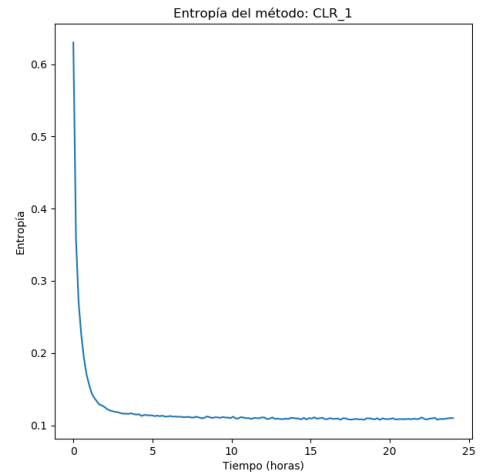
Para la Figura 1. podemos apreciar lo que se comentaba con respecto a la evolución de la variabilidad, este método tomó un carácter muy elitista y por ello desde el comienzo descendió bastante rápido su variabilidad (5 primeras horas).

Pero después la misma empezó a oscilar por el resto del tiempo lo cual puede ser un poco raro ya que al principio se tornó muy elitista siempre quedándose con lo mejor que obtenía en las poblaciones pero después tuvo varios cambios poblacionales, esto podría indicar que al llegar a un óptimo local intensificó mucho la búsqueda en esa zona. Para este caso puede ser bueno ver la gráfica de la evolución de su entropía poblacional para una ejecución en particular, la cual se muestra a continuación.

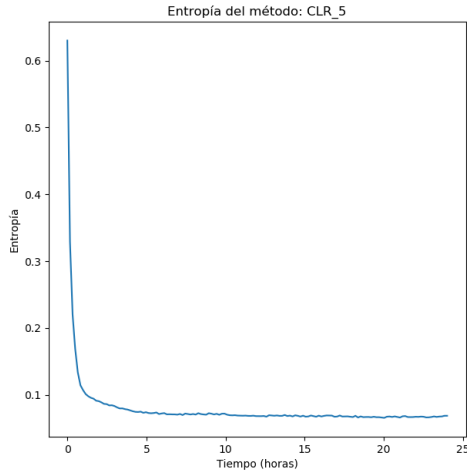


(b) Figura 2. Evolución de la entropía poblacional del método CHC (corrida individual).

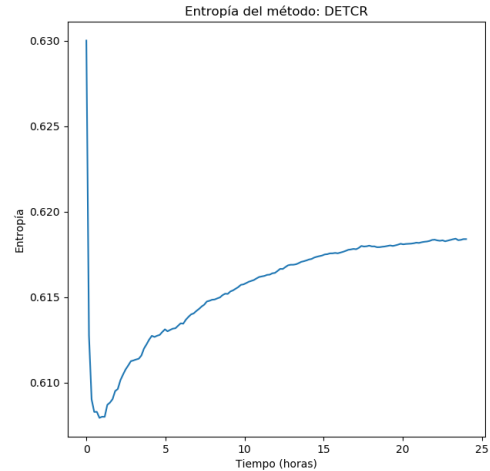
En la Figura 2. podemos apreciar la evolución de la entropía poblacional para una sola corrida de este método, aquí se puede apreciar que en realidad la magnitud de la oscilación que genera la variabilidad era mucho más grande que la descrita en la gráfica anterior, por lo que ahora es más objetivo el asumir que el método no se estanca fácilmente en algún óptimo local encontrado en etapas tempranas, sino que hace más búsquedas en otras zonas del dominio de la función o realiza reinicios parciales.



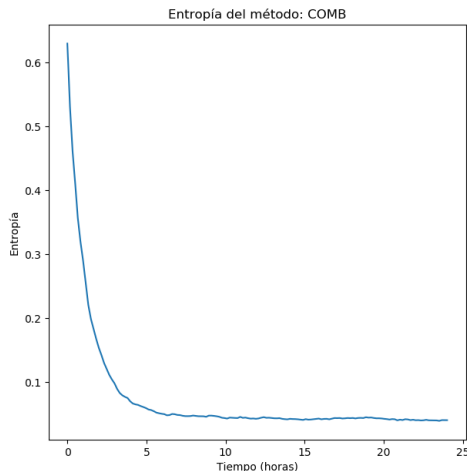
(c) Figura 3. Evolución de la entropía poblacional del método CLR 1.



(d) Figura 4. Evolución de la entropía poblacional del método CLR 5.



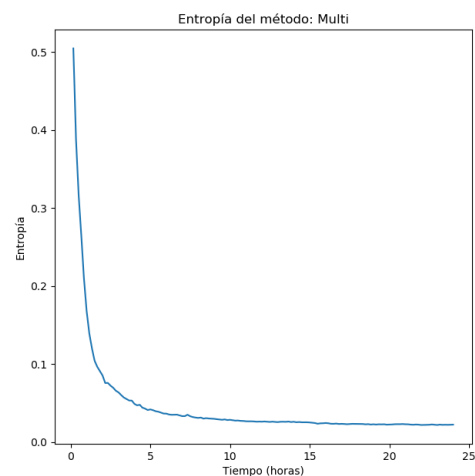
(f) Figura 6. Evolución de la entropía poblacional del método DETCR.



(e) Figura 5. Evolución de la entropía poblacional del método COMB.

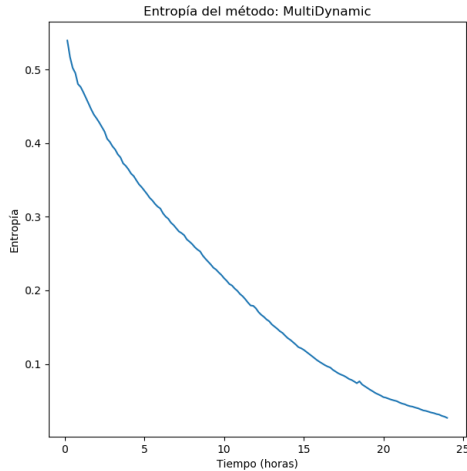
Para las Figuras 3, 4 y 5 generadas por los métodos CLR 1, CLR 5 y COMB podemos hacer una generalización ya que las gráficas son muy similares. Se puede ver que estos métodos tienden a descender su variabilidad poblacional de manera rápida (primeras 2-3 horas de la corrida) y posteriormente esta se mantiene casi constante, esto porque se quedan en una zona referente a un óptimo local, y hacen una búsqueda intensiva en dicha zona. Esto tiene una ventaja y una desventaja, por un lado tenemos el hecho que el método realiza una buena exploración en toda una zona referente a un óptimo local, pero por otro lado vemos que el método se va por el camino del primer (o de los primeros) óptimo local que encuentre, de ahí no hace más exploración en otras zonas que podrían dar mejores resultados.

En esta gráfica podemos ver un comportamiento extraño de la diversidad poblacional ya que inicialmente desciende muy rápido pero al pasar el tiempo el comportamiento se invierte ya que la diversidad va aumentando, lo cual puede indicar que este método aplica técnicas que evitan estancarse directamente en un óptimo local y recorrer otras zonas del espacio de búsqueda, pero se puede apreciar que la gráfica no vuelve a descender en las últimas horas de su ejecución, lo que puede indicar que no encontró algún buen óptimo local que podría ser candidato para su solución final. Para estar seguros de esto último hay que analizar la evolución de la calidad al rededor del tiempo, cosa que se hará en la siguiente sección del documento.



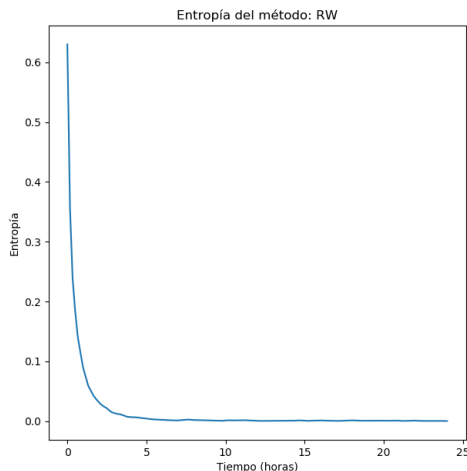
(g) Figura 7. Evolución de la entropía poblacional del método Multi.

No hay mucho que comentar de esta gráfica ya que tiene un comportamiento muy similar a las mostradas en las Figuras 3, 4 y 5.



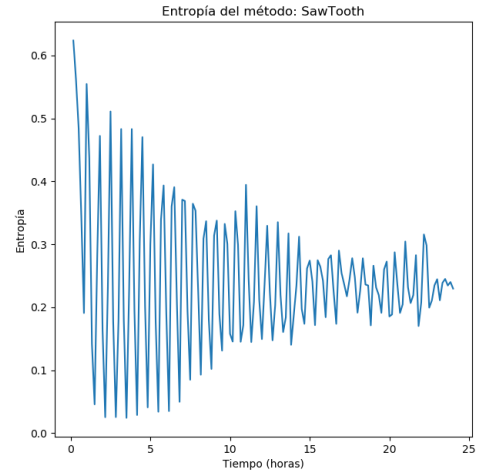
(h) Figura 8. Evolución de la entropía poblacional del método MultiDynamic.

En la Figura 8 podemos observar un comportamiento de la variabilidad poblacional muy interesante ya que este no desciende tan rápido como lo hacen algunos de los métodos vistos anteriormente. El descenso en la gráfica fue de manera casi lineal lo cual indica que el algoritmo no converge tan rápido hacia un óptimo local sino que realiza una búsqueda extensa por gran parte del espacio de la función analizando varios óptimos locales y al final generando una buena solución.



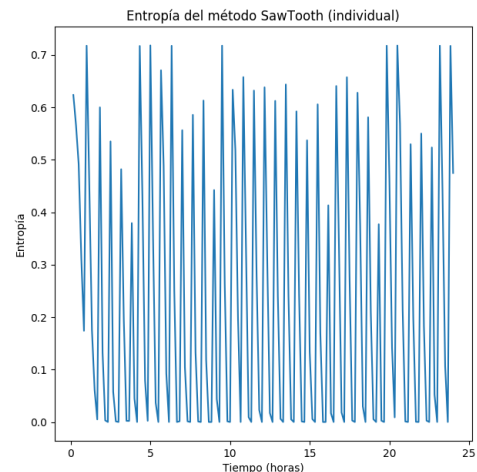
(i) Figura 9. Evolución de la entropía poblacional del método RW.

En la Figura 9. vemos el mismo comportamiento que en las Figuras 3, 4 y 5.



(j) Figura 10. Evolución de la entropía poblacional del método SawTooth.

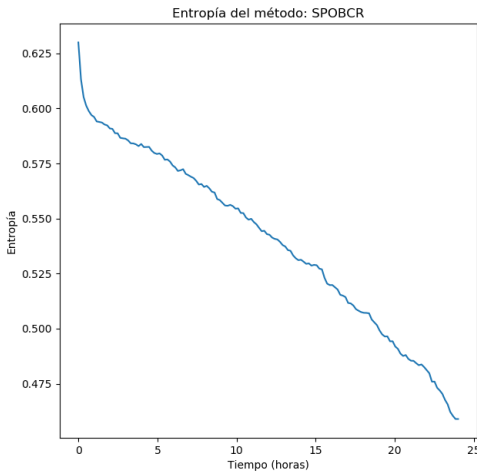
Para la Figura 10 vemos un comportamiento diferentes a los ya vistos. Podemos ver que la entropía poblacional oscila bastante alrededor del tiempo, lo que indica que no converge hacia cierto óptimo local de manera rápida como lo hacen otros métodos. Para este método también se mostrará la gráfica de la evolución de la entropía para una sola ejecución.



(k) Figura 11. Evolución de la entropía poblacional del método SawTooth (corrida individual).

En la Figura 11. vemos la evolución de la entropía con respecto a los datos de la primer corrida. Se puede observar que pese a que el comportamiento sigue siendo el mismo que en la gráfica anterior, referente al promedio, aquí vemos que la variabilidad es mayor, la gráfica anterior mostraba una cierta convergencia con respecto a las últimas horas de ejecución del método, en la presente se puede ver que no existe dicha convergencia, lo cual pone en duda si se encontraron buenas soluciones o no, de la misma forma se hace la suposición que este método hace reinicios globales para no converger en un

cierto óptimo local. Para ello es necesario analizar la evolución de la calidad con respecto al tiempo.



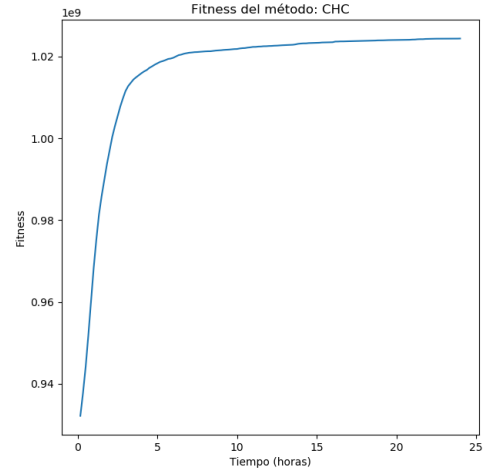
(l) Figura 12. Evolución de la entropía poblacional del método SPOBCR.

Para la gráfica mostrada en la Figura 12, vemos un comportamiento muy similar a la gráfica de la Figura 8, referente al método *MultiDynamic* por lo cual se puede realizar el mismo análisis.

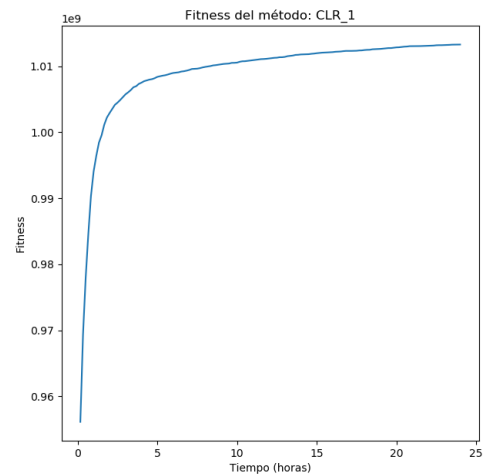
#### IV. EVOLUCIÓN DE LA CALIDAD CON RESPECTO AL TIEMPO

En esta parte se mostrarán las gráficas obtenidas analizando la calidad de los diferentes métodos con respecto al tiempo, para ello se extrajeron los mejores fitness tras cada periodo de tiempo en las diferentes corridas de un método y posteriormente se hizo un promedio, como resultado se obtuvo una lista de la evolución promedio de la calidad con respecto a un método. De esta forma podemos complementar el análisis de la sección anterior con la presente. A continuación se mostrarán las gráficas obtenidas.

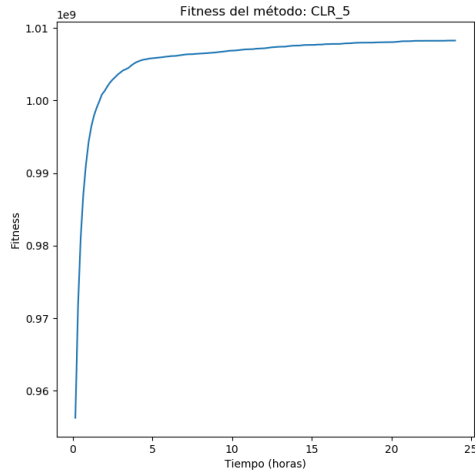
Como nota previa a la revisión de las gráficas hay que mencionar que la escala en la parte del tiempo es la misma que se manejó en las gráficas de entropía, en la parte de la entropía la escala manejará valores con magnitud de  $10^9$ .



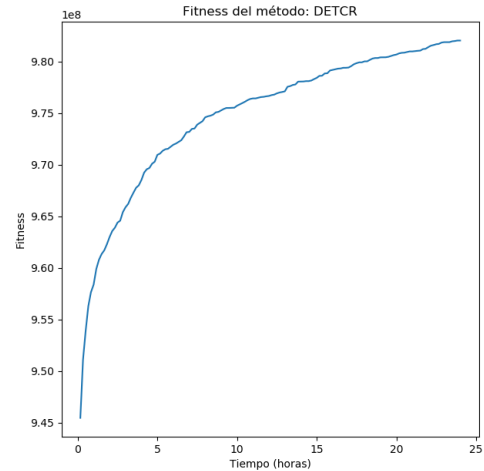
(m) Figura 13. Evolución de la calidad a lo largo del tiempo del método CHC.



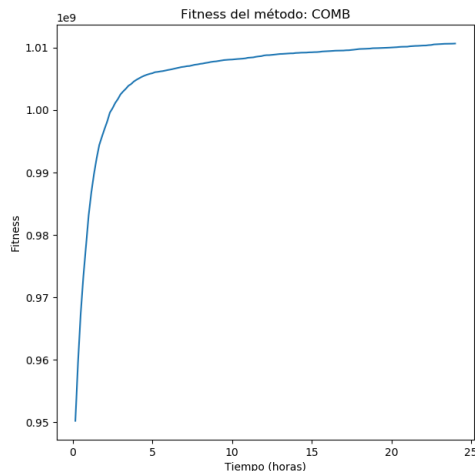
(n) Figura 14. Evolución de la calidad a lo largo del tiempo del método CLR 1.



(ñ) Figura 15. Evolución de la calidad a lo largo del tiempo del método CLR 5.



(p) Figura 17. Evolución de la calidad a lo largo del tiempo del método DETCR.

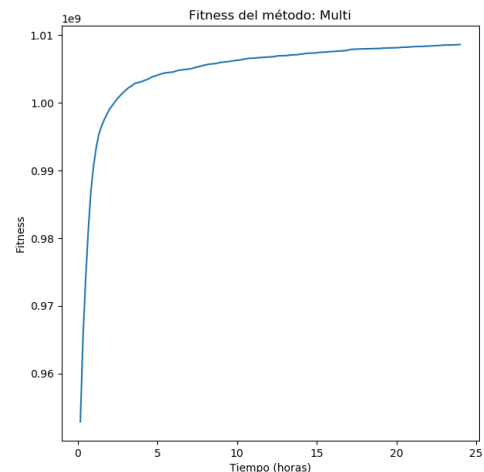


(o) Figura 16. Evolución de la calidad a lo largo del tiempo del método COMB.

Para las cuatro primeras gráficas de evolución de calidad que corresponden a los métodos CHC, CLR 1, CLR 5 y COMB un comportamiento bastante similar, vemos que el fitness crece de manera veloz en las primeras horas de la ejecución del método. Y posteriormente después de la hora 5 en promedio los métodos ya no aumentan tanto su fitness, esto es probablemente por el hecho que estaban en un óptimo local y simplemente se estaba realizando una búsqueda intensiva en esa zona.

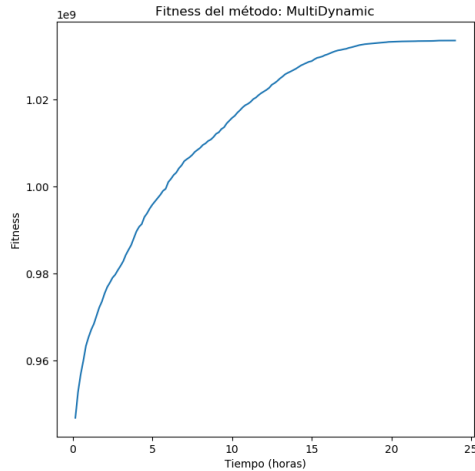
De todas estas gráficas podemos notar que la que mayor fitness dio al final de su ejecución fue la generada por el método CHC ya que como vimos en la sección anterior este método es un poco más exploratorio, en cambio los métodos CLR 1 y CLR 5 tendrían más a irse por el primer óptimo local que encontraban, por eso llegaron a un fitness no tan alto como el CHC.

Para el método DETCR podemos apreciar la gráfica de la evolución de su calidad en la Figura 17. Ahí se aprecia un comportamiento un poco distinto en su fitness ya que este no sube rápidamente como tiende a hacerlo en los métodos anteriores, sin embargo en las últimas horas de su ejecución este método no llega a ofrecer un buen puntaje, lo cual concuerda con lo que se apreció en su gráfica de evolución de diversidad. Podemos recordar que este método descendió muy rápido su variabilidad poblacional en las primeras horas pero a medida que avanzó el tiempo su variabilidad fue creciendo por lo que podemos concluir que no convergió hacia un buen óptimo local.



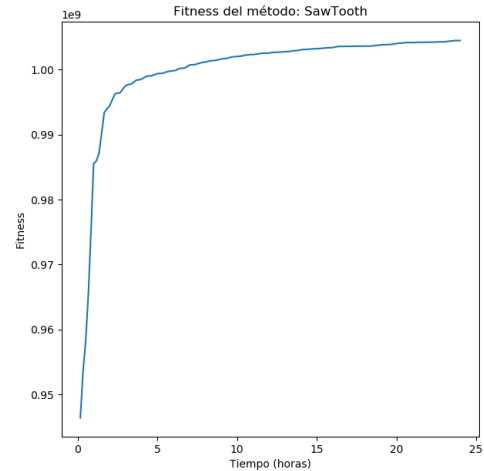
(q) Figura 18. Evolución de la calidad a lo largo del tiempo del método Multi.

Para la gráfica mostrada en la Figura 18. referente al método Multi podemos ver un comportamiento similar a las mostradas en los métodos (CLR 1 y CLR 5) por lo que podemos hacer una misma conclusión en torno a dichas gráficas.



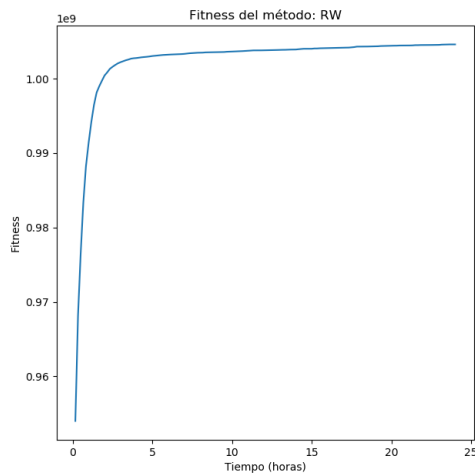
(r) Figura 19. Evolución de la calidad a lo largo del tiempo del método MultiDynamic.

En la Figura 19. podemos apreciar la gráfica de la evolución de calidad del método MultiDynamic. Para este caso podemos ver que la evolución mostrada en torno a su calidad corresponde ampliamente con la gráfica de la evolución de su diversidad poblacional. El fitness no fue creciendo muy rápido pero en etapas avanzadas de su ejecución llegó a valores muy grandes, por lo cual se puede decir que este método es muy exploratorio en torno a su espacio de búsqueda. Actualmente este método es el que mejores resultados dió en torno a su calidad.



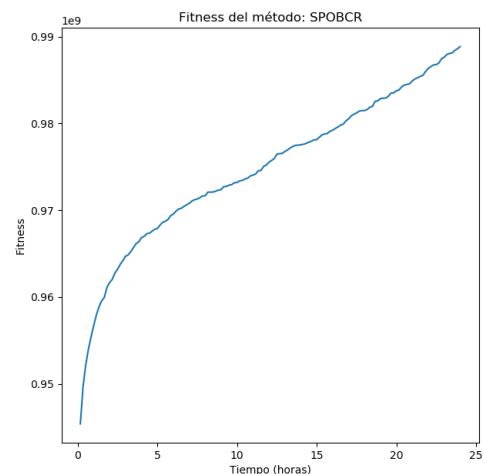
(t) Figura 21. Evolución de la calidad a lo largo del tiempo del método SawTooth.

Para la gráfica mostrada en la Figura 21. referente al método SawTooth podemos apreciar un comportamiento similar al de las gráficas mostradas en los métodos (CLR 1 y CLR 5) pero en este método se hace un pequeño paréntesis ya que recordando su gráfica de variabilidad poblacional la cual oscilaba mucho y parecía no converger en las últimas horas de su ejecución, la calidad ofrecida por este método parece no ser mala pero tampoco muy buena como la vista en el método anterior, y esto probablemente a que no hacía búsquedas intensivas en zonas óptimas que encontraba, más bien se enfocó en ser un método muy exploratorio.



(s) Figura 20. Evolución de la calidad a lo largo del tiempo del método RW.

Para la gráfica mostrada en la Figura 20. referente al método RW podemos ver un comportamiento similar a las mostradas en los métodos (CLR 1 y CLR 5) por lo que podemos hacer una misma conclusión en torno a dichas gráficas.



(u) Figura 22. Evolución de la calidad a lo largo del tiempo del método SPOBCR.

Para la gráfica mostrada en la Figura 22. referente al método SPOBCR podemos apreciar un comportamiento similar al de la mostrada en el método MultiDynamic, a primera vista se puede decir que este método tiende ser más exploratorio que intensivo pero al final nos damos cuenta

que no ofreció una buena calidad como lo hizo MultiDynamic.

Una vez que hemos realizado el análisis tanto de diversidad poblacional como de evolución de calidad podemos comentar algunos puntos importantes que llevan al segundo supuesto realizado en la sección *Motivación*. Se puede apreciar que varios de los métodos anteriores manejaban su diversidad de una forma muy distinta a otros, en general vimos métodos que toman una conducta muy elitista ya que siempre actualizan su población con la que mejor fitness da y esto a corto plazo ofrece buenos resultados pero cuando estos métodos continúan ejecutándose por gran cantidad de tiempo su fitness no mejora mucho ya que llegaron a un óptimo local y solamente están realizando una búsqueda intensiva.

Por otro lado se tienen los métodos exploratorios que hacen una gran búsqueda en el dominio de la función y probablemente usan técnicas como búsqueda por recocido simulado que tiende a aceptar resultados de empeoramiento para no quedarse en un sólo óptimo local y a largo plazo encontrar mejores resultados que los métodos descritos anteriormente. Por lo mencionado anteriormente podemos aceptar como cierta la segunda hipótesis, la cual establece:

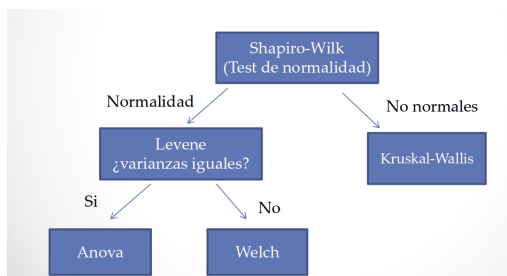
- Se cree que diferentes métodos manejan la diversidad de forma drásticamente diferente y que esto puede influir en los resultados a corto y largo plazo.

## V. ANÁLISIS ESTADÍSTICOS

En esta parte se mostrarán los resultados obtenidos al realizar análisis estadísticos a los métodos trabajados anteriormente, esto para poder hacer un rating de qué métodos fueron mejores en general.

Para esto anterior se usaron no solo los datos de la instancia *GECCO* sino que se analizaron los datos de otras 9 instancias divididas en: pequeñas, medianas y grandes, aquí cabe aclarar que la instancia *GECCO* está contenida en el conjunto de instancias grandes. En general se tienen 3 instancias pequeñas, 3 medianas y 3 grandes.

Para hacer una comparación objetiva se usaron varios métodos estadísticos y determinísticos que se basan en pruebas de hipótesis para los cuales se miden tres aspectos de un conjunto de datos, como son: Normalidad, varianza y media, la aplicación de dichos test se realizó de la siguiente manera:



(v) Figura 23. Jerarquía de los test estadísticos.

En la Figura 23. vemos la jerarquía usada al aplicar los test estadísticos anterior mencionados. A continuación se explicará para qué sirve cada test:

1. **Test de Shapiro-Wilk:** Este test verifica si un conjunto dado proviene de una distribución normal. Para nuestro caso debemos ver que los dos conjuntos (dos métodos) provengan de una distribución normal.
2. **Test de Levene:** Verifica que dos conjuntos provenientes de distribuciones normales tengan varianzas iguales.
3. **Test de Anova:** Verifica que dos conjuntos provenientes de distribuciones normales y de varianzas iguales tengan medias iguales.
4. **Test de Welch:** Verifica que dos conjuntos provenientes de distribuciones normales pero no de varianzas iguales tengan medias iguales.
5. **Test de Kruskal-Wallis:** Verifica que dos conjuntos que no son de distribuciones normales tengan medias iguales.

Podemos observar que los test que están en los nodos hoja del árbol mostrado en la Figura 23 siempre son referentes a verificar las medias de dos conjuntos, de esta forma si dichos test indican igualdad en medias diremos que los métodos son iguales y llegan a un empate, de caso contrario la última verificación que se hace es la comparación de las medias. Para el caso del problema 2DPP como es un problema de maximización se verifica que si un método tiene media más grande que otro se dice que el primer método es mejor que el segundo y viceversa.

Ahora para que lo anterior tenga sentido, los conjuntos de datos son obtenidos extrayendo el mejor fitness de cada corrida en el método, como cada instancia cuenta con 30 corridas por método entonces se genera un vector de 30 elementos refiriéndose a los 30 mejores fitness que el método generó en la misma instancia.

Cuadro I: Resumen de los test estadísticos aplicados a los métodos con instancias pequeñas

	Instancia pequeña 1 (10x10)				Instancia pequeña 2 (10x10)				Instancia pequeña 3 (10x10)			
	↑	↓	↔	Puntaje	↑	↓	↔	Puntaje	↑	↓	↔	Puntaje
CHC	4	3	2	1	5	3	1	2	4	3	2	1
CLR 1	4	4	1	0	3	4	2	-1	4	3	2	1
CLR 5	1	6	2	-5	1	6	2	-5	1	6	2	-5
COMB	6	1	2	5	7	1	1	6	6	1	2	5
DETCR	1	6	2	-5	2	5	2	-3	2	6	1	-4
Multi	6	1	2	5	7	1	1	6	7	1	1	6
MultiDynamic	9	0	0	9	9	0	0	9	9	0	0	9
RW	0	9	1	-9	0	9	0	-9	0	9	0	-9
SawTooth	1	6	2	-5	1	7	1	-6	1	7	1	-6
SPOBCR	5	1	3	4	4	3	2	1	4	2	3	2

Cuadro II: Resumen de los test estadísticos aplicados a los métodos con instancias medianas

	Instancia mediana 1 (15x15)				Instancia mediana 2 (15x15)				Instancia mediana 3 (15x15)			
	↑	↓	↔	Puntaje	↑	↓	↔	Puntaje	↑	↓	↔	Puntaje
CHC	6	3	0	3	6	2	1	4	6	2	1	4
CLR 1	4	4	1	0	4	4	1	0	4	4	1	0
CLR 5	2	6	1	-4	3	6	0	-3	2	6	1	-4
COMB	4	4	1	0	4	4	1	0	4	4	1	0
DETCR	0	9	0	-9	0	9	0	-9	0	9	0	-9
Multi	7	2	0	5	6	1	2	5	6	2	1	4
MultiDynamic	9	0	0	9	9	0	0	9	9	0	0	9
RW	1	8	0	-7	1	7	1	-6	1	7	1	-6
SawTooth	2	6	1	-4	1	7	1	-6	1	6	2	-5
SPOBCR	8	1	0	7	7	1	1	6	8	1	0	7



Cuadro III: Resumen de los test estadísticos aplicados a los métodos con instancias grandes

	Instancia mediana 1 (20x20)			Instancia mediana 2 (20x20)			Instancia mediana 3 (20x20)			Instancia grande 4 (20x20)		
	+	+	Puntaje	+	+	Puntaje	+	+	Puntaje	+	+	Puntaje
CHC	8	1	0	7	8	1	0	7	8	1	0	7
CLR 1	5	3	1	2	5	3	1	2	5	3	1	2
CLR 5	4	5	0	-1	4	5	0	-1	4	5	0	-1
COMB	5	3	1	2	5	3	1	2	5	3	1	2
DETCR	0	8	1	-8	0	9	0	-9	0	8	1	-8
Multi	7	2	0	5	7	2	0	5	7	2	0	5
MultiDynamic	9	0	0	9	9	0	0	9	9	0	0	9
RW	3	6	0	-3	2	6	1	-4	3	6	0	-3
Sawtooth	2	7	0	-5	2	6	1	-4	2	7	0	-5
SPOBCR	0	8	1	-8	1	8	0	-7	0	8	1	-8

Podemos ver en las Tablas 1, 2 y 3 los resultados aplicados a los test estadísticos antes mencionados. La explicación de la tabla es la siguiente: Por cada instancia se tienen 4 columnas, el método gana, el método pierde, el método empatiza y el puntaje final del método. Para obtener un punto en alguna de las primeras tres categorías se hace una comparación del método actual contra algún otro siguiendo la jerarquía de test estadísticos antes mencionada, si se llega a algún test que corresponde a un nodo hoja y este acepta la hipótesis que las medias de los métodos son iguales entonces se considera empate, y se añade un punto en esta categoría, de caso contrario se comparan las medias de los métodos y la que sea más grande es la que gana. Finalmente el puntaje se calcula restando las victorias contra las derrotas.

Lo primero que es necesario destacar es que el método *MultiDynamic* fue el que mejor se desempeñó en todos los casos, tal así que no perdió contra ningún método. De esta forma podemos concluir que este método es invariante al tamaño de la instancia.

Por otra parte el método *DETCR* fue el que peor desempeño tuvo en todas las instancias y esto concuerda con el comportamiento analizado en su evolución de calidad y entropía.

Hay métodos que tuvieron un desempeño bueno con ciertas instancias y malo en otras, tal es el caso del método *SPOBCR* en donde su mejor desempeño se apreció en las instancias medianas pero en las grandes fue superado por bastantes métodos. Otro caso similar es con el método *CHC* en donde su desempeño se vio muy bien reflejado en las instancias grandes pero en las pequeñas y medianas no tanto.

Por lo mencionado anteriormente podemos aceptar como cierta la primer hipótesis realizada, la cual establece:

- Se cree que el comportamiento relativo de los métodos puede depender del tamaño de la instancia.

## VI. CONCLUSIONES

En el presente reporte se pudo analizar el desempeño de 10 métodos estocásticos que se enfocaron en resolver el problema de optimización conocido como 2DPP. Primero se hizo un análisis detallado de la variabilidad poblacional generada por cada método en la instancia GECCO donde se apreció que varios métodos manejan su variabilidad de forma diferente con respecto al tiempo de la ejecución. Y de la misma forma esto se ve reflejado en la evolución de su fitness. Hay métodos que prefieren realizar una convergencia rápida por lo cual su variabilidad se reduce en un tiempo corto ya que llega a un óptimo local que puede no ser necesariamente una buena solución pero este tipo de métodos funcionan muy bien para ejecuciones cortas.

Por el otro lado tenemos los métodos que deciden hacer una búsqueda más global y tardan en realizar su convergencia ya que en cierto momento hacen reinicios o toman movimientos de empeoramiento, y a largo plazo encuentran buenas soluciones, este tipo de métodos son buenos para ejecuciones largas. Finalmente se hizo una comparación de todos los métodos usando los test estadísticos de normalidad, varianza y media para generar un ranking de los mismos. Se observó que había métodos que dieron resultados muy buenos en ciertas instancias pero en otras empeoraron de manera significativa por lo que se concluyó que el tamaño de la instancia afecta a la calidad de las soluciones del método.