

Santafé de Bogotá, Junio de 2003

**Universidad de los Andes**  
**Facultad de Ingeniería**  
**Departamento de Ingeniería Civil y Ambiental**



**Diseño de Redes de Distribución de Agua Potable de Mínimo  
Costo con Algoritmos Genéticos**

**PROYECTO DE GRADO**

Carlos Santiago López Giraldo

Asesor de Tesis: Juan G. Saldarriaga  
Profesor Titular

## TABLA DE CONTENIDO

<b>INTRODUCCIÓN.....</b>	<b>4</b>
<b>OBJETIVOS .....</b>	<b>6</b>
OBJETIVO GENERAL: .....	6
OBJETIVOS ESPECÍFICOS:.....	6
<b>PARTE 1: MARCO TEÓRICO.....</b>	<b>7</b>
<b>1. TEORÍA BASE DEL FLUJO DE AGUA EN TUBERÍAS.....</b>	<b>7</b>
1.1. <i>Tipos de flujo</i> .....	7
1.2. <i>Pérdidas de presión por fricción</i> .....	8
1.3. <i>Pérdidas menores</i> .....	10
<b>2. ANÁLISIS DE REDES CERRADAS.....</b>	<b>11</b>
2.1. <i>Definición y teoría base</i> .....	11
2.2. <i>Método del Gradiente</i> .....	18
<b>3. DISEÑO DE REDES CERRADAS .....</b>	<b>22</b>
3.1. <i>Descripción del problema</i> .....	22
<b>4. ALGORITMOS GENÉTICOS.....</b>	<b>24</b>
4.1. <i>Definición</i> .....	24
4.2. <i>Algoritmos genéticos dentro de problemas de optimización</i> .....	26
4.3. <i>Manejo de restricciones en un problema de optimización</i> .....	27
<b>5. ADAPTACIÓN DE LOS ALGORITMOS GENÉTICOS AL DISEÑO DE REDES CERRADAS.....</b>	<b>29</b>
5.1. <i>Objetivo del algoritmo</i> .....	29
5.2. <i>Espacio de búsqueda e individuos</i> .....	29
5.3. <i>Funciones de reproducción y mutación</i> .....	31
5.4. <i>Función de costo (sin manejo de restricciones)</i> .....	32
5.5. <i>Manejo de las restricciones del problema</i> .....	33
5.6. <i>Método de reproducción</i> .....	36
5.7. <i>Operadores genéticos adicionales – Reacondicionamiento de la probabilidad de reproducción</i> .....	37
5.8. <i>Operadores genéticos adicionales – Reducción del error estocástico de la “Rueda de Ruleta”</i> .....	39

6.	BIBLIOGRAFÍA.....	40
<b>PARTE 2: DOCUMENTACIÓN DEL PROGRAMA.....</b>		<b>41</b>
1.	REQUERIMIENTOS GENERALES.....	41
1.1.	Generalidades.....	41
1.2.	Curso Normal de Acciones.....	41
1.3.	Casos de Excepción.....	42
1.4.	Precondiciones.....	43
1.5.	Poscondiciones.....	43
2.	FUNCIONES.....	44
2.1.	Funciones Básicas.....	44
2.2.	Funciones de Costo.....	44
2.3.	Funciones de Reproducción.....	45
2.4.	Funciones Estadísticas.....	45
2.5.	Funciones de Salida.....	46
3.	DIAGRAMA DE SECUENCIA DEL SISTEMA.....	47
3.1.	Uso típico del programa.....	47
4.	CASOS DE USO.....	48
4.1.	Casos de uso.....	48
4.2.	Descripción.....	48
5.	REQUERIMIENTOS.....	50
5.1.	Abrir módulo de diseño.....	50
5.2.	Ingresar configuración/Variables Hidráulicas.....	51
5.3.	Ingresar configuración/Variables de Algoritmo.....	52
5.4.	Ver configuración Default/Variables Hidráulicas.....	54
5.5.	Ver configuración Default/Variables de Algoritmo.....	55
5.6.	Correr el Algoritmo.....	56
5.7.	Ver calificadores de Diseño.....	57
5.8.	Guardar Diseño.....	58
5.9.	Guardar Estadísticas.....	59
6.	ARQUITECTURA DEL SOFTWARE.....	60
7.	DIAGRAMAS DE CLASE.....	61
7.1.	TAlelo.....	61
7.2.	TIndividuo.....	61
7.3.	TGeneración.....	63

7.4.	<i>TCalificador</i> .....	64
7.5.	<i>THerencia</i> .....	66
7.6.	<i>TEstadísticas</i> .....	67
7.7.	<i>TControlDisenoGenetico</i> .....	68
8.	<b>BIBLIOGRAFÍA</b> .....	70
	<b>PARTE 3: PRUEBAS Y RESULTADOS</b> .....	<b>71</b>
9.	<b>FUNCIONAMIENTO DEL PROGRAMA</b> .....	71
9.1.	<i>Recuento del desarrollo</i> .....	71
9.2.	<i>Uso Típico</i> .....	73
10.	<b>REVISION DE REQUERIMIENTOS</b> .....	79
	<i>Requerimiento 1: Abrir Módulo de Diseño</i> .....	79
	<i>Requerimiento 2: Ingresar Configuración/VARIABLES Hidráulicas</i> .....	79
	<i>Requerimiento 3: Ingresar Configuración/VARIABLES de Algoritmo</i> .....	80
	<i>Requerimiento 4: Ver Configuración Default/VARIABLES Hidráulicas</i> .....	81
	<i>Requerimiento 5: Ver Configuración Default/VARIABLES de Algoritmo</i> .....	81
	<i>Requerimiento 6: Correr el Algoritmo</i> .....	81
	<i>Requerimiento 7: Ver Calificadores de Diseño</i> .....	82
	<i>Requerimiento 8: Guardar Diseño</i> .....	82
	<i>Requerimiento 9: Guardar Estadísticas</i> .....	82
11.	<b>ANALISIS DE RESULTADOS</b> .....	84
11.1.	<i>Influencia de las variables sobre el resultado del algoritmo</i> .....	84
11.2.	<i>Desempeño sobre redes de mayor tamaño</i> .....	94
11.3.	<i>Superficies de presión óptimas</i> .....	96
12.	<b>BIBLIOGRAFIA</b> .....	104
	<b>CONCLUSIONES</b> .....	<b>105</b>
	<b>RECOMENDACIONES – TRABAJOS FUTUROS</b> .....	<b>107</b>
	<b>ANEXOS</b> .....	<b>110</b>
1.	REDES EJEMPLO.RED, TOP1/R1.RED, TOP3/R28.RED. ....	110
2.	EJEMPLO DE ARCHIVO DE ESTADÍSTICAS.....	118
3.	VARIACIÓN DE LOS RESULTADOS CON LAS DISTINTAS VARIABLES. ....	120
4.	CONTENIDO DEL CD-ROM ADJUNTO.....	122



## INTRODUCCIÓN

El ser humano viene diseñando sus redes de distribución de agua potable desde ya más de mil años, comenzando con los acueductos de las ciudades romanas. Hasta principios de este siglo no se contaba sin embargo con las bases teóricas necesarias para diseñar estas redes de distribución, y por tanto todos los diseños se realizaban de manera empírica. Con el desarrollo de las ecuaciones de la mecánica de fluidos y del flujo de agua en tuberías, se abre el camino para poder diseñar redes de distribución de una manera analítica.

El problema básico de diseño consiste en tener una serie de nodos, donde se conectan los usuarios a la red, y una serie de tuberías uniendo estos nodos entre sí y uniendo la fuente de agua principal (un embalse, por lo general) a la red. Las restricciones que pesan sobre la red son mantener un caudal de demanda mínimo en cada uno de los nodos y una presión mínima que, por lo general, es la misma para toda la red. Esta presión mínima es necesaria para que el agua llegue a niveles más altos al nivel del nodo de conexión (los pisos altos de un edificio, por ejemplo).

Una vez se dispone la localización de los tubos y nodos, conocida como la topología de la red, el diseño tiene dos variables básicas: material y diámetros de los tubos en la red. Usualmente, el material se fija en función del costo, disponibilidad, o método constructivo. Resta a determinar la combinación de diámetros de la red que permita cumplir con las restricciones de presión y demanda.

El problema no tiene, sin embargo, una solución única. Hay un número considerable del total posible de combinaciones de diámetros que permite a la red cumplir con las restricciones impuestas. Es así como el enfoque real del problema no es encontrar una solución cualquiera sino buscar la solución óptima desde el punto de vista de costos. El problema de diseño se torna por tanto en un problema de optimización.

En esta tesis se adaptará un método de optimización conocido como algoritmos genéticos. Estos algoritmos no son precisamente métodos de optimización ya que su objetivo real no es la búsqueda de un mínimo global; sin embargo, son métodos de búsqueda cuyo objetivo es buscar en el espacio posible de soluciones, visto biológicamente como conjuntos de individuos (cada solución posible se mira como un individuo), un individuo altamente adaptado a la solución. En este caso el nivel de adaptación se mide como la minimización del costo de cada solución.

Se espera así poder desarrollar un algoritmo que permita diseñar una red de manera que el costo de esta sea mínimo. Se estudiarán los resultados que se obtengan para distintas redes a la luz de otros criterios de diseño como el de Featherstone.

## OBJETIVOS

### Objetivo General:

Crear un programa que, dada la topología, material y coeficientes de pérdidas menores de una red de distribución de agua potable cualquiera y teniendo en cuenta las restricciones de demanda y de presión mínima en los nodos de la red, encuentre un diseño (combinación de diámetros) que se aproxime a la solución de mínimo costo, por medio de un algoritmo genético.

### Objetivos Específicos:

- Verificar qué tan bien se adaptan los algoritmos genéticos al complejo problema de optimización del costo de una red de distribución de agua potable.
- Estudiar la manera de introducir las restricciones de diseño de un acueducto (demandas y presiones mínimas en los nodos) en un algoritmo genético.
- Relacionar, en el algoritmo genético, el costo de un diseño determinado con su probabilidad de reproducirse a la siguiente generación.
- Evaluar la efectividad del programa una vez desarrollado.

## PARTE 1: MARCO TEÓRICO

### 1. TEORÍA BASE DEL FLUJO DE AGUA EN TUBERÍAS

---

#### 1.1. Tipos de flujo

Hacia 1880 el ingeniero Osborne Reynolds hace un aporte muy importante a la teoría del flujo en conductos cerrados desarrollando de manera sólida la teoría de los diferentes tipos de flujo: flujo laminar, flujo turbulento y un punto medio entre los dos denominado flujo en transición.

El flujo laminar se caracteriza porque las líneas de corriente, o de flujo, siguen una trayectoria paralela entre ellas. En el flujo turbulento hay un intercambio significativo de partículas entre líneas de corriente. Mientras que en el flujo laminar las pérdidas de energía provienen solamente del rozamiento entre partículas en distintas líneas de corriente y entre el fluido y las paredes de la tubería, en el flujo turbulento se presentan pérdidas de energía adicionales debido al intercambio de partículas entre capas. De ello se desprende necesariamente que, al presentarse un mayor rozamiento intrínseco al flujo, las pérdidas de energía sean (sustancialmente) mayores para el flujo turbulento.

Con el fin de describir adecuadamente los tres tipos de flujo, Reynolds toma como base la teoría de similitud de fuerzas para relacionar las fuerzas inerciales y viscosas en una cantidad, adimensional, que caracterice el flujo: dos flujos que se encuentren bajo diferentes condiciones, pero cuyo valor de esta cantidad no difiera significativamente, se comportarán de manera similar [Streeter, pp. 260]. Esta cantidad se conoce hoy en día como el número de Reynolds:

$$Re = \frac{vd\rho}{\mu} \quad [1-1]$$

donde  $v$  es la velocidad del flujo,  $d$  una característica geométrica del flujo conocida como longitud significativa (que, en el caso de flujo en tuberías, corresponde al diámetro) y  $\rho$  y  $\mu$  la densidad y viscosidad del flujo.

El numerador de la expresión representa las fuerzas inerciales del flujo, y el denominador las fuerzas viscosas. Un número de Reynolds bajo indica una predominancia de las fuerzas viscosas, aumentando la tendencia a un flujo sin intercambio de partículas entre líneas de corriente y por lo tanto laminar. Un número alto indica predominancia de las fuerzas inerciales, caracterizando el flujo turbulento. Como límites usuales se toma  $Re < 2200$  para flujo laminar,  $Re > 4500$  para flujo turbulento. El intervalo intermedio ( $2200 < Re < 4500$ ) corresponde al flujo en transición. Estos intervalos de flujo son empíricos y difusos, y varían dependiendo de las condiciones específicas en que se midan. Sin embargo, para el flujo en tuberías, otorgan una buena caracterización del tipo de flujo presente.

## 1.2. Pérdidas de presión por fricción

Mientras el agua fluye a través de una tubería, como se mantiene (en general) un diámetro constante, se mantiene una velocidad media de flujo constante y, por lo tanto, la energía cinética del flujo es constante. Por consiguiente, las pérdidas de energía del flujo se dan en forma de presión.

Tales pérdidas están dadas por la siguiente ecuación:

$$h_f = f \frac{l \bar{v}^2}{d 2g} \quad [1-2]$$

donde  $h_f$  denota las pérdidas de presión por fricción (en unidades de metros de cabeza de agua) del fluido a velocidad media  $\bar{v}$  en un tramo de tubería de longitud  $l$ , diámetro  $d$  y con factor de fricción  $f$  ( $g$  denota la aceleración de la gravedad).

Esta ecuación, físicamente basada, fue deducida simultáneamente por Henry Darcy, ingeniero francés, y Julius Weisbach, científico e ingeniero alemán, en el s. XIX (ver [Saldarriaga, 2001], p. 45). Se conoce como la ecuación de Darcy-Weisbach.

Vale la pena señalar que a partir de esta ecuación, si se conociera el factor de fricción, sería posible diseñar una tubería simple ya que son parámetros del problema la presión con la que se cuenta inicialmente, la presión que se quiere tener en la salida y el caudal que se quiere distribuir (el único parámetro desconocido es el diámetro dado que la velocidad es función del diámetro y el caudal).

Es por ello que las investigaciones subsecuentes se orientaron hacia el estudio de la naturaleza de este factor de fricción adimensional  $f$ . El mismo Weisbach pudo determinar este factor en el caso de flujo laminar:

$$f = \frac{64}{\text{Re}} \quad \mathbf{[1-3]}$$

El siguiente paso lo dio el ingeniero alemán Johann Nikuradse, quien varió de manera controlada el factor de rugosidad absoluta de la tubería ( $k_s$ ) y el número de Reynolds para diagramar el valor del factor de fricción en función de estas dos variables. El valor  $k_s$  es una medida de la fricción que produce la tubería debido a las irregularidades en su material. Posteriormente, C.F. Moody produjo un diagrama similar al de Nikuradse para tuberías comerciales.

Por otro lado, L. Prandtl desarrolló la base teórica necesaria para expresar el factor de fricción para flujos turbulentos. Junto con su alumno T. Von Karman, dedujeron ecuaciones válidas en flujo turbulento mas no en flujo en transición, predominante en el flujo en tuberías comerciales [Saldarriaga, 2001]. Posteriormente C.F. Colebrook y H. White, investigadores ingleses, completaron los trabajos de Prandtl y Von Karman al deducir una fórmula que expresa el factor de fricción implícitamente en función de la rugosidad relativa de un tubo ( $k_s/d$ ) y el número de Reynolds del fluido. Esta fórmula es válida tanto para flujo

turbulento como para el flujo en transición. Se conoce como la ecuación de Colebrook-White:

$$\frac{1}{\sqrt{f}} = -2 \log_{10} \left( \frac{k_s}{3.7d} + \frac{2.51}{\text{Re} \sqrt{f}} \right) \quad [1-4]$$

### 1.3. Pérdidas menores

En una red de distribución de agua hay accesorios como codos, diferentes tipos de válvulas, Tes, expansiones o contracciones, entre otros, necesarios para el correcto funcionamiento hidráulico o para la construcción de la red. Estos accesorios causan un segundo tipo de pérdidas de energía conocido como pérdidas menores.

Para un accesorio determinado las pérdidas menores que este induzca se expresan como una porción de energía cinética así:

$$h_m = k_m \cdot \frac{\bar{v}^2}{2g} \quad [1-5]$$

donde  $h_m$  denota las pérdidas de presión (en unidades de metros de cabeza de agua) por el accesorio con coeficiente de pérdidas menores  $k_m$ . De nuevo,  $\bar{v}$  denota la velocidad media del fluido y  $g$  la aceleración de la gravedad.

## 2. ANÁLISIS DE REDES CERRADAS

---

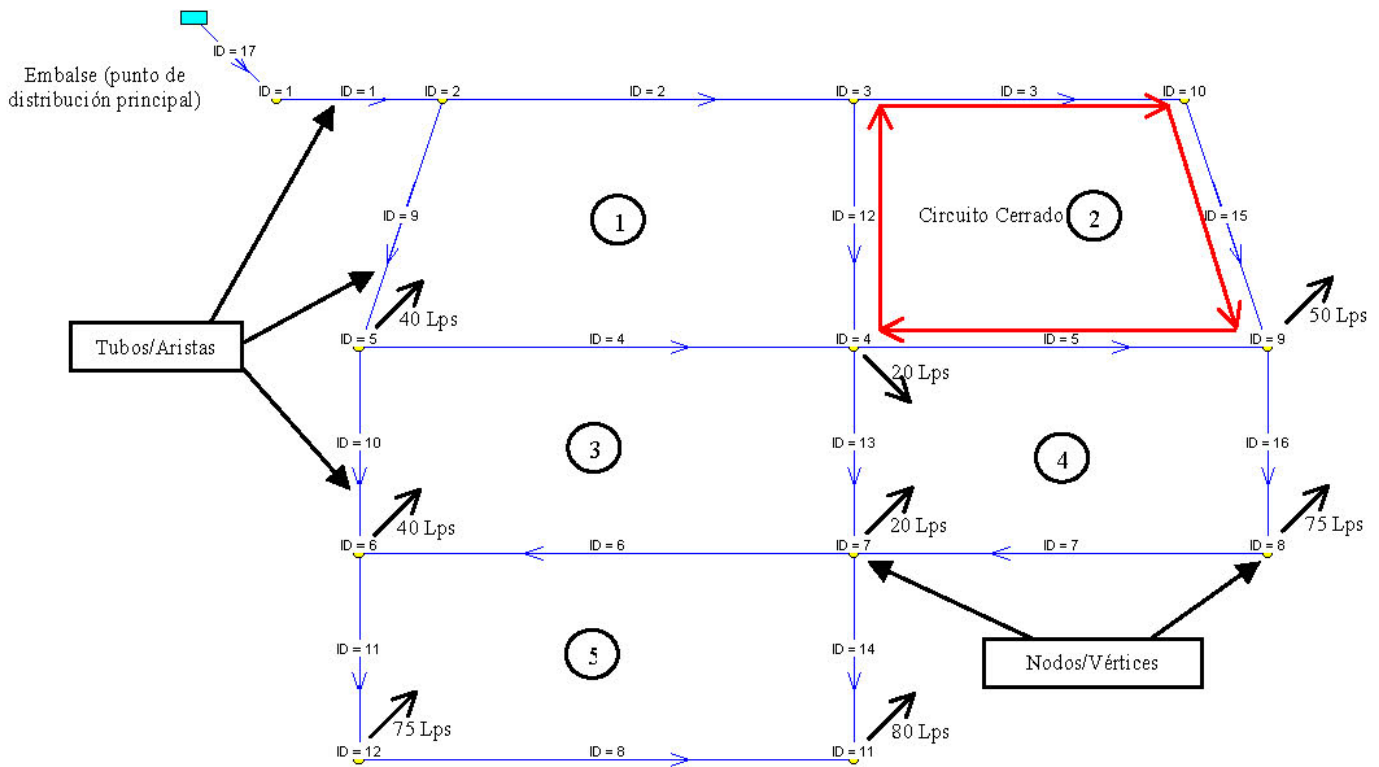
### 2.1. Definición y teoría base

Uno de los requisitos básicos que debe cumplir una red de abastecimiento de agua potable es ser redundante; es decir, que exista más de un camino por el cual el agua pueda moverse de un punto dado de la red a cualquier otro (incluso si ambos puntos se localizan sobre la misma tubería). Es claro porqué es necesario cumplir este requisito, ya que de no cumplirse, el cierre de una tubería por mantenimiento, daño o cualquier otra razón podría causar un corte de agua total para uno o más sectores de la ciudad.

Para que una red sea de alguna manera redundante, debe contar con un circuito cerrado (y viceversa, si una red cuenta con circuitos cerrados es hasta cierto punto redundante). Las redes con circuitos cerrados se conocen como redes cerradas, en contraste con las redes abiertas, donde el agua tiene un único camino para moverse entre dos puntos cualesquiera de la red (v. [Saldarriaga, 2001], p. 255).

En la figura 2-1 se muestra un ejemplo simple de red cerrada de distribución de agua potable (realizado sobre el software REDES). Las flechas azules indican la dirección del flujo.





**Figura 2-1. Ejemplo de una red cerrada.**

En una red de distribución real suelen ramificarse conexiones domiciliarias sobre toda la longitud de los tubos. Por simplicidad, y ya que el cambio en la hidráulica de la red es mínimo, se modela la red con salidas de caudal solo en los nodos, correspondientes a uniones de tuberías. De esta forma, la red de distribución de agua potable se modela como un conjunto de nodos o vértices (amarillos en la figura precedente) y uniones o aristas (correspondientes a las tuberías). Las salidas de caudal para el ejemplo anterior se indican sobre ciertos nodos con flechas y el valor del caudal demandado.

En rojo (en la figura 2-1) se resalta el circuito 2 de los cinco circuitos cerrados señalados en la red. Este circuito está compuesto por los tubos 3, 15, 5 y 12. Obsérvese que, por ejemplo, para el tubo 5 se deben cumplir las ecuaciones [1-2] y [1-5] de pérdida de cabeza:

$$\begin{aligned}
\Delta H = H_4 - H_9 &= -(h_f + h_m) \\
&= -\left[ f_{4,9} \frac{l_{4,9}}{d_{4,9}} \frac{\bar{v}_{4,9}^2}{2g} + \sum_{i=1}^{n_5} \left( k_{m_i} \frac{\bar{v}_{4,9}^2}{2g} \right) \right] \\
&= -\left[ f_{4,9} \frac{l_{4,9}}{d_{4,9}} \frac{\bar{v}_{4,9}^2}{2g} + \frac{\bar{v}_{4,9}^2}{2g} \cdot \sum k_{m4,9} \right] \\
&= -\frac{\bar{v}_{4,9}^2}{2g} \left( f_{4,9} \frac{l_{4,9}}{d_{4,9}} + \sum k_{m4,9} \right)
\end{aligned}$$

donde  $H_4$  y  $H_9$  se refieren a la presión en los nodos terminales del tubo 5, y el resto de los términos están indexados respecto a estos nodos también. Los coeficientes de pérdidas menores, si bien se refiere cada  $k_{mi}$  a un accesorio diferente sobre el tubo 5 (donde en total hay  $n_5$  accesorios) suelen expresarse simplemente como  $\sum k_{m4,9}$ . Esta generalización se debe a que el modelo hidráulico que se implanta es equivalente, y simplifica la formulación de las ecuaciones.

En general, la ecuación de pérdidas para un tubo que va del nodo  $i$  al nodo  $j$  es:

$$\begin{aligned}
\Delta H_{i,j} = H_j - H_i &= -\frac{\bar{v}_{i,j}^2}{2g} \left( f_{i,j} \frac{l_{i,j}}{d_{i,j}} + \sum k_{m_{i,j}} \right) \\
&= -\bar{v}_{i,j} \frac{|\bar{v}_{i,j}|}{2g} \left( f_{i,j} \frac{l_{i,j}}{d_{i,j}} + \sum k_{m_{i,j}} \right) \\
&= -Q_{i,j} \frac{|Q_{i,j}|}{2g \cdot A_{i,j}} \left( f_{i,j} \frac{l_{i,j}}{d_{i,j}} + \sum k_{m_{i,j}} \right) \quad [2-1]
\end{aligned}$$

El término de velocidad al cuadrado se expresa de esta manera para tener en cuenta la dirección de movimiento del flujo. Así, si el flujo va del nodo  $i$  al  $j$  se presentará una pérdida de energía del nodo  $i$  al  $j$ , siendo  $H_j - H_i$  en este caso negativo (lo que concuerda con el caudal positivo  $Q_{i,j}$  multiplicado por  $-1$ ).

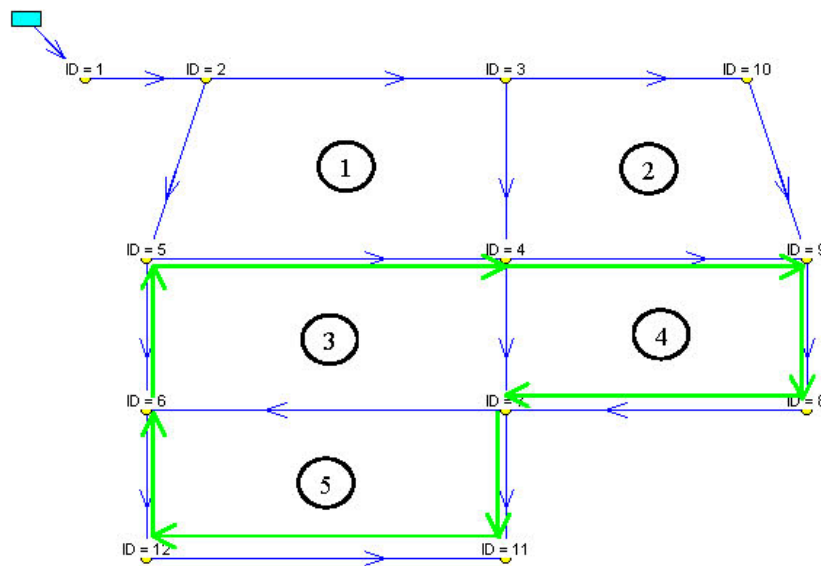
Suponiendo que los nodos  $i_1, i_2, \dots, i_n$  conforman un circuito, es obvio que:

$$\begin{aligned}
0 &= (H_{i_2} - H_{i_1}) + (H_{i_3} - H_{i_2}) + \dots + (H_{i_n} - H_{i_{n-1}}) + (H_{i_1} - H_{i_n}) \\
&= \sum_{k=1}^n \Delta H_{i_k, j_{k+1}} \\
&= \sum_{k=1}^n \left[ -Q_{i_k, j_{k+1}} \frac{|Q_{i_k, j_{k+1}}|}{2g \cdot A_{i_k, j_{k+1}}} \left( f_{i_k, j_{k+1}} \frac{l_{i_k, j_{k+1}}}{d_{i_k, j_{k+1}}} + \sum k_{m_{i_k, j_{k+1}}} \right) \right] \\
\Rightarrow 0 &= \sum_{k=1}^n \left[ Q_{i_k, j_{k+1}} \frac{|Q_{i_k, j_{k+1}}|}{2g \cdot A_{i_k, j_{k+1}}} \left( f_{i_k, j_{k+1}} \frac{l_{i_k, j_{k+1}}}{d_{i_k, j_{k+1}}} + \sum k_{m_{i_k, j_{k+1}}} \right) \right] \quad [2-2]
\end{aligned}$$

Siguiendo la convención de que  $i_{n+1}$  se refiere a  $i_1$ . Esta última ecuación se conoce como la ecuación de conservación de energía en el circuito.

Obsérvese que las ecuaciones de los cinco circuitos resaltados en la figura son linealmente independientes. Es decir, ninguna de las 5 ecuaciones puede ser deducida a partir de las otras 4 ecuaciones. Además nótese que, para cualquier otro circuito sobre la red, la respectiva ecuación de circuito se puede expresar como la suma de algunas de las 5 ecuaciones de circuito ya deducidas. Por ejemplo, la ecuación del circuito cerrado verde resaltado en la figura 2.2 se puede expresar como:

$$\begin{aligned}
0 &= (H_4 - H_5) + (H_9 - H_4) + (H_8 - H_9) + (H_7 - H_8) + (H_{11} - H_7) \\
&\quad + (H_{12} - H_{11}) + (H_6 - H_{12}) + (H_5 - H_6) \\
&= [(H_4 - H_5) + (H_7 - H_4) + (H_6 - H_7) + (H_5 - H_6)] + [(H_9 - H_4) + (H_8 - H_9) + (H_7 - H_8) + (H_4 - H_7)] \\
&\quad + [(H_7 - H_6) + (H_{11} - H_7) + (H_{12} - H_{11}) + (H_6 - H_{12})] \\
0 &= \mathbf{C}_3 + \mathbf{C}_4 + \mathbf{C}_5
\end{aligned}$$



**Figura 2-2. Circuito cerrado compuesto por más de un circuito simple.**

donde  $C_3$ ,  $C_4$ , y  $C_5$  representan las ecuaciones de conservación de energía de los circuitos 3, 4 y 5.

Así, en una red cerrada dada hay  $C$  ecuaciones de circuito, donde  $C$  corresponde al número de circuitos cerrados simples en la red.

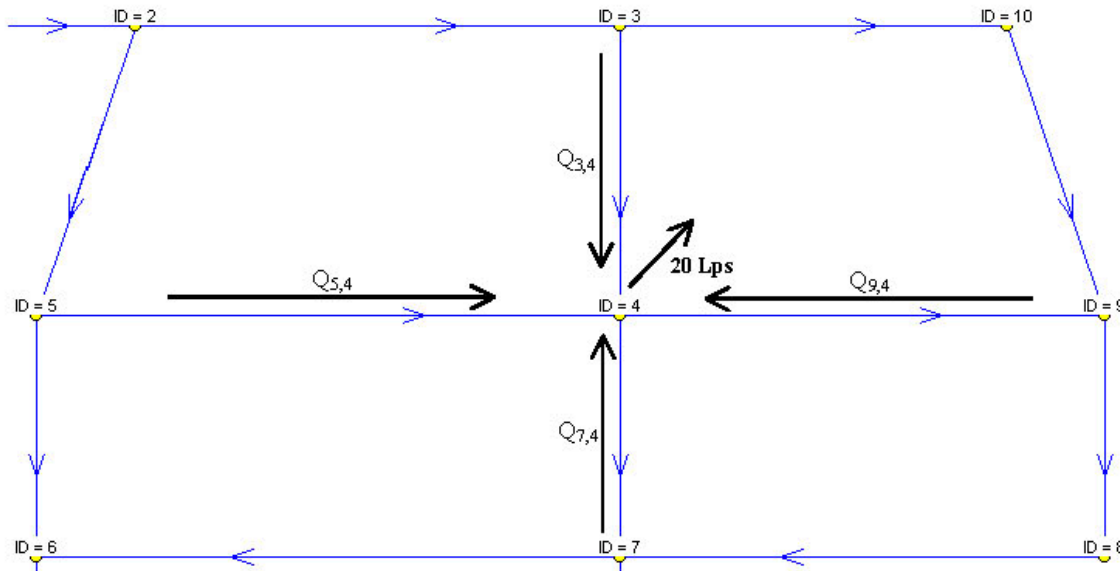
Por otro lado, para un nodo cualquiera, por ejemplo el nodo 4, se debe cumplir la ecuación de conservación de masa:

$$Qe_4 - Qs_4 = 0$$

$$\Rightarrow \sum_j Q_{j,4} - Qd_4 = 0$$

donde  $Q_{j,4}$  representa los caudales que van de los distintos nodos  $j$  al nodo 4 (en este caso de los nodos 3, 5, 7 y 9 al nodo 4), y  $Qd_4$  representa el caudal de salida (el caudal demandado) en el nodo 4 (20 Lps, ver figura 2-3 a continuación). Vale la pena remarcar

que los caudales  $Q_{j,4}$  hacen parte tanto de los caudales  $Qe_4$  de entrada (en este caso  $Q_{3,4}$ ,  $Q_{5,4}$ , que por ser de entrada tienen signo positivo) como de los caudales  $Qs_4$  de salida ( $Q_{7,4}$ ,  $Q_{9,4}$ , que, acá, tendrían signo negativo). En este caso en la ecuación no hay términos análogos a  $Qd_4$  que indiquen una entrada externa de caudal al nodo 4.



**Figura 2-3. Caudales de entrada y salida a un nodo particular.**

Análogamente, se tiene para cada nodo  $i$ :

$$\begin{aligned}
 Qe_i - Qs_i &= 0 \\
 \Rightarrow \sum_{j \in N_i} Q_{j,i} + Qe_i - Qd_i &= 0
 \end{aligned}
 \tag{2-3}$$

donde  $N_i$  es el conjunto de nodos que están conectados por algún tubo al nodo  $i$ . En una red cerrada de  $N$  nodos se tienen  $N$  de estas ecuaciones de conservación de masa. En este caso, más general, sí se consideran entradas netas externas de caudal ( $Qe_i$ ).

Al sumar todas estas ecuaciones se obtiene la ecuación de conservación de masa para el circuito entero:

$$\sum_{i=1}^N \left[ \sum_{j \in N_i} Q_{j,i} + Qe_i - Qd_i \right] = 0$$

$$\Rightarrow \sum_{i=1}^N Qe_i - \sum_{i=1}^N Qd_i = 0 \quad [2-4]$$

Ya que los términos  $Q_{j,i}$  de movimiento de agua internamente en la red se cancelan entre si ( $Q_{j,i} + Q_{i,j} - Q_{j,i} - Q_{j,i} = 0$ ).

Dado que esta ecuación se conoce previamente (los caudales demandados en la red son datos conocidos y, en general, el o los caudales de entrada lo son también) y es independiente de los diámetros u otras características de la red, no posee información adicional. Sin embargo, como ella se puede expresar como la suma de las demás ecuaciones de conservación de masa de la red se puede concluir que cualquiera de las ecuaciones de masa de la red se puede expresar como la suma de las demás quitándole a esta suma (de ecuaciones) la ecuación del circuito. Por lo tanto, habrá  $N - 1$  ecuaciones independientes de conservación de masa.

Por otro lado, sobre cualquier red cerrada arbitraria se puede demostrar por inducción matemática sobre el número de nodos  $N$  y de circuitos  $C$  que el número de uniones  $U$  (obviamente igual al número de caudales dentro de la red) es:

$$U = N + C - 1 \quad [2-5]$$

por lo que el sistema tiene el mismo número de incógnitas que de ecuaciones, aunque desafortunadamente estas sean altamente no lineales y por tanto no se pueda resolver por los métodos estándar del álgebra lineal. En consecuencia, se han desarrollado diversos métodos, de los cuales se presenta el Método del Gradiente, por tratarse del más eficaz en términos computacionales (tiempo de cálculo).

## 2.2. Método del Gradiente

El método del gradiente es, como los demás métodos de cálculo hidráulico existentes, un método iterativo de cálculo de los caudales y presiones de la red. El método comienza con unos caudales y presiones supuestas que convergen después de, a lo sumo,  $N$  (número de nodos de la red) iteraciones (v. [Saldarriaga, 2001] p. 369).

La suposición principal del método es que las pérdidas menores y las pérdidas por fricción se pueden aproximar como sigue:

$$h_f + h_m = \alpha \cdot Q^n + \beta + \frac{\gamma}{Q}$$

siendo  $\alpha$  un parámetro característico del tubo y,  $\beta$  y  $\gamma$  parámetros característicos del tubo, las válvulas y bombas presentes.  $n$  es un entero positivo que depende de la ecuación de fricción utilizada. Para la ecuación de Darcy-Weisbach este exponente es 2. Se estudiará solo el caso en el cual no se tienen válvulas ni bombas, por lo que la ecuación toma la siguiente forma:

$$\begin{aligned} h_f + h_m &= \alpha \cdot Q^n \\ h_f + h_m &= \alpha \cdot Q^2 \end{aligned} \quad [2-6]$$

Dada la naturaleza iterativa del método, el parámetro  $\alpha$  para la  $k$ -ésima iteración vendría a ser:

$$\alpha^{k+1} = \frac{h_f^{k+1} + h_m^{k+1}}{(Q^k)^2} \quad [2-7]$$

Los superíndices denotan la iteración del parámetro respectivo (con la excepción, por supuesto, del exponente 2 sobre  $Q^k$ ). En lo que resta de esta sección se denotarán los superíndices y exponentes como en la ecuación [2-7].

En esta ecuación, tanto  $h_f^{k+1}$  como  $h_m^{k+1}$ , si bien son los utilizados en la iteración  $k+1$ , son calculados también con  $Q^k$  (como en los demás métodos).

Con este supuesto, la ecuación de pérdida de energía para el tubo que va del nodo  $i$  al nodo  $j$  se puede escribir como:

$$\begin{aligned}\Delta H_{i,j} &= H_j - H_i = -(h_f + h_m)_{i,j} \\ \Rightarrow (h_f + h_m)_{i,j} - H_i + H_j &= 0\end{aligned}$$

Lo que motiva la siguiente definición de la función  $f_{i,j}$ :

$$f_{i,j}(Q) = \alpha_{i,j} \cdot (Q)^2 - H_i + H_j \quad [2-8]$$

Para cualquier caudal  $Q$ . De la definición de  $f_{i,j}$  y la ecuación precedente a ésta es claro que  $f_{i,j}(Q_{i,j}) = 0$ .

Se recuerda la ecuación de conservación de masa y se define otra función  $g$  a partir de ella,

$$\begin{aligned}\sum_{j \in N_i} Q_{j,i} - q_i &= 0 \\ g_i(Q_1, Q_2, \dots, Q_{n_i}) &= \sum_{j=1}^{n_i} Q_j - q_i\end{aligned} \quad [2-9]$$

donde  $q_i$  indica la demanda neta de caudal ( $Qd_i - Qe_i$ ) en el nodo  $i$ . El término  $n_i$  indica la cantidad de tubos que llegan al nodo  $i$ , siendo entonces  $n_i = |N_i|$  (este conjunto  $N_i$  es el descrito en la ecuación 2-3), y los caudales  $Q_1, Q_2, \dots, Q_{n_i}$  son caudales cualesquiera. Así, cada función  $g_i$  tiene un número diferente de argumentos, pero bien definido en cada caso. Es claro también de la ecuación de conservación de masa que cuando los caudales son  $Q_{j,i}$  con  $j$  en  $N_i$  (es decir, los caudales que van de todos los tubos conectados al nodo  $i$ ) la ecuación [2-9] viene a ser:



$$g_i \left( \left\langle Q_{j,j} \right\rangle_{j \in N_i} \right) = \sum_{j \in N_i} Q_{j,j} - q_i = 0$$

siendo la ecuación [2-9] en este caso la misma ecuación de conservación de masa para el nodo  $i$ . Para simplificar la notación se adoptará la siguiente convención:

$$g_i(N_i) \stackrel{def}{=} g_i \left( \left\langle Q_{j,j} \right\rangle_{j \in N_i} \right)$$

El método del gradiente consiste en solucionar las funciones  $f_{i,j}$  y  $g_i$  para los caudales y presiones de la red ayudándose de su expansión de Taylor truncada (ignorando los términos de orden 2 en adelante). Teniendo en cuenta que  $f_{i,j}(Q_{i,j}) = 0$  y  $g_i(N_i) = 0$ , se deduce de manera idéntica al método de Newton-Rhapson (v. [Saldarriaga, 2001] p. 326) que:

$$\frac{df_{i,j}(Q_{i,j}^k)}{dQ_{i,j}^{k+1}} dQ_{i,j}^{k+1} = -f_{i,j}(Q_{i,j}^k) \quad [2-10]$$

$$\frac{dg_i(N_i^k)}{dQ_{j_1,j}^{k+1}} dQ_{j_1,j}^{k+1} = -g_i(N_i^k) \quad [2-11]$$

donde, de nuevo, el superíndice indica la iteración. La ecuación [2-11] se está derivando respecto a  $Q_{j_1,j}$ , siendo este caudal el primero del conjunto  $N_i$ . Desarrollando ambas expresiones,

$$\begin{aligned} & \left( 2\alpha_{i,j}^k \cdot Q_{i,j}^k - \frac{dH_i^{k+1}}{dQ_{i,j}^{k+1}} + \frac{dH_j^{k+1}}{dQ_{i,j}^{k+1}} \right) dQ_{i,j}^{k+1} = -\left( \alpha_{i,j}^k \cdot (Q_{i,j}^k)^2 - H_i^{k+1} + H_j^{k+1} \right) \\ \Rightarrow & 2\alpha_{i,j}^k \cdot Q_{i,j}^k \cdot dQ_{i,j}^{k+1} - dH_i^{k+1} + dH_j^{k+1} = -\left( \alpha_{i,j}^k \cdot (Q_{i,j}^k)^2 - H_i^{k+1} + H_j^{k+1} \right) \quad [2-13] \end{aligned}$$

$$\begin{aligned} & \frac{dg_i \left( \left\langle Q_{j,j}^k \right\rangle_{j \in N_i} \right)}{dQ_{j_1,j}^{k+1}} dQ_{j_1,j}^{k+1} = \left( \sum_{j \in N_i} \frac{dQ_{j,j}^k}{dQ_{j_1,j}^{k+1}} \right) dQ_{j_1,j}^{k+1} \quad [2-12] \\ \Rightarrow & \sum_{j \in N_i} dQ_{j,j}^{k+1} = - \left( \sum_{j \in N_i} Q_{j,j}^k - q_i \right) \end{aligned}$$

Y los términos  $dQ_{i,j}^{k+1}$  y  $dH_i^{k+1}$  se aproximan de manera natural así:

$$dQ_{i,j}^{k+1} = Q_{i,j}^{k+1} - Q_{i,j}^k \quad [2-14]$$

$$dH_i^{k+1} = H_i^{k+1} - H_i^k \quad [2-15]$$

Finalmente, las ecuaciones se plantean de manera matricial:

$$\begin{bmatrix} F' & C \end{bmatrix} \begin{bmatrix} dQ^{k+1} \\ dH^{k+1} \end{bmatrix} = \begin{bmatrix} \Delta E^k \\ \Delta M^k \end{bmatrix} \quad [2-16]$$

donde,

$$F' = \begin{bmatrix} 2\alpha_1^k \cdot Q_1^k & 0 & \dots & 0 \\ 0 & 2\alpha_2^k \cdot Q_2^k & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 2\alpha_T^k \cdot Q_T^k \end{bmatrix} \quad [2-17]$$

Dado que cada caudal  $Q_{i,j}$  corresponde a exactamente un tubo se hace el cambio de notación de nodos a tubos ( $U$  es el número de tubos de la red). La matriz  $C$  se conoce como matriz de conectividad, y es una matriz de  $T \times N$  donde la entrada  $[C]_{i,j}$  es  $-1$  si el tubo  $i$  viene del nodo  $j$ ,  $1$  si llega al nodo  $j$  y  $0$  de lo contrario (v. [Saldarriaga, 2001] p. 366).

Los vectores  $\Delta E^k$  y  $\Delta M^k$  son los lados derechos de las ecuaciones [2-13] y [2-12] respectivamente (los errores en las ecuaciones de conservación de masa y energía para la iteración  $k$ ).

La solución al sistema consiste en invertir la matriz de la ecuación [2-16]. Esta matriz es invertible ya que las ecuaciones allí descritas son las de conservación de energía y masa (en su expansión de Taylor), que son independientes como se mostró hacia el final de la sección 2.1.

### 3. DISEÑO DE REDES CERRADAS

---

#### 3.1. Descripción del problema

Durante el desarrollo de las diferentes ecuaciones presentadas en la sección anterior se tomó el diámetro de los tubos de la red como un dato del problema, y así al solucionar las ecuaciones de conservación de masa y energía (por el método del gradiente por ejemplo) se realizó una comprobación de diseño: se calcularon los caudales en los tubos y cabezas en los nodos de la red especificada.

Al diseñar una red cerrada se parte por supuesto de una situación bastante diferente. Se tiene una topología determinada por la localización de los puntos de consumo y, a lo sumo, un material (rugosidad) ya escogido para construir la red y la localización de válvulas y accesorios (coeficientes de pérdidas menores). En lo que consiste el diseño es, precisamente, en determinar los diámetros de todos los tubos de la red.

El diseño que se escoja para la red debe cumplir con los requisitos básicos para satisfacer al usuario: llevar como mínimo el caudal demandado en cada nodo y hacerlo con una presión mínima requerida (para que, por ejemplo, llegue adecuadamente a los pisos altos de un edificio). Las presiones y caudales mínimos de la red se escogen según diferentes metodologías (v. [Saldarriaga, 2001], pp. 413, 414).

Existe sin embargo, para una red dada, un número considerable de diseños que satisfacen estas restricciones mínimas. Por esto, el problema de fondo no es encontrar un diseño cualquiera para la red, sino encontrar un diseño óptimo desde un punto de vista, usualmente, económico.

La solución al problema de fondo es compleja: para una red cualquiera el tamaño del espacio solución (sin tener en cuenta las restricciones) es de  $|D|^U$ , donde  $D$  es el conjunto de diámetros comerciales que se pueden escoger para los tubos,  $|D|$  representa el tamaño de

este conjunto (por ejemplo,  $|D| = 17$  para  $D = \{ \frac{1}{2} \text{ in}, 1 \text{ in}, 1\frac{1}{2} \text{ in}, 2 \text{ in}, 3 \text{ in}, 4 \text{ in}, 5 \text{ in}, 6 \text{ in}, 8 \text{ in}, 10 \text{ in}, 12 \text{ in}, 16 \text{ in}, 18 \text{ in}, 24 \text{ in}, 30 \text{ in}, 60 \text{ in}, 72 \text{ in} \}$ , un conjunto de diámetros usual) y  $U$  es el número de tubos de la red (ver sección 2.1). Tan solo en el caso de una red simple como la de la figura 2-1, estaría compuesto de  $17^{17}$  (más de  $10^{18}$ ) diseños posibles. Esto descarta instantáneamente el uso de algoritmos de búsqueda exhaustiva para diseñar, y se debe recurrir por lo tanto a algoritmos de optimización más sofisticados.

Los algoritmos genéticos se adaptan adecuadamente al objetivo del diseño optimizado de redes y a los métodos de comprobación de diseño existentes, ya que lo único que su funcionamiento requiere, en relación con este problema, es un método que calcule las presiones en los nodos y los caudales en los tubos de la red para un diseño determinado y así “ranquearlo” (ver capítulos 4 y 5). Por ello, con base en las ecuaciones desarrolladas en la sección previa, en un método de solución de éstas (el método del gradiente) y a través del módulo de cálculo del programa REDES se cuenta con los requisitos hidráulicos necesarios para la solución del problema presentado.

## 4. ALGORITMOS GENÉTICOS

---

### 4.1. Definición

En 1953, los trabajos conjuntos de varios investigadores y, en especial, de James Watson y Francis Crick, cuyo resultado fue el descubrimiento de la estructura química del ADN dieron lugar a la complementación de uno de los conceptos biológicos más importantes: el de fenotipo de un organismo con el de genotipo de un organismo.

El genotipo de un organismo se refiere a la composición de su ADN como tal; es decir, la secuencia exacta de bases nitrogenadas que lo conforman (la secuencia exacta de Adenina, Guanina, Citosina y Uracil que compone su ADN). Por su parte, el fenotipo se refiere a las características físicas de un organismo (color de piel, forma de las extremidades, dureza de sus huesos, rata de producción de ciertas hormonas, entre otros).

A partir de ese momento se desarrollaron las implicaciones de este gran hallazgo y poco a poco se confirmó la relación íntima y totalmente intuitiva existente entre genotipo y fenotipo. Los mecanismos de transcripción del ADN son los responsables, por intermedio del ARN (ácido ribonucleico) y de un gran número de proteínas, de construir a partir del genotipo (como un código) las diversas proteínas responsables del fenotipo. En consecuencia, se replanteó la definición de fenotipo como la expresión del genotipo de un organismo conjuntamente con los efectos de los diversos factores ambientales<sup>1</sup>.

Paralelo a lo anterior, se ligó la teoría Darwiniana de la evolución de las especies al concepto de genotipo. Así, el individuo cuyo fenotipo sea el mejor adaptado a su entorno es el que tiene mayor probabilidad de reproducirse, y dado que los mecanismos de reproducción perpetúan (en su totalidad para organismos asexuales y parcialmente para organismos sexuales) el genotipo de un individuo, entre más adaptado se encuentre el

---

<sup>1</sup> Para mayor información en lo concerniente a los aspectos químicos y biológicos del ADN y su transcripción (en proteínas) y la relación entre el genotipo y fenotipo referirse a [Purves et al., 1998], caps. 11 y 12.

fenotipo de un individuo en particular a su entorno, mayor será la probabilidad de perpetuarse del genotipo correspondiente. Esto implica que las cadenas de ADN que codifican fenotipos más exitosos tenderán a aumentar su frecuencia relativa en una población dada de individuos.

Con base en estos conceptos y en el obvio éxito de este mecanismo (que ha logrado evolucionar especies altamente exitosas y adaptadas a su entorno), se desarrolla el modelo computacional conocido como Algoritmos Genéticos.

Estos modelos están compuestos por los siguientes elementos, análogos a los correspondientes conceptos biológicos:

- Espacio de búsqueda: autodescriptivo, espacio donde residen las soluciones candidatas al problema. Los casos típicos son  $\{0, 1\}^n$  (para un problema de  $n$  caracteres binarios),  $R^n$  ( $n$ -vectores sobre los reales) o  $\Sigma^n$ , donde  $\Sigma$  es un alfabeto predeterminado. Es análogo al concepto biológico de entorno o medio ambiente.
- Individuos: cada individuo es un candidato solución al problema, un vector dentro del espacio de búsqueda. Conceptualmente, lo caracterizan su genotipo y fenotipo.
- Función de costo: función que relaciona la aptitud del individuo para solucionar el problema con un valor numérico. Entre más apto sea un individuo, se espera que sea menor el valor que tome la función de costo en él o viceversa. Esta función es el puente entre la adaptabilidad del individuo y su fenotipo.
- Genotipo: es el individuo visto como vector o como una cadena de caracteres. Estos caracteres tienen como equivalente biológico las bases nitrogenadas del ADN.
- Fenotipo: el fenotipo de un individuo es la representación que puede llegar a tomar la cadena de caracteres que conforma su genotipo dependiendo de cada problema (se ejemplificará este concepto en la sección 5.2). En algunas ocasiones será posible calcular el valor que toma un individuo en la función de costo a partir de su genotipo (tornando irrelevante el fenotipo en el contexto del algoritmo), mientras que en otras será necesario recurrir al fenotipo del individuo.

- **Función de reproducción:** función que asigna a cada individuo una probabilidad de reproducirse de acuerdo con el valor que éste toma en la función de costo.
- **Método de reproducción:** algoritmo que a partir de dos individuos (padres o ascendientes) produce un tercero (descendiente). El método usual se conoce como “crossover” (recombinación), donde el genotipo del individuo descendiente es idéntico, hasta una posición escogida al azar, al genotipo de uno de los padres, y de ahí en adelante al genotipo del segundo.
- **Generación:** todos los individuos presentes en una iteración del algoritmo.

Adicionalmente, el modelo incluye, por lo general, el fenómeno de mutación y, con menor frecuencia, el de “inversión” del ADN. La mutación consiste en cambiar por un valor aleatorio una posición escogida al azar del genotipo de un individuo. La inversión consiste en transcribir en orden inverso una sección del genotipo de un individuo escogida también al azar.

La primera generación de individuos se construye tomando el genotipo de cada uno de manera aleatoria sobre el espacio de búsqueda. Posteriormente, se asigna a cada uno de estos individuos su costo, y una probabilidad de reproducción inversamente proporcional a este (o relacionada a este de alguna otra manera dependiendo de la función de reproducción). Luego se reproducen al azar los individuos con la probabilidad asignada en el paso previo, para generar la siguiente generación que tendrá, con casi total seguridad, un menor valor promedio de costo. El procedimiento se repite hasta llegar a un número máximo de generaciones o hasta encontrar un individuo cuya función de costo esté por debajo de un cierto umbral predeterminado.

#### **4.2. Algoritmos genéticos dentro de problemas de optimización**

Los algoritmos genéticos constituyen, en un problema de optimización, un método de búsqueda ordenada por el espacio de solución. Es decir, los algoritmos genéticos no garantizan de ninguna manera el encontrar el mínimo global y, ni siquiera, los mínimos

locales de una función determinada. Simplemente estos algoritmos van buscando por el espacio solución localizando los individuos de cada generación progresivamente más cerca de los mínimos (locales o globales) de la función de costo. Es de esperar que después de un número razonable de generaciones se encuentre una solución candidata cuyo valor en la función de costo sea relativamente bajo, o se halle relativamente “cerca” al mínimo.

Aunque lo anterior muestre que los algoritmos genéticos no son formalmente algoritmos de optimización, su eficacia como tal se ha visto comprobada tanto por la resolución efectiva de diversos problemas desde su desarrollo (para ver algunos ejemplos referirse a [Michell, 1999], p. 17), como por el éxito que han tenido como mecanismo de optimización dentro de los sistemas biológicos.

### **4.3. Manejo de restricciones en un problema de optimización**

Los algoritmos genéticos no controlan de manera “natural” las restricciones dentro de un problema de optimización. Es decir, la búsqueda ordenada que proporcionan estos algoritmos garantiza constantemente una perpetuación de los genotipos que tienen ciertas secuencias (cadenas de vectores de menor tamaño al genotipo entero contenidos en éste) que contribuyen a tener una baja función de costo (ver [Goldberg, 1989], cap. 2), pero en ningún momento garantizan que los individuos de una generación dada se mantengan dentro del subconjunto del espacio de búsqueda delimitado por las restricciones del problema.

Existen básicamente dos razones que lo explican: es natural que al crear la primera generación de manera aleatoria haya individuos que no cumplan con las restricciones, y aún si se tienen individuos que cumplen con las restricciones el mecanismo de reproducción (recombinación) hace posible que los descendientes de estos individuos no cumplan con las restricciones.



Como está explicado en [Goldberg, 1989] (pp. 85) para muchos problemas prácticos resulta casi tan difícil encontrar individuos que no violen las restricciones como encontrar el óptimo. Por consiguiente, muchas veces la opción de simplemente descartar los individuos que no cumplen con las restricciones dificultaría bastante el funcionamiento de los algoritmos genéticos.

[Goldberg, 1989] recomienda, por tanto, introducir las restricciones dentro de la función de costo. Así, suponiendo que se busca minimizar la función de costo  $f(x)$ , donde  $x$  representa un individuo dentro del espacio de búsqueda, y que las restricciones están dadas por una serie de desigualdades (las restricciones de igualdad se modificarían por un par de desigualdades):

$$\begin{aligned} g_1(x) &\leq 0 \\ g_2(x) &\leq 0 \\ &\vdots \\ g_n(x) &\leq 0 \end{aligned} \quad [4-1]$$

(es claro que toda desigualdad se puede expresar en la forma propuesta), se buscaría minimizar la siguiente función:

$$h(x) = f(x) + \sum_{i=1}^n \alpha_i \cdot \Phi_i(g_i(x)) \quad [4-2]$$

donde los  $\alpha_i$  son coeficientes que le dan una importancia relativa a cada función de restricción  $g_i$ , y los  $\Phi_i$  son funciones de penalización (por ejemplo,  $\Phi_i(g_i(x)) = g_i(x)^2$ ).

Por consiguiente, al avanzar las generaciones los individuos tenderán a localizarse dentro de las restricciones puesto que serán estos los que obtendrán una mayor probabilidad de reproducción. Alternativamente, si un individuo no cumple con las restricciones pero tiene un valor muy bajo de función de costo es probable que perpetúe las características que lo hacen tener tal valor tendiendo a la vez a localizarse dentro de las restricciones (puesto que es más probable que se reproduzca con un individuo que sí cumple con ellas).

## 5. ADAPTACIÓN DE LOS ALGORITMOS GENÉTICOS AL DISEÑO DE REDES CERRADAS

---

### 5.1. Objetivo del algoritmo

El objetivo principal del algoritmo a implantar es, partiendo de las características topológicas<sup>2</sup>, topográficas y ciertas características físicas (materiales, ubicación de accesorios/coeficientes de pérdidas menores) e hidráulicas (caudales demandados y presiones mínimas en los nodos) de una red de acueducto determinada, encontrar la combinación de diámetros de los tubos de la red que (1) lleve el caudal demandado especificado a todos los nodos de la red (2) con no menos de la presión mínima especificada a priori y (3) que sea de mínimo costo según los criterios de la sección 5.4.

La condición (1) se cumple siempre que el caudal de entrada a la red sea mayor que la suma de los caudales demandados, afirmación que se considera cierta en general (ya que el caudal de entrada proviene de uno o varios embalses cuya capacidad se modela como infinita). La condición (2) depende de la solución de la ecuación [2-16] (ecuaciones de conservación de masa y energía resueltas a través del método del gradiente) para cada combinación de diámetros determinada, y la condición (3) depende directamente de los diámetros que se usen y su posición en la red (no es función de las características hidráulicas de la red).

### 5.2. Espacio de búsqueda e individuos

El espacio de búsqueda es el conjunto que contiene todos los posibles diseños que se pueden escoger para la red especificada, sin importar que estos cumplan con las restricciones hidráulicas o no (las restricciones limitan el espacio entero a un subconjunto de soluciones factibles). Es por tanto fácil ver que este espacio es  $D^U$  (el conjunto de  $U$ -

vectores sobre un alfabeto  $D$ ), donde, como se especificó previamente en la sección 3,  $D$  es el conjunto de diámetros comerciales que se pueden escoger para el diseño y  $U$  es el número de tubos que tiene la red. Evidentemente, el tamaño de este espacio es  $|D|^U$ .

Cada individuo  $X$  en  $D^U$  es un vector que contiene  $U$  diámetros comerciales, cada posición del vector  $X$  correspondiente al diámetro de un tubo de la red. Cada posición, independientemente del individuo, corresponde a un tubo fijo ya que no sería útil en el contexto de los algoritmos genéticos que la  $i$ -ésima posición de un vector  $X$  correspondiera a un tubo diferente que la  $i$ -ésima de otro vector  $Y$ . La necesidad de mantener fijas las posiciones respecto a los tubos proviene de la noción de “esquema” responsable (teóricamente) del buen funcionamiento de los algoritmos genéticos<sup>3</sup>.

El genotipo de cada individuo es, por supuesto, él mismo visto como vector. Por su parte, el fenotipo de cada individuo es bastante más interesante. Como cada individuo representa un diseño, el fenotipo corresponde a la hidráulica del individuo; es decir, a cada individuo  $X$ , vector de diámetros, corresponde las presiones resultantes en los  $N$  nodos de la red, siendo así el fenotipo un vector en  $R^N$ . Este es un caso bastante interesante de fenotipo, donde para encontrarlo se debe procesar la información del genotipo en un paquete de software especializado (el programa REDES) a través de métodos matemáticos con cierto nivel de complejidad (el método del gradiente).

Los caudales en los  $U$  tubos de la red pueden también ser vistos como parte del fenotipo. Sin embargo, estos caudales no serán de interés, al menos en un caso inicial, mientras que las presiones sí serán importantes dentro del algoritmo (ver sección 5.5).

---

<sup>2</sup> Se refiere a la localización espacial de embalses, nodos y tubos.

<sup>3</sup> Para ver más sobre los fundamentos teóricos de los algoritmos genéticos, referirse a [Goldberg, 1989], cap. 2, o a [Michell, 1999], cap. 4.

### 5.3. Funciones de reproducción y mutación

Se implementará el método de reproducción descrito en la sección 4.1: crossover (recombinación) clásico. Para realizar la recombinación se escogerán dos individuos padres de la generación  $k$  según el método descrito en la sección 5.6 y se escogerá una posición  $i$  entre 0 y  $U$  (inclusive). Esta posición indica el sitio de recombinación: si es 0 el individuo nuevo será exactamente igual al padre 2, si es  $U$  será exactamente igual al padre 1 y para cualquier otro sitio  $i$  los primeros  $i$  caracteres serán copia del padre 1 y los siguientes  $U-i$  caracteres serán copia del padre 2.

Además, se implementará una función de reproducción que tenga en cuenta la posición física de cada uno de los tubos (coordenadas  $(x, y)$  en la topografía de la red) para realizar la recombinación. En este método, que se denominará recombinación hiperplanar, se seleccionará un punto al azar entre los extremos de la red y una dirección (entre  $0^\circ$  y  $360^\circ$ ), al azar también, para determinar un plano que corte la red (física) en dos (en vez de un sitio que corte el genotipo en dos). Es probable que este método, aunque más complicado, aumente sustancialmente la eficacia del algoritmo puesto que conserva los distintos esquemas (bloques de caracteres útiles dentro del genotipo de cada individuo por mejorar la aptitud de su fenotipo) que se pueden presentar en la red.

Durante la reproducción de dos individuos operará la función de mutación. Esta función afectará cada carácter que se esté copiando en el individuo descendiente con una probabilidad fija  $p_m$  de cambiar el carácter (diámetro) proveniente del padre por un carácter escogido aleatoriamente del conjunto  $D$ . Para simplificar el funcionamiento del operador, y como la descripción previa del operador de mutación muestra que es una variable binomial (con probabilidad de éxito  $p_m$  y de falla  $1-p_m$ ), se evaluará previo a comenzar la reproducción la probabilidad que  $n$  de los  $U$  caracteres del descendiente sean mutados en el proceso:

$$P(\text{mutados} = n) = \binom{U}{n} \cdot (p_m)^n \cdot (1-p_m)^{U-n} \quad \text{[5-1]}$$

Y se seleccionarán posteriormente estos  $n$  caracteres a mutar al azar sobre los  $U$  caracteres posibles.

#### 5.4. Función de costo (sin manejo de restricciones)

Siendo el objetivo del algoritmo encontrar un diseño que minimice el costo de construcción de la red, la función de costo será evidentemente el costo constructivo de cada diseño. Se considerará, inicialmente, el costo constructivo de las tuberías de la red sin tener en cuenta el costo de accesorios y otros costos propios a la construcción de un acueducto.

El costo constructivo de las tuberías se adapta de la función encontrada por [Gutiérrez, 2002]. Esta función se basa en un Análisis de Precio Unitario que tiene en cuenta el costo de la tubería (física), el material de soporte y relleno de la tubería dentro de una zanja de tamaño fijo predeterminado (1 m de profundidad y 80 cm de longitud) y el costo de la mano de obra. El APU está dado por unidad de longitud de la tubería, y se resume en la siguiente función de costo de tubería ajustada por regresión lineal:

$$C_{tubería} = \$734,450 \cdot L_{tubería} \cdot (D_{tubería})^{1.0138} \quad [5-2]$$

donde  $L_{tubería}$  y  $D_{tubería}$  (longitud y diámetro de la tubería, respectivamente) están dados en metros. El costo resultante está dado en pesos (colombianos) del 2002. La regresión lineal de la cual se dedujo esta ecuación fue realizada con valores de diámetro entre 4 y 12 pulgadas. La función de costo (constructivo) de un individuo  $X$  es por tanto:

$$C_c(X) = \$734,450 \cdot \sum_{i=1}^U L_i \cdot \pi_i(X)^{1.0138} \quad [5-3]$$

siendo  $\pi_i$  la función  $i$ -ésima proyección; dado que  $X$  es un vector con los  $U$  diámetros de la red,  $\pi_i(X)$  viene a ser el diámetro del  $i$ -ésimo tubo de la red, cuya longitud (fija para el diseño) es  $L_i$ .

Aunque esta función no tiene en cuenta todos los factores que generan costos en la construcción de una red de distribución de agua potable, está limitada en los diámetros sobre los cuales es válida y al estar ligada a precios del 2002 tiene cada vez menos vigencia, a través de ella se obtiene un criterio objetivo para comparar el costo de los distintos individuos (diseños) generados por el algoritmo genético. Esta ecuación se constituye más en un medio relativo de comparación que en un indicador absoluto del costo de la red, y mientras los precios en los costos de las tuberías de distintos diámetros se mantengan similares en sus magnitudes el individuo que obtenga el menor costo según esta función será el individuo de menor costo en la práctica (sin tener en cuenta los costos de accesorios y otros), aunque ambos costos difieran.

### 5.5. Manejo de las restricciones del problema

Como ya se ha especificado previamente, cualquier diseño debe cumplir con las condiciones hidráulicas de llevar el caudal demandado a cada uno de los nodos con la presión mínima especificada por el usuario. Haciendo referencia a la sección 5.2, se denotará el fenotipo de un individuo  $X_i^k$ , como  $F_i^k$ . Este vector  $F_i^k$  es un elemento de  $R^N$ , puesto que está compuesto de las  $N$  presiones de la red en cada uno de los  $N$  nodos, valores que pueden tomar cualquier valor (teóricamente, al menos) sobre los reales.

Utilizando esta notación, las  $N$  restricciones vendrían a ser:

$$\begin{aligned} \pi_1(F_i^k) &\geq p_1 \\ \pi_2(F_i^k) &\geq p_2 \\ &\vdots \\ \pi_N(F_i^k) &\geq p_N \end{aligned} \quad [5-4]$$

donde  $p_j$  es la presión mínima para el nodo  $j$ , y  $\pi_j$  es la  $j$ -ésima proyección del vector  $F_i^k$ , es decir la presión del  $j$ -ésimo nodo para el individuo (diseño)  $i$  de la generación  $k$ . Para simplificar un poco la notación, y hacerla uniforme con la de la ecuación [4-1], se definen las funciones  $g_i : D^U \rightarrow R$  así<sup>4</sup>:

$$g_i(X) = p_i - \pi_i(F(X)) \quad [5-5]$$

$F(X)$  denota el fenotipo del individuo  $X$  (de hecho, esta función  $F$  corresponde a evaluar el diseño codificado por el individuo  $X$  en REDES para obtener el correspondiente vector de presiones  $F(X)$ ). A partir de esta definición, nuestras restricciones de la ecuación [5-4] toman exactamente la misma forma de la ecuación [4-1]:

$$\begin{aligned} g_1(X) &= p_1 - \pi_1(F(X)) \leq 0 \\ g_2(X) &= p_2 - \pi_2(F(X)) \leq 0 \\ &\vdots \\ g_N(X) &= p_N - \pi_N(F(X)) \leq 0 \end{aligned} \quad [5-6]$$

Para el algoritmo se utilizará un valor fijo de  $p_i$ ,  $i=1, \dots, N$ , que podrá ser definido por el usuario.

Finalmente, haciendo referencia a la ecuación [4-2], solo bastaría definir las constantes  $\alpha_i$  y las funciones  $\Phi_i$  para obtener nuestra función objetivo. En una primera implantación, se utilizará:

$$\Phi_i(c) = \begin{cases} c, & \text{si } c > 0 \\ 0, & \text{si } c \leq 0 \end{cases}$$

<sup>4</sup> Esta notación indica que cada una de las funciones  $g_i$  toman un parámetro en  $D^U$  y devuelven un valor en  $R$ .

Se tomarán las constantes  $\alpha_i = 1$ , con lo que se dará la misma importancia a la restricción de presión mínima en cada nodo de la red. Se define el costo hidráulico, o la violación de las restricciones [5-6], como:

$$C_h(X) = \sum_{i=1}^N \Phi(g_i(X)) \quad [5-7]$$

Por lo tanto, la función objetivo (función de adaptabilidad)  $h(X)$  será descrita por

$$h(X_i^k) = \begin{cases} \frac{C_c(X_i^k)}{\mu[C_c^k]} + \omega_h \cdot \frac{C_h(X_i^k)}{\mu[C_h^k]}, & \text{si } \mu[C_h^k] \neq 0 \\ \frac{C_c(X_i^k)}{\mu[C_c^k]}, & \text{si } \mu[C_h^k] = 0 \end{cases} \quad [5-8]$$

donde  $\mu[C_c^k]$  denota el promedio del costo constructivo para la generación  $k$  y  $\mu[C_h^k]$  el promedio del costo hidráulico. La constante  $\omega_h$  denota el peso relativo que se le da al costo hidráulico sobre el costo constructivo (que, por supuesto, puede ser menor que 1 haciendo que influya más el costo constructivo sobre el algoritmo restándole importancia al cumplimiento de las restricciones).

Como el costo constructivo de cualquier diseño siempre va a ser un valor positivo, el término  $\mu[C_c^k]$  será siempre distinto de cero por lo que la expresión

$$\frac{C_c(X_i^k)}{\mu[C_c^k]}$$

estará bien definida. Y, en el caso que  $\mu[C_h^k]$  sea cero para una generación (todos los diseños estén cumpliendo con las restricciones)  $h(X)$  es el segundo caso de la ecuación [5-8], por lo que  $h$  estará siempre bien definida.



De la ecuación de  $h(X)$  se puede ver que un individuo promedio recibirá una puntuación de  $1 + \alpha_c$ , 1 por situarse en el promedio del costo constructivo y  $\alpha_h$  por situarse en el promedio del costo hidráulico. Adicionalmente, si un individuo se encuentra totalmente adentro de las restricciones su valor de costo hidráulico será 0 y la contribución de la parte hidráulica a su función de adaptabilidad será nula. Es fácil ver además que la contribución máxima de las partes hidráulica ó constructiva a la función de adaptabilidad es  $n$  (el número de individuos por generación).

Finalmente, vale la pena remarcar que aunque para comparar los individuos dentro de cada generación se recurra a la función de adaptabilidad, los valores tanto de costo constructivo como de costo hidráulico siguen siendo la base para la comparación de los mejores individuos de cada generación y, por tanto, para la determinación del mejor individuo encontrado por el algoritmo. Para esto, se almacenarán cada iteración  $k$  del algoritmo los mejores individuos de la generación  $k$  en una “generación” que contiene a los mejores individuos hasta la generación  $k-1$  y se compararán dentro de esta generación con los individuos ya almacenados calculando de nuevo su adaptabilidad utilizando la función [5-8] (con los promedios de costo constructivo e hidráulico de esta “generación” de mejores individuos).

## 5.6. Método de reproducción

Se comenzará utilizando un modelo donde la probabilidad de reproducción asignada a cada individuo es inversamente proporcional (puesto que se trata de un problema de minimización de costos – ver sección 5.4) al valor de su función de costo.

Para un individuo dado  $X_i^k$  (el superíndice se refiere a la generación o iteración a la que pertenece el vector) en  $D^U$ , su probabilidad de ser escogido para reproducirse será entonces:

$$P_k(X_i^k) = \frac{\frac{1}{h(X_i^k)}}{\sum_{j=1}^n \frac{1}{h(X_j^k)}} \quad [5-9]$$

donde  $h(X)$  es la función de costo y  $X_1^k, X_2^k, \dots, X_n^k$  son los  $n$  individuos de la  $k$ -ésima generación. El denominador asegura que la suma de las probabilidades de reproducción para cada generación sea 1.

### 5.7. Operadores genéticos adicionales – Reacondicionamiento de la probabilidad de reproducción

Es común que en las primeras generaciones de un algoritmo genético haya individuos cuyo valor de función de costo ( $h(X)$ , sección 5.5) sea relativamente mucho menor al de los demás y que tengan por consiguiente una probabilidad de reproducción mucho mayor. Esto puede resultar en una disminución considerable de la variabilidad genética de la población en las primeras generaciones y, por tanto, de la sección del espacio de búsqueda que abarca cada generación. En muchas ocasiones esto puede ser dañino al resultado del algoritmo si los primeros individuos dominantes llevan a la población a localizarse alrededor de mínimos locales, impidiendo al algoritmo aproximarse al mínimo global.

Para evitar esto, se han desarrollado operadores de reacondicionamiento de la probabilidad de reproducción. El más simple de ellos, el operador lineal (v. [Goldberg, 1989], p. 76), funciona transformando la probabilidad de reproducción  $P_k(X)$  en  $P_k'(X)$  como sigue:

$$P_k'(X) = P_k(X) \cdot a^k + b^k \quad [5-10]$$

donde  $a^k$  y  $b^k$  son constantes para cada generación  $k$ . Estas constantes son calculadas de tal manera que (1) mantengan el promedio de  $P_k'(X)$  igual al de  $P_k(X)$  (igual a  $1/n$ , de la

ecuación [5-9]) para que un individuo promedio siga contribuyendo 1 individuo a la siguiente generación; y que (2) el valor de

$$P_{MAX}^k = \underset{X \in k\text{-ésima Generación}}{\text{Max}} (P_k(X)) \quad [5-11]$$

sea un múltiplo predefinido por una constante  $C^k$  (generalmente entre 1.2 y 2) del promedio de  $P_k(X)$ . Así, se asegura que el individuo dominante de las generaciones tempranas no tenga (probabilísticamente) más de  $C^k$  descendientes en la siguiente generación. Las restricciones (1) y (2) se pueden expresar como

$$\begin{aligned} \frac{1}{n} \cdot \sum_{i=1}^n P_k(X_i^k) &= \frac{1}{n} \cdot \sum_{i=1}^n P_k'(X_i^k) \\ &= \frac{1}{n} \cdot \sum_{i=1}^n [P_k(X_i^k) \cdot a_k + b_k] \\ \sum_{i=1}^n P_k(X_i^k) &= a_k \cdot \left[ \sum_{i=1}^n P_k(X_i^k) \right] + n \cdot b_k \end{aligned} \quad [5-12]$$

$$P_{MAX}^k \cdot a_k + b_k = C^k \cdot \frac{1}{n} \cdot \sum_{i=1}^n P_k(X_i^k) \quad [5-13]$$

La solución al sistema de ecuaciones aquí planteado está dado por:

$$\begin{aligned} a^k &= \frac{(1 - C^k) \cdot \left( \frac{1}{n} \cdot \sum_{i=1}^n P_k(X_i^k) \right)}{\frac{1}{n} \cdot \sum_{i=1}^n P_k(X_i^k) - P_{MAX}^k} \\ b^k &= C^k \cdot \frac{1}{n} \cdot \sum_{i=1}^n P_k(X_i^k) - a^k \cdot P_{MAX}^k \end{aligned} \quad [5-14]$$

Para el algoritmo, se utilizará un  $C^k$  de 2 inicialmente.

### **5.8. Operadores genéticos adicionales – Reducción del error estocástico de la “Rueda de Ruleta”**

En el esquema clásico de selección de individuos para el apareamiento, donde la probabilidad de ser escogido de cada individuo para la siguiente generación es proporcional al inverso de su función de costo, hay errores estocásticos inherentes. En este método, conocido popularmente como la Rueda de Ruleta, el número real de descendientes de un individuo va a variar considerablemente y no va a ser igual al número esperado de éstos.

Para reducir este error, [De Jong] considera un sistema de reproducción donde se lleva un registro que tiene el valor del número esperado de descendientes de un individuo  $X$ ,  $n * P_k(X)$ , menos 0.5 por el número de veces que este individuo ya haya sido escogido para reproducirse. En el momento que el valor de este registro sea negativo el individuo no podrá volver a ser escogido, lo que asegura que el número de descendientes de un individuo  $X$  no sobrepase  $n * P_k(X) + 0.5$ .

Ha sido ya comprobado varias veces (v. [Goldberg, 1989] p. 115 y [De Jong]) que este operador de reproducción incrementa considerablemente el desempeño del algoritmo.

## 6. BIBLIOGRAFÍA

---

- [De Jong] De Jong, K. A. “An analysis of the behavior of a class of genetic adaptive systems. (Doctoral dissertation, University of Michigan)”. Dissertation Abstracts International 36 (10), 5140B (University Microfilms No. 76-9381).
- [Goldberg, 1989] Goldberg, David, (1989, ed. rev.). “Genetic algorithms in search, optimization and machine learning”. Reading, Mass.: Addison-Wesley Pub Co.
- [Gutiérrez, 2002] Gutiérrez, Gabriel, (2002). “Diseño óptimo de redes bajo ambiente de fugas”. Tesis de grado no publicada. Departamento de Ingeniería Civil y Ambiental, Universidad de los Andes, Bogotá, Colombia.
- [Michell, 1999] Michell, Melanie, (1999). “An introduction to genetic algorithms (Complex Adaptive Systems)”. London: MIT Press.
- [Purves et al., 1998] Purves, W., Orians, G., Heller, C., Sadava, D., (1998, 4ta ed.). “Life, the science of biology”. E.E.U.U.: W.H. Freeman and Co.
- [Saldarriaga, 2001] Saldarriaga, Juan, (2001). “Hidráulica de Tuberías”. Bogotá, Colombia: McGraw-Hill Interamericana S.A.
- [Streeter et al., 1998] Streeter, V., Wylie, B., Bedford, K., (1998, 9na ed.). “Fluid Mechanics”. E.E.U.U.: WCB/McGraw-Hill.

## PARTE 2: DOCUMENTACIÓN DEL PROGRAMA

### 1. REQUERIMIENTOS GENERALES

---

#### 1.1. Generalidades

El programa se realizará en el lenguaje Object Pascal, sobre la plataforma Borland Delphi 6. Se escogió esta plataforma puesto que en ella se desarrolló el programa REDES, necesario para el funcionamiento del algoritmo genético, como se especifica en la sección **¡Error! No se encuentra el origen de la referencia.** del Marco Teórico. Más aún, el programa de Diseño Óptimo Genético se desarrollará como un módulo del programa REDES. Funcionará desde la interfaz de este abriendo el programa como una opción dentro del menú “Calcular”.

El desarrollo del programa se realizará de acuerdo a lo especificado en esta Documentación, siguiendo los parámetros técnicos del Marco Teórico.

#### 1.2. Curso Normal de Acciones

Antes de escoger la opción de diseñar, el usuario debe abrir un archivo en el programa REDES. Al escoger la opción de diseñar, una forma nueva se abre. En esta forma el usuario digita los diámetros sobre los cuales se puede diseñar y la presión de nodo mínima. En otra pestaña, dentro de la misma forma, el usuario escoge las opciones concernientes al funcionamiento del algoritmo genético: tamaño de la población, número de generaciones, peso relativo del déficit de presión, probabilidad de mutación  $p_m$ , método de recombinación, método de asignación de probabilidad y constante de reproducción  $C$ .

El usuario presiona luego sobre el botón de diseñar. El algoritmo corre y, después del número predeterminado de iteraciones, muestra al usuario en otra pestaña los resultados. Se debe mostrar el valor de costo constructivo e hidráulico y el número de nodos que incumplen la restricción de presión para cada uno de los 10 mejores diseños según lo descrito en el numeral 5.5 del Marco Teórico. El usuario tiene la opción de salvar cualquiera de estos 10 diseños como archivos de REDES y las estadísticas resultantes del algoritmo en un archivo de texto.

### **1.3. Casos de Excepción**

1. El archivo de REDES no es válido: el programa REDES detecta esto automáticamente.
2. Los tipos de datos ingresados por el usuario (diámetros, presión mínima, tamaño de la población, número de generaciones, peso relativo del déficit de presión, probabilidad de mutación  $p_m$ , y constante de reproducción  $C$ ) no son correctos: un cuadro de diálogo debe avisar al usuario en el momento de correr el algoritmo.
3. Los datos de diámetros, presión mínima, tamaño de la población, peso relativo del déficit de presión, probabilidad de mutación o constante inicial de reproducción son negativos: se debe avisar al usuario en un cuadro de diálogo que debe ingresar valores positivos.
4. La probabilidad de mutación ingresada por el usuario no esta entre 0 y 1 (incluso): un cuadro de diálogo avisará al usuario en el momento que sea ingresada.
5. Falta algún dato al presionar el botón de correr el algoritmo: un cuadro de diálogo debe avisar que falta alguno de los datos para correr el algoritmo (sin especificar cuál).

#### **1.4. Precondiciones**

Solo se podrá diseñar sobre un número finito y discreto de diámetros. Es decir, solo se contempla el caso de diseño de redes sobre diámetros comerciales; no se implementa el caso de diseño sobre diámetros no comerciales.

El tamaño mínimo de población para el algoritmo es de 2 individuos (necesarios para el cálculo de la desviación estándar no sesgada, según la sección 7.3).

#### **1.5. Poscondiciones**

Dada la naturaleza de los algoritmos genéticos, no se garantiza que los 10 mejores diseños que se encuentren sean, en orden descendente, los 10 mejores diseños posibles. Tampoco se garantiza que el mejor diseño sea el óptimo global. Sin embargo, se espera que el mejor diseño se acerque considerablemente a éste óptimo.



## 2. FUNCIONES

---

### 2.1. Funciones Básicas

Ref. #	Función
R1.1	Recibir la solicitud de diseño del usuario y revisar que la red esté correctamente ingresada a REDES.
R1.2	Recibir los datos de entrada hidráulicos (diámetros sobre los que se puede diseñar y presión de nodo mínima).
R1.3	Rellenar los datos de entrada hidráulicos con los parámetros por default.
R1.4	Recibir los datos de entrada concernientes al funcionamiento del algoritmo genético (tamaño de la población, número de generaciones, peso relativo del déficit de presión, probabilidad de mutación $p_m$ , método de recombinación, método de asignación de probabilidad y constante de reproducción $C$ ).
R1.5	Rellenar los datos de entrada concernientes al funcionamiento del algoritmo genético con los parámetros por default.

### 2.2. Funciones de Costo

Ref. #	Función
R2.1	Cálculo de los fenotipos de cada individuo.
R2.2	Cálculo del costo constructivo de cada individuo.
R2.3	Cálculo del costo hidráulico de cada individuo.
R2.4	Número de nodos donde cada individuo incumple las restricciones.
R2.5	Cálculo de la adaptabilidad de cada individuo (usando R4.2 y R4.3).

### 2.3. Funciones de Reproducción

Ref. #	Función
R3.1	Constitución de la primera generación.
R3.2	Probabilidad bruta de reproducción de cada individuo.
R3.3	Probabilidad escalada de reproducción de cada individuo.
R3.4	Selección de individuos padres para un nuevo individuo.
R3.5	Determinación de procedencia de un alelo de padre 1 o padre 2.
R3.6	Constitución de la siguiente generación.

### 2.4. Funciones Estadísticas

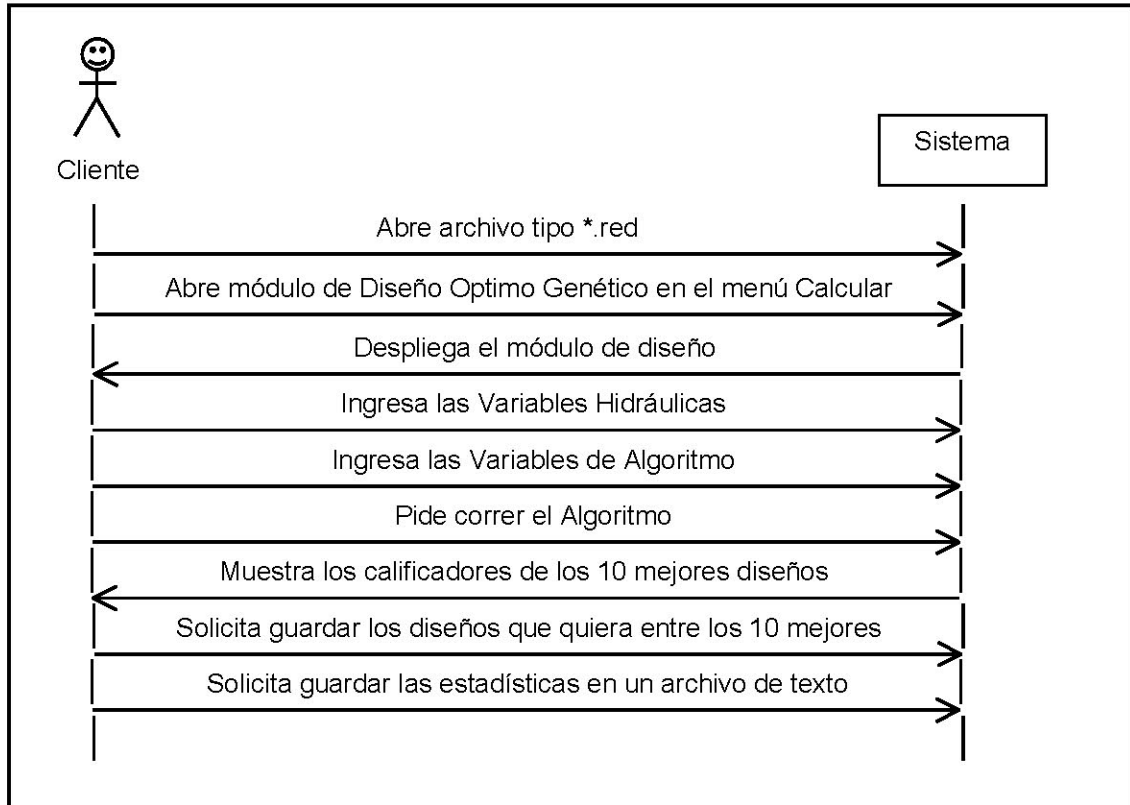
Ref. #	Función
R4.1	Almacenamiento de los 10 mejores individuos producidos por el algoritmo (según el numeral 5.4 del Marco Teórico).
R4.2	Promedio y desviación estándar del costo constructivo de la generación.
R4.3	Promedio y desviación estándar del costo hidráulico de la generación.
R4.4	Promedio y desviación estándar de la adaptabilidad de la generación.
R4.5	Número promedio de nodos donde no se cumplen las restricciones de la generación.
R4.6	Número de individuos que no cumplen las restricciones de la generación.

## 2.5. Funciones de Salida

Ref. #	Función
R5.1	Obtención del $n$ -ésimo individuo ( $n$ entre 1 y 10) de menor costo de los individuos almacenados en R4.1, con su costo constructivo, costo hidráulico y respectivo número de generación.
R5.2	Exhibición de los 10 mejores individuos en pantalla.
R5.3	Posibilidad de almacenamiento de los individuos de menor costo constructivo y menor costo hidráulico.
R5.4	Almacenamiento en texto de los datos estadísticos (R4.1 – R4.6) por generación.
R5.5	Almacenamiento del tiempo de cálculo por generación (opcional).
R5.6	Tiempo de cálculo total (opcional).

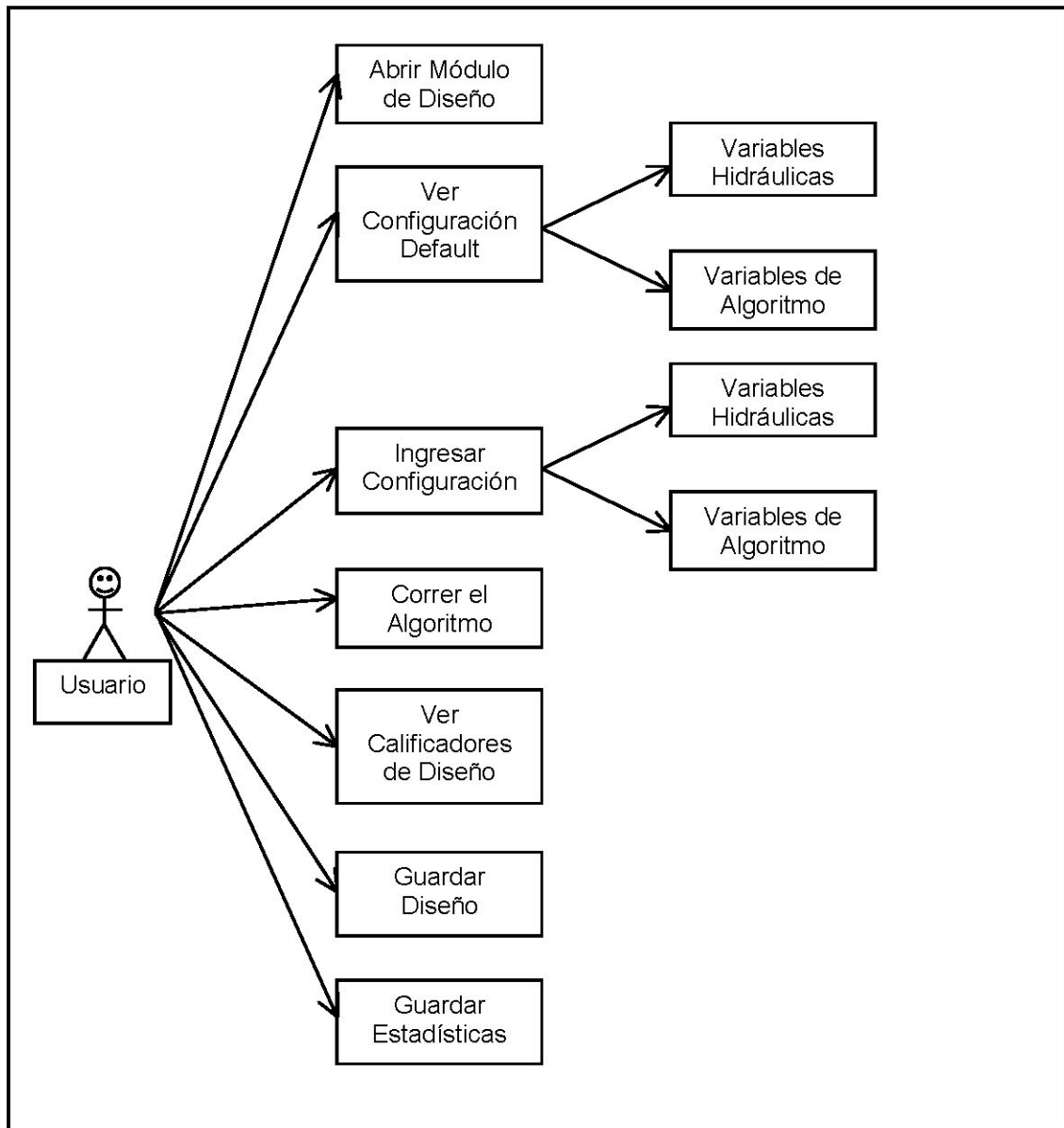
### 3. DIAGRAMA DE SECUENCIA DEL SISTEMA

#### 3.1. Uso típico del programa



## 4. CASOS DE USO

### 4.1. Casos de uso



### 4.2. Descripción

- Abrir Módulo de Diseño: con un archivo de redes abierto, el usuario escoge la opción de Diseño Genético con lo cual se abre la forma de Diseño.

- Ver Configuración Default/VARIABLES Hidráulicas: muestra los diámetros sobre los que se va a diseñar por default, y la presión de nodo mínima default.
- Ver Configuración Default/VARIABLES de Algoritmo: muestra el valor por default del tamaño de población, el número de generaciones, el peso relativo del déficit de presión, la probabilidad de mutación y la constante de reproducción, y activa las opciones de recombinación hiperplanar y asignación de descendencia escalada.
- Ingresar Configuración/VARIABLES Hidráulicas: el usuario ingresa los diámetros sobre los que él quiere que se diseñe y la presión de nodo mínima.
- Ingresar Configuración/VARIABLES de Algoritmo: el usuario ingresa el tamaño de la población, el número de generaciones, el peso relativo del déficit de presión, la probabilidad de mutación, la constante de reproducción y selecciona el tipo de recombinación y asignación de descendencia que quiera usar.
- Correr el Algoritmo: el sistema corre el algoritmo genético durante el número de generaciones especificadas por el usuario, según los parámetros ingresados (por default o por el usuario).
- Ver Calificadores de Diseño: muestra el costo constructivo, costo hidráulico y número de generación de los 10 mejores diseños escogidos entre los mejores diseños de cada generación.
- Guardar Diseño: guarda, entre los 10 mejores diseños, el diseño escogido por el usuario en un archivo de REDES (\*.red).
- Guardar Estadísticas: guarda las estadísticas descritas en el numeral 2.4, todas en un mismo archivo texto

## 5. REQUERIMIENTOS

---

### 5.1. Abrir módulo de diseño

<b>Requerimiento:</b>	1
<b>Nombre:</b>	Abrir Módulo de Diseño
<b>Descripción:</b> Sobre la plataforma principal del programa REDES, abre la forma de Diseño Genético	
<b>Casos de excepción:</b> 1. Ausencia de red a diseñar 2. Red con errores	
<b>Prueba:</b> 1. El usuario no abre ningún archivo. En el menú "Calcular", escoge la opción "Diseño Genético".  2. El usuario abre un archivo de redes con errores.  3. El usuario abre el archivo "Cazuca.red". En el menú "Calcular", debe escoger la opción "Diseño Genético".	
<b>Resultado:</b> 1. Un cuadro de diálogo muestra el mensaje "No hay ninguna red presente para diseño". No se abre la forma (Tform) "Diseño Genético".  2. El programa REDES no permite abrir este archivo.  3. La forma (Tform) "Diseño Genético" se abre con los valores por default de todos los campos y las opciones por default seleccionadas (ver requerimientos 2 y 3 para los valores de estas opciones).	



## **5.2. Ingresar configuración/Variables Hidráulicas**

<b>Requerimiento:</b>	2
<b>Nombre:</b>	Ingresar Configuración/Variables de Usuario
<b>Descripción:</b>	
En la pestaña "Variables Hidráulicas y de Usuario", en la forma "Diseño Genético", el usuario puede llenar los campos correspondientes a diámetros con los cuales se puede diseñar y presión de nodo mínima para toda la red. Puede decidir si activar o no la opción de Guardar Estadísticas	
<b>Casos de excepción:</b>	
1. Datos de tipo erroneo 2. Valores de diámetros o presión negativos 3. Valores de diametro igual a cero	
<b>Prueba:</b>	
Una vez realizada la prueba 3 del requerimiento 1, el usuario selecciona la pestaña "Variables Hidráulicas y de Usuario". 1. Se ingresan strings no correspondientes a enteros al campo de presión de nodo mínima. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético". 2. Se ingresan strings no correspondientes a enteros en varias posiciones del TstringGrid de "Diametros de Diseño". El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético". 3. Se ingresan valores negativos aleatorios al campo de presión de nodo mínima. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético". 4. Se ingresan valores negativos aleatorios en varias posiciones del TstringGrid de "Diametros de Diseño". El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético". 5. Se ingresan valores iguales a cero (0) en varias posiciones del TstringGrid de "Diametros de Diseño". El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético". 6. Se ingresan valores enteros positivos (correctos) en todos los campos. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético".	
<b>Resultado:</b>	
1. Un cuadro de diálogo muestra el mensaje "Tipo de datos erroneo". El campo de presión de nodo mínima queda con la presión por default. Se selecciona de nuevo la pestaña "Variables Hidráulicas y de Usuario". 2. Un cuadro de diálogo muestra el mensaje "Tipo de datos erroneo". Los campos de diámetro con valores no enteros quedan con el valor "0". Se selecciona de nuevo la pestaña "Variables Hidráulicas y de Usuario". 3. Un cuadro de diálogo muestra el mensaje "Tipo de datos erroneo". El campo de presión de nodo mínima queda con la presión por default. Se selecciona de nuevo la pestaña "Variables Hidráulicas y de Usuario". 4. Un cuadro de diálogo muestra el mensaje "Tipo de datos erroneo". Los campos de diámetro con valores negativos quedan con el valor "0". Se selecciona de	

nuevo la pestaña "Variables Hidráulicas y de Usuario".

5. Un cuadro de diálogo muestra el mensaje "Diámetros no pueden ser 0". Se selecciona de nuevo la pestaña "Variables Hidráulicas y de Usuario".

6. Se selecciona la pestaña que el usuario haya seleccionado.

### **5.3. Ingresar configuración/Variables de Algoritmo**

Este requerimiento toma dos hojas.

<b>Requerimiento:</b>	3
<b>Nombre:</b>	Ingresar Configuración/Variables de Algoritmo
<b>Descripción:</b>	
En la pestaña "Variables de Algoritmo", en la forma "Diseño Genético", el usuario puede llenar los campos correspondientes a tamaño de la población, número de generaciones, probabilidad de mutación. Puede activar la opción de Método de Crossover "Crossover Clásico" ó "Crossover Hiperplanar" y la opción de Asignación de descendencia "Sin Escalar" ó "Escalada" (con lo que se activa la constante inicial de reproducción).	
<b>Casos de excepción:</b>	
<ol style="list-style-type: none"><li>1. Datos de tipo erroneo</li><li>2. Valores de tamaño de la población, número de generaciones, probabilidad de mutación o constante inicial de reproducción negativos</li><li>3. Valores de tamaño de la población, número de generaciones, probabilidad de mutación o constante inicial de reproducción igual a cero</li><li>3. Valor de tamaño de población menor de 10.</li> <li>4. Valor de probabilidad de mutación mayor que 1.</li> <li>5. (Warning) Valor de probabilidad de mutación mayor que 0.1</li></ol>	
<b>Prueba:</b>	
Una vez realizada la prueba 3 del requerimiento 1, el usuario selecciona la pestaña "Variables de Algoritmo". <ol style="list-style-type: none"><li>1. Se ingresan strings no correspondientes a enteros a los campos de tamaño de la población, número de generaciones, probabilidad de mutación y constante inicial de reproducción (activando la opción "Escalada" de Asignación de Descendencia). El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético".</li><li>2. Se ingresan valores reales pero no enteros a los campos de tamaño de la población, número de generaciones y constante inicial de reproducción. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético".</li><li>3. Se ingresan enteros negativos aleatorios a los campos de tamaño de la población, número de generaciones y constante inicial de reproducción. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético".</li><li>4. Se ingresan valores de cero (0) a los campos de tamaño de la población, número de generaciones y constante inicial de reproducción. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético".</li><li>5. Se ingresan valores reales aleatorios menores que 0 y mayores que 1 al campo de probabilidad de mutación. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético".</li><li>6. Se ingresan valores enteros menores a 10 al campo de tamaño de la población. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético".</li></ol>	

7. Se ingresan valores aleatorios mayores a 0.1 y menores a 1 al campo de probabilidad de mutación. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético".

8. Se ingresan valores enteros positivos (correctos) en los campos de tamaño de la población, número de generaciones y constante inicial de reproducción, y valores reales positivos menores que 0.1 (correctos) en el campo de probabilidad de mutación. El usuario selecciona cualquier otra pestaña de la forma "Diseño Genético".

**Resultado:**

1. Un cuadro de diálogo muestra el mensaje "Tipo de datos erroneo. Por favor trate de nuevo". Los campos que tengan datos de tipo erroneo quedan con sus valores por default. Se selecciona de nuevo la pestaña "Variables de Algoritmo".

2. Un cuadro de diálogo muestra el mensaje "Tipo de datos erroneo. Por favor trate de nuevo". Los campos que tengan datos de tipo erroneo quedan con sus valores por default. Se selecciona de nuevo la pestaña "Variables de Algoritmo".

3. Un cuadro de diálogo muestra el mensaje "Debe especificar valores estrictamente positivos. Por favor trate de nuevo". Los campos con valores negativos quedan con su valor por default. Se selecciona de nuevo la pestaña "Variables de Algoritmo".

4. Un cuadro de diálogo muestra el mensaje "Debe especificar valores estrictamente positivos. Por favor trate de nuevo". Los campos con valores de 0 quedan con su valor por default. Se selecciona de nuevo la pestaña "Variables de Algoritmo".

5. Un cuadro de diálogo muestra el mensaje "Probabilidad de Mutación debe estar entre 0 y 1. Por favor trate de nuevo". El campo de Probabilidad de Mutación queda con su valor por default. Se selecciona de nuevo la pestaña "Variables de Algoritmo".

6. Un cuadro de diálogo muestra el mensaje "Población debe ser mayor a 10. Por favor trate de nuevo". El campo de Tamaño de Población queda con su valor por default. Se selecciona de nuevo la pestaña "Variables de Algoritmo".

7. Un cuadro de diálogo muestra el mensaje "Probabilidad de Mutación puede afectar el funcionamiento del algoritmo". Se selecciona la pestaña que el usuario haya seleccionado.

8. Se selecciona la pestaña que el usuario haya seleccionado.

#### **5.4. Ver configuración Default/Variables Hidráulicas**

<b>Requerimiento:</b>	4
<b>Nombre:</b>	Ver Configuración Default/Variables de Usuario
<b>Descripción:</b>	En la pestaña "Variables de Hidráulicas y de Usuario", en la forma "Diseño Genético", el sistema llena los campos correspondientes a diámetros con los cuales se puede diseñar y presión de nodo mínima para toda la red. También activa la opción de Guardar Estadísticas
<b>Casos de excepción:</b>	Ninguno
<b>Prueba:</b>	1. Una vez realizada la prueba 6 del requerimiento 2, el usuario oprime el botón de "Parámetros por Default" sobre la forma "Diseño Genético", pestaña "Variables Hidráulicas y de Usuario".
<b>Resultado:</b>	1. Se deben llenar: la tabla de "Diámetros de Diseño" (TstringGrid) con el vector de diámetros {3, 4, 6, 8, 10, 12, 16, 20, 24}. El campo de "Presión Mínima de Nodo" (Tedit) con 15 (metros). El check box (Tcheckbox) de "Grabar Rugosidades" se debe activar.



### **5.5. Ver configuración Default/Variables de Algoritmo**

<b>Requerimiento:</b>	5
<b>Nombre:</b>	Ver Configuración Default/Variables de Algoritmo
<b>Descripción:</b>	
En la pestaña "Variables de Algoritmo" de la forma "Diseño Genético" el sistema llena los campos correspondientes a tamaño de población, número de generaciones, probabilidad de mutación y constante de reproducción. Escoge además como Método de Reproducción la opción de "Crossover Hiperplanar" y como método de asignación de descendencia "Sin Escalar".	
<b>Casos de excepción:</b>	
Ninguno	
<b>Prueba:</b>	
1. Una vez realizada la prueba 8 del requerimiento 3, el usuario oprime el botón de "Parámetros por Default" sobre la forma "Diseño Genético", pestaña "Variables de Algoritmo".	
<b>Resultado:</b>	
1. Se deben llenar: el campo de tamaño de población (Tedit) con 50, el campo de número de generaciones (Tedit) con 500, el campo de probabilidad de mutación (Tedit) con 0.05. Se activará la opción (Tradiobutton) "Crossover Hiperplanar" y la opción (Tradiobutton) "Sin Escalar" dejando el campo de constante de reproducción C (Tedit) con un valor de 2 pero desactivado.	

## **5.6. Correr el Algoritmo**

<b>Requerimiento:</b>	6
<b>Nombre:</b>	Correr el Algoritmo
<b>Descripción:</b>	
El programa corre el algoritmo genético con las características especificadas (diámetros de diseño, presión de nodos mínima, número de generaciones, tamaño de población, probabilidad de mutación y constante inicial de reproducción). Además para cada generación almacena los datos de: individuo de mayor adaptabilidad, de menor costo constructivo y de menor costo hidráulico, número promedio de nodos donde se cumplen las restricciones, número de individuos que cumplen todas las restricciones, promedio y desviación estándar del costo constructivo e hidráulico de cada generación y desviación estándar de la adaptabilidad de cada generación.	
<b>Casos de excepción:</b>	
1. Falta alguna de las Variables Hidráulicas y de Usuario o Variables de Algoritmo.	
<b>Prueba:</b>	
1. Se realiza la prueba 3 del requerimiento 1 y luego se dejan vacíos alternativamente cada uno de los campos mencionados en los requerimientos 2 y 3, presionando cada vez el botón de "Diseñar" en la pestaña "Variables de Algoritmo".	
2. Una vez realizadas las pruebas 6 del requerimiento 2 y 8 del requerimiento 3 el usuario en la pestaña de "Variables de Algoritmo" presiona el botón "Diseñar".	
3. Una vez realizadas las pruebas 1 del requerimiento 4 y 1 del requerimiento 5 el usuario en la pestaña de "Variables de Algoritmo" presiona el botón "Diseñar".	
<b>Resultado:</b>	
1. Un cuadro de diálogo muestra el mensaje "Faltan datos para realizar el diseño". Se selecciona la pestaña "Variables de Algoritmo".	
2. El sistema corre el algoritmo y muestra la pestaña "Resultados".	
3. El sistema corre el algoritmo y muestra la pestaña "Resultados".	

## **5.7. Ver calificadores de Diseño**

<b>Requerimiento:</b>	7
<b>Nombre:</b>	Ver Calificadores de Diseño
<b>Descripción:</b>	Muestra, en la pestaña de "Resultados", después de haber cumplido el requerimiento 6 (correr el programa), los 10 mejores diseños ranqueados según la función $h_{max}$ mencionada en el numeral 5.5 del Marco Teórico. Para cada individuo muestra su número de generación, su costo constructivo y su costo hidráulico.
<b>Casos de excepción:</b>	Ninguno
<b>Prueba:</b>	1. Una vez realizada la prueba del requerimiento 6 el sistema localiza al usuario en la pestaña "Resultados".
<b>Resultado:</b>	1. Deben mostrarse los 10 mejores diseños con sus respectivos número de generación, costo constructivo y costo hidráulico.

## **5.8. Guardar Diseño**

<b>Requerimiento:</b>	8
<b>Nombre:</b>	Guardar Diseño
<b>Descripción:</b>	
Guarda, entre los 10 diseños mostrados después de haber corrido el programa, el diseño escogido por el usuario en un archivo de REDES (*.red). El usuario puede repetir este procedimiento tantas veces como quiera.	
<b>Casos de excepción:</b>	
1. El número de diseño a guardar no es un entero entre 1 y 10 (inclusive). 2. No se especificó un número de diseño a guardar.	
<b>Prueba:</b>	
1. Una vez realizada la prueba del requerimiento 7 se llena el campo (Tedit) "Diseño a Guardar" con un string que no sea un entero y se presiona el botón "Guardar Diseño". 2. Una vez realizada la prueba del requerimiento 7 se llena el campo "Diseño a Guardar" con un entero menor que 1 o mayor que 10 y se presiona el botón "Guardar Diseño". 3. Una vez realizada la prueba del requerimiento 7 se deja vacío el campo "Diseño a Guardar" y se presiona el botón "Guardar Diseño". 4. Una vez realizada la prueba del requerimiento 7 se llena el campo "Diseño a Guardar" con un entero mayor que 1 y menor que 10 (correcto) y se presiona el botón "Guardar Diseño".	
<b>Resultado:</b>	
1. Un cuadro de diálogo muestra el mensaje "Tipo de datos erróneo". El campo de "Diseño a Guardar" queda vacío. 2. Un cuadro de diálogo muestra el mensaje "Tipo de datos erróneo". El campo de "Diseño a Guardar" queda vacío. 3. Un cuadro de diálogo muestra el mensaje "Seleccione un número de diseño a guardar". El campo de "Diseño a Guardar" sigue vacío. 4. Se abre una ventana de "Guardar", donde el usuario puede seleccionar el nombre de archivo y carpeta para guardar el diseño que escogió.	



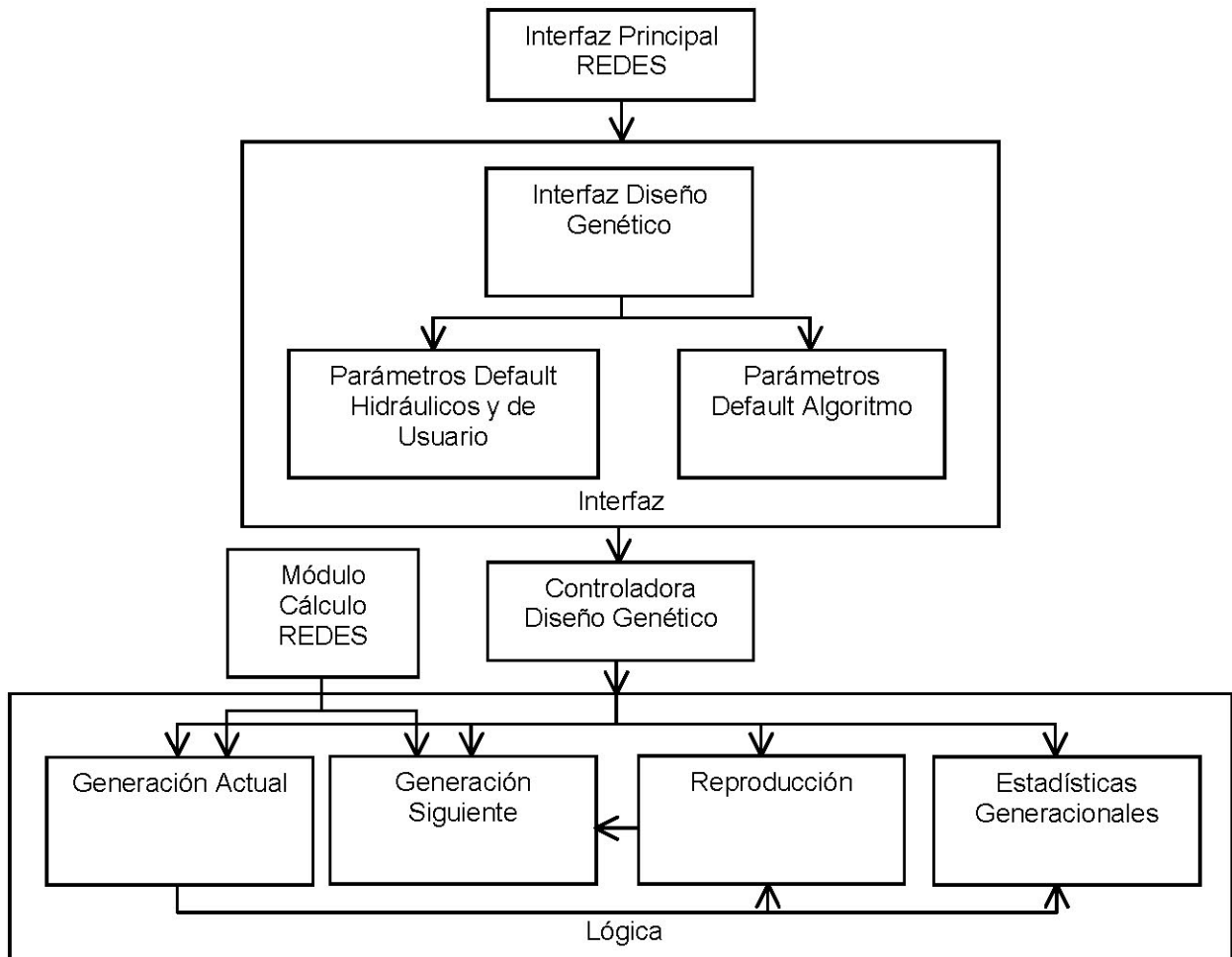
## **5.9. Guardar Estadísticas**

<b>Requerimiento:</b>	9
<b>Nombre:</b>	Guardar Estadísticas
<b>Descripción:</b>	
Si la opción de "Guardar Estadísticas" fue activada en la pestaña "Variables Hidráulicas y de Usuario", después de correr el programa el sistema guarda en un archivo de texto con el mismo nombre del archivo de REDES, para cada generación, lo siguiente: individuo de mayor adaptabilidad, de menor costo constructivo y de menor costo hidráulico, número promedio de nodos donde se cumplen las restricciones, número de individuos que cumplen todas las restricciones, promedio y desviación estándar del costo constructivo e hidráulico de cada generación y desviación estándar de la adaptabilidad de cada generación.	
<b>Casos de excepción:</b>	
Ninguno	
<b>Prueba:</b>	
1. Durante los requerimientos 2 o 4 <b>no</b> se activa la opción "Guardar Estadísticas" de la pestaña "Variables Hidráulicas y de Usuario". Se realiza la prueba 2 o 3 del requerimiento 6. 2. Durante los requerimientos 2 o 4 <b>si</b> se activa la opción "Guardar Estadísticas" en la pestaña "Variables Hidráulicas y de Usuario". Se realiza la prueba 2 o 3 del requerimiento 6.	
<b>Resultado:</b>	
1. No se debe haber creado un archivo de texto con el mismo nombre del archivo de REDES en el directorio donde se encuentra este archivo. 2. Hay un archivo con el mismo nombre del archivo de REDES pero con extensión txt en el mismo directorio donde se encuentra este archivo.	

## 6. ARQUITECTURA DEL SOFTWARE

---

A continuación se presenta el diagrama de Arquitectura del Software.



## 7. DIAGRAMAS DE CLASE

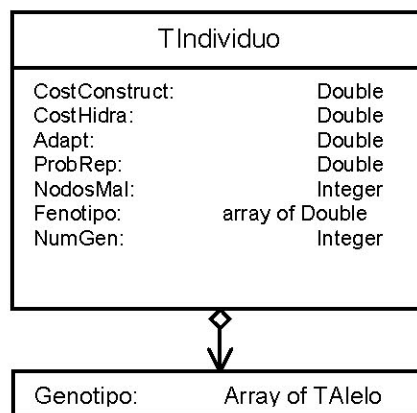
---

### 7.1. TAlelo



Descripción: como su equivalente biológico, el tipo alelo es el bloque básico del genotipo. Cada alelo contiene un valor, correspondiente al diámetro del tubo que codifica, en su variable *Valor*. Al construirse se debe especificar su valor o se debe pasar un vector de tipo Double para escoger su valor aleatoriamente entre los valores de este vector; esto con el fin de poder crear aleatoriamente el genotipo de la primera generación y mutar cuando un nuevo individuo esté siendo generado (por reproducción de dos padres). Para esto se ayuda de una instancia de números aleatorios, tipo TUnif01Persist<sup>5</sup>.

### 7.2. TIndividuo



<sup>5</sup> Esta clase sirve para generar números aleatorios. Fue creada por el ing. Daniel Salas Useche como parte de su proyecto de grado [Salas, 2002].

Descripción: un TIndividuo contiene un Genotipo, un Fenotipo, todos los diversos costos (Constructivo, Hidráulico y su Adaptabilidad), la Probabilidad de Reproducción que se le asigna dependiendo de su adaptabilidad y el número de nodos en los que no cumple con las restricciones. Dada la complejidad de los procedimientos de cálculo de los diversos costos, y como en su mayoría requieren de un conocimiento de parámetros dependientes de toda la generación (por ejemplo, promedios de costos para calcular la adaptabilidad), estos procedimientos no están especificados dentro de esta clase sino en una clase TCalificador que se maneja desde el TControl.

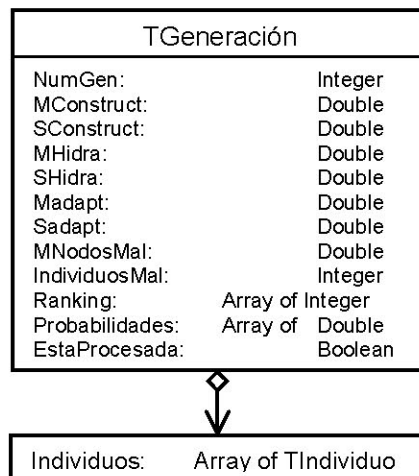
El Genotipo es una cadena (array of) de alelos de longitud fija  $U$  (número de tubos). El TIndividuo tiene procedimientos para devolver el valor de uno de sus alelos, cambiarlo por un valor especificado o cambiarlo al azar.

El Fenotipo contiene las presiones en toda la red cuando se soluciona la hidráulica de algún diseño (arreglo con las  $N$  presiones de la red).

TIndividuo tiene constructores:

- Con el número de generación, la longitud del genotipo y un vector de diámetros, con el cual se crea un individuo nuevo con todos sus campos vacíos excepto el Genotipo, que es generado al azar (útil para la primera generación). El vector de diámetros es opcional; en su ausencia se le asigna la longitud al genotipo y se le asigna un valor de cero a todos los alelos.
- Con dos TIndividuos, una lista de  $U$  booleanos, un vector de diámetros y la probabilidad de mutación. Es un método que crea un nuevo individuo, a partir de dos padres y una lista que especifica de qué padre proviene cada alelo, obtenida a través de THerencia. Los alelos del nuevo individuo pueden también mutar tomando un valor sobre el vector de diámetros al azar. El número de generación del nuevo individuo es uno más que el del primer padre.

### 7.3. TGeneración



Descripción: esta clase, además de contener todos los individuos de una generación (cuyo número, determinado por el usuario, debe ser mayor que 2), contiene la desviación estándar de la Adaptabilidad y los promedios y desviaciones estándar no sesgadas del Costo Constructivo y del Costo Hidráulico.

Al proceso de construcción de la primera generación se pasa el número de individuos, la longitud del genotipo y un vector de Diámetros, en cuyo caso se construye la generación aleatoriamente (con número de generación 1). Posteriormente, se puede construir una generación nueva a partir de una generación “padre”, reproduciendo los individuos de ésta. A este constructor se debe pasar la probabilidad de mutación, el vector de diámetros sobre el cual mutan los alelos, la generación padre y una instancia de THerencia (ver 7.5). Se pasa una instancia de THerencia puesto que cada vez que se reproducen dos individuos se requiere una nueva lista de booleanos que indique la procedencia de cada alelo.

El control de todos los procedimientos de cálculo de costos y de probabilidades se lleva a cabo a través de una instancia de TCalificador (ver 7.4 a continuación), a la cual se le pasa una instancia de TGeneracion desde la controladora (TControlDisenoGenetico, sección 7.7). TGeneracion calcula internamente los promedios y desviaciones estándar no sesgadas<sup>6</sup>

<sup>6</sup> La media de una muestra  $X_1, \dots, X_n$  se define como (sigue en la hoja siguiente):

de los calificadores para la generación, mientras que TCalificador calcula los calificadores de cada individuo. El resultado de este procedimiento es el almacenamiento de las probabilidades de reproducción de todos los individuos de la generación en Probabilidades, arreglo que servirá para seleccionar, a través de una instancia de TUnif01Persist, los padres que se van a utilizar para la Reproducción (para la construcción de una nueva TGeneracion). Al haber sido calificada la generación la variable booleana EstaProcesada toma el valor Verdadero.

Esta clase tiene también un procedimiento para determinar los  $n$  mejores individuos (ordenados en el arreglo de enteros Ranking según su probabilidad de reproducción), útil para determinar el mejor individuo de cada generación y los 10 mejores individuos de la última generación.

Finalmente, hay procedimientos útiles para comunicar los distintos parámetros estadísticos (ver sección 2.4) a la controladora, añadir individuos si se requiere y eliminar el  $i$ -ésimo mejor individuo.

#### 7.4. TCalificador

TCalificador	
Pmin:	Double
ControlRedGenetico:	TControlRed
Metodo:	Boolean
CReprod:	Double
PesoRelativoHidra:	Double

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

La desviación estándar no sesgada  $s^*$ , al cuadrado, es:

$$(s^*)^2 = \frac{n}{n-1} \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n} = \frac{n}{n-1} \left[ \sum_{i=1}^n \frac{X_i^2}{n} - \bar{X}^2 \right]$$

Tomado de [Faur, 2000-2], p. 28.

Descripción: el objetivo único de esta clase es contener los diferentes procedimientos de cálculo que requiere un individuo. Así, tiene procedimientos que calculan:

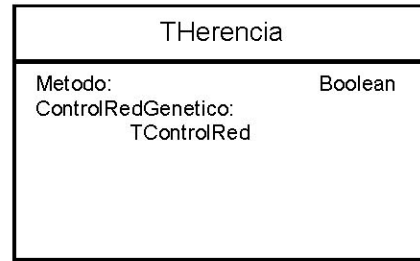
- El fenotipo de un individuo y su Costo Hidráulico a partir de la presión mínima de nodo para la red. Con este fin, esta clase se comunica con el módulo de cálculo de REDES a través de la instancia de TControlRed.
- El Costo Constructivo de un individuo a partir de su genotipo.
- La adaptabilidad de un individuo dadas las medias de costos Constructivo e Hidráulico (obtenidas a través de la instancia de TGeneración que se pase) y el peso relativo del costo hidráulico (PesoRelativoHidra, ecuación [5-8] del Marco Teórico).
- La probabilidad de reproducción según el método usual (ecuación [5-9] del Marco Teórico) ó escalando las probabilidades (ecuación [5-10]) de acuerdo al coeficiente de reproducción inicial  $C$  (almacenado en CReprod), al número de generación y a la desviación estándar de la Adaptabilidad de la generación. Para cualquiera de los métodos escogidos se requiere el tamaño de la población (para asignar al individuo su número de descendientes esperado). El método escogido se almacena en Metodo.

Todos estos procedimientos funcionan sobre todos los individuos de una generación que se pasa por referencia y se devuelve con la información de todos los individuos (fenotipo, y costos) y de la generación (probabilidades de reproducción de cada uno de sus individuos y promedios de costos) completa.

La construcción de una instancia de esta clase se realiza pasando la presión mínima  $P_{min}$ , el peso relativo del costo hidráulico  $PesoRelHidra$ , el método de asignación de probabilidad  $Metodo$ , y la referencia a la instancia de TControlRed que controla la Red. Estos atributos son fijos para la instancia una vez creada, ya que son únicos para cada corrida del algoritmo.



## 7.5. THerencia



Descripción: Similarmente a TCalificador, el atributo Método especifica el método de reproducción que se va a utilizar y se fija desde la construcción de un objeto de este tipo. El procedimiento principal de esta clase tiene como argumentos dos TIndividuo (padres) y como salida una lista de Boolean. Esta lista especifica cuales de los  $U$  alelos de un individuo nuevo van a provenir del padre 1 (True) y cuales del padre 2 (False). Utiliza como procedimientos intermedios los de selección del punto de recombinación (o de hiperplano), para los que utiliza una instancia de TUnif01Persist.

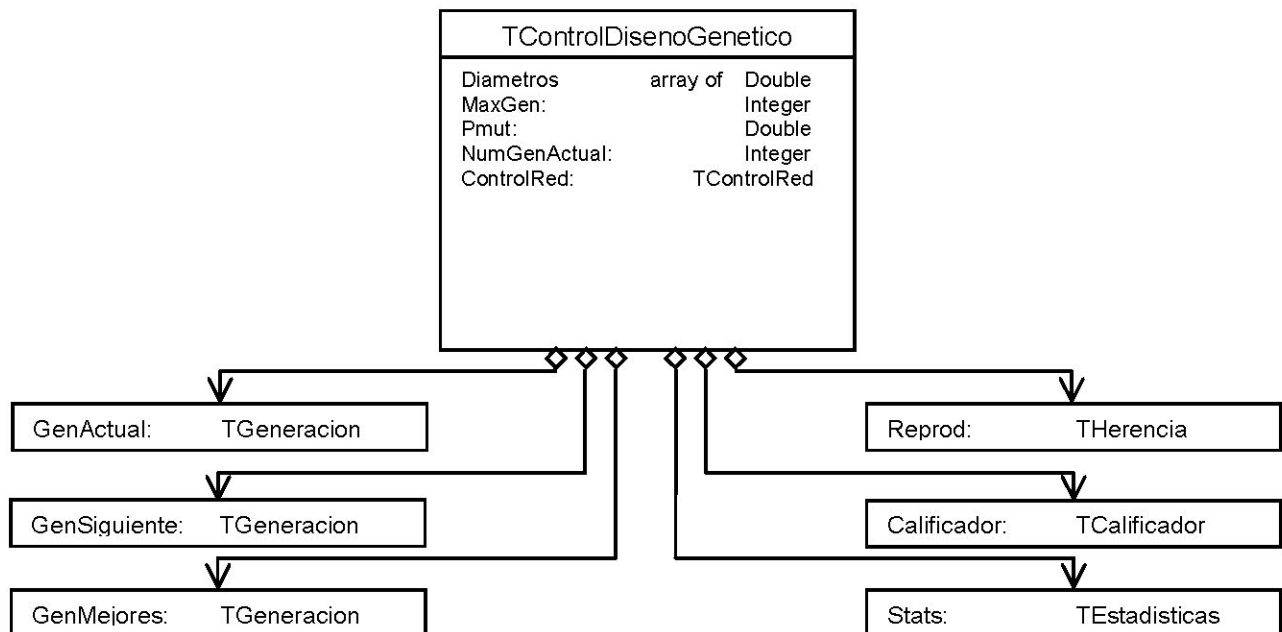
## 7.6. TEstadísticas

TEstadísticas	
Diametros:	Array of Double
PMin:	Double
PesoRelativoHidra:	Double
PMut:	Double
PopSize:	Integer
MaxGen:	Integer
CReprod:	Double
MetodoReprod:	Boolean
MetodoProb:	Boolean
NumGen:	Array of Integer
MConstruct:	Array of Double
SConstruct:	Array of Double
MHidra:	Array of Double
SHidra:	Array of Double
SAdapt:	Array of Double
MNodosMal:	Array of Double
IndividuosMal:	Array of Integer
MejorConstruct:	Array of Double
MejorHidra:	Array of Double
MejorNodosMal:	Array of Integer
NumGenMejores:	Array of Integer
ConstructMejores:	Array of Double
HidraMejores:	Array of Double
NodosMalMejores:	Array of Integer
PresionesMejorIndividuo:	Array of Double

Descripción: esta clase se utiliza para almacenar todos los parámetros del algoritmo y las estadísticas que se hallan generación por generación en arreglos, para su posterior impresión en un archivo de texto. El primer grupo de atributos en el diagrama de clase son las variables hidráulicas y de algoritmo con las que se corrió el algoritmo, el segundo grupo son las estadísticas generacionales, el tercer grupo son las estadísticas de los 10 mejores diseños resultantes del algoritmo y el vector PresionesMejorIndividuo que contiene las presiones de nodo del mejor diseño encontrado por el algoritmo.

Las estadísticas MejorConstruct y MejorNodosMal se refieren al costo constructivo y al número de nodos mal del mejor individuo de cada generación. Las demás se explican ellas mismas.

### 7.7. TControlDisenoGenetico



Descripción: esta clase es la controladora principal del algoritmo. Sus atributos son los parámetros necesarios para reproducir una generación (los diámetros y la probabilidad de mutación), el número de generación actual y el número de generaciones que se van a correr (MaxGen). Las características hidráulicas y del algoritmo también son necesarias para construir un objeto de este tipo, utilizadas para la inicialización de las instancias de TEstadisticas, TCalificador y THerencia. Todas las variables hidráulicas y de algoritmo son por tanto fijas durante el algoritmo.

El control del algoritmo se realiza mandando a GenActual (la generación actual) a crearse aleatoriamente si es la primera, y a generarse a partir de la generación anterior si no lo es. Luego se manda a Calificador para que califique los individuos y complete todas las estadísticas de la generación, las cuales se extraen y se guardan en Stats. Se piden además los 10 mejores individuos de la generación para ser comparados con los individuos en GenMejores con el fin de almacenarlos si son superiores. Este procedimiento se repite hasta que transcurra el número de generaciones especificado.

Es TControlDisenoGenetico quien almacena también los 10 mejores individuos en la “generación” GenMejores, y quien guarda cualquiera de estos 10 individuos en un archivo de REDES si el usuario lo requiere. Es la encargada también de la salida del resto de la información (pasa los parámetros a la interfaz para su exhibición por pantalla y pide grabar el archivo de texto con las estadísticas a Stats).

## 8. BIBLIOGRAFÍA

---

[Charte, 2000]Charte, Francisco (2000). “Programación en Delphi 5”. Madrid, España: Ediciones Anaya Multimedia.

[Faur, 2000-2]Faur, C. (2000), “Statistique, Notes de Cours”, Nantes, Francia: Notas de Curso, Ecole des Mines de Nantes.

[Salas, 2002] Salas Useche, Daniel (2002). “Algoritmo genético modificado para calibración de redes de acueducto”. Tesis de grado no publicada. Departamento de Ingeniería Civil y Ambiental, Universidad de los Andes, Bogotá, Colombia.

## PARTE 3: PRUEBAS Y RESULTADOS

### 9. FUNCIONAMIENTO DEL PROGRAMA

---

#### 9.1. Recuento del desarrollo

Como se especificó en la Documentación, el programa se desarrolló en el lenguaje de programación Object Pascal, en el entorno Borland Delphi 6. Se desarrollaron las 7 clases especificadas: TAlelo, TIndividuo, TGeneracion, TCalificador, THerencia, TEstadisticas y TControlDisenoGenetico, repartidas en 6 unidades: TOperadoresGeneticos (contiene TAlelo y TIndividuo), TGeneracionClass, TCalificadorClass, THerenciaClass, TEstadisticasClass y TControlDisenoGeneticoClass.

La clase THerencia no fue desarrollada por completo, puesto que faltó el método de Recombinación Hiperplanar. Se espera que en desarrollos subsecuentes se desarrolle este método.

Se desarrollo una clase adicional, TProgressRedesDummy, clase que hereda de TProgressRedes (que a su vez hereda de TProgressBar, una componente gráfico estándar de Delphi cuya función es indicar en la interfaz el progreso del programa) pero que tiene sin embargo todos sus métodos vacíos. Esta clase fue necesaria para el funcionamiento de TCalificador cuando debía evaluar la hidráulica de los TIndividuo; solo debía tomar el lugar de TProgressRedes pero sin hacer nada. Por esto no se describe esta clase, siendo irrelevante al funcionamiento del algoritmo en si.

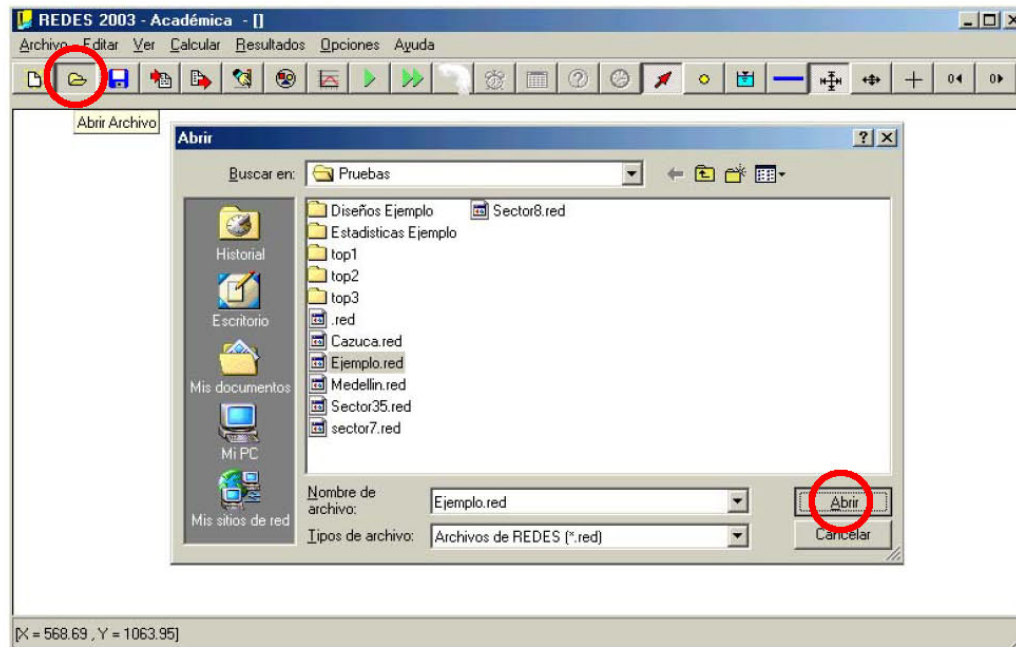
Se desarrolló también la interfaz de usuario. Esta interfaz recibe los parámetros del usuario, revisa que sean del tipo correcto y que estén dentro del rango correcto, y los pasa a TControlDisenoGenetico, encargada de manejar el módulo “Lógico” del programa (ver

numeral 6 de la Documentación). Así mismo, TControlDisenoGenetico pasa los resultados del algoritmo y guarda los diseños o estadísticas si el usuario lo requiere a través de la interfaz.

## 9.2. Uso Típico

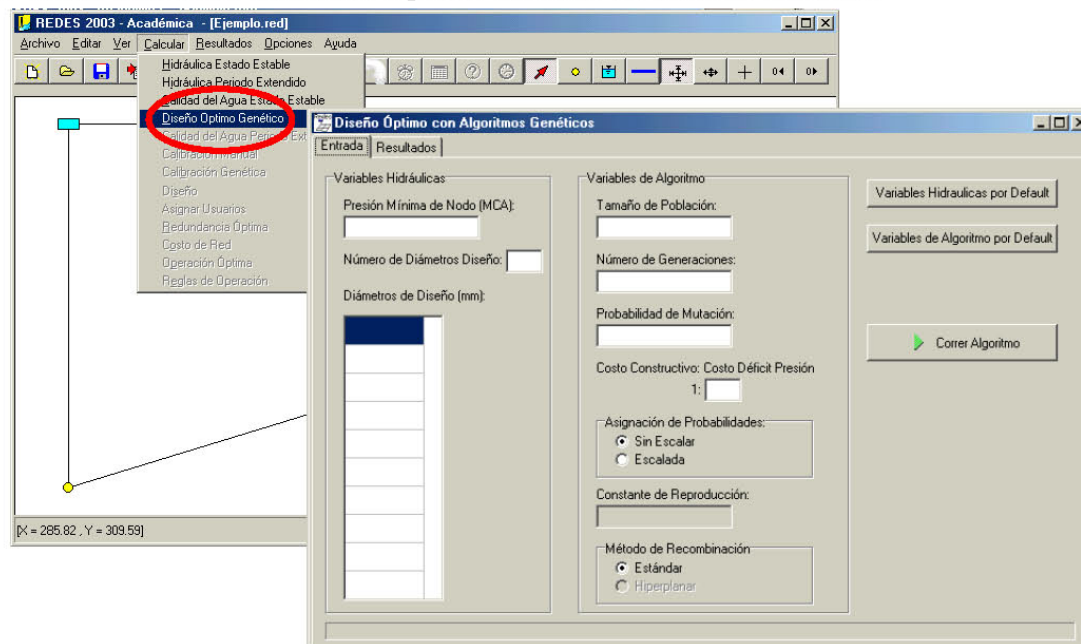
A continuación se muestra una secuencia de uso típico del programa.

### 1. Abrir la red a diseñar en REDES.



Seleccione “Abrir Archivo”, busque el archivo de REDES que desea abrir y presione el botón “Abrir”. En este caso se trabajará con el archivo “Ejemplo.red”, descrito en el ejemplo 7.1 de [Saldarriaga, 2001] y en el Anexo 1.

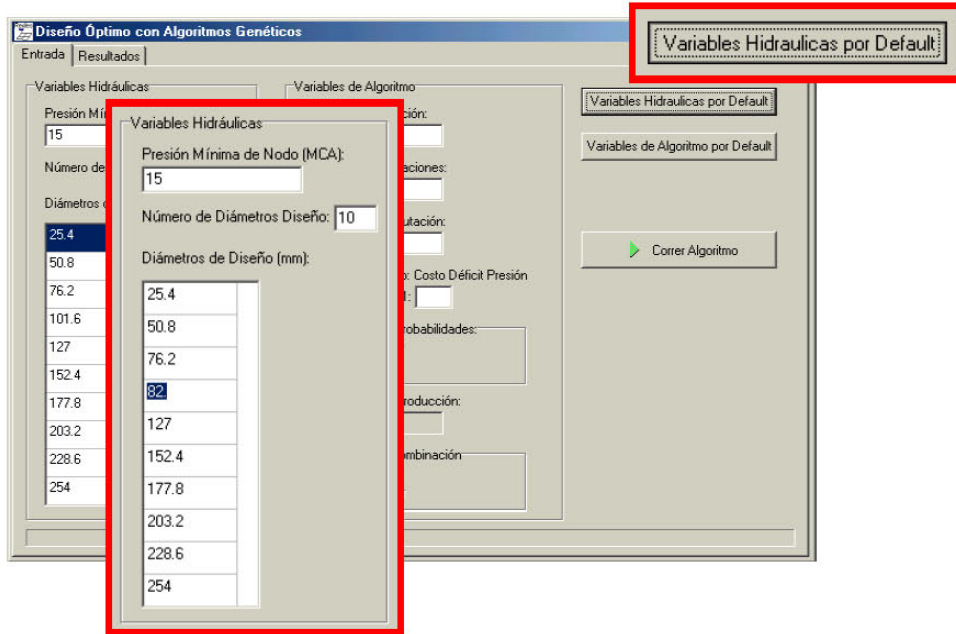
### 2. Abrir el módulo de Diseño Óptimo Genético en el menú Calcular.





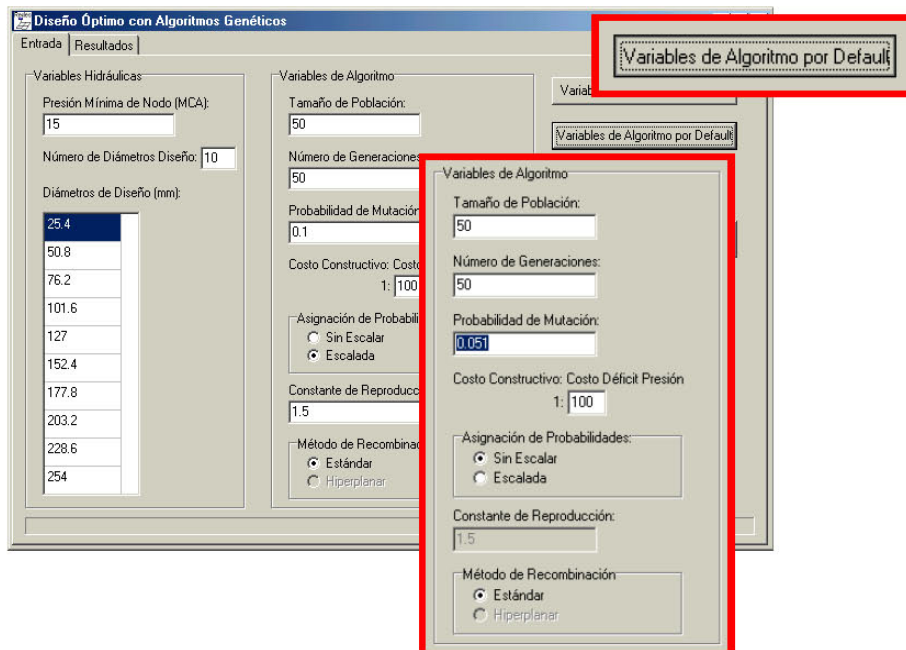
Seleccione la opción “Diseño Óptimo Genético” en el menú “Calcular”. La ventana “Diseño Óptimo con Algoritmos Genéticos” aparece.

3. Ingresar las variables hidráulicas.



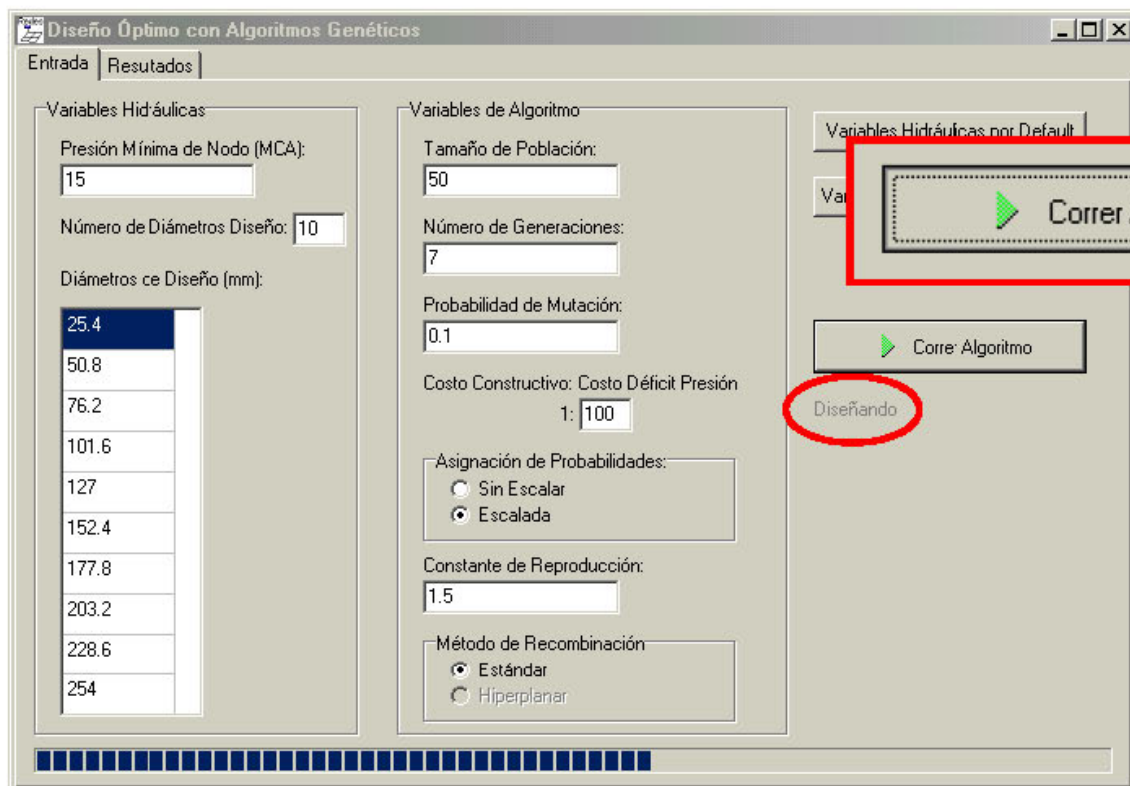
Presione “Variables Hidráulicas por Default” para llenar los campos con los valores por default de estas variables (ver Requerimiento 4). También puede modificar usted mismo los valores presionando sobre ellos e ingresando el valor con el teclado.

4. Ingresar las variables de algoritmo.



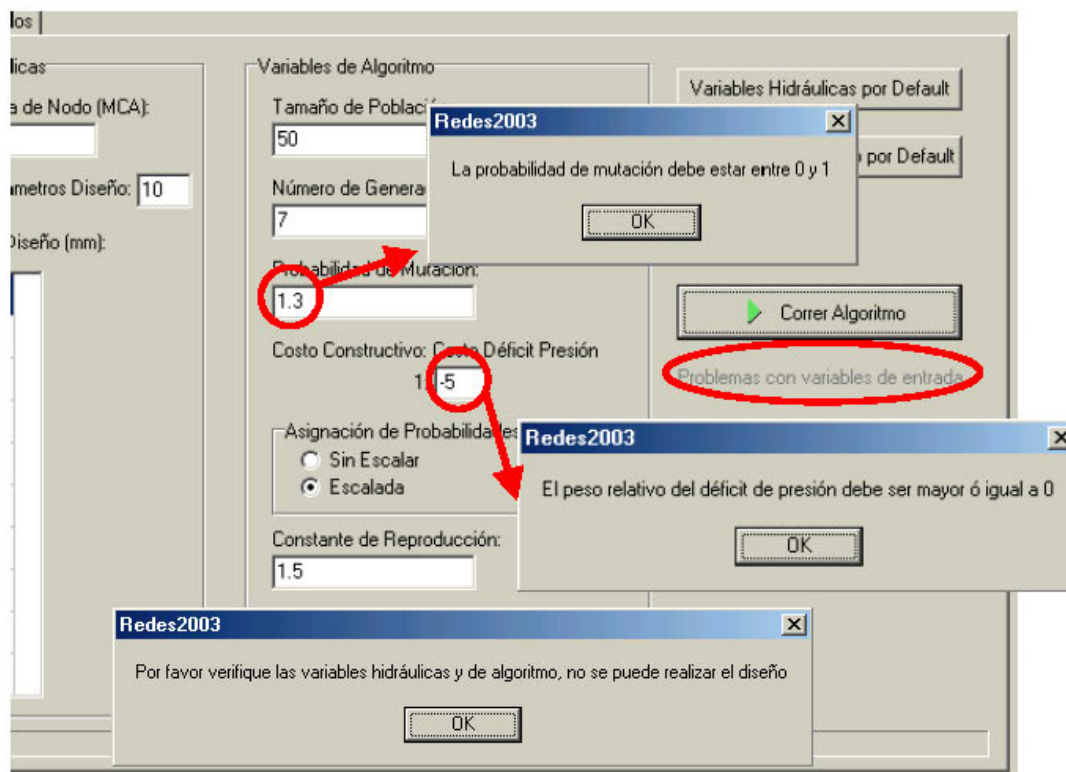
Presione “Variables de Algoritmo por Default” para llenar los campos con los valores por default de estas variables (ver Requerimiento 5). También puede modificar usted mismo los valores presionando sobre ellos e ingresando el valor con el teclado y activar las opciones de algoritmo que quiera, a excepción de Método de Recombinación “Hiperplanar”.

## 5. Correr el Algoritmo.



Presione “Correr Algoritmo” (en este caso se corrió con las variables por default excepto que se realizaron solo 7 generaciones). Un aviso de “Diseñando” aparece bajo el botón, y la barra azul en la parte inferior de la ventana comienza a llenarse a medida que avanza el algoritmo.

En caso de ingresar mal una variable se muestra un mensaje de error (que depende específicamente de la variable que esté mal – ver figura a continuación y requerimientos 2 y 3), se muestra el mensaje “Por favor verifique las variables hidráulicas y de algoritmo, no se puede realizar el diseño” y el aviso “Problemas con variables de entrada” aparece bajo el botón de “Correr Algoritmo”.



## 6. Resultados del Diseño.

Al terminar de diseñar el programa pasa automáticamente a la pestaña “Resultados” donde aparece el cuadro con las estadísticas de los 10 mejores diseños encontrados.

Mejores Diseños:

	Generación	Costo Constructivo	Déficit Presión (m)	Nodos con déficit Presión
1	7	221851726.46420	0	0
2	5	250793536.89013	0	0
3	4	250793536.89013	0	0
4	7	252799003.04786	0	0
5	4	258355723.07598	0	0
6	3	263910655.34160	0	0
7	7	272727431.36899	0	0
8	4	274718845.47806	0	0
9	7	276700781.89737	0	0
10	1	278606796.36973	0	0

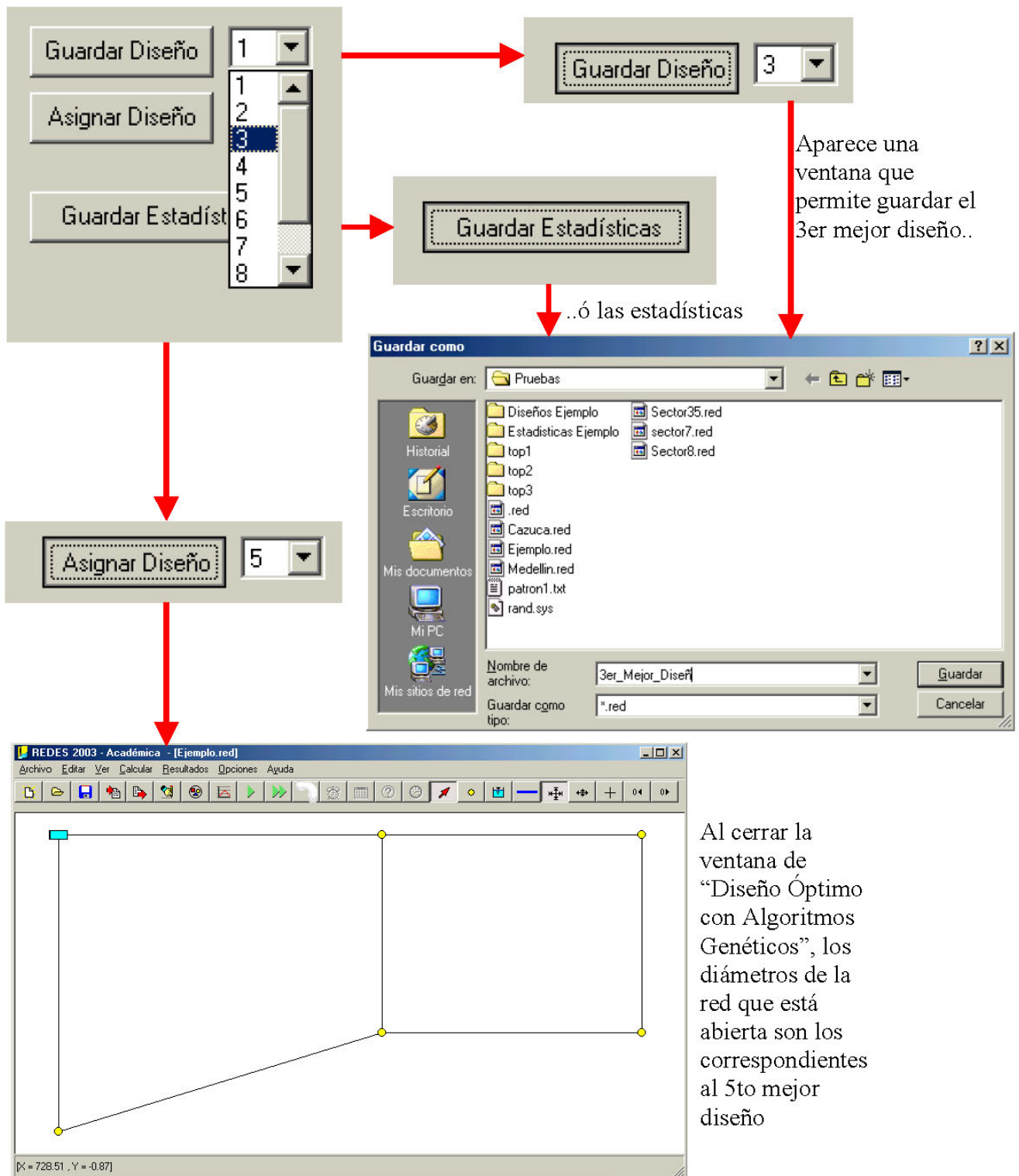
Guardar Diseño 1

Asignar Diseño 1

Guardar Estadísticas

## 7. Guardar Resultados.

Si el usuario lo desea, puede presionar el botón “Guardar Diseño”, seleccionando previamente un número de diseño de la lista a la derecha del botón. Así mismo puede guardar las estadísticas del diseño (con el botón “Guardar Estadísticas”) o asignar directamente un diseño en REDES con la opción de “Asignar Diseño”. Al cerrar la ventana de “Diseño Óptimo con Algoritmos Genéticos” el número de diseño que el usuario haya asignado aparece directamente en la red que esté abierta.



## 10. REVISIÓN DE REQUERIMIENTOS

---

Después de finalizado el programa, se cumplen casi todos los requerimientos. A continuación se presenta un informe detallado (para seguir este informe referirse al capítulo 5 de la Documentación):

### **Requerimiento 1: Abrir Módulo de Diseño.**

Prueba 1: Al presionar sobre “Correr Algoritmo”, el sistema levanta una excepción con el mensaje “[TSparseSolver.InitHydraulicSystem]: El sistema es vacío”. No permite diseñar.

Pruebas 2 y 3: Se cumplen como se indica en los requerimientos.

### **Requerimiento 2: Ingresar Configuración/Variables Hidráulicas.**

Pruebas 1 y 2: Se muestra el mensaje “*cadena* is not a valid floating point value”, donde *cadena* es la cadena de caracteres errónea que entró el usuario. Todos los parámetros quedan con el mismo valor que ingresó el usuario. El algoritmo no corre.

Pruebas 3 y 4: Se muestra el mensaje “El valor de los diámetros debe ser mayor estricto que 0” seguido del mensaje “Por favor verifique las variables hidráulicas y de algoritmo, no se puede realizar el diseño”. El algoritmo no corre y aparece el aviso “Problemas con variables de entrada” bajo el botón de Correr Algoritmo.

Prueba 5: Se cumple como se indica en los requerimientos.



**Requerimiento 3: Ingresar Configuración/VARIABLES de Algoritmo.**

Prueba 1: Se muestra el mensaje “*cadena* is not a valid floating point value” para el caso de Probabilidad de Mutación, Peso Relativo del Déficit de Presión y Constante de Reproducción, y el mensaje, “*cadena* is not a valid integer value” para el Tamaño de Población y el Número de Generaciones, donde *cadena* es la cadena de caracteres errónea que entró el usuario. Todos los parámetros quedan con el mismo valor que ingresó el usuario. El algoritmo no corre.

Pruebas 2, 3 y 6: Se muestra el mensaje “El tamaño de población debe ser mayor o igual a 2” y el mensaje, “El número de generaciones debe ser mayor estricto que 0” para el Tamaño de Población y el Número de Generaciones respectivamente, seguido del mensaje “Por favor verifique las variables hidráulicas y de algoritmo, no se puede realizar el diseño”. Todos los parámetros quedan con el mismo valor que ingresó el usuario. El algoritmo no corre y aparece el aviso “Problemas con variables de entrada” bajo el botón de Correr Algoritmo.

Prueba 4: Se muestra el mensaje “La constante de reproducción debe ser mayor estricta que 0”, seguido del mensaje “Por favor verifique las variables hidráulicas y de algoritmo, no se puede realizar el diseño”. Todos los parámetros quedan con el mismo valor que ingresó el usuario. El algoritmo no corre y aparece el aviso “Problemas con variables de entrada” bajo el botón de Correr Algoritmo.

Prueba 5: Se muestra el mensaje “La probabilidad de mutación debe estar entre 0 y 1”, seguido del mensaje “Por favor verifique las variables hidráulicas y de algoritmo, no se puede realizar el diseño”. Todos los parámetros quedan con el mismo valor que ingresó el usuario. El algoritmo no corre y aparece el aviso “Problemas con variables de entrada” bajo el botón de Correr Algoritmo.

Prueba 7: Se muestra el mensaje “Advertencia: un valor de constante de reproducción debajo de 1 conlleva a diseños deficientes”. El algoritmo corre normalmente.

Prueba 8: Se cumple como se indica en los requerimientos.

#### **Requerimiento 4: Ver Configuración Default/Variables Hidráulicas.**

Prueba 1: Se llena la tabla “Diámetros de Diseño” con el equivalente en milímetros de los diámetros 1, 2, 3, 4, 5, 6, 7, 8, 9 y 10 pulgadas. El campo de “Presión Mínima de Nodo” toma el valor 15 metros.

#### **Requerimiento 5: Ver Configuración Default/Variables de Algoritmo.**

Prueba 1: Se cumple como se indica en los requerimientos, excepto que se selecciona la opción “Estándar” como método de recombinación. Esto se debe a que la opción “Hiperplanar” no alcanzó a ser implementada. En caso que esta opción sea seleccionada, aparece el mensaje “THerencia.RecombinacionHiperplanar: Método aún no implementado”, y el algoritmo no corre.

#### **Requerimiento 6: Correr el Algoritmo.**

Prueba 1: Se muestra el mensaje “ ‘ ’ is not a valid floating point value” o el mensaje “ ‘ ’ is not a valid integer value” dependiendo si el parámetro omitido es un real o un entero (respectivamente). Esto no sucede si se omite la Constante de Reproducción y se ha seleccionado el método de Asignación de Probabilidad “Sin Escalar”. Cuando no se ingresa ningún diámetro en la tabla de Diámetros de Diseño se muestra el mensaje “Debe asignar diámetros de diseño”. En todos los casos todos



los parámetros quedan con el mismo valor que ingresó el usuario y el algoritmo no corre.

Pruebas 2 y 3: Se cumplen como se indica en los requerimientos.

### **Requerimiento 7: Ver Calificadores de Diseño**

Prueba 1: Se cumple como se indica en los requerimientos, indicándose adicionalmente el número de nodos donde no se cumple la restricción de presión.

### **Requerimiento 8: Guardar Diseño.**

Pruebas 1, 2 y 3: ya no es necesario realizarlas puesto que se cambió el cuadro de texto donde se escribía el número de diseño a guardar por una lista desplegable de donde solo se puede escoger un número del 1 al 10.

Prueba 4: Previo a que el algoritmo corra el botón de “Guardar Diseño” se encuentra apagado, por lo que no se puede intentar guardar ningún diseño sin haber corrido el algoritmo al menos una primera vez después de abrir el módulo de diseño.

Prueba 5: Se cumple como se indica en los requerimientos.

### **Requerimiento 9: Guardar Estadísticas**

Prueba 1: Previo a que el algoritmo corra el botón de “Guardar Estadísticas” se encuentra apagado, por lo que no se puede intentar guardar las estadísticas sin haber corrido el algoritmo al menos una primera vez después de abrir el módulo de diseño.

Prueba 2: Se cumple como se indica en los requerimientos, con la excepción que no se guarda la información del individuo de menor costo constructivo ni de menor costo hidráulico. Solo se guarda la información (costos constructivo, hidráulico y nodos que no cumplen las restricciones) del mejor individuo de cada generación ranqueado según probabilidad de reproducción. Adicionalmente, se imprimen todas las variables de entrada al algoritmo, las estadísticas de los 10 mejores diseños y las presiones de red del mejor diseño.

En el Anexo 2 se encuentra un ejemplo de archivo de estadísticas.

## 11. ANALISIS DE RESULTADOS

---

### 11.1. Influencia de las variables sobre el resultado del algoritmo

Con el fin de investigar el funcionamiento del algoritmo, se realizaron varias corridas sobre la red Ejemplo.red (Anexo 1). Se variaban sistemáticamente las variables de algoritmo para evaluar cómo se veían afectados (o si se afectaban del todo) los resultados del algoritmo. Para hacer esta evaluación se comparó el costo constructivo del mejor diseño encontrado, considerado el resultado principal del algoritmo, con la variación de los distintos parámetros.

La herramienta de análisis utilizada fue el ANOVA (ANalysis Of VAriance – Análisis de la Varianza)<sup>7</sup>. Esta herramienta estadística permite asegurar, sujeto a un nivel de confianza, que la variación sobre un cierto parámetro influye sobre el resultado de otro, bajo ciertas hipótesis estadísticas dadas. Como su nombre lo dice, el análisis se efectúa teniendo en cuenta la varianza y el error intrínsecos a los resultados, siendo mucho más confiable por tanto que, por ejemplo, una simple comparación de los promedios de los resultados bajo variaciones en los parámetros.

Las hipótesis estadísticas del análisis son (1) que la muestra (la playa de resultados) se distribuya según una variable normal (gaussiana) y (2) que la varianza de la muestra cuando se varía el valor del parámetro estudiado permanezca igual. El primer supuesto parecería ser cierto en el caso de los algoritmos genéticos, cuyo resultado final depende de qué tan cerca se localice la primera generación del óptimo global y de otros factores aleatorios como las mutaciones y las recombinaciones. En lo concerniente al segundo supuesto aunque, por ejemplo, una disminución en la probabilidad de mutación disminuye considerablemente la variabilidad de cada generación y por tanto la desviación estándar *dentro* de la población, ésta no tendría porque afectar la desviación estándar *entre*

---

<sup>7</sup> Para más información sobre este tipo de análisis, ver [Faur, 2000-1], o en general cualquier libro de estadística o de Análisis de la Varianza.

*diferentes corridas del algoritmo*, por lo que se cumpliría este supuesto. La anterior afirmación, aunque requiere un estudio más profundo, se asumirá cierta para la realización del análisis.

El análisis de varianza permite adicionalmente decidir (de nuevo bajo un cierto nivel de confianza) si el promedio del resultado cuando un parámetro toma cierto valor es diferente al promedio del resultado bajo un valor diferente; esto es, permite evaluar si los contrastes, o diferencia de promedios para diferentes valores, son iguales o no a 0.

Durante todas las pruebas se dejaron fijas las variables hidráulicas, diseñando sobre los diámetros por default y con una presión mínima de nodo de 15 mca. Las variables de algoritmo “estándar” son: tamaño de población: 100, número de generaciones: 50, peso relativo del déficit de presión: 100, probabilidad de mutación: 0.1, constante de reproducción: 1.5, asignación de probabilidad escalada y método de recombinación estándar. Se realizaron cuatro pruebas con estos valores, contra las cuales se comparan el resto de las pruebas.

Es importante tener en cuenta que, aunque los resultados que arroja este análisis son significativos, pueden variar si se utilizan otras redes. Es probable que ciertos parámetros influyeran de manera diferente en redes de distinto tamaño. Por ejemplo, para redes más grandes puede que sea necesaria una constante de reproducción inicial mucho menor para no caer en mínimos locales y poder explorar de manera más completa el espacio de búsqueda (que crece exponencialmente con el número de tubos de la red).

Finalmente, se presentan los resultados del ANOVA hecho para las distintas variables. La totalidad de las pruebas hechas sobre el algoritmo se presenta en el Anexo 3.

## Influencia de la Probabilidad de Mutación

Se realizaron 12 pruebas variando la probabilidad de mutación: 4 con una probabilidad de 0.025, 4 con 0.05 y 4 con 0.2. Los resultados del ANOVA son los siguientes (las variables X1, X2, X3, X4 se refieren a los resultados de cada una de las 4 pruebas hechas con cada valor):

Variación Costo Diseño según

Pmutación

Pmutación	0.025	0.05	0.1	0.2
X1	2.33E+08	2.02E+08	2.10E+08	2.29E+08
X2	2.11E+08	2.22E+08	2.20E+08	2.51E+08
X3	1.96E+08	2.00E+08	2.22E+08	2.23E+08
X4	2.07E+08	2.07E+08	2.11E+08	2.40E+08

Análisis de varianza de un factor

$\alpha =$

0.05

RESUMEN

Grupos	Cuenta	Suma	Promedio	Varianza
0.025	4	8.46E+08	2.11E+08	2.34E+14
0.05	4	8.30E+08	2.07E+08	1.00E+14
0.1	4	8.63E+08	2.16E+08	3.53E+13
0.2	4	9.43E+08	2.36E+08	1.48E+14

ANÁLISIS DE VARIANZA

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	1.88E+15	3	6.28E+14	4.85	0.0195	3.49
Dentro de los grupos	1.55E+15	12	1.29E+14			
Total	3.44E+15	15				

El primer cuadro indica los resultados de las 12 pruebas ya indicadas, el segundo es un resumen de las estadísticas de estas pruebas, agrupado por el valor de Pmutación. El tercer cuadro es el ANOVA en sí: el factor F es la razón entre el promedio de la suma de las diferencias al cuadrado entre grupos (diferencia del promedio de cada grupo con el promedio de *la muestra entera*) y el promedio de la suma de las diferencias al cuadrado dentro de los grupos (diferencia de cada resultado con el promedio de *su grupo*). Este parámetro es un indicador por tanto de que tan grande es la influencia de variar el parámetro en comparación con el error que ocurre naturalmente en la muestra (a mayor F mayor variación debida al parámetro que está bajo estudio).

Para que se considere que el parámetro estudiado influye sobre el resultado, con un nivel de confianza de  $1-\alpha$  (es decir, con riesgo de equivocarse de  $\alpha$ ), el valor de F debe ser mayor que el valor crítico para F allí indicado. Este valor crítico corresponde a una variable aleatoria de Fisher con parámetros  $p-1$  y  $n-p$  ( $p$  es el número de grupos o valores que toma el parámetro – en este caso 4 – y  $n$  es el tamaño de la muestra).

En este caso vemos que, con una confiabilidad de 95% (ya que  $\alpha = 5\%$ ), la probabilidad de mutación tiene una influencia significativa sobre el resultado del algoritmo.

El hecho que el resultado del ANOVA sea positivo indica que hay al menos dos de los grupos sobre los que se puede asegurar que sus promedios son diferentes. Se realiza entonces el análisis de contrastes:

**ESTUDIO DE CONTRASTES**

S = raiz((p-1)*F(0.05, p-1, n-p))	3.24
sigma ( $\sigma$ )	1.14E+07
S* $\sigma$ *raiz(1/nj + 1/nk)	2.60E+07

DIFERENCIA PROMEDIOS,  xj-xk	0.025	0.05	0.1	0.2
0.025	0.00E+00	4.00E+06	4.22E+06	2.42E+07
0.05	4.00E+06	0.00E+00	8.21E+06	2.82E+07
0.1	4.22E+06	8.21E+06	0.00E+00	2.00E+07
0.2	2.42E+07	2.82E+07	2.00E+07	0.00E+00

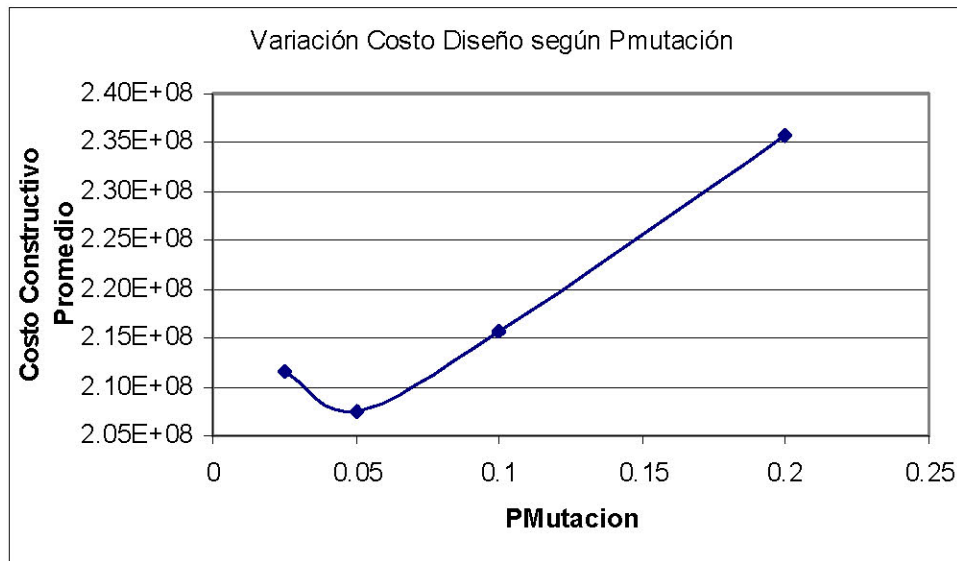
¿CONTRASTE SIGNIFICATIVO?	0.025	0.05	0.1	0.2
0.025	NO	NO	NO	NO
0.05	NO	NO	NO	SI
0.1	NO	NO	NO	NO
0.2	NO	SI	NO	NO

Se afirma que hay un contraste significativo (con un nivel de confianza de  $1-\alpha$ ) si la diferencia entre los promedios del grupo  $i$  y el grupo  $k$  es mayor que:

$$\sqrt{(p-1) \cdot F_{\alpha, p-1, n-p}} \cdot \sigma \cdot \sqrt{\frac{1}{n_i} + \frac{1}{n_k}}$$

donde  $F_{\alpha, p-1, n-p}$  es la variable de Fisher mencionada anteriormente,  $\sigma$  es la desviación estándar de la muestra y  $n_i, n_k$  son el número de efectivos del grupo  $i$  y el grupo  $k$ .

Se puede observar del cuadro de contrastes significativos que el promedio de resultados obtenidos con probabilidad de mutación de 0.05 es distinto al obtenido con una probabilidad de 0.2. Esto se ve con mayor claridad a continuación.



Aparentemente los mejores resultados se encuentran con una probabilidad de mutación de 0.05. Esto se debe probablemente a que valores por encima hacen que se pierda demasiada información valiosa de los mejores diseños de generación a generación, mientras que valores menores no permiten una variación suficiente como para que el algoritmo no quede estancado en ciertos mínimos locales, alejando por tanto los resultados del mínimo global.

### Influencia del Peso Relativo del Déficit de Presión

A continuación se presentan los resultados del ANOVA.

Variación Costo Diseño según

WRelHidra

WRelHidra	1	3	50	100	200
X1	2.69E+08	2.42E+08	2.20E+08	2.10E+08	2.22E+08
X2	2.79E+08	2.29E+08	2.16E+08	2.20E+08	2.22E+08
X3	2.45E+08	2.36E+08	2.29E+08	2.22E+08	2.18E+08

**Análisis de varianza de un factor**

$\alpha =$  0.05

RESUMEN

Grupos	Cuenta	Suma	Promedio	Varianza
1	3	7.93E+08	2.64E+08	2.98E+14
3	3	7.07E+08	2.36E+08	4.13E+13
50	3	6.65E+08	2.22E+08	4.71E+13
100	3	6.52E+08	2.17E+08	3.66E+13
200	3	6.61E+08	2.20E+08	4.29E+12

ANÁLISIS DE VARIANZA

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	4.53E+15	4	1.13E+15	13.24	0.0005	3.48
Dentro de los grupos	8.54E+14	10	8.54E+13			
Total	5.38E+15	14				

Dado que  $F = 13.24 > F_{crítico}$  se puede concluir, con un 95% de confiabilidad, que hay una influencia significativa del valor del peso relativo del déficit de presión sobre el resultado del algoritmo. Se realiza el análisis de diferencia de contrastes.

ESTUDIO DE CONTRASTES

$S = \text{raiz}((p-1) \cdot F(0.05, p-1, n-p))$	3.23
$\text{sigma } (\sigma)$	9.24E+06
$S \cdot \sigma \cdot \text{raiz}(1/n_j + 1/n_k)$	2.44E+07

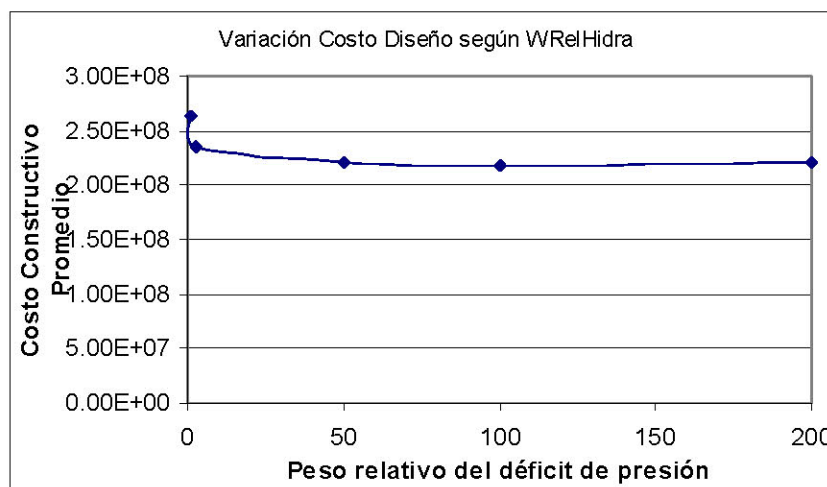
DIFERENCIA PROMEDIOS, $ x_j - x_k $	1	3	50	100	200
1	0.00E+00	2.86E+07	4.26E+07	4.69E+07	4.38E+07
3	2.86E+07	0.00E+00	1.40E+07	1.83E+07	1.52E+07
50	4.26E+07	1.40E+07	0.00E+00	4.34E+06	1.21E+06
100	4.69E+07	1.83E+07	4.34E+06	0.00E+00	3.13E+06
200	4.38E+07	1.52E+07	1.21E+06	3.13E+06	0.00E+00

¿CONTRASTE SIGNIFICATIVO?	1	3	50	100	200
1	NO	SI	SI	SI	SI
3	SI	NO	NO	NO	NO
50	SI	NO	NO	NO	NO
100	SI	NO	NO	NO	NO
200	SI	NO	NO	NO	NO



Como se puede ver, solo hay diferencias significativas entre un peso relativo de 1 y el resto de los valores. Además, la gráfica a continuación y el valor de los promedios ya indicado en el ANOVA indican que los resultados con un peso relativo de 1 son mucho peores que con valores mayores, es decir, cuando se da un mismo peso al costo constructivo y al costo hidráulico los resultados a los cuales llega el algoritmo no son buenos.

Un ejemplo claro es la segunda prueba hecha con peso relativo de 1, donde el valor de costo constructivo fue de  $2.79 \cdot 10^8$ , valor que, además, no correspondía al mejor diseño encontrado por el algoritmo sino al segundo puesto, ya que el mejor diseño no cumplía con las restricciones de presión. Aparentemente, una vez se hace suficientemente fuerte la restricción de presión (que aparentemente ya sucede con un peso relativo de 3) los resultados del algoritmo no varían mucho. Sin embargo, el mínimo pareciera encontrarse en el valor usado por default: 100.



## Influencia de la Constante de Reproducción

A continuación se presentan los resultados del ANOVA.

Variación Costo Diseño según

CReprod

CReprod	1	1.5	3	6
X1	2.23E+08	2.10E+08	2.09E+08	2.25E+08
X2	2.20E+08	2.20E+08	2.00E+08	1.98E+08
X3	2.22E+08	2.22E+08	1.96E+08	2.40E+08
	3.03E+08	2.11E+08	2.03E+08	2.02E+08

Análisis de varianza de un factor

$\alpha =$  0.05

RESUMEN

Grupos	Cuenta	Suma	Promedio	Varianza
1	4	9.68E+08	2.42E+08	1.65E+15
1.5	4	8.63E+08	2.16E+08	3.53E+13
3	4	8.08E+08	2.02E+08	2.93E+13
6	4	8.65E+08	2.16E+08	3.96E+14

ANÁLISIS DE VARIANZA

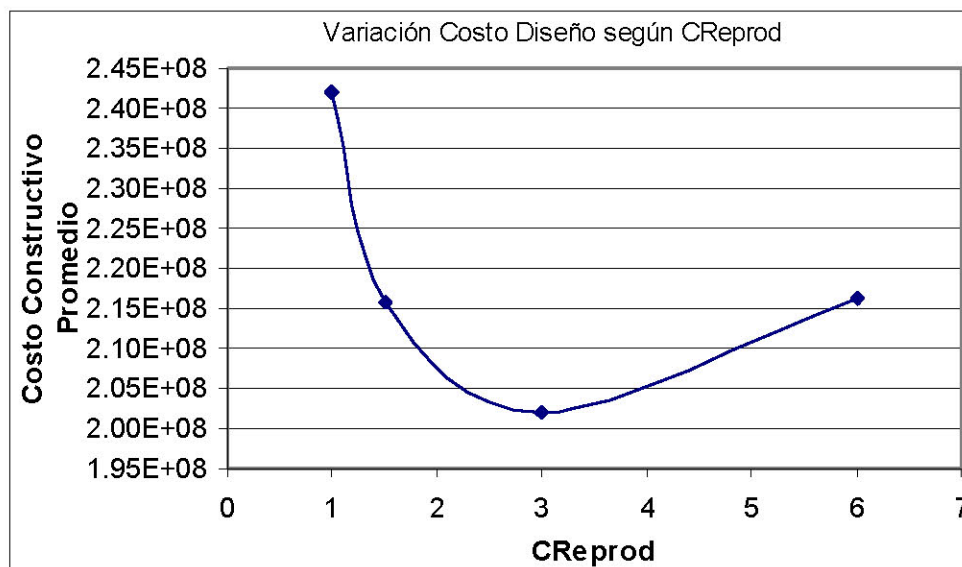
Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	3.34E+15	3	1.11E+15	2.11	0.1525	3.49
Dentro de los grupos	6.33E+15	12	5.27E+14			
Total	9.67E+15	15				

Los resultados de este ANOVA son bastante particulares ya que, si bien la diferencia entre diferentes valores de la constante de reproducción es bastante alta, también la variación al interior de cada uno de los valores lo es. Cuando la constante es 1 hay diseños cerca de los “usuales” pero también hay diseños con costos muy elevados (de hecho el diseño correspondiente al valor de X4, además de tener un costo muy elevado, ni siquiera cumplía con las restricciones de presión).

Similarmente a lo que sucedió con la probabilidad de mutación, valores muy bajos de constante de reproducción no permiten una perpetuación de los “buenos” diseños o de las “buenas” características de generación en generación. Sin embargo, cuando se hace muy grande el valor de la constante de reproducción disminuye considerablemente la variabilidad y el algoritmo comienza a depender demasiado de la localización de los mejores individuos de las primeras generaciones, pudiendo caer rápidamente en mínimos locales (como en el caso de los valores X1 y X3) dando resultados no muy buenos, o

especializarse en un mínimo global (como los valores X2 y X4) dando muy buenos resultados. Este factor hace que la variabilidad para valores grandes de constante de reproducción sea, al igual que para valores pequeños, bastante alta.

Aunque el ANOVA arrojó resultados negativos (no hay diferencia significativa en el resultado del algoritmo debido a los cambios en la constante de reproducción), basados en la siguiente gráfica se considera conveniente tomar la constante de reproducción alrededor de 3. Se debe tener en cuenta que si se utiliza una menor probabilidad de mutación (0.05 o 0.025) tomar una constante de reproducción de 3 puede disminuir demasiado la variabilidad en los individuos de cada generación, haciendo conveniente en este caso utilizar una constante de reproducción no tan alta (1.5, por ejemplo).



### Influencia del Tamaño de la Población

Al querer determinar si el tamaño de la población influía sobre el resultado del algoritmo se encontró un problema obvio: si se disminuye o aumenta el tamaño de la población y se mantienen constantes los demás parámetros la cantidad de individuos que se va a evaluar durante el algoritmo disminuirá o aumentará consecuentemente. Así, no se sabrá realmente

si la variación fue inducida porque cambiar el tamaño de población hace más o menos eficiente la búsqueda o si fue inducida porque simplemente había más o menos individuos sobre los cuales se realizaba la búsqueda.

Por tanto, se realizó un ajuste complementario al tamaño de población: si se modificaba éste se modificaba el número de generaciones de tal forma que el número total de individuos del algoritmo fuera constante. Así, se evaluaron los siguientes pares de tamaño de población/número de generaciones: 25/200; 50/100; 75/67; 100/50.

#### Variación Costo Diseño según

#### Tamaño de Población

Tamaño de Población	25	50	75	100
X1	2.20E+08	2.10E+08	2.16E+08	2.18E+08
X2	2.22E+08	2.20E+08	2.20E+08	2.12E+08
X3	2.22E+08	2.22E+08	2.20E+08	2.22E+08

#### Análisis de varianza de un factor

$\alpha =$  0.05

#### RESUMEN

Grupos	Cuenta	Suma	Promedio	Varianza
25	3	6.63E+08	2.21E+08	1.30E+12
50	3	6.52E+08	2.17E+08	3.66E+13
75	3	6.56E+08	2.19E+08	4.67E+12
100	3	6.52E+08	2.17E+08	2.15E+13

#### ANÁLISIS DE VARIANZA

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	2.81E+13	3	9.37E+12	0.59	0.6415	4.07
Dentro de los grupos	1.28E+14	8	1.60E+13			
Total	1.56E+14	11				

Como se puede observar,  $F < F_{\text{crítico}}$ , con lo que, con 95% de confiabilidad, se puede asegurar que no hay una influencia real del tamaño de población sobre el resultado del algoritmo. Dado que el resultado es negativo no hay necesidad de proceder a un análisis de contrastes.

#### Análisis

Se observaron resultados similares en lo que se refiere a Probabilidad de Mutación y Constante de Reproducción, operadores cuyo fin es aumentar o disminuir la diversidad

dentro del algoritmo. Cuando se toman valores que tienden a aumentar mucho la diversidad (constante de reproducción baja y probabilidad de mutación alta), los mejores diseños de cada generación no pesan lo suficiente como para dirigir al algoritmo hacia los mínimos (locales o globales) donde se localizan. Sin embargo, si se toman valores que disminuyen demasiado la diversidad, el algoritmo tiende a arrojar resultados que son buenos pero no son óptimos. Al tender demasiado rápido hacia los mínimos donde se localizan los mejores diseños de las primeras generaciones se obvian otros mínimos que podrían llevar a mejores resultados.

Por otro lado, se encontró que el peso relativo del déficit de presión era un factor importante para reducir efectivamente el espacio de búsqueda del algoritmo al conjunto factible (conjunto donde se satisfacen todas las restricciones, en este caso, de presión mínima en cada uno de los nodos). Sin embargo, una vez el peso relativo cumplía con este objetivo, aumentarlo subsecuentemente no afectaba el desempeño del algoritmo.

Finalmente, se halló que era irrelevante el tamaño de población (mientras no se tomen valores extremadamente bajos) siempre que se mantuviera igual el número total de individuos que se evaluarán en el algoritmo. No se puede descartar que esta característica dependa del tamaño de la red que se está diseñando, y que para redes más grandes sea importante tener poblaciones grandes al menos en las primeras generaciones.

## **11.2. Desempeño sobre redes de mayor tamaño**

El algoritmo se probó sobre redes con un mayor número de nodos/tubos, como la red Cazuca.red (145 nodos/150 tubos), la red top1/r1.red (41/42) y la red top3/r28 (39/67). Ambas redes se encuentran en el anexo 1.

El principal problema que se encontró al tratar de diseñar sobre estas redes fue un consumo excesivo de memoria por parte del programa. Al parecer, hay una fuga (“leak”) de memoria en algún punto del algoritmo, por lo que a medida que corre el algoritmo ésta se

va agotando hasta, en ciertos casos, agotar los recursos del sistema. Esto sucedía en la red Cazucá después de aproximadamente 30 generaciones de 50 individuos, en la red top1/r1.red después de 95 generaciones de 50 individuos.

El programa llegó a soluciones factibles en los casos de las redes top1/r1 y top3/r28. Sólo para la red top1/r1 fue necesario ajustar el peso relativo del déficit de presión a 1000 (en vez de 100) para que el diseño se realizara sobre el conjunto factible. En el siguiente numeral se encuentran los planos de presiones de ambas redes.

Para el caso de la red top3/r28 se realizaron 3 corridas del algoritmo, cada una de 50 generaciones y 50 individuos por generación. En ninguna de ellas el diseño resultado estaba “saturado” en ninguna de las restricciones, es decir, la presión de ningún nodo estaba cerca de la presión mínima de 15 mca. Esto y el hecho que los mejores diseños se encontraban en las últimas generaciones del algoritmo, hace pensar que teniendo mayores recursos computacionales para correr el programa durante más generaciones mejorarían estos diseños.

Sobre la red top1/r1 se realizaron solo dos diseños. El mejor de ellos tampoco se encontraba “saturado”, pero sí se encontraba más cerca de las restricciones, con la menor presión de nodo en 35 mca.

Para la red de Cazucá, no se encontraron diseños dentro del conjunto factible en las 5 pruebas que se realizaron y documentaron. De hecho, los resultados obtenidos mejoraban bastante cuando se aumentaba considerablemente el peso relativo del déficit de presión, lo que hace pensar que el valor óptimo de este factor puede depender del tamaño de la red. Así mismo, es posible que para diseñar efectivamente una red de este tamaño se requiera una población mucho más grande para cubrir de manera más completa el espacio de búsqueda y poder localizarse dentro del conjunto factible desde las primeras generaciones. El número de generaciones que se podían correr era una restricción considerable al desempeño del algoritmo.

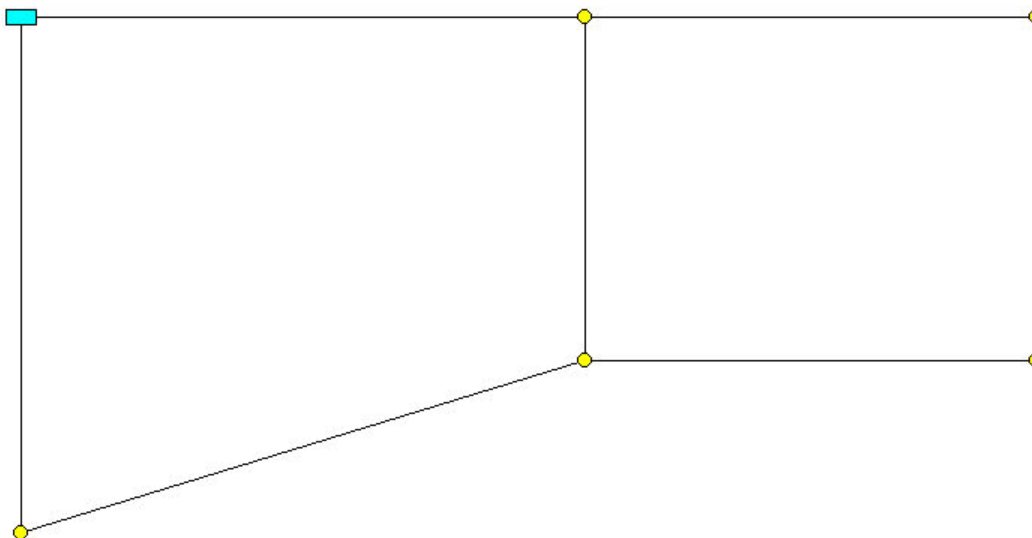
### 11.3. Superficies de presión óptimas

A continuación se presentarán las superficies de presiones obtenidas para las redes Ejemplo.red, top1/r1.red y top3/r28.red. Para una mejor percepción de la superficie de presión, se presentan dos vistas por cada diseño. Vale la pena remarcar que las tres redes que se presentan son planas en su topografía.

Las imágenes fueron realizadas con una versión beta de REDES desarrollada en el CIACUA por el ingeniero Germán Villalba<sup>8</sup>.

#### Red Ejemplo.red

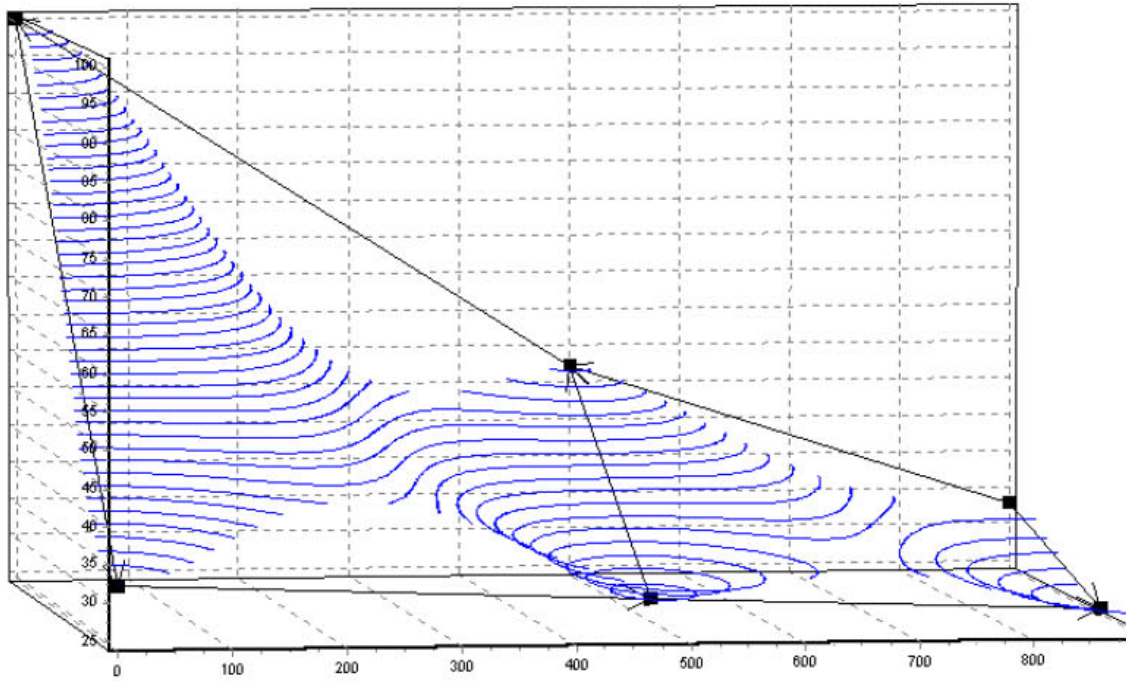
Se presentan las superficies de presiones obtenidas para el mejor de los diseños (el diseño encontrado en la prueba número 21 – ver Anexo 3) y para el quinto mejor diseño (encontrado en la prueba 58), que era el primero que difería relativamente en su superficie con el mejor diseño (las superficies de presión del segundo al cuarto mejor diseño eran bastante similares a la del mejor diseño). Como se puede apreciar, la superficie del mejor diseño se encuentra por debajo de lo que sería un plano entre el embalse (con cabeza de 100m) y el menor nodo (con cabeza de 25m). Por el contrario, la superficie del quinto mejor diseño se encuentra por encima de este plano.



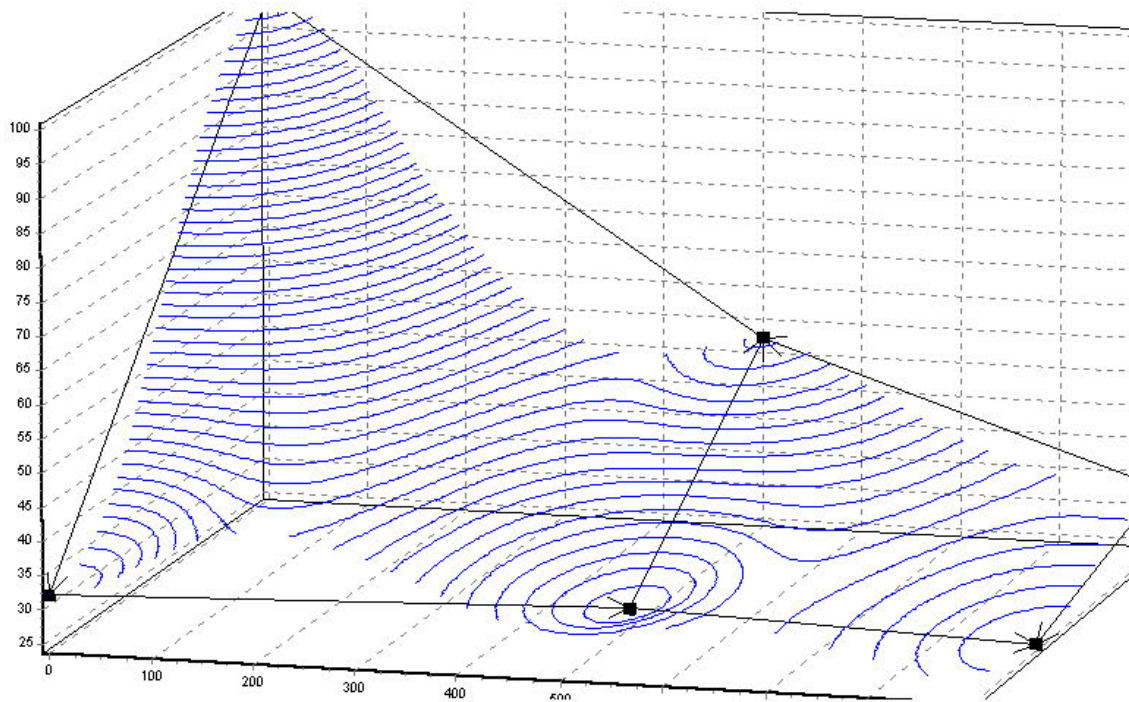
**Red Ejemplo.red**

<sup>8</sup> Ingeniero Germán Villalba, g-villal@uniandes.edu.co



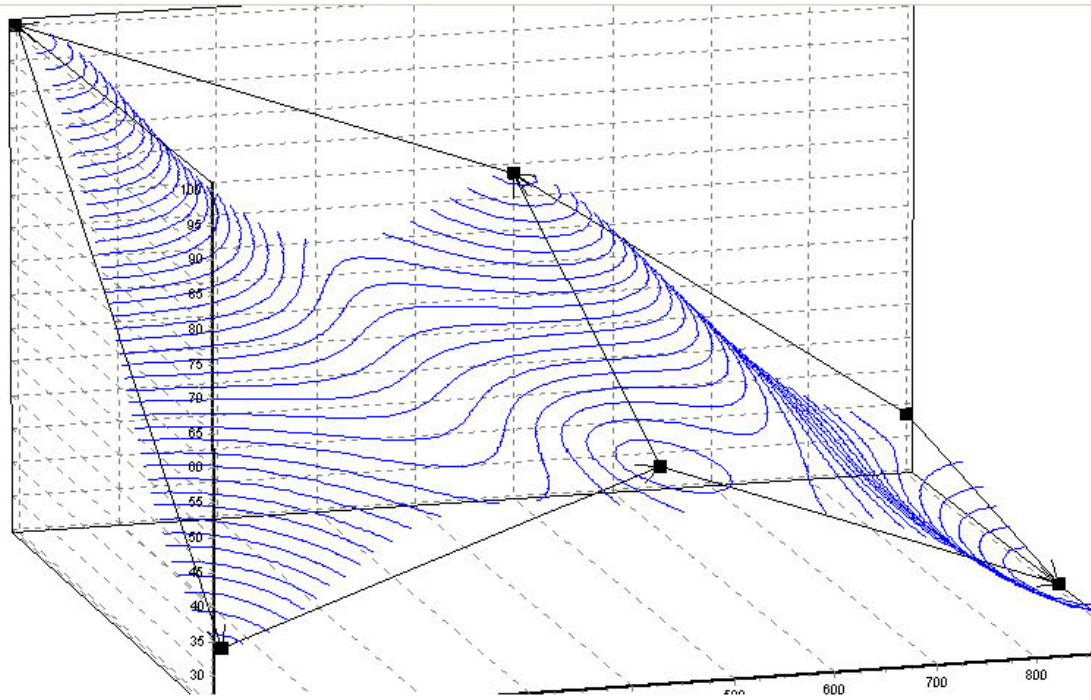


**Mejor Diseño, prueba 21. Primera Vista.**

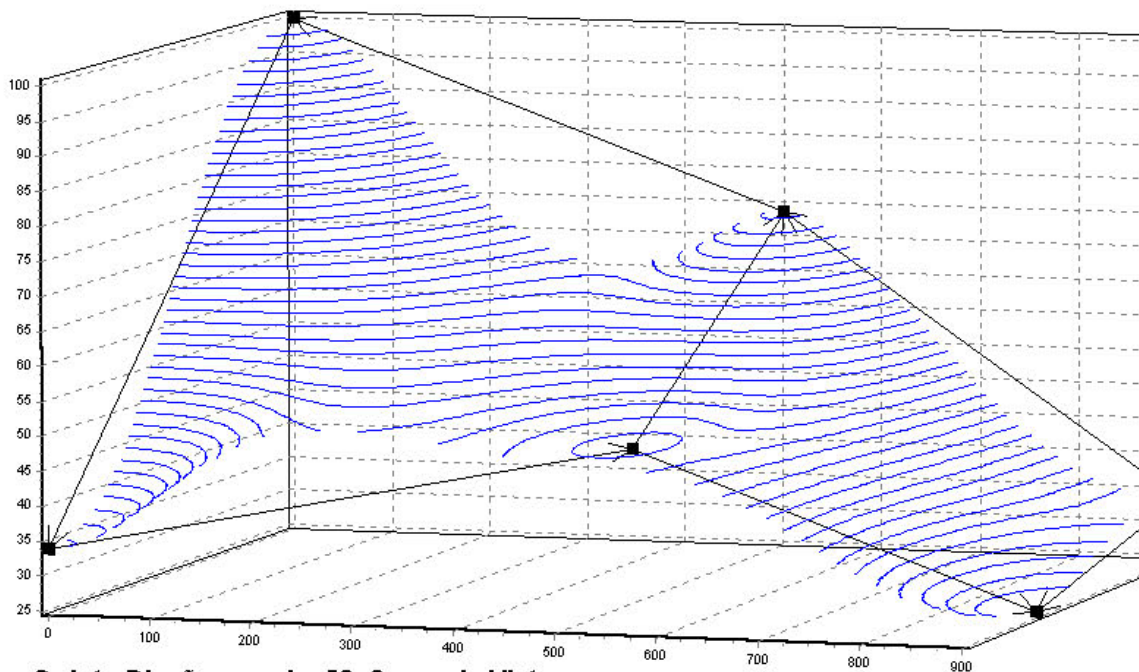


**Mejor Diseño, prueba 21. Segunda Vista**





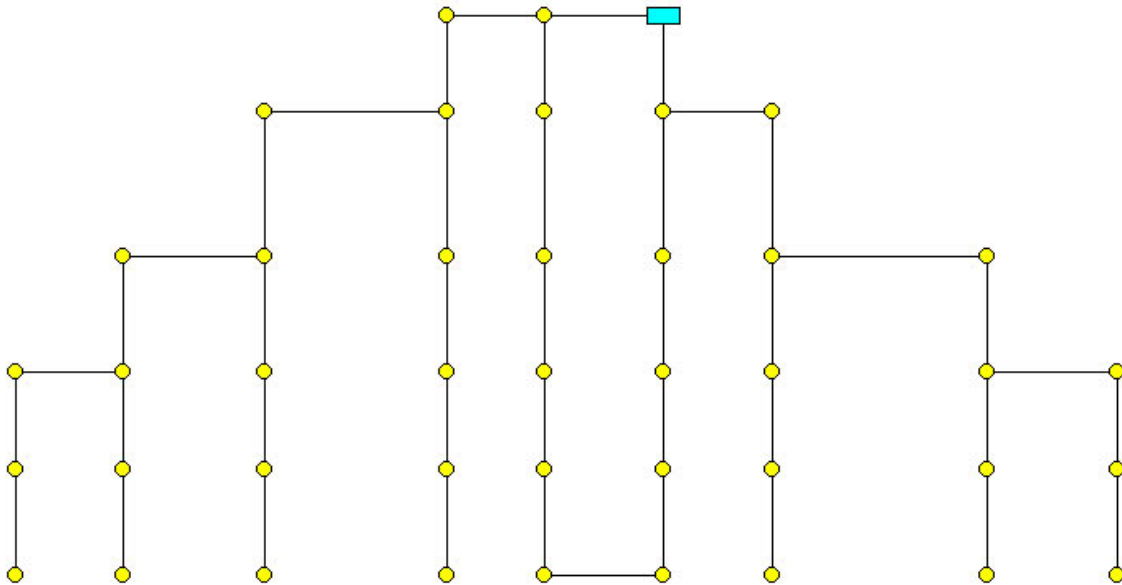
**Quinto Diseño, prueba 58. Primera Vista**



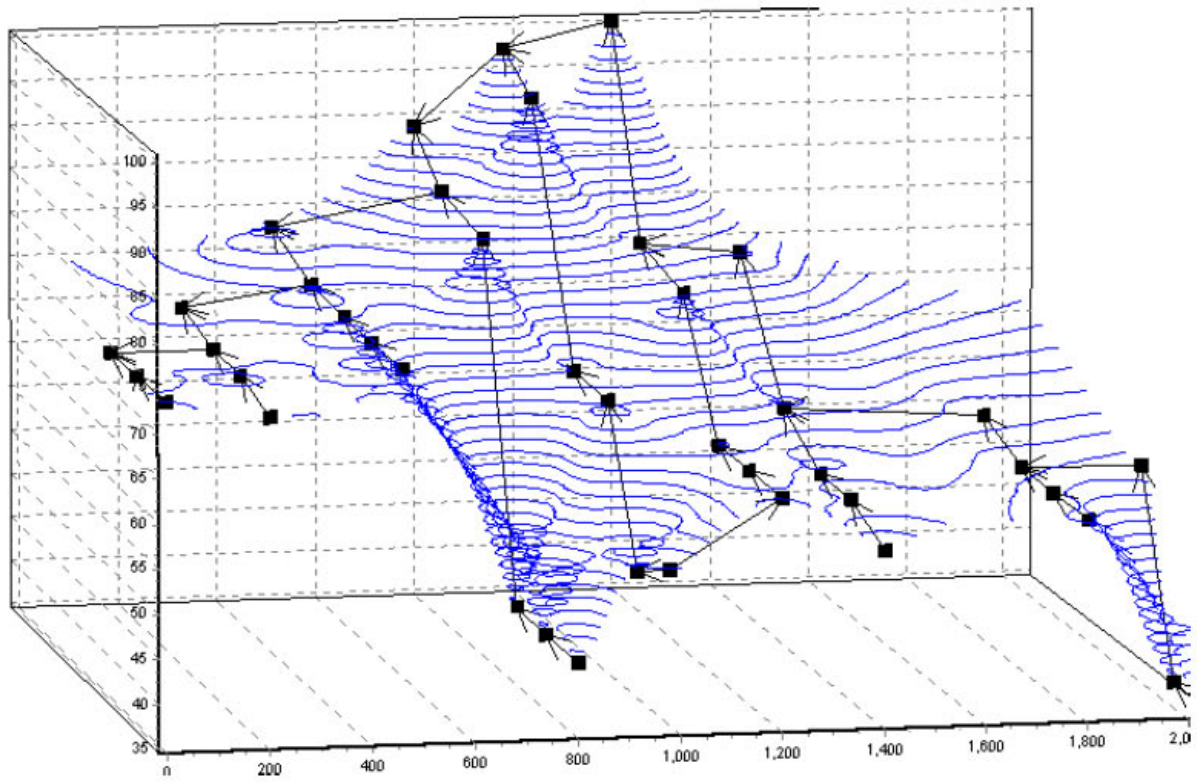
**Quinto Diseño, prueba 58. Segunda Vista**

### Red top1/r1.red

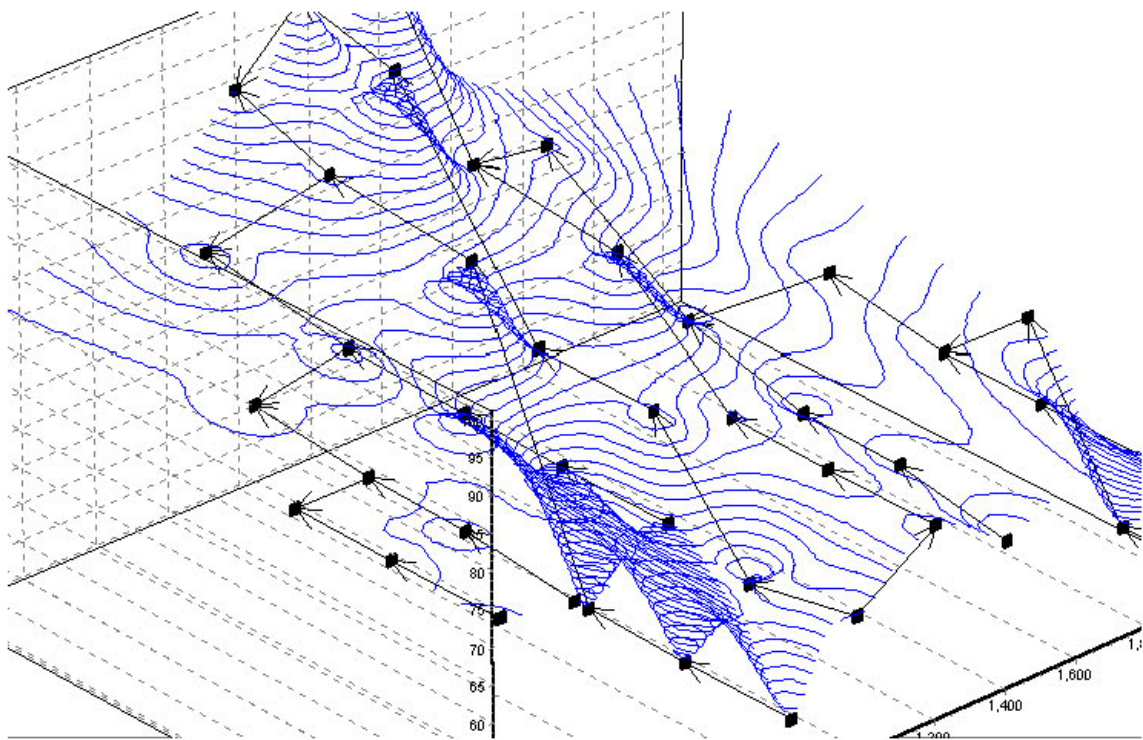
Se presentan las superficies de presiones obtenidas para el mejor de los diseños encontrados, en la segunda de las pruebas hechas. Como se mencionó en la sección 11.2, las restricciones de presión no están saturadas, la menor de ellas encontrándose en 35 mca (nodo más a la derecha, abajo). Las presiones parecen seguir una especie de “plano” sostenido entre el embalse y el nodo más alejado a cada uno de los lados. Sin embargo, estos resultados no son tan significativos ya que es probable que este diseño se encuentre lejos de lo que sería el diseño óptimo.



*Red top1/r1.red*



*Mejor Diseño, Primera Vista*

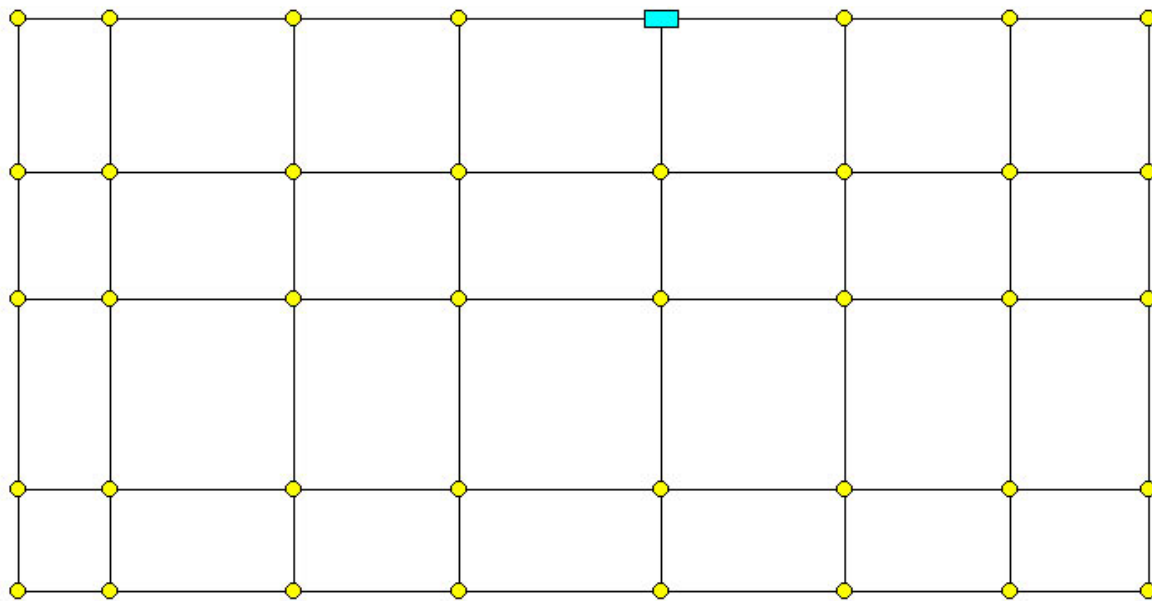


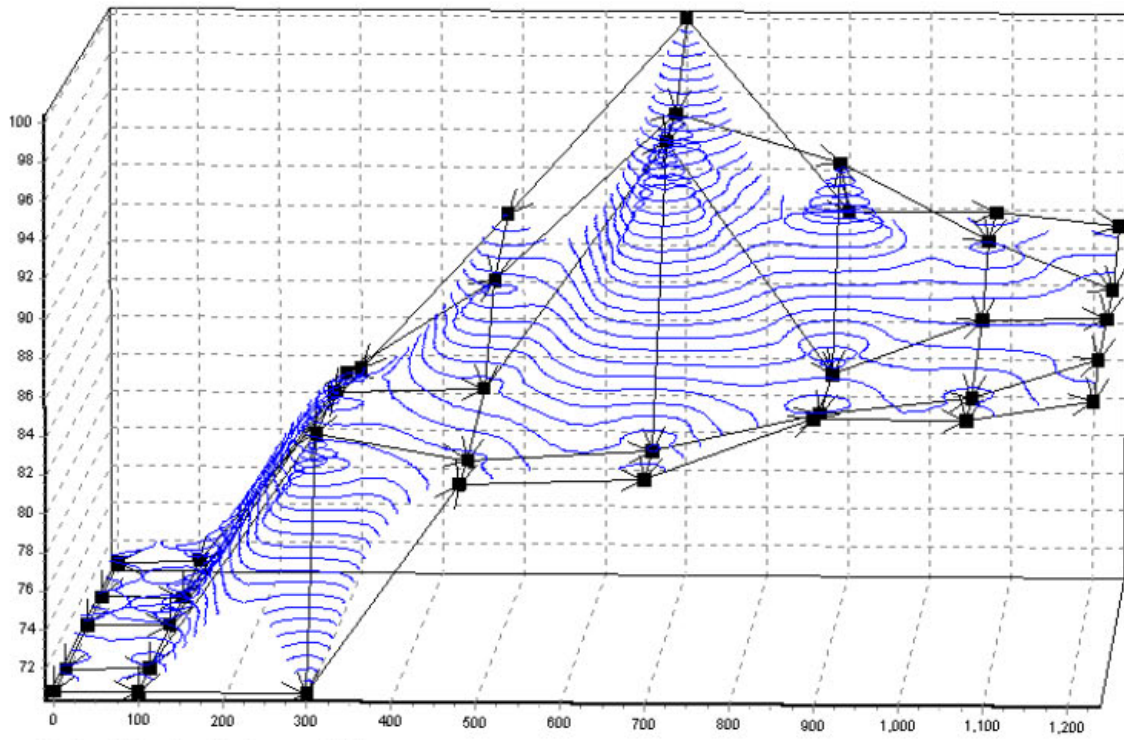
*Mejor Diseño, Segunda Vista*



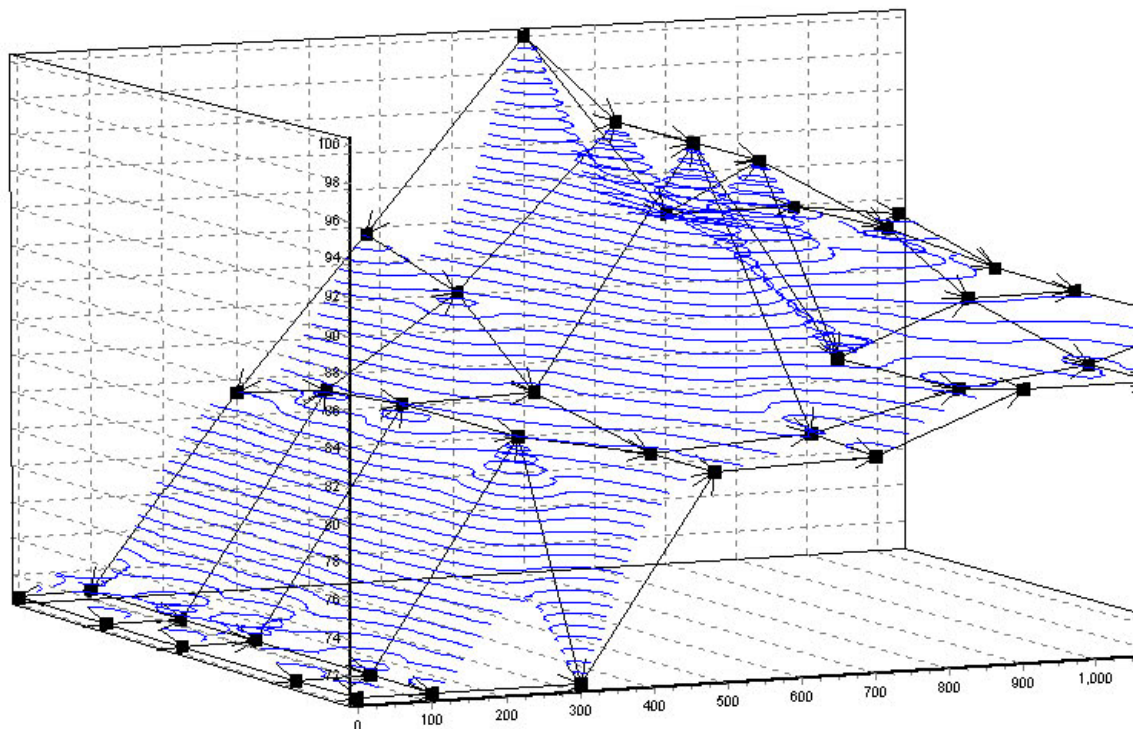
**Red top3/r28.red**

Se presentan las superficies de presiones obtenidas para el mejor diseño encontrado, en la tercera de las pruebas hechas. Como se mencionó en el numeral 11.2 y como se puede apreciar en las figuras, las restricciones de presión no están del todo saturadas. Las más bajas están en alrededor de 70 mca, cuando la restricción de presión estaba en 15 mca. La superficie de presión pareciera seguir un plano entre el embalse y los nodos más alejados a cada lado; de nuevo es probable que el diseño se encuentre lejos de lo que puede ser el diseño óptimo siendo estos resultados no tan significativos.

*Red top3/r28.red*



*Mejor Diseño, Primera Vista*



*Mejor Diseño, Segunda Vista*

## Análisis

El criterio de diseño de Featherstone, estudiado extensamente para diseño de tuberías en serie, sostiene que el diseño más económico es el logrado cuando los diámetros de las tuberías se toman de tal manera que la línea de gradiente hidráulico se encuentre ligeramente por debajo de una línea sostenida entre la tubería inicial (con altura igual a la cabeza inicial) y la tubería final (con altura igual a la cabeza final requerida)<sup>9</sup>. Se ha planteado la posibilidad de que este criterio, extendido a redes de distribución, siga siendo efectivo en el diseño de redes de mínimo costo. Por supuesto, ya no se hace referencia a una “línea” de presión sino a un plano de presión sostenido entre el embalse (con altura igual a la cabeza del embalse) y el nodo más alejado de éste (con altura igual a la cabeza mínima requerida).

En la red Ejemplo.red la similitud del plano de presiones del mejor diseño encontrado con el criterio de Featherstone es impresionante. Efectivamente, las presiones en los nodos se encuentran ligeramente por debajo de lo que sería un plano sostenido entre el nodo correspondiente al embalse y el nodo más alejado de éste. Más aún, el siguiente diseño que difiere de esta configuración tiene un mayor costo y se encuentra ligeramente *por encima* de este plano.

Esta similitud permanece en los diseños de las redes top1/r1 y top3/r28. En la red top3/r28 se puede observar el plano de manera mejor definida. Es sin embargo lamentable que no se puedan estudiar con mayor detalle los diseños de estas dos redes dado el problema de memoria actual del programa para poder ratificar o desmentir la similitud con el criterio de Featherstone de manera más contundente.

---

<sup>9</sup> Este criterio se explica con detalle en [Saldarriaga, 2001].

## 12.BIBLIOGRAFIA

---

[Faur, 2000-1] Faur, C. (2000), "Méthodologie Experimentale, Analyse de la Variance, Regression Linéaire Multiple", Nantes, Francia: Notas de Curso, Ecole des Mines de Nantes.

[Saldarriaga, 2001] Saldarriaga, Juan, (2001). "Hidráulica de Tuberías". Bogotá, Colombia: McGraw-Hill Interamericana S.A.

## CONCLUSIONES

- Se considera que la adaptación de los algoritmos genéticos al problema de diseño óptimo de redes de distribución de agua potable fue exitoso. La codificación intrínseca al problema de diseño (las cadenas de diámetros que integran el diseño) y el hecho que *dentro de las aplicaciones prácticas reales* se diseñe sobre un conjunto discreto (y obviamente finito) de diámetros parece hacer este problema naturalmente apropiado a los algoritmos genéticos. Particularmente, el operador de mutación se adapta bien a esta codificación afectando de manera importante los resultados que se obtienen en el algoritmo.
- Aunque, por ejemplo, de las 60 corridas que se hicieron sobre la red Ejemplo1.red en la mayoría de los casos se encontraban resultados por encima del mínimo encontrado (que se encontró en solo dos de estas corridas), en general, una vez se encontraban las variables de algoritmo que optimizaban el desempeño para esta red, los resultados se localizaban todos cerca de este mínimo. Como se afirmó inicialmente, los algoritmos genéticos no son un método de optimización ya que no se asegura encontrar el mínimo global y, ni siquiera, un mínimo local; sin embargo, una vez se adaptó el algoritmo a la red que se pretende diseñar los resultados son consistentemente buenos. El programa no es por tanto una solución inmediata al problema de diseño, ya que primero se debe adaptar éste a la red que se pretenda diseñar.
- Este programa se convierte en una herramienta útil para el diseño sin sustituir al diseñador. Es decir, sigue siendo necesario el uso del criterio al diseñar ya que el único criterio de diseño no es el costo. El diseñador debe evaluar según su criterio los resultados que se obtengan del algoritmo, ya que éste no tiene en consideración factores como la existencia de tuberías madres que deben tener diámetros considerablemente mayores por seguridad, entre muchos otros factores involucrados en el diseño de un acueducto.



- Los resultados encontrados por el algoritmo en las redes que se estudiaron parecen confirmar, o al menos “amistar” con el criterio de diseño conocido como el criterio de Featherstone. Aunque esta apreciación se basa en un número limitado de resultados, este comportamiento, muy interesante desde el punto de vista teórico, merece ser analizado con mayor profundidad ya que se puede constituir en un soporte estadístico contundente a una teoría que no ha sido demostrada de manera definitiva.
- Se espera que, con desarrollos posteriores que hagan del programa una herramienta más robusta, éste se constituya en un aporte al desarrollo del país. El programa debe permitir mayor investigación en cuanto a los factores que optimizan el diseño de redes de manera analítica, ya que el programa como tal no arroja luz respecto a ello, siendo una “caja negra” de diseño. Además, éste se debe constituir en un aporte a las entidades encargadas del diseño de redes de acueducto, haciendo posible el diseño de acueductos a menores precios y el aumento de la cobertura de agua potable a nivel nacional.

## RECOMENDACIONES – TRABAJOS FUTUROS

Durante el desarrollo de la tesis fueron muchos los comentarios recibidos respecto a las posibles mejoras y otras aplicaciones que se le podían dar al trabajo desarrollado. Entre ellas están la implementación de nuevos operadores para el algoritmo y la realización de modificaciones ligeras para poder aplicarlo en otras áreas. Se mencionan acá las propuestas que el autor considera deberían ser seguidas.

- Durante el desarrollo del programa no se alcanzó a implementar el operador de recombinación hiperplanar. Se espera que este operador mejore el desempeño del algoritmo de manera considerable, en especial sobre redes de gran tamaño haciendo más “ordenada” y consistente la recombinación.
- La adaptación de un método de optimización local al algoritmo podría mejorar considerablemente su desempeño. Tal método actuaría sobre el diseño resultante del algoritmo, modificándolo para mejorar su función de costo de tal manera que convierta el diseño en un “mínimo” local. Un método como el del gradiente sería fácilmente adaptable, consistiendo en cambiar sistemáticamente los diámetros de la red por el diámetro anterior de los diámetros de diseño, hasta que la disminución de cualquiera de los diámetros resulte en incumplir las restricciones de presión en alguno de los nodos.
- Se puede investigar mucho para desarrollar mejores métodos para asegurar el cumplimiento de las restricciones. El actual, asignándole un peso considerablemente grande al déficit de presión funciona relativamente bien pero podría ser mejorado. Se podría, por ejemplo, en una generación en la que ya se tenga una porción considerable de los individuos cumpliendo las restricciones, eliminar todos los individuos que no las cumplan de ahí en adelante y tomar como costo hidráulico la suma de las diferencias (positivas) de la cabeza en el nodo con la cabeza mínima. Así, el mejor individuo será el

que tenga menor costo constructivo y se encuentre más cerca de las restricciones de presión.

- Podría ser conveniente variar la constante de reproducción a medida que avanza el algoritmo. Esta podría ser una función del número de generación y de la desviación estándar de la probabilidad sin escalar. Esto con el fin que, a medida que las generaciones avanzan y se hacen más homogéneas se de menos peso al mejor individuo para que vuelvan a tomar diversidad y exploren otros sectores del espacio de búsqueda. Algo similar se puede plantear con la probabilidad de mutación.
- Se pueden hacer grandes mejoras a la función de costo constructivo. Esta podría incluir, por ejemplo, el costo de accesorios cuando hay un cambio de diámetro en la tubería y, en especial, los costos por fugas. Esta adaptación requeriría trabajo considerable inclusive sobre el módulo de cálculo de REDES, pero se constituiría en una mejora apreciable siendo los costos por fugas uno de los principales problemas de las redes de distribución hoy en día.
- La restricción de presión puede no ser única para cada nodo. Esto permitiría diseñar tratando de seguir una superficie de presión predeterminada, lo que permitiría al diseñador influir el algoritmo con su criterio y experiencia de cómo suelen ser los diseños óptimos. Con esta modificación se podría hacer tender al algoritmo para diseñar, por ejemplo, con tubos madre de diámetro alto para que la pérdida de presión allí no fuera muy alta.
- Siguiendo la idea del punto anterior se podría, con una ligera modificación del algoritmo que ya no tuviera en cuenta el costo constructivo, calibrar los diámetros de la red para que ésta tendiera a acercarse a un plano de presión dado. Esto sería útil en el contexto de calibración de redes, área en la cual se desarrolla investigación extensiva en este momento.

- Como se menciona en las Conclusiones, se debe seguir la investigación en el sentido de reafirmar o desmentir el criterio de Featherstone en base a los resultados del algoritmo genético sobre varias redes. Esto se constituiría en un avance importante en cuanto a la teoría de diseño de redes de distribución. Así mismo, dado que tanto el programa REDES como el programa resultante de esta tesis lo permiten, el diseño sobre redes con topografía no uniforme debe ser analizado.

## ANEXOS

### 1. REDES EJEMPLO.RED, TOP1/R1.RED, TOP3/R28.RED.

#### Ejemplo.red

[UNION]

<ID	X	Y	Z>
2	500	300	0
3	900	300	0
4	900	100	0
5	500	100	0
6	0	0	0
1	0	300	0

[FIN UNION]

[NODO]

<ID	DEMANDA>
2	60
3	40
4	30
5	30
6	40

[FIN NODO]

[EMBALSE]

<ID	LGH>
1	100

[FIN EMBALSE]

[TUBO]

<ID	NODO1	NODO2	LONGITUD	DIAMETRO	RUGOSIDAD
COEF_MENORES>					
1	2	1	500	0.254	6E-5 0
2	3	2	400	0.1524	6E-5 10
3	3	4	200	0.1016	6E-5 0
4	5	4	400	0.1524	6E-5 0
5	5	2	200	0.1016	6E-5 0
6	6	5	600	0.2032	6E-5 0
7	1	6	300	0.254	6E-5 0

[FIN TUBO]

[UNIDADES]

<ID	NOMBRE	UNIDAD>
-----	--------	---------

```

1   NODO_X      Metros
2   NODO_Y      Metros
3   NODO_Z      Metros
4   NODO_DEMANDA  Lps
5   EMBALSE_X  Metros
6   EMBALSE_Y  Metros
7   EMBALSE_Z  Metros
8   EMBALSE_LGH  Metros
9   TUBO_DIAMETRO  Metros
10  TUBO_LONGITUD  Metros
11  TUBO_RUGOSIDAD  Metros
[FIN UNIDADES]

```

### Top1/r1.red

```

[UNION]
<ID  X      Y      Z>
1   0      0      0
2   0      220   0
3   0      420   0
4   200    0      0
5   200    220   0
6   200    420   0
7   200    660   0
8   460    0      0
9   460    220   0
10  460    420   0
11  460    660   0
12  460    960   0
13  800    0      0
14  800    220   0
15  800    420   0
16  800    660   0
17  800    960   0
18  800    1160  0
19  980    0      0
20  980    220   0
21  980    420   0
22  980    660   0
23  980    960   0
24  980    1160  0
25  1200   0      0
26  1200   220   0
27  1200   420   0
28  1200   660   0

```

29	1200	960	0
30	1400	0	0
31	1400	220	0
32	1400	420	0
33	1400	660	0
34	1400	960	0
35	1800	0	0
36	1800	220	0
37	1800	420	0
38	1800	660	0
39	2040	0	0
40	2040	220	0
41	2040	420	0
42	1200	1160	15

[FIN UNION]

[NODO]

<ID	DEMANDA>
1	1.4
2	2.7
3	1.3
4	3.3
5	6.3
6	5
7	2
8	4.3
9	8.2
10	8.6
11	8
12	3.3
13	3.7
14	7.1
15	7.4
16	9.1
17	6.3
18	1.2
19	2.9
20	5.5
21	5.7
22	7
23	6.5
24	4
25	3
26	5.7
27	6
28	7.4
29	5.5

```

30  4.3
31  8.2
32  8.6
33  6.6
34  2
35  4.6
36  8.8
37  7.3
38  3.1
39  1.7
40  3.3
41  1.6
[FIN NODO]
    
```

```

[EMBALSE]
<ID  LGH>
42  15
[FIN EMBALSE]
    
```

```

[TUBO]
<ID  NODO1  NODO2  LONGITUD  DIAMETRO  RUGOSIDAD
COEF_MENORES>
1  1  2  220  76.2  0.0015  0
2  2  3  200  101.6  0.0015  0
3  4  5  220  101.6  0.0015  0
4  5  6  200  152.4  0.0015  0
5  6  7  240  203.2  0.0015  0
6  8  9  220  101.6  0.0015  0
7  9  10  200  152.4  0.0015  0
8  10  11  240  203.2  0.0015  0
9  11  12  300  254  0.0015  0
10  13  14  220  101.6  0.0015  0
11  14  15  200  152.4  0.0015  0
12  15  16  240  152.4  0.0015  0
13  16  17  300  203.2  0.0015  0
14  17  18  200  304.8  0.0015  0
15  19  20  220  101.6  0.0015  0
16  20  21  200  50.8  0.0015  0
17  21  22  240  101.6  0.0015  0
18  22  23  300  152.4  0.0015  0
19  23  24  200  152.4  0.0015  0
20  25  26  220  152.4  0.0015  0
21  26  27  200  152.4  0.0015  0
22  27  28  240  203.2  0.0015  0
23  28  29  300  203.2  0.0015  0
24  29  42  200  355.6  0.0015  0
25  30  31  220  101.6  0.0015  0
    
```



26	31	32	200	152.4	0.0015	0
27	32	33	240	152.4	0.0015	0
28	33	34	300	254	0.0015	0
29	35	36	220	101.6	0.0015	0
30	36	37	200	152.4	0.0015	0
31	37	38	240	203.2	0.0015	0
32	39	40	220	76.2	0.0015	0
33	40	41	200	101.6	0.0015	0
34	41	37	240	152.4	0.0015	0
35	38	33	400	254	0.0015	0
36	34	29	200	304.8	0.0015	0
37	42	24	220	355.6	0.0015	0
38	24	18	180	355.6	0.0015	0
39	17	12	340	304.8	0.0015	0
40	11	7	260	203.2	0.0015	0
41	6	3	200	152.4	0.0015	0
42	19	25	220	152.4	0.0015	1

[FIN TUBO]

**Top3/r28.red**

[UNION]

<ID	X	Y	Z>
1	0	0	0
2	0	80	0
3	0	230	0
4	0	330	0
5	0	450	0
6	100	450	0
7	100	330	0
8	100	230	0
9	100	80	0
10	100	0	0
11	300	0	0
12	300	80	0
13	300	230	0
14	300	330	0
15	300	450	0
16	480	450	0
17	480	330	0
18	480	230	0
19	480	80	0
20	480	0	0
21	700	0	0
22	700	80	0

23	700	230	0
24	700	330	0
25	900	450	0
26	900	330	0
27	900	230	0
28	900	80	0
29	900	0	0
30	1080	0	0
31	1080	80	0
32	1080	230	0
33	1080	330	0
34	1080	450	0
35	1230	450	0
36	1230	330	0
37	1230	230	0
38	1230	80	0
39	1230	0	0
40	700	450	15

[FIN UNION]

[NODO]

<ID	DEMANDA>
1	0.3
2	0.7
3	0.8
4	0.7
5	0.4
6	1.2
7	2.1
8	2.4
9	2.2
10	0.8
11	1
12	2.8
13	3.1
14	2.7
15	1.5
16	2.4
17	2.9
18	3.3
19	3
20	1
21	1.1
22	3.1
23	3.4
24	3
25	2.3

```

26  2.7
27  3.1
28  2.8
29  1
30  0.9
31  2.5
32  2.7
33  2.4
34  1.3
35  0.6
36  1.1
37  1.2
38  1.1
39  0.4
[FIN NODO]

```

```

[EMBALSE]
<ID  LGH>
40  100
[FIN EMBALSE]

```

```

[TUBO]
<ID  NODO1      NODO2      LONGITUD  DIAMETRO  RUGOSIDAD
   COEF_MENORES>
1   2    1    80    25.4 0.0015    0
2   3    2   150   101.6 0.0015    0
3   4    3   100   76.2 0.0015    0
4   5    4   120   177.8 0.0015    0
5   6    5   100   101.6 0.0015    0
6   6    7   120   152.4 0.0015    0
7   7    8   100   101.6 0.0015    0
8   8    9   150   203.2 0.0015    0
9   9   10   80    50.8 0.0015    0
10  15   6   200   101.6 0.0015    0
11  15  14   120   203.2 0.0015    0
12  14  13   100   101.6 0.0015    0
13  13  12   150   203.2 0.0015    0
14  12  11   80    50.8 0.0015    0
15  16  15   180   254 0.0015    0
16  16  17   120   228.6 0.0015    0
17  17  18   100   152.4 0.0015    0
18  18  19   150   50.8 0.0015    0
19  19  20   80    101.6 0.0015    0
20  40  16   220   228.6 0.0015    0
21  40  24   120   254 0.0015    0
22  24  23   100   203.2 0.0015    0
23  23  22   150   177.8 0.0015    0

```

24	22	21	80	101.6	0.0015	0
25	40	25	200	254	0.0015	0
26	25	26	120	203.2	0.0015	0
27	26	27	100	228.6	0.0015	0
28	27	28	150	254	0.0015	0
29	28	29	80	228.6	0.0015	0
30	25	34	180	254	0.0015	0
31	34	33	120	152.4	0.0015	0
32	33	32	100	127	0.0015	0
33	32	31	150	76.2	0.0015	0
34	31	30	80	228.6	0.0015	0
35	34	35	150	152.4	0.0015	0
36	35	36	120	50.8	0.0015	0
37	36	37	100	177.8	0.0015	0
38	37	38	150	25.4	0.0015	0
39	38	39	80	25.4	0.0015	0
40	30	39	150	254	0.0015	0
41	21	29	200	127	0.0015	0
42	20	21	220	177.8	0.0015	0
43	11	20	180	177.8	0.0015	0
44	1	10	100	228.6	0.0015	0
45	10	11	200	127	0.0015	0
46	29	30	180	254	0.0015	0
47	2	9	100	127	0.0015	0
48	31	38	150	254	0.0015	0
49	9	12	200	25.4	0.0015	0
50	28	31	180	101.6	0.0015	0
51	12	19	180	203.2	0.0015	0
52	22	28	200	127	0.0015	0
53	19	22	220	254	0.0015	0
54	3	8	100	254	0.0015	0
55	32	37	150	76.2	0.0015	0
56	8	13	200	76.2	0.0015	0
57	27	32	180	76.2	0.0015	0
58	13	18	180	228.6	0.0015	0
59	23	27	200	177.8	0.0015	0
60	18	23	220	152.4	0.0015	0
61	4	7	100	101.6	0.0015	0
62	33	36	150	76.2	0.0015	0
63	7	14	200	254	0.0015	0
64	26	33	180	25.4	0.0015	0
65	14	17	180	127	0.0015	0
66	24	26	200	254	0.0015	0
67	17	24	220	228.6	0.0015	0

[FIN TUBO]

## 2. EJEMPLO DE ARCHIVO DE ESTADÍSTICAS

Se presenta el archivo de estadísticas del diseño de la red Ejemplo.red, en una prueba no incluida en el Anexo 3.

### Resultados Algoritmo Genético

Diametros (mm): 0.0254 0.0508 0.0762 0.1016 0.127 0.1524 0.1778 0.2032 0.228 0.254  
 Presión: 15  
 Min (m):  
 PopSize: 100  
 Generaciones: 15  
 Peso Relativo: 100  
 Déficit Presión:  
 Prob Mut: 0.05  
 Probabilidad Escalada Const: 1.5  
 Reprod: Recombinación Simple

### Estadísticas Generacionales

Generación	MConstructiva	SConstructiva	MHidra	SHidra	SAdaptabilidad	MNodos Malos	Individuos Malos	Mejor Individuo Generación:		
								Constructivos	Hidra	Nodos Malos
1	258115153	51582875	87199	457663	525	3.36	83	239612423	38	1
2	263601165	57530056	60695	446642	736	3	74	238452606	0	0
3	277713794	49600095	46335	366677	791	2.85	73	238452606	0	0
4	279725105	55251669	35944	269861	751	2.5	66	234477705	9	2
5	291171576	53846435	127566	581864	456	2.2	56	254488227	35	4
6	295169984	51874793	66172	379297	573	2.02	56	217656111	82	2
7	298975660	45562370	32119	270081	841	1.7	52	262052537	0	0
8	307910799	43194479	29484	269230	913	1.23	41	260122095	0	0
9	308474578	41413035	3582	24743	691	1.35	43	251133267	0	0
10	304537495	41886064	2907	22610	778	1.23	36	254722247	0	0
11	311786062	40796474	4375	35850	819	0.85	26	250838911	0	0
12	317450679	40410166	4741	31675	668	0.64	19	245499045	0	0
13	318295455	41821706	1014	4510	445	0.8	27	284196837	0	0
14	327143434	32287929	458	4164	908	0.51	15	272849272	0	0
15	326767490	37383001	29741	269284	905	0.52	14	281954726	0	0

### Mejores Diseños

Diseño	Generación	Costo Constructivo	Déficit Presión (m)	Nodos con Déficit Presión
1	3	238452606	0	0
2	2	238452606	0	0
3	2	243336496	0	0

4	4	243691639	0	0
5	12	245499045	0	0
6	11	250838911	0	0
7	9	251133267	0	0
8	2	253020569	0	0
9	10	254722247	0	0
10	4	256453415	0	0

## Presiones Mejor Diseño

Nodo	Presión (m)
0	84.70
1	31.71
2	37.17
3	69.85
4	96.01

### 3. VARIACIÓN DE LOS RESULTADOS CON LAS DISTINTAS VARIABLES

A continuación se presentan los resultados de las pruebas hechas sobre la red Ejemplo.red.

Id Prueba	Tam Población	Num Gen	Wrel Hidra	Prob Mut	C Reprod	Costo Constructivo Mejor Diseño
10	50	100	100	0.1	1.5	2.105E+08
11	50	100	100	0.1	1.5	2.198E+08
12	50	100	100	0.1	1.5	2.218E+08
13	50	100	100	0.1	1.5	2.108E+08
14	50	100	100	0.1	1	2.232E+08
15	50	100	100	0.1	1	2.200E+08
16	50	100	100	0.1	1	2.216E+08
17	50	100	100	0.2	1.5	2.287E+08
18	50	100	100	0.05	1.5	2.016E+08
19	50	100	100	0.025	1.5	2.325E+08
20	50	100	100	0.025	1.5	2.105E+08
21	50	100	100	0.025	1.5	1.960E+08
22	50	100	100	0.2	1.5	2.507E+08
23	50	100	100	0.05	1.5	2.218E+08
24	50	100	100	0.05	1.5	1.997E+08
25	25	200	100	0.1	1.5	2.198E+08
26	25	200	100	0.1	1.5	2.216E+08
27	75	67	100	0.1	1.5	2.161E+08
28	75	67	100	0.1	1.5	2.200E+08
29	100	50	100	0.1	1.5	2.180E+08
30	100	50	100	0.1	1.5	2.124E+08
31	100	50	100	0.1	1.5	2.216E+08
32	50	100	100	0.2	1.5	2.234E+08
33	25	200	100	0.1	1.5	2.219E+08
34	75	67	100	0.1	1.5	2.197E+08
35	50	100	100	0.1	3	2.087E+08
36	50	100	100	0.1	3	1.997E+08
37	50	100	100	0.1	3	1.960E+08
38	50	100	50	0.1	1.5	2.198E+08
39	50	100	50	0.1	1.5	2.160E+08
40	50	100	50	0.1	1.5	2.293E+08
41	50	100	200	0.1	1.5	2.218E+08
42	50	100	200	0.1	1.5	2.215E+08
43	50	100	200	0.1	1.5	2.181E+08
44	50	100	1	0.1	1.5	2.690E+08
45	50	100	100	0.025	1.5	2.069E+08
46	50	100	100	0.05	1.5	2.069E+08
47	50	100	100	0.2	1.5	2.400E+08

48	50	100	1	0.1	1.5	2.787E+08
49	50	100	1	0.1	1.5	2.452E+08
50	50	100	10	0.1	1.5	2.251E+08
51	50	100	5	0.1	1.5	2.233E+08
52	50	100	5	0.1	1.5	2.233E+08
53	50	100	3	0.1	1.5	2.417E+08
54	50	100	3	0.1	1.5	2.289E+08
55	50	100	3	0.1	1.5	2.364E+08
56	50	50	100	0.1	6	2.142E+08
57	50	100	100	0.1	6	2.253E+08
58	50	100	100	0.1	6	1.979E+08
59	50	100	100	0.1	6	2.398E+08
60	50	100	100	0.1	6	2.016E+08
61	50	100	100	0.1	3	2.034E+08
62	50	100	100	0.1	1	3.028E+08

Diseño 48: *No es el mejor diseño. El mejor diseño no cumple con las restricciones de presión mínima*

Diseño 62: *Ni este ni ningún otro de los 10 mejores diseños cumple con las restricciones de presión mínima*

Mejores diseños	Costo	Id Prueba
1	1.960E+08	21
2	1.960E+08	37
3	1.979E+08	58
4	1.997E+08	24
5	1.997E+08	36
6	2.016E+08	18
7	2.016E+08	60
8	2.034E+08	61
9	2.069E+08	45
10	2.069E+08	45



#### 4. CONTENIDO DEL CD-ROM ADJUNTO

El CD-ROM adjunto contiene:

- Una copia académica (apta para redes de hasta 50 tubos) del programa REDES con el módulo de diseño genético.
- Las redes Ejemplo.red, top1/r1.red, top3/r28.red.