

**MODELO DIGITAL DE ELEVACIÓN PARA LA CUENCA DEL RÍO DE ORO**

**WILSON GONZALO LÓPEZ  
RAFAEL MAURICIO OSORIO AMAYA**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO MECÁNICAS  
ESCUELA DE INGENIERÍA CIVIL  
BUCARAMANGA  
2005**

**MODELO DIGITAL DE ELEVACIÓN PARA LA CUENCA DEL RÍO DE ORO**

**WILSON GONZALO LÓPEZ  
RAFAEL MAURICIO OSORIO AMAYA**

**Proyecto de Grado para optar el título de Ingeniero Civil**

**Director  
Jorge Alberto Guzmán.  
I.C. M.Sc.**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO MECÁNICAS  
ESCUELA DE INGENIERÍA CIVIL  
BUCARAMANGA  
2005**

A mi madre, por su amor y apoyo incondicional  
A Diógenes y Angela por su cariño y amistad  
A Eliana por su eterno amor.

Wilson

A **Dios**, por ser la esencia y luz de mi vida...

A mi madre Emma y mi padre Luis, por creer y apoyar mis sueños...

A mis hermanos y amigos, por respaldar y orientar mi trabajo...

A Elizabeth por todo su amor y comprensión...

**Rafael**

## **AGRADECIMIENTOS**

Agradecemos a todas las personas que intervinieron en la realización de nuestro proyecto, a nuestras familias, profesores y amigos por depositar la confianza necesaria para sacarlo adelante.

Al personal del grupo de predicción y modelamiento hidroclimático de la Universidad Industrial de Santander, especialmente al ingeniero Jorge Guzmán por su incondicionalidad entorno al trabajo. Al ingeniero Edward Niño por sus aportes y disposición laboral.

Igualmente queremos agradecer a entidades como la CDMB y el acueducto metropolitano de Bucaramanga por su cooperación e información suministrada para la elaboración de nuestro trabajo.

## **TABLA DE CONTENIDO**

	PAG.
INTRODUCCIÓN	1
1. UBICACIÓN Y DESCRIPCIÓN GENERAL DE LA ZONA DE ESTUDIO	3
1.1 UBICACIÓN DE LAS CUENCAS DE ESTUDIO	3
1.1.1 Cuenca del río de oro	3
1.1.2 Cuenca de río frío	5
1.1.3 Municipio de bucaramanga	5
1.2 ARCHIVOS DIGITALES DE LAS ZONAS DE ESTUDIO	7
1.2.1 Archivos digitales escala 1:25000	7
1.2.2 Archivos digitales escala 1:2000	9
2. MODELOS DIGITALES DE ELEVACIÓN (DEM)	11
2.1 DEFINICIÓN	11
2.2 CREACIÓN DE UN MODELO DIGITAL DE ELEVACIONES	11
2.3 RECOPIACIÓN DE DATOS	12

2.4	ESTRUCTURA DE UN DEM	13
2.4.1	estructura vectorial	14
2.4.2	estructura ráster	15
2.4.3	comparación entre estructuras ráster y vectorial	15
2.5	SELECCIÓN DEL MÉTODO DE INTERPOLACIÓN	18
2.6	VALIDACIÓN DEL MODELO DIGITAL DE ELEVACIÓN	19
3.	LIMPIEZA Y COMPILACIÓN DE ARCHIVOS .DWG	22
3.1	DESCRIPCIÓN DE ARCHIVOS DIGITALES	22
3.2	SUPRESIÓN DE OBJETOS DUPLICADOS	23
3.3	ELIMINACIÓN DE OBJETOS CORTOS	24
3.4	AGRUPACIÓN DE NODOS	24
3.5	DISOLUCIÓN DE NODOS	25
3.6	SPELL MAP	26
4.	ANÁLISIS DE INTERPOLADORES PARA LA CONSTRUCCIÓN DEL DEM SOBRE EL ÁREA METROPOLITANA DE BUCARAMANGA	30

4.1	ESCOGENCIA DE REGIONES REPRESENTATIVAS PARA EL MUNICIPIO DE BUCARAMANGA	30
4.1.1	Mapas de pendiente	32
4.2	VALIDACIÓN CRUZADA	35
4.3	ALGORITMOS MATEMÁTICOS DE INTERPOLACIÓN	36
4.3.1	Krigeaje ordinario	37
4.3.2	Funciones de base radial	47
4.3.3	Inverso de la distancia	49
4.3.4	Vecino más cercano	51
4.3.5	Triangulación con interpolación lineal	52
4.3.6	Mínima curvatura	53
5	ANÁLISIS DE INTERPOLADOR FINAL PARA EL DEM DE BUCARAMANGA	55
5.1.	DETERMINACIÓN DE LA FUNCIÓN DE BASE RADIAL	55
5.2.	NÚMERO DE PUNTOS NECESARIOS PARA LA VALIDACIÓN CRUZADA	57
5.3.	ESTIMACIÓN DEL FACTOR $R^2$ , RELACIÓN DE	

ANISOTROPÍA Y DIRECCIÓN	58
5.4. PRUEBA DE SENSIBILIDAD PARA EL FACTOR $R^2$	63
5.5. DISTRIBUCIÓN DE ERRORES	64
5.6. VALIDACION DEL DEM PARA BUCARAMANGA	65
5.6.1. Interpolación bilineal	66
5.6.2. Puntos tomados con equipo gps	67
5.6.3. Tamaño de la celda	69
5.7. DEM PARA EL ÁREA METROPOLITANA DE BUCARAMANGA	72
6. CONSTRUCCIÓN DEL DEM PARA LA CUENCA DE RIO DE ORO	74
6.1 ANÁLISIS DE INTERPOLADORES	74
6.1.1 Krigeadó	74
6.1.2 Función de base radial multicuadrática	80
6.1.3 Inverso de la distancia	84
6.1.4 Vecino más cercano	85
6.1.5 Triangulación con interpolación lineal	85

6.2	DISTRIBUCION DE ERRORES	86
6.3	TAMAÑO DE CELDA PARA EL DEM SOBRE RÍO DE ORO	87
6.4	DEM PARA LA CUENCA DE RIO DE ORO	88
6.5	REMOCION DE PITS	89
6.6	RED DE DRENAJES	90
7.	CONCLUSIONES	91
	BIBLIOGRAFIA	93
	ANEXOS	95

## **LISTADO DE FIGURAS**

	PAG.
<i>Figura 1.</i> Limite de la cuenca del río de Oro y sub cuencas	4
<i>Figura 2.</i> Cuenca de río Frío; sub cuenca de la cuenca de río de Oro	5
<i>Figura 3.</i> Municipio de Bucaramanga	6
<i>Figura 4.</i> Planchas requeridas para construcción del DEM (1/25000)	8
<i>Figura 5.</i> Planchas necesarias para la construcción del DEM 1/2000	9
<i>Figura 6.</i> Representación del terreno mediante una estructura TIN	14
<i>Figura 7.</i> Tipos de estructura ráster. Representación en 2D y 3D	15
<i>Figura 8.</i> Malla con los puntos a partir de los cuales se genera el MDE	17
<i>Figura 9</i> Supresión de objetos duplicados	23
<i>Figura 10</i> Eliminación de objetos cortos	24
<i>Figura 11</i> Agrupación de nodos	25
<i>Figura 12</i> Disolución de seudonodos	25
<i>Figura 13</i> Cuadro de dialogo (ACAD 3D 2005) para limpieza de archivos .dwg	26

<i>Figura 14</i>	Procedimiento para la extracción de información de archivos .dxf	28
<i>Figura 15</i>	Regiones representativas del área metropolitana de Bucaramanga	31
<i>Figura 16</i>	Mapa de pendientes para las 10 regiones representativas	32
<i>Figura 17</i>	Variograma	39
<i>Figura 18</i>	Relación de anisotropía para búsqueda de datos muestrales	41
<i>Figura 19</i>	Cuadro de dialogo (surfer V 8.0) para construcción del variograma	42
<i>Figura 20</i>	Variogramas experimentales y teóricos para las 10 regiones representativas	42
<i>Figura 21</i>	Variograma teórico final representativo para construir el modelo por Krigeado	47
<i>Figura 22</i>	RMSE y error máximo para diferentes valores del parámetro $R^2$	49
<i>Figura 23</i>	Efecto “smoothing” o de suavizado en la generación de DEM	50
<i>Figura 24</i>	Radio menor de búsqueda para el método del vecino mas cercano	52
<i>Figura 25</i>	Triangulación por el método de Delaunay	53

<i>Figura 26</i>	Tiempo de computo para las funciones de base radial	57
<i>Figura 27</i>	Número de puntos necesarios para realizar la validación cruzada	58
<i>Figura 28</i>	RMSE Vs factor $R^2$ para diferentes anisotropías en Bucaramanga	59
<i>Figura 29</i>	Cambio del RMSE respecto al radio de anisotropía	63
<i>Figura 30</i>	Prueba de sensibilidad para ajustar el factor $R^2$ en la función de base radial	63
<i>Figura 31</i>	Distribución de valores residuales	65
<i>Figura 32</i>	Interpolación bilineal	67
<i>Figura 33</i>	Puntos tomados con equipo GPS Trimble	68
<i>Figura 34</i>	Punto validado con GPS	69
<i>Figura 35</i>	Tamaño de celda óptimo para construcción del DEM en Bucaramanga	71
<i>Figura 36</i>	DEM para Bucaramanga utilizando función de base radial multicuadrática	73
<i>Figura 37</i>	Variograma experimental para la cuenca de río de Oro	75
<i>Figura 38</i>	Variogramas lineales para la cuenca de río de Oro	75
<i>Figura 39</i>	Condiciones de anisotropía para ajuste lineal de	

variograma teórico	76
<i>Figura 40</i> Variogramas gaussianos para la cuenca de río de Oro	77
<i>Figura 41</i> Relación de anisotropía, escala y longitud de intervalo para ajuste gaussiano de variograma experimental	78
<i>Figura 42</i> Variogramas exponenciales para la cuenca de río de Oro	78
<i>Figura 43</i> Relación de anisotropía, escala y longitud del intervalo para ajuste exponencial de variograma experimental	79
<i>Figura 44</i> RMSE Vs factor $R^2$ para diferentes anisotropías en río de Oro	80
<i>Figura 45</i> Sensibilidad del parámetro $R^2$ para río de Oro	84
<i>Figura 46</i> Distribución de valores residuales para la cuenca de río de Oro	87
<i>Figura 47</i> Rango de celda para la construcción del DEM para la cuenca de río de Oro	88
<i>Figura 48</i> DEM para la cuenca de río de Oro utilizando la función de base radial	89

## LISTADO DE TABLAS

	PAG.
<i>Tabla 1</i> Coordenadas IGAC de cartografía 1:25000	7
<i>Tabla 2</i> Estructura ráster Vs estructura vectorial	16
<i>Tabla 3</i> Clasificación de las áreas piloto según pendiente en porcentaje (%)	34
<i>Tabla 4</i> Estadísticas básicas de las regiones piloto	34
<i>Tabla 5</i> Pendiente promedio de variogramas teóricos	45
<i>Tabla 6</i> Cálculo de la relación de anisotropía y ángulo para el variograma lineal	46
<i>Tabla 7</i> Resumen del valor de RMSE para todos los interpoladores analizados en la construcción del DEM sobre Bucaramanga	54
<i>Tabla 8</i> Comparación del RMSE para las funciones de base radial	56
<i>Tabla 9</i> Estadísticas de los errores residuales para Bucaramanga	64
<i>Tabla 10</i> Coordenadas de los puntos GPS	68
<i>Tabla 11</i> Errores residuales para puntos tomados con GPS	70
<i>Tabla 12</i> RMSE para diferentes tamaño de celda	71
<i>Tabla 13</i> RMSE para ajustes diferentes del variograma experimental	80
<i>Tabla 14</i> Errores cuadráticos en todas las direcciones para la función IDW	84
<i>Tabla 15</i> Valores de RMSE para río de Oro con vecino natural	85
<i>Tabla 16</i> RMSE para río de Oro utilizando triangulación de Delaunay	86
<i>Tabla 17</i> Resumen de interpoladores para río de Oro	86
<i>Tabla 18</i> Valores de RMSE para diferentes tamaños de celda en río de Oro	88

## RESUMEN

**TITULO:** MODELO DIGITAL DE ELEVACIÓN PARA LA CUENCA DEL RIO DE ORO\*

**AUTORES:** Wilson Gonzalo López y Rafael Mauricio Osorio A\*\*.

**PALABRAS CLAVE:** Modelo + elevación, DEM, Río + Oro, DEM + Bucaramanga.

**DESCRIPCIÓN:** Este proyecto consiste en la creación y validación de un modelo digital de elevación (DEM) para el área metropolitana de Bucaramanga y para la cuenca del río de Oro teniendo como base, planchas cartográficas obtenidas mediante restitución fotogramétrica.

Con base en esta información se procedió a buscar el método de interpolación que proporcionará un mejor grado de ajuste a los datos de entrada. Para la selección del método de interpolación se valoró el error medio cuadrático RMSE utilizando validación cruzada sobre todos los datos muestrales de la cuenca del río de oro y para Bucaramanga sobre los datos muestrales de 10 áreas piloto seleccionadas debido a la gran cantidad puntos con los que se contaba para generar el modelo completo.

Como resultado del análisis se encontró que el mejor ajuste para el modelo del área metropolitana lo proporcionó la función de base radial multicuadrática no solo desde el punto de vista estadístico sino también visual, superando a otros métodos considerados como el krigeado, la triangulación, o el vecino más cercano.

Se analizó también la selección de los parámetros óptimos en la función de base radial multicuadrática, el efecto del tamaño de la celda para la grilla, el tiempo de cálculo del modelo y el número de puntos tomados en la validación cruzada.

Por último se hizo una validación comparando los datos de elevación del modelo con puntos tomados con un sistema de posicionamiento global GPS en varios puntos dentro del área metropolitana, haciendo un análisis estadístico descriptivo de los datos residuales.

---

\* Trabajo de Grado.

\*\* Facultad De Ingenierías Físico Mecánicas Escuela De Ingeniería Civil. Director I.C. M.Sc. Jorge Alberto Guzmán.

Este mismo procedimiento se hizo para el análisis del modelo en la cuenca del río de oro, donde se encontró que los mejores resultados se dieron con el método de la función de base radial multicuadrática.

## ABSTRACT

**TITTLE:** DIGITAL ELEVATION MODEL FOR RIO DE ORO BASIN

**AUTHORS:** Wilson Gonzalo López and Rafael Mauricio Osorio A.

**KEY WORDS:** Modelo + elevación, DEM, Río + Oro, DEM + Bucaramanga.

The project aim was the creation and validation of a digital elevation model for the Río de Oro basin area and the Bucaramanga metropolitan area. The basis of the models was cartographic maps obtained from photogrammetric methods to scale of 1:25000 and 1:2000 respectively.

Based on this information, the interpolation method for fitting input data was made computing the root mean square error RMSE by means of cross validation over all of the Río de Oro basis sample points and ten testing areas in the Bucaramanga metropolitan area selected due to the large amount of sample points. As a result of the analysis, the best choice for Bucaramanga model was made using a multiquadratic radial basis function not only from a statistical point of view but also visual, surpassing other considered methods like kriging, triangulation with linear interpolation or nearest neighborhood.

It was also analysed the optimal parameters selection in the radial basis function, the cell size effect in the grid, the elapsed time to generate the model and the number of sample points to use in cross validation.

Then it was made a model validation comparing grid elevation data with elevation data obtained by means of a global positioning system GPS on Bucaramanga metropolitan area.

The RMSE value and the overall statistics for the residual data were computed. A similar procedure was made to analyze the Río de Oro basin model, where the best result was given with the radial basis function.

## INTRODUCCIÓN

El presente proyecto tiene como finalidad la creación y validación de un modelo digital que represente la superficie del terreno para la cuenca del río de oro, y que sirva de base para posteriores estudios hidrológicos al interior del grupo de predicción y modelamiento hidroclimático GPH de la Universidad Industrial de Santander.

Este modelo es especialmente útil en los estudios hidrológicos, ya que muchos de los fenómenos analizados toman como variables datos de la forma y la altitud del terreno, además por su carácter digital, permite la simulación de procesos dinámicos, resultando valioso en la toma de decisiones y otros tipos de aplicaciones.

Actualmente los estudios hidrológicos se fundamentan en la interpretación de datos obtenidos a partir de planchas cartográficas con curvas de nivel a diferentes escalas. Esto crea el inconveniente que los valores de elevación del terreno se limiten a la trayectoria de las curvas y que los puntos de elevación entre ellas se desconozcan, obligando así, a que estos datos se obtengan frecuentemente mediante una simple interpolación lineal.

Este trabajo propone como alternativa la creación de un modelo digital del terreno que represente la cuenca en forma tridimensional. Este proceso de modelamiento se basa en la discretización del sitio de interés en filas y columnas formando una malla regular. De esta manera se obtiene una mejor aproximación de los datos de elevación y forma, logrando una equivalencia entre el modelo digital y la superficie real.

Por tal motivo, se hace realmente necesario que el grupo de modelamiento y predicción hidroclimático de la universidad, trabaje bajo estos modelos tridimensionales digitales; ya que las investigaciones futuras concernientes a la cuenca en estudio podrán ser manejadas con eficiencia, realismo, y fundamentadas en novedosas técnicas de interpolación suministradas por el software.

El presente proyecto aportará al grupo de modelamiento y predicción hidroclimática un modelo digital validado de la cuenca del río de oro y Bucaramanga; que le servirá para estudiar la respuesta de dicha cuenca ante fenómenos naturales a los que esta sometida.

En el futuro, con la ayuda de este informe los estudiantes de ingeniería civil de la Universidad Industrial de Santander interesados en el tema, abrirán las puertas hacia al manejo de la información tridimensional en superficies. La elaboración de modelos de superficies digitales, no solo intervendrán en estudios de tipo hidrológico; también estarán relacionados con proyectos topográficos y viales que hasta el día de hoy han sido manejados en formatos bidimensionales.

## **1. UBICACIÓN Y DESCRIPCION GENERAL DE LA ZONA DE ESTUDIO**

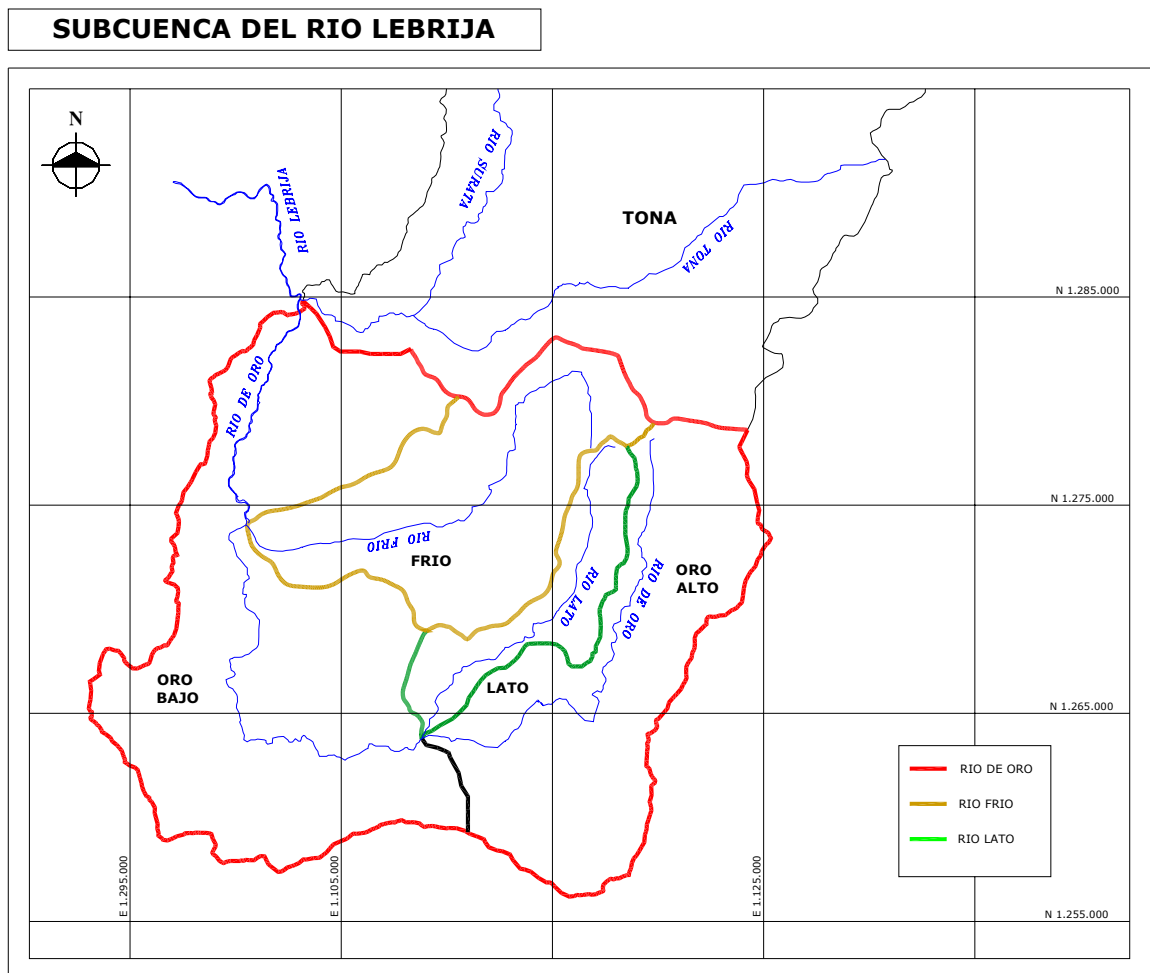
### **1.1 UBICACIÓN DE LAS CUENCAS DE ESTUDIO**

Los DEM (Digital Elevation Model) que se presentan en este proyecto, tienen su marco geográfico en el departamento de Santander (Colombia). La ubicación geográfica de la cuenca del río de oro, como subcuenca del río Lebrija se describirá a continuación:

**1.1.1 Cuenca del Río de Oro.** Se localiza en el extremo nordeste del departamento de Santander, sobre la vertiente occidental de la cordillera oriental, entre las coordenadas 6 55`y 7 27`de latitud norte, y 72 51`y 73 18`de longitud al oeste de Greenwich. El área de captación hídrica comprende las subcuencas de los ríos de oro, frío y lato. El relieve presenta un sector nororiental montañoso con paisaje de cordillera y fuertes pendientes, separado por la falla de Bucaramanga, de una zona sur occidental conformada por plataformas con topografía suave y ondulada en estado avanzado de erosión. Este relieve montañoso y variado crea una gran cantidad de microclimas, siendo la altura promedio 1735 msnm, la temperatura media es de 18 °C, y la precipitación media anual; de acuerdo a la distribución de isoyetas, de 1475 mm. El 50 % del área se encuentra dentro del piso premontano o clima medio, el 35 % se ubica dentro de la vida de bosque húmedo premontano, y el 15 % dentro de la zona de vida de bosque húmedo-montano bajo. Por ser parte de la cuenca del río Lebrija, podría considerarse como buena productora de agua. Sin embargo, lo montañoso del territorio, las fuertes pendientes, la deforestación y la gran actividad humana que allí se desarrolla, no permiten el adecuado almacenamiento del agua, ni su disponibilidad con buena calidad. Este hecho es agravado además por el excesivo arrastre de sedimentos desde las partes más altas, (Oro alto). *(Plan de desarrollo cuenca superior del río Lebrija, CDMB, 1991 )*

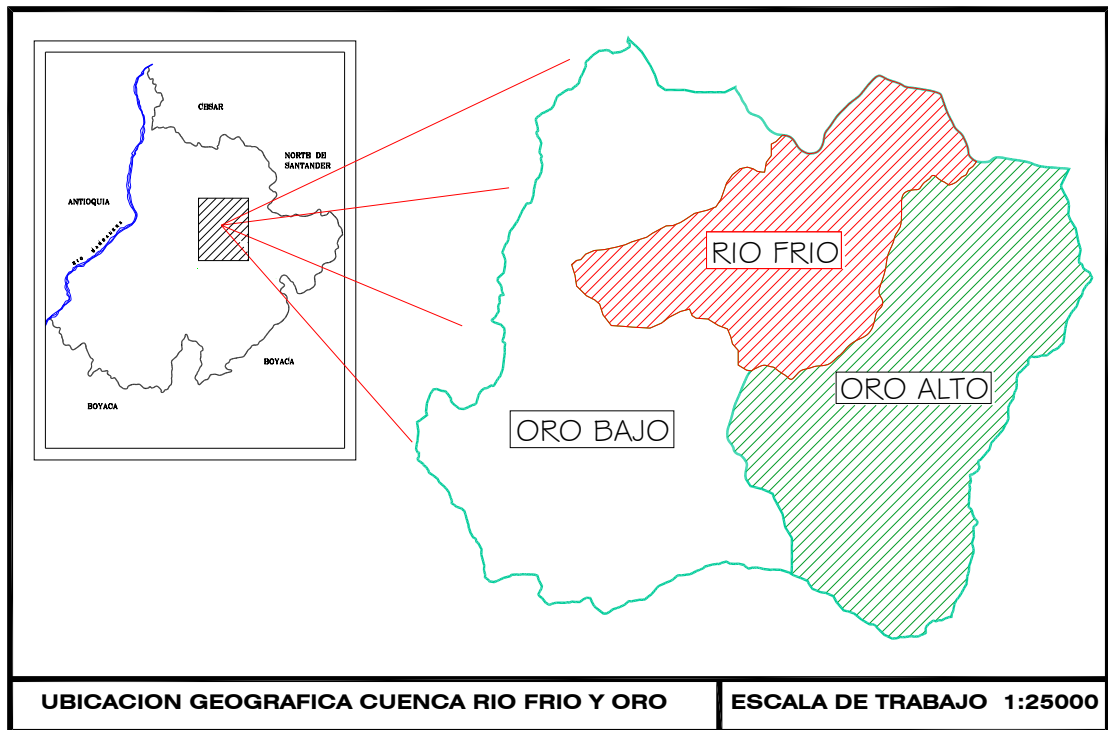
El Río de Oro contempla un amplio recorrido por zonas pobladas recibiendo descargas directas de los alcantarillados de los municipios de Bucaramanga, Piedecuesta y Girón e indirectas del municipio de Floridablanca, también recibe vertimientos de la zona industrial del Palenque - Café Madrid, y descargas no puntuales de las áreas agroindustriales.

Aguas abajo del área urbana del municipio de Piedecuesta, el Río de Oro recibe a la Quebrada Grande, Quebrada Soratá y el Río Lato. En La zona del Municipio de Girón los afluentes son el Río Frío, Quebrada La Iglesia, Quebrada Chimitá y las corrientes de la escarpa de la Meseta de Bucaramanga.



**Figura 1: Limite de la cuenca del río de Oro y sub cuencas**

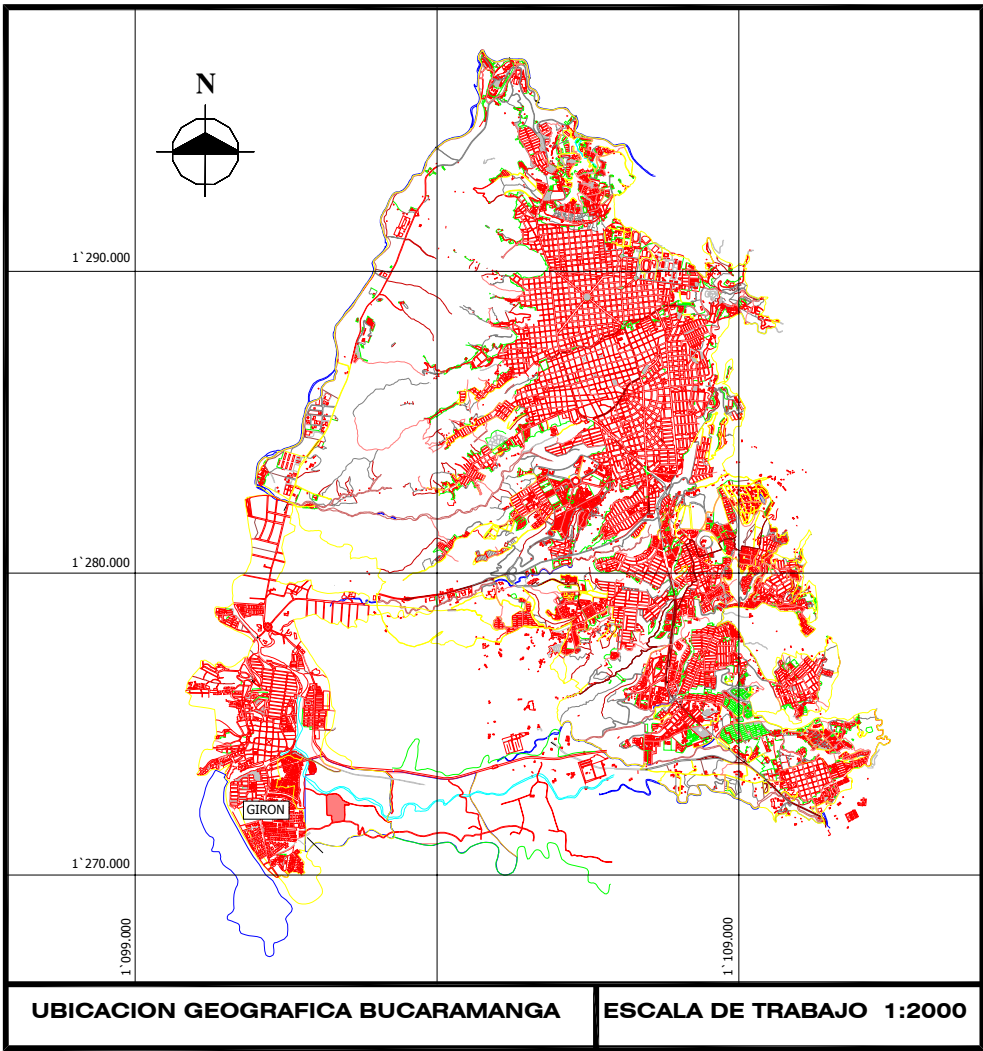
**1.1.2 Cuenca de Río Frío.** Se localiza en la zona central de la cuenca del río Lebrija y pertenece a la cuenca del río de Oro. Esta justo entre la cuenca del río de Oro alto y el bajo. Sus coordenadas 5 35`y 5 55`de latitud norte, y 69 47`y 70 18`de longitud al oeste de Greenwich. (*Plan de desarrollo cuenca superior del río Lebrija, CDMB, 1991*).



**Figura 2: Cuenca de río Frío; sub cuenca de la cuenca de río de Oro.**

**1.1.3 Municipio de Bucaramanga.** Bucaramanga, capital del departamento de Santander limita por el Norte con el municipio de Ríonegro; por el Oriente con los municipios de Matanza, Charta y Tona; por el Sur con el municipio de Floridablanca y; por el Occidente con el municipio de Girón. Se encuentra en una terraza inclinada de la Cordillera Oriental a los 7 08' de latitud norte con respecto al Meridiano de Bogotá y 73° 08' de longitud al Oeste de Greenwich. El área metropolitana formada por Bucaramanga, Piedecuesta, Florida y Girón está

ubicada sobre el Valle del Río de Oro. Se distingue en ella dos sectores de diferente conformación física: uno formado por la meseta y otro por el valle. El área municipal es de 165 kilómetros cuadrados, su altura promedio sobre el nivel del mar es de 959m y sus pisos térmicos se distribuyen en: cálido 55 kilómetros cuadrados: medio 100 kilómetros cuadrados y frío 10 kilómetros cuadrados. Su temperatura media es 23° C y su precipitación media anual es de 1.041 mm. La topografía de Bucaramanga es en promedio un 15% plana, 30% ondulada y el restante 55% es quebrada. Tres grandes cerros se destacan a lo largo del territorio: Morro rico, Alto de San José, y el Cacique.



**Figura 3: Municipio de Bucaramanga**

Los ríos principales son: El Río de Oro y el Suratá y las quebradas: La flora, Tona, La Iglesia, Quebrada Seca, Cacique, El Horno, San Isidro, Las Navas, La Rosita, Bucaramanga. (*Micro zonificación sísmica del área metropolitana de Bucaramanga, Gobernación de Santander, 2001*).

## 1.2 ARCHIVOS DIGITALES DE LAS ZONAS DE ESTUDIO

Para la construcción de los modelos digitales de elevación planteados en este proyecto, se utilizaron diferentes planchas cartográficas suministradas por varias entidades interesadas en trabajo.

- Acueducto metropolitano de Bucaramanga: Suministró las planchas escala 1:2000 del área urbana de Bucaramanga. La cartografía se realizó por medio de restitución fotogramétrica. Total de archivos digitales: 96 planchas.
- Grupo de predicción y modelamiento hidroclimático (GPH) UIS: Suministró las planchas digitalizadas escala 1:25000 pertenecientes a los linderos de la cuenca del río de Oro. Estas planchas provienen de la cartografía IGAC. Total de archivos: 12 planchas.

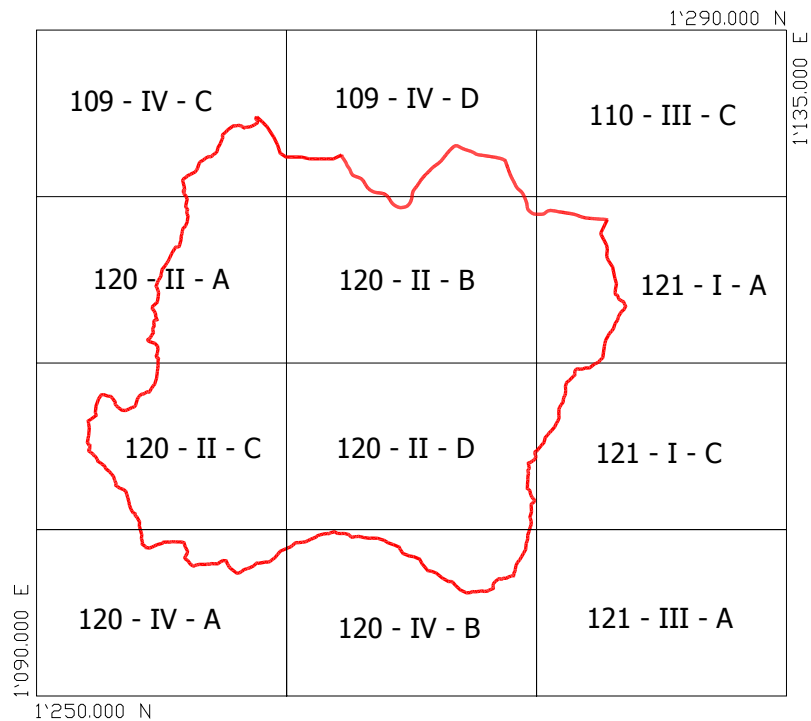
**1.2.1 Archivos digitales escala 1:25000.** Las planchas necesarias para la construcción del DEM para la cuenca de río de Oro, escala 1:25000 se describirán en la *tabla (1)* según nomenclatura IGAC:

PLANCHA IGAC	ESCALA	COORDENADAS			
		ESTE		NORTE	
		MINIMO	MAXIMO	MINIMO	MAXIMO
120-IV-A	1: 25000	1090000	1105000	1290000	1300000
120-IV-B		1105000	1120000	1290000	1300000
121-III-A		1120000	1135000	1290000	1300000

109-IV-C	1090000	1105000	1280000	1290000
109-IV-D	1105000	1120000	1280000	1290000
110-III-C	1120000	1135000	1280000	1290000
120-II-A	1090000	1105000	1270000	1280000
120-II-B	1105000	1120000	1270000	1280000
121-I-A	1120000	1135000	1270000	1280000
120-II-C	1090000	1105000	1260000	1270000
120-II-D	1105000	1120000	1260000	1270000
121-I-C	1120000	1135000	1260000	1270000

**Tabla 1** Coordenadas IGAC de cartografía 1:25000

Las planchas necesarias se esquematizan en la figura (4):

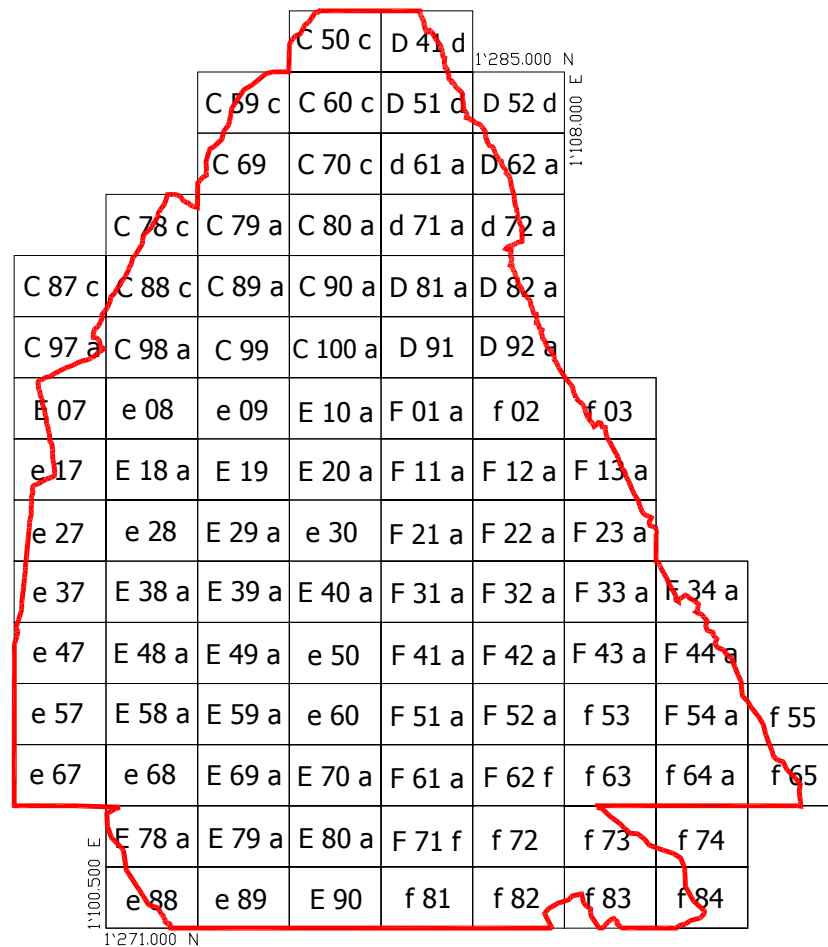


**figura 4:** Planchas requeridas para construcción del DEM (1:25000)

El DEM para la cuenca de río de Oro se elaborara con la información completa de todas las planchas 1:25000. Es decir, el DEM abarcará topografía mas allá de la divisoria de aguas de la cuenca con el objetivo de que los puntos cercanos al

lindero no queden truncados ni con errores en el proceso de interpolación debido a la ausencia de datos alrededor. Las curvas de nivel presentes en estas planchas se encuentran cada 50 m.

**1.2.2 Archivos digitales escala 1:2000.** El DEM para el municipio de Bucaramanga requerirá de 96 planchas digitalizadas escala 1:2000 suministradas por el acueducto metropolitano de Bucaramanga. La nomenclatura de dichas planchas se describirá a continuación:



**Figura 5: Planchas necesarias para la construcción del DEM 1:2000**

Algunas de las planchas restituidas 1:2000 ubicadas en la zona sur occidental no se pudieron obtener a la escala requerida. Esta zona corresponde a una muy pequeña porción del río de Oro. Sin embargo el área metropolitana y el 95 % de Girón presentan información completa 1:2000. Las curvas de nivel para este tipo de información se encuentran cada 2 m.

## **2. MODELOS DIGITALES DE ELEVACION (DEM)**

### **2.1 DEFINICIÓN**

Un modelo digital de elevaciones es una estructura discreta de datos que representa una variable distribuida en el espacio. La variable representada puede ser cualquiera que cumpla con el requisito fundamental de la misma, es decir, que se pueda representar espacialmente. (*FELICISIMO, A.M.. Modelos Digitales del Terreno. Oviedo, 1994*)

Cuando esta variable hace referencia a la elevación de la superficie del terreno el Modelo Digital de Elevación (DEM por sus siglas en ingles) toma el nombre de Modelo Digital del Terreno (MDT). La finalidad del DEM es modelar una variable distribuida y continua en el espacio mediante un arreglo de valores discretos que permitan el tratamiento numérico de los datos a través de algoritmos matemáticos y lo hagan útil para representar la superficie del terreno.

### **2.2 CREACIÓN DE UN MODELO DIGITAL DE ELEVACIONES (DEM)**

La creación de un DEM involucra una serie de pasos previos, cada uno de los cuales genera algún tipo de error que debe ser considerado con cuidado para que el modelo final se ajuste lo más posible a la realidad que se pretende representar. Los pasos para la creación de un DEM pueden resumirse en:

- Recopilación de datos.
- Selección de tipo de estructura del DEM (vectorial o ráster).
- Selección de la técnica de interpolación.
- Validación del modelo.

## 2.3 RECOPIACION DE DATOS

La recopilación de los datos que van a servir de base para la generación del DEM constituye la primera y tal vez la mas importante etapa del proceso, ya que la calidad de estos determina en gran medida la utilidad del modelo.

Se han desarrollado varias técnicas para la captura de esta información, clasificándose de acuerdo a si los datos son extraídos directamente del terreno utilizando levantamientos topográficos, GPS's, altimetría generada por radar o láser, o si son tomados indirectamente mediante la digitalización de planchas cartográficas o restitución fotogramétrica a partir de fotografías aéreas tomadas por aviones o satélites. El empleo de técnicas indirectas es mas generalizado por su economía comparada con las técnicas directas y por la facilidad para obtener gran cantidad de información en corto tiempo, sin embargo, la utilización de GPS's para la toma de datos podría considerarse cuando se trate de áreas pequeñas que requieran un gran nivel de detalle o en la etapa de validación del modelo.

En la actualidad se cuenta con GPS's diferenciales de gran exactitud que dan información de la elevación en centímetros, proporcionando datos que podrían considerarse como "reales" al momento de calcular la exactitud del modelo.

Los datos recopilados que van a servir de base para la creación del DEM, pueden contener uno o varios de los siguientes errores verticales que deben ser eliminados o minimizados antes de utilizarlos para generar el modelo (*USGS DEM Standards, 1997*):

1. *Blunders*: Los blunders o equivocaciones, son errores en la elevación generalmente mayores a 3 veces el máximo error permitido para el DEM.

debido a su magnitud, estos errores son fácilmente identificables y pueden ser fácilmente eliminados. Las causas de los blunders se deben principalmente a una mala interpretación de las curvas de nivel, transposición de valores numéricos u observaciones sin cuidado de los datos de entrada para el modelo.

2. *Errores sistemáticos*: Estos errores se caracterizan por seguir una tendencia específica, generalmente de magnitud o signo constante. Los errores sistemáticos pueden causar sesgo o artefactos en el DEM, es decir, introducir características ficticias como cimas, riscos o estrías donde en realidad no existen. En los procesos fotogramétricos estos errores se pueden generar debido a una mala interpretación del relieve a causa de árboles, edificios o sombras.
3. *Errores aleatorios*: Son aquellos que permanecen después de que los blunders y los errores sistemáticos son eliminados o minimizados. Estos errores son considerados como accidentales, desconocidos y fuera del control del observador.

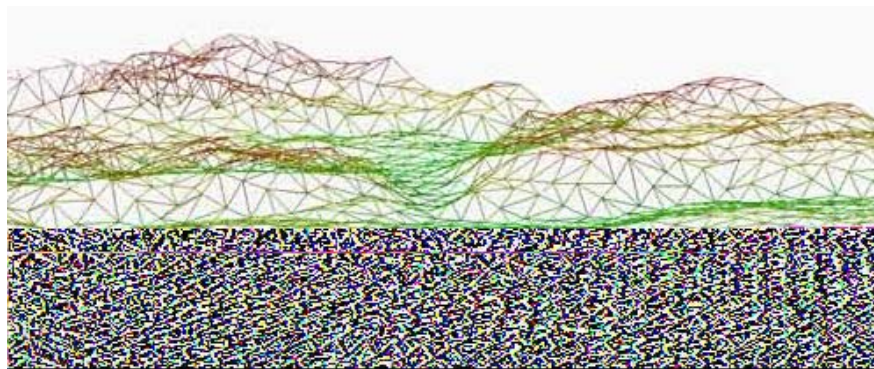
## 2.4 ESTRUCTURA DE UN DEM

Aunque todos los modelos de elevación digital son representaciones de una variable que se distribuye en el espacio, la forma en la que esta información está organizada constituye su estructura. Se tienen dos estructuras básicas de representación de datos:

- **Estructura vectorial**
- **Estructura ráster**

**2.4.1 Estructura vectorial.** En una estructura vectorial la representación del terreno se hace mediante una serie de puntos, líneas o polígonos a cada uno de los cuales le corresponde su valor de elevación, no necesariamente de magnitud constante. Cuando las líneas o polígonos describen trayectorias con un único valor de altitud, se generan un mapa de contornos o de curvas de nivel, similar a una plancha cartográfica donde cada trayectoria representa un mismo valor de elevación, con separación entre curvas de nivel a intervalos constantes.

Dentro de las estructuras vectoriales se encuentran también las Redes Irregulares de Triángulos o TIN (por sus siglas en inglés). Esta estructura representa la superficie del terreno por medio de planos de forma triangular (*figura 6*). Los datos recopilados previamente o puntos muestra son utilizados como vértices para los triángulos y los valores a interpolar son los puntos interiores de cada triángulo. Ya que los valores de los vértices son conocidos, se puede deducir la ecuación de un plano para hallar los valores de elevación en los puntos interiores. (*FELICISIMO, A.M.. Modelos Digitales del Terreno. Oviedo, 1994*)

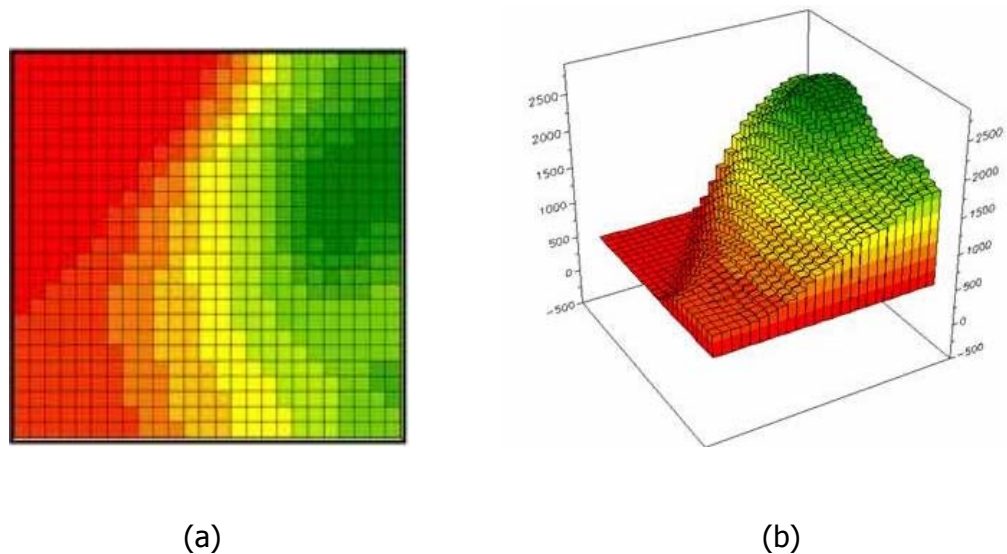


***Figura 6: representación del terreno mediante una estructura TIN.***

La principal ventaja de esta estructura es que el valor de la elevación de los datos recopilados permanece en el modelo, lo que no sucede con los modelos basados en una estructura raster, además es útil para representar características

específicas del terreno tales como líneas de elevación máxima en donde se producen cambios bruscos de pendiente, líneas de quiebre como riveras de ríos, orillas de lagos, escarpas o alteraciones hechas por el hombre.

**2.4.2 Estructura ráster.** Esta estructura representa la superficie del terreno mediante la discretización del sitio de interés en filas y columnas formando una malla regular. Cada celda dentro de la malla representa un área elemental conocida como píxel a la que se le asigna un valor del atributo, que en este caso es el valor de la elevación (*figura 7 b*). Las coordenadas de cada píxel están definidas implícitamente de acuerdo a su posición relativa dentro de la matriz o malla es decir, para localizar espacialmente un píxel solo es necesario darle coordenadas a un punto de la malla y definir la longitud de los lados de las celdas o lo que es lo mismo definir la resolución del modelo.



**Figura 7: Estructura ráster**  
**(a) Representación en 2D (b) Representación en 3D**

**2.4.3 Comparación entre estructura ráster y vectorial.** Las características propias de cada tipo de estructura, así como las ventajas y desventajas

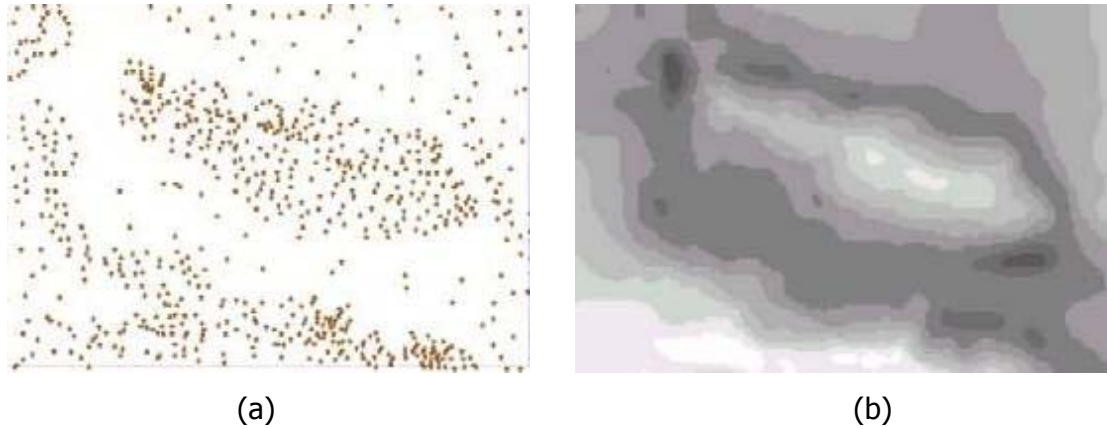
operacionales de cada sistema al ser aplicado en cualquier medio de información se resumen en la *tabla (2)*:

<b>COMPARACION ENTRE ESTRUCTURA RASTER Y VECTORIAL</b>		
	<b>VENTAJAS</b>	<b>DESVENTAJAS</b>
<b>VECTOR</b>	<ul style="list-style-type: none"> <li>▪ Buena similitud de las formas.</li> <li>▪ Cálculo más preciso de áreas, perímetros y longitudes.</li> <li>▪ Análisis de redes más consistentes.</li> <li>▪ Se pueden adaptar bajo bases de datos orientadas a objetos.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Estructura más compleja.</li> <li>▪ Las superposiciones exigen más verificación de errores y pueden ser más lentas.</li> <li>▪ La diferente topología de las unidades espaciales dificulta los ejercicios de simulación.</li> </ul>
<b>RASTER</b>	<ul style="list-style-type: none"> <li>▪ Estructura más simple.</li> <li>▪ Asimilación más directa a datos de sensores remotos.</li> <li>▪ Operaciones de superposición más sencillas.</li> <li>▪ Análisis espaciales como distribución, densidad y de superficie, mas eficientes.</li> <li>▪ La unidad espacial conserva la misma forma y tamaño de modo que se facilitan las simulaciones.</li> <li>▪ Útil para análisis de grandes extensiones con baja precisión de propiedades espaciales.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Alto nivel de error en estimaciones de área, perímetro y longitud.</li> <li>▪ Gran espacio de almacenamiento a medida que aumenta la resolución.</li> <li>▪ Desperdicio de espacio de almacenamiento para datos espaciales muy esparcidos.</li> <li>▪ Análisis de redes muy complejos y menos consistentes.</li> <li>▪ Transformación de coordenadas menos eficiente.</li> </ul>

***Tabla (2) Estructura ráster Vs estructura vectorial***

Las fuentes de datos para la creación de modelos digitales son principalmente curvas de nivel extraídas mediante la digitalización de planchas cartográficas, o a partir de ortofotos mediante restitución fotogramétrica. Estos datos se encuentran en una estructura vectorial y es necesaria su rasterización para la creación del modelo, es decir, crear una malla con una resolución determinada, asignándole valores de elevación a los píxeles por los cuales pase una curva de nivel. El

modelo se genera interpolando los restantes píxeles de la malla utilizando para ello los puntos rasterizados (*figura 8*).



**Figura 8**

**(a) Información vectorial a partir de la cual se genera el DEM. Los espacios en blanco son los puntos a interpolar (b) malla completamente interpolada.**

Los dos parámetros más importantes que se deben seleccionar cuando se utiliza una estructura ráster son la resolución y la precisión. En conjunto estos dos parámetros determinan el nivel de detalle con el cual se representara el terreno. La resolución es la longitud de cada celda o píxel dentro de la matriz y su valor se relaciona la escala horizontal que tendrá el DEM. La precisión es el número (entero o decimal) con el que se especifican los valores de elevación en el modelo.

Existe una estrecha relación entre precisión y resolución cuando los datos del modelo se usan para hallar parámetros topográficos que involucran diferencias en la elevación. La pendiente se define como la diferencia de elevación entre dos píxeles adyacentes dividida por la distancia entre ellos. Así por ejemplo, al crear un modelo con una resolución de 30 metros y una precisión ajustada a valores enteros, se tienen valores discretos de pendiente de 0.033, 0.066 o múltiplos de estos. Esto podría ser adecuado para representar terrenos de alto relieve o

montañosos pero inadecuado para representar terrenos con superficies planas o donde se tengan variaciones de pendiente entre píxeles adyacentes menores a 0.033 (*Garbrecht y Martz, 1994*).

Las características hidrológicas extractadas de un DEM dependen en gran parte de su resolución. Esta relación fue estudiada por *Garbrecht y Martz (1994)*, quienes utilizaron un DEM base con una resolución de 30 metros para comparar características de una cuenca como longitud de canales, número de canales, densidad de drenaje, etc., a medida que se reducía la longitud de la celda en el modelo. A partir de este estudio los autores concluyen que para reproducir características de drenaje con una exactitud de aproximadamente el 10%, el área de la celda o píxel en el modelo debe ser menor que el 5% del área de drenaje mas pequeña que se quiera representar. Esto se debe a la imposibilidad del DEM para reproducir adecuadamente características de drenaje que están a la misma escala que la resolución del modelo.

## **2.5 SELECCIÓN DEL MÉTODO DE INTERPOLACIÓN**

Para la generación de la malla interpolada a partir de los puntos de muestra se tienen varios algoritmos de interpolación divididos en estocásticos o determinísticos.

Los métodos determinísticos tienen en cuenta en el proceso de interpolación solamente criterios geométricos, es decir, la distancia relativa del punto a interpolar con los datos muestrales. Aunque todos los métodos utilizan un factor de peso durante la interpolación, los determinísticos no tienen en cuenta la tendencia de los datos muestrales usados en los algoritmos matemáticos para calcular estos pesos durante la interpolación de la grilla. Los métodos de interpolación determinísticos que se utilizan en el presente trabajo con el propósito de hallar el que produzca los mejores resultados son: El inverso de la distancia a

una potencia, Mínima curvatura, Vecino más cercano, Vecino natural y Funciones de base radial. (Moral, F.J. 2004. *Aplicación de la geoestadística en las ciencias ambientales. Ecosistemas 2004*)

Un método de interpolación que no solamente tiene en cuenta criterios geométricos sino también criterios basados en el análisis directo de la información disponible se denomina geoestadístico. El método que se utiliza en este trabajo es el método Kriging Ordinario.

Las principales características que hacen del kriging un buen método de estimación son:

- Mientras que los métodos tradicionales utilizan el concepto euclidiano de la distancia para el cálculo de los pesos que se aplicarán a cada dato muestral, el kriging considera tanto la distancia como la geometría de la localización de las muestras.
- Mediante el kriging se minimiza la varianza del error esperado (diferencia entre el valor real y el estimado). Como el valor real en un punto no muestral es desconocido, el kriging emplea un modelo conceptual con una función aleatoria asociada a los valores reales.

## **2.6 VALIDACIÓN DEL MODELO DIGITAL DE ELEVACIÓN**

La validación es el último paso en la creación de un DEM e involucra la evaluación de la exactitud entre los datos del modelo y la superficie real del terreno. El *U.S. geological survey* evalúa la exactitud vertical de los modelos digitales de elevación mediante el cálculo de un error igual a la raíz cuadrada del promedio de los errores al cuadrado o RMSE (por sus siglas en inglés), tomado para varios puntos de prueba distribuidos irregularmente en el terreno y el cual tiene en cuenta los

errores aleatorios y sistemáticos introducidos durante la recopilación de los datos (*USGS DEM Standards, 1997*). Esto es:

$$RMSE = \sqrt{\frac{1}{n} \sum (Z_i - Z_t)^2} \quad \text{Eq 1}$$

Donde:

*Z<sub>i</sub>* = valor generado por el MDE en un punto de prueba.

*Z<sub>t</sub>* = valor real en el punto de prueba.

*n* = número de puntos de prueba.

Se recomienda un mínimo de 28 puntos de prueba para el cálculo del error, de los cuales 20 se deben localizar en el interior y 8 en los bordes o esquinas del modelo, ya que estas fronteras presentan problemas al momento de la interpolación.

Los puntos de prueba deben estar bien distribuidos y ser representativos del terreno, en general se debería muestrear puntos donde se presenten grandes variaciones en la topografía. Los valores de elevación que se van a tomar como reales para el cálculo del error deben ser hallados con métodos que aseguren la mayor exactitud posible, para esto se puede hacer uso de GPS o si es difícil acceder directamente al terreno, mediante el uso de planchas cartográficas de la menor escala existente.

El RMSE permitido para un DEM cuyos valores de elevación han sido procesados, eliminando puntos o regiones anómalas o editados, removiendo errores sistemáticos identificables, es de la mitad del intervalo entre curvas de nivel del mapa original, y donde no hay errores en los puntos de muestra individuales mas grandes que el valor de un intervalo del mapa original. Es decir, para un DEM generado a partir de una plancha cartográfica con curvas de nivel cada 2 metros,

el máximo RMSE permitido es de 1 metro, donde cada diferencia  $(Z_i - Z_t)$  en la ecuación, no debe ser mayor que 2 metros.

Además del cálculo del RMSE se debe hacer una verificación visual del modelo en busca de artefactos generados durante el proceso de interpolación. Estos artefactos se presentan generalmente en zonas con tendencia plana o en cuerpos de agua donde el nivel de la superficie de agua tiene una misma elevación. En estos casos se debe editar el modelo, introduciendo las elevaciones que describan las características reales del terreno. Las regiones con tendencia plana no deben tener estrías y los cuerpos de agua deben tener su forma real y una elevación constante.

### **3. LIMPIEZA Y COMPILACION DE ARCHIVOS .DWG**

#### **3.1 DESCRIPCIÓN DE ARCHIVOS DIGITALES.**

Toda la información cartográfica usada en este proyecto se encuentra en formatos .dwg (ACAD, Autodesk®), por tal razón, la terminología usada en el presente capítulo esta orientada a comandos propios de dicho software.

Las planchas digitalizadas (Río de Oro 1:25000) y las rasterizadas (Bucaramanga 1:2000) están compuestas por puntos, poli-líneas, líneas, arcos, polígonos y splines. La extracción de la información necesaria en la construcción de los DEM para Bucaramanga y Río de Oro, debe desarrollarse bajo formatos completamente limpios y continuos; es decir, el formato final de las planchas dentro del archivo tipo .dxf solo estará compuesto de poli-líneas que estarán juntas en toda su extensión longitudinal y libres de cualquier partición.

Los problemas más comunes presentados en las planchas digitalizadas son los siguientes:

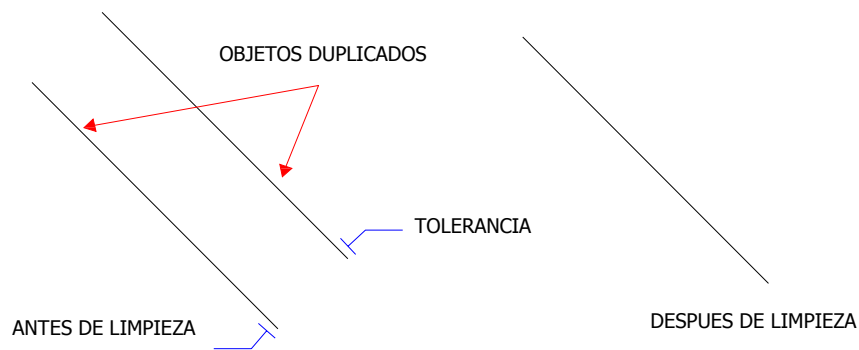
- Objetos insuficientes, objetos excesivos u objetos sueltos no deseados
- Líneas digitalizadas dos veces (duplicados)
- Objetos cortos
- Seudonodos

Por tal razón, es necesario utilizar una herramienta cartográfica para el manejo y ajuste de los archivos digitales. Autodesk® map, es tal vez la mejor alternativa para llevar a cabo este proceso de limpieza y pegado de curvas. Las opciones mas

utilizadas para la limpieza de las planchas son idénticas a las utilizadas en la creación de topologías para la creación de sistemas de información geográfica. Algunas están descritas en detalle a continuación:

### 3.2 SUPRESIÓN DE OBJETOS DUPLICADOS

Los cálculos de supresión de duplicados *no* analizan los puntos, el texto ni los bloques. Se tienen en cuenta las posiciones de coordenadas de los objetos y el número de vértices, pero los objetos con direcciones, objetos de tipos diferentes (por ejemplo, líneas y poli- líneas) y los objetos con propiedades diferentes (por ejemplo, tipos de línea y colores) pueden considerarse objetos duplicados. Los objetos que tengan la misma geometría pero que se encuentren en capas diferentes se consideran duplicados. Todos los objetos que se encuentren el uno del otro dentro de la distancia de tolerancia pueden considerarse duplicados, tal como se muestra en la *figura (9)*. En la mayoría de los casos, se tomaron valores de tolerancia muy bajos cuando se eliminaron objetos duplicados.

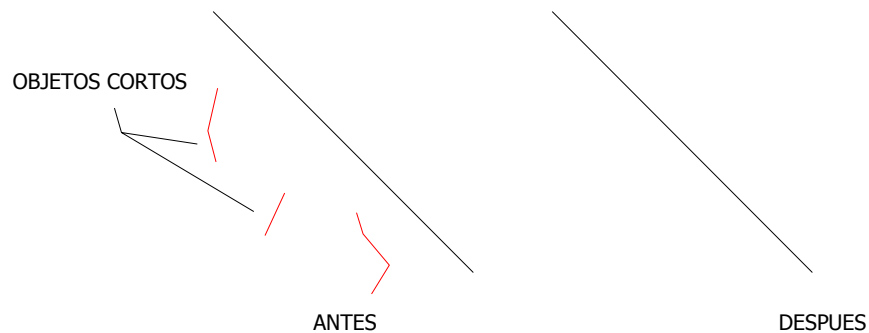


**Figura 9. Supresión de objetos duplicados**

La función sólo elimina aquellos objetos que tengan las mismas propiedades y geometría, aun cuando se encuentren en capas distintas. No se tiene en cuenta la dirección. (*Manual de autocad map 3D, 2005.pdf*).

### 3.3 ELIMINACIÓN DE OBJETOS CORTOS

Puede reducir el número de objetos lineales y nodos innecesarios de un mapa borrando los objetos lineales cortos que se encuentren aislados o los que formen parte de una poli- línea, como se muestra en la *figura (10)*:

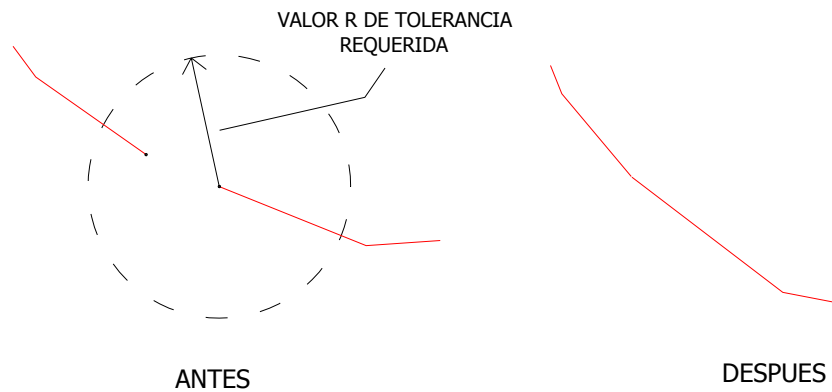


**Figura 10 Eliminación de objetos cortos**

La eliminación aumenta el rendimiento del sistema y reduce el tamaño del archivo. Esta función es similar a la de disolver seudonodos, salvo que borrar los objetos cortos elimina los objetos lineales y los nodos asociados. (*Manual de autocad map 3D, 2005.pdf*).

### 3.4 AGRUPACIÓN DE NODOS:

Es necesario cambiar la tolerancia a un valor ligeramente superior al radio del círculo que engloba los nodos. Todos los nodos se agrupan en la posición del nodo que se utiliza en más vectores dentro del círculo de tolerancia. Se conserva un único nodo en la nueva intersección. (*Manual de autocad map 3D, 2005.pdf*).

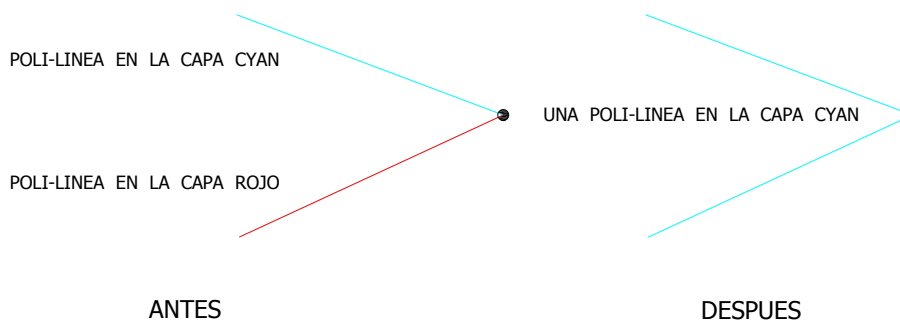


**Figura 11 Agrupación de nodos**

Con esta opción, la tolerancia es la distancia radial de búsqueda en la que se detectan los errores geométricos durante el proceso de limpieza.

### 3.5 DISOLUCIÓN DE SEUDONODOS

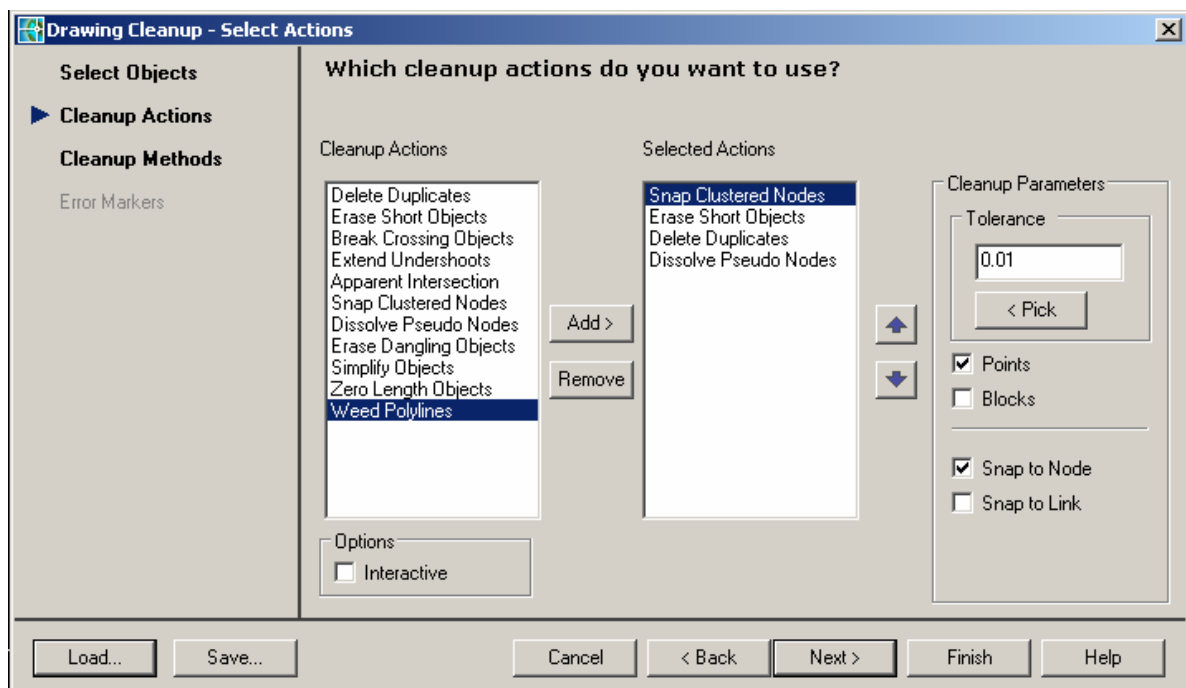
Esta opción elimina los nodos que se encuentran en las intersecciones de dos objetos lineales, pero deja el vértice en su lugar, como se muestra en la *figura (12)*:



**Figura 12 Disolución de seudonodos**

La opción Dissolve pseudonodos une objetos cuando hay dos objetos lineales que terminan en la misma ubicación. Se puede utilizar también esta opción para eliminar los pseudonodos junto con la simplificación de objetos lineales. De esta forma convertirá en una sola polilínea un conjunto de segmentos de línea, arco y polilínea que se encuentren dentro de la misma tolerancia.

El cuadro de diálogo propio de autocad map 3D 2005 que resume lo anteriormente dicho, se presenta en la *figura (13)*. Es importante utilizar esta herramienta para el manejo de la cartografía debido a la gran cantidad y densidad de información que presenta cada una de las planchas necesarias para la construcción de los DEM. (*Manual de autocad map 3D, 2005.pdf*).



**Figura 13 Cuadro de diálogo (ACAD 3D 2005) para limpieza de archivos .dwg**

### 3.6 SPELL MAP:

El grupo de predicción y modelamiento hidroclimático y específicamente el Ingeniero Jorge Alberto Guzmán desarrolló una herramienta de software llamada

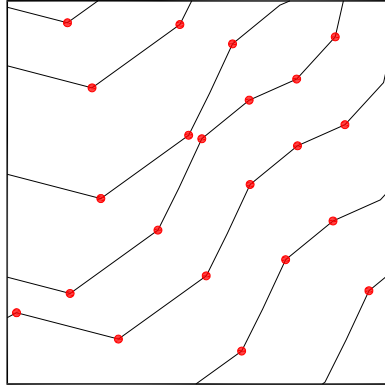
spell map utilizando como plataforma de programación en lenguaje delphi. El software se presenta en su version beta 1.0 en la página del grupo: [www.gph-uis.info](http://www.gph-uis.info) para descargar gratuitamente.

Para poder generar los archivos .GRID dentro del proceso de construcción del DEM, es necesario extraer los atributos de la información base (archivos .dxf) para poder enlistarlos en un procesador de texto y posteriormente incluirlos en surfer para desarrollar el objetivo del presente proyecto.

La poli línea en los archivos .dxf es una sucesión de puntos distribuidos en el espacio que tienen un valor de elevación constante (curva de nivel). El objetivo de *Spell map* es extraer los datos referentes a las coordenadas espaciales propias de cada poli línea. Esto es, sacar todos los puntos que conforman dicha poli línea como parejas espaciales (x,y,z) y agruparlas en un archivo .txt. Spell tiene la capacidad de extraer archivos de líneas, poli líneas o simplemente puntos desde un archivo base .dxf.

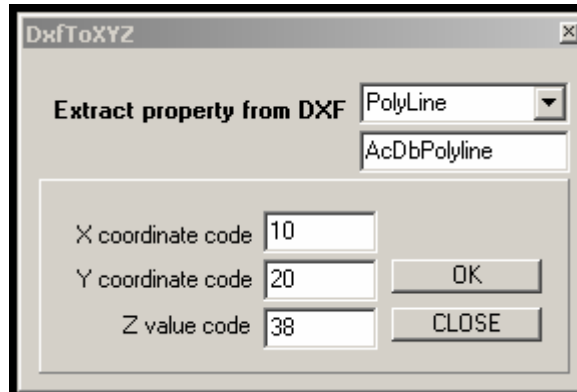
En la *figura 14 (a), (b) y (c)* se resumen los pasos para la extracción de información necesaria par la generación del DEM.

- Archivos .dxf depurados y pegados en toda su extensión, listos para extraer los atributos de información:



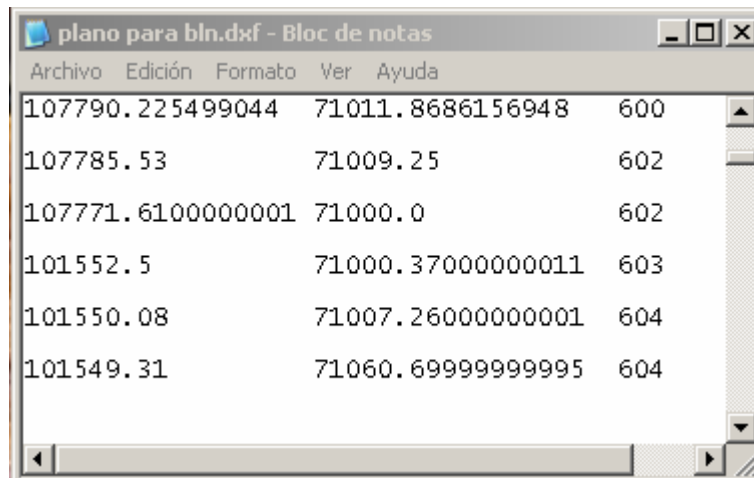
**figura 14(a) Archivo .dxf con poli líneas constituidas por series de puntos (x,y,z)**

- Determinar en *spell map* el tipo de línea o puntos a los que se les desea extraer las coordenadas espaciales:



**figura 14 (b) Cuadro de dialogo de spell map para escoger los elementos del archivo.  
base**

- Finalmente se obtiene un archivo tipo .txt con todos los puntos (x,y,z) de las poli líneas que se encontraban en el archivo .dxf y queda listo el archivo muestral de entrada para surfer.



**Figura 14(c) archivo de entrada de surfer tipo texto generado por spell map**

De esta manera las 96 planchas digitales depuradas y pegadas en toda su extensión quedan convertidas en un archivo .txt de 390 MB en promedio. Este archivo contiene todos los puntos (x,y,z) necesarios para la construcción del DEM para el área metropolitana de Bucaramanga. De igual manera, el archivo para la cuenca de río de Oro contiene alrededor de 350000 puntos en promedio.

## **4. ANÁLISIS DE INTERPOLADORES PARA LA CONSTRUCCIÓN DEL DEM SOBRE EL ÁREA METROPOLITANA DE BUCARAMANGA**

### **4.1 ESCOGENCIA DE REGIONES REPRESENTATIVAS PARA EL MUNICIPIO DE BUCARAMANGA**

Debido a los requerimientos de hardware, necesarios para el procesamiento de todos los datos (12.5 millones aproximadamente), se realizó una división del área metropolitana en 10 áreas piloto de 1km<sup>2</sup>, representativas de toda la información 1:2000.

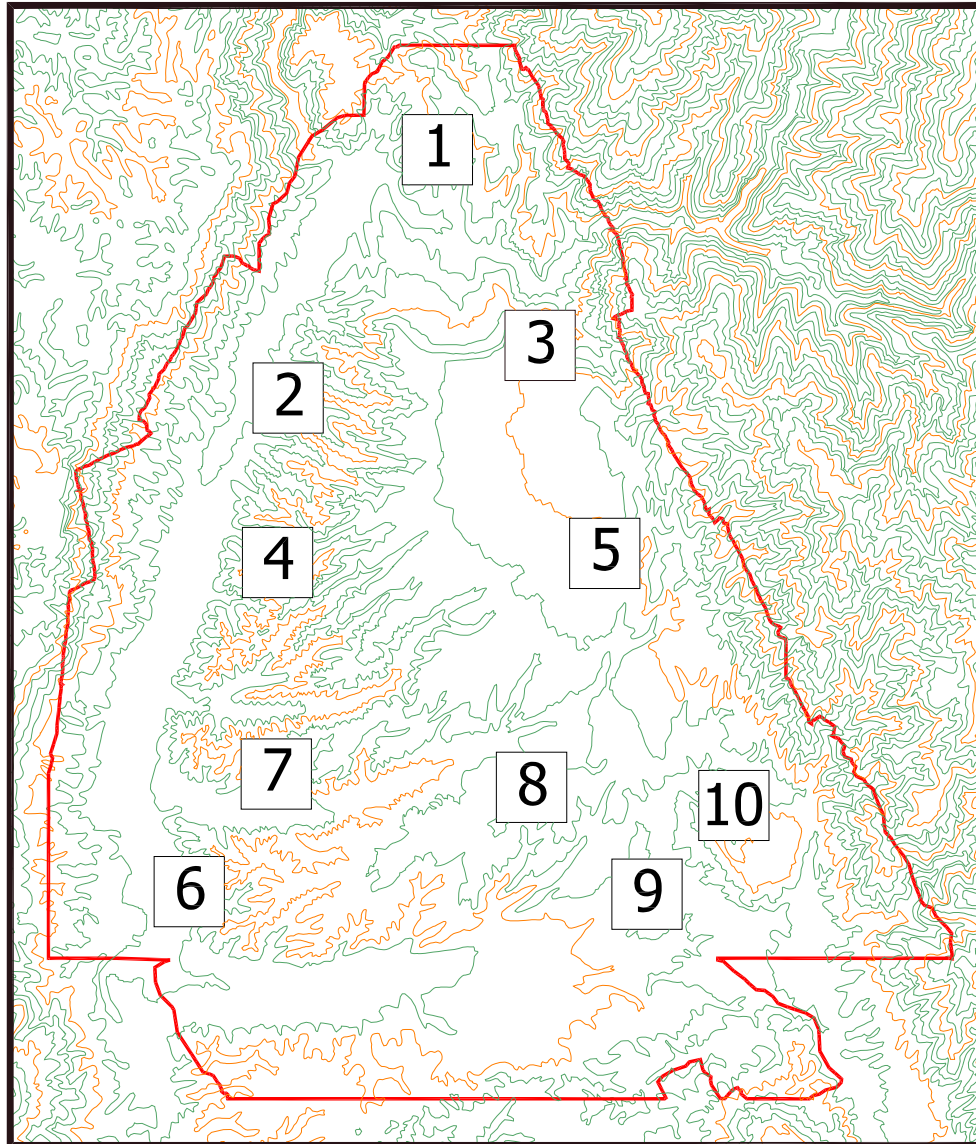
El hardware utilizado para el cómputo de los modelos se describe a continuación:

- Procesador Intel Pentium IV de 3200 MHz
- Memoria RAM física 1000 MB (1 Giga)
- Bus de datos de 800 MHz

Con la velocidad de procesamiento que proporciona cada uno de los elementos anteriormente descritos el modelo para Bucaramanga con cerca de 12.6 millones de puntos requiere de 200 horas para el cómputo completo del DEM. Por tal razón se hace casi imposible analizar todo el archivo completo de puntos y es necesario realizar una división representativa de la zona para calcular los errores de interpolación.

Las regiones representativas se escogieron, tomado como criterios la pendiente, la densidad de puntos y su distribución espacial alrededor del área metropolitana incluyendo así, zonas planas, escarpadas, onduladas y de media pendiente. Esta decisión se tomo, teniendo en cuenta la cantidad de tiempo necesario para

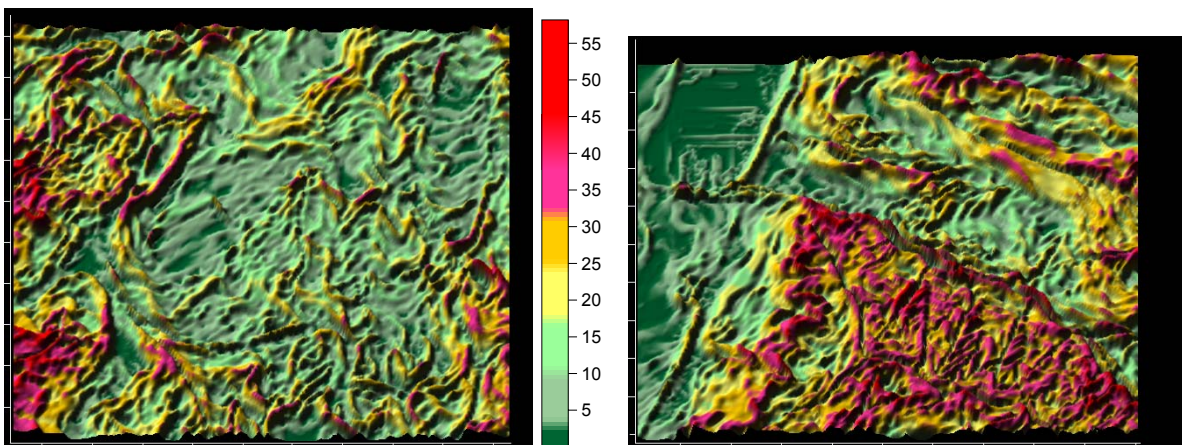
generar el archivo de grilla del modelo. Las representaciones de las áreas piloto se muestran en la *figura (15)*:



**Figura 15** *Regiones representativas del área metropolitana de Bucaramanga*

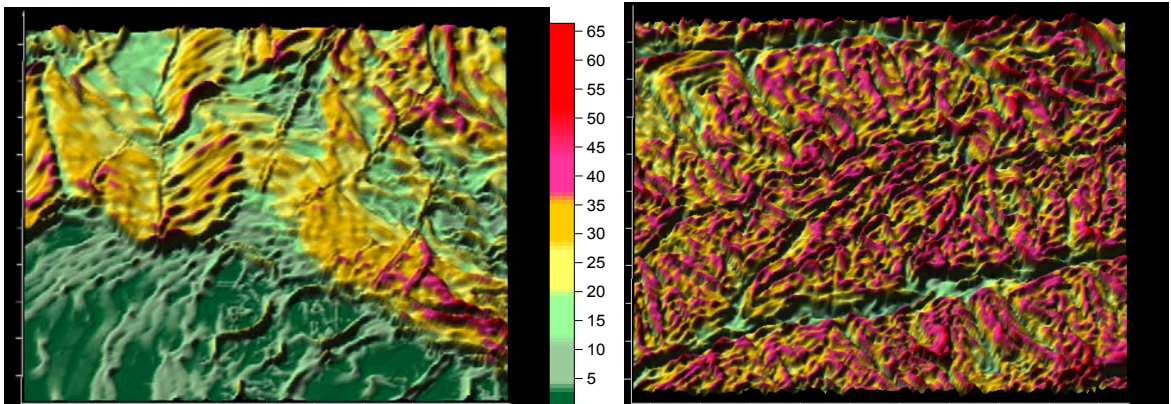
Es conveniente analizar cada una de las áreas piloto anteriores con el objetivo de garantizar que sean totalmente representativas de toda el área metropolitana. Para esto, es necesario construir un mapa de pendientes para las 10 regiones que determine claramente los cambios que ocurren en la topografía y así, cuantificar y clasificar cada una de las áreas según corresponda su variación espacial.

**4.1.1 Mapas de pendiente.** Este procedimiento arroja una medida acerca de que tanta variación topográfica hay en cada una de las regiones; además, se demuestra así la inclusión de todas las posibles variaciones espaciales en cada una de las regiones piloto. La *figura (16)* muestra el mapa de pendientes de las áreas en estudio por medio de colores asignados para cada gradiente. La paleta de colores en la *figura (16)* muestra indica la variación en grados de la topografía.



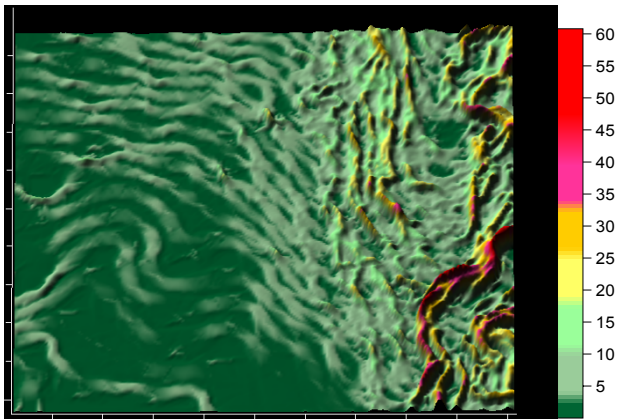
**Figura 16 (a) Área 1**

**Figura 16 (b) Área 2**

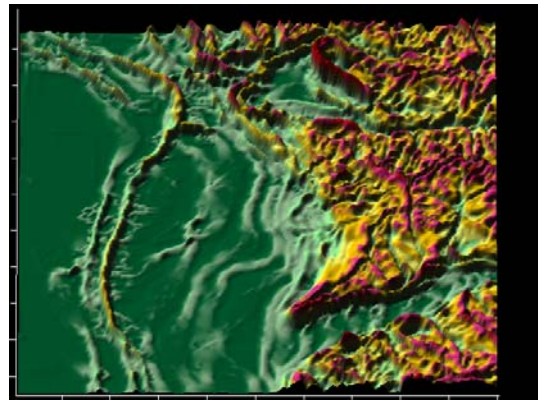


**Figura 16 (c) Área 3**

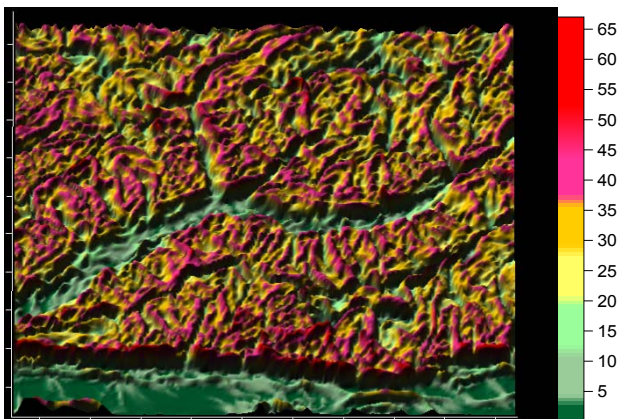
**Figura 16 (d) Área 4**



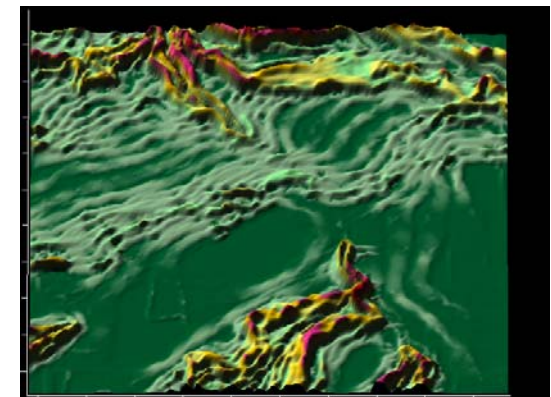
**Figura 16 (e) Área 5**



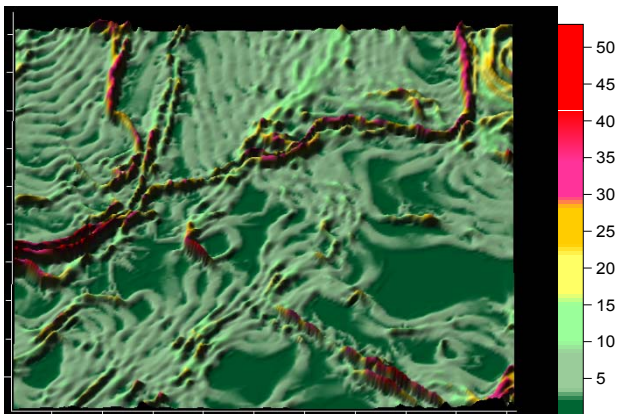
**Figura 16 (f) Área 6**



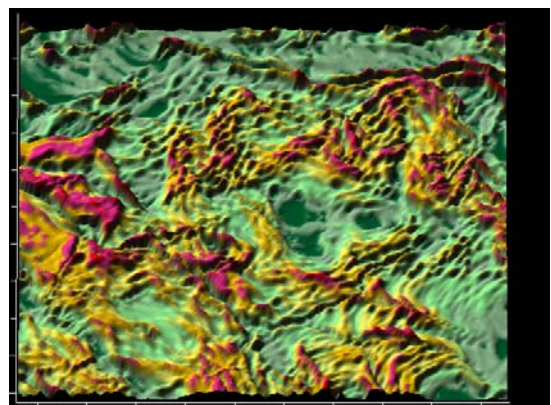
**Figura 16 (g) Área 7**



**Figura 16 (h) Área 8**



**Figura 16 (i) Área 9**



**Figura 16 (j) Área 10**

Las variaciones de pendiente expresadas en porcentaje (%), incluidas en cada una de las regiones piloto se resumen en la *tabla (3)*:

	PLANA	LIGERAMENTE INCLINADO	MODERADO	MODERADAMENTE INCLINADO	EMPINADO	MUY EMPINADO	ESCARPA
<b>AREA</b>	0°-2°	2°-10°	10°-18°	18°-26°	26°-34°	34°-42°	42° ...
1 (a)	2.0454	30.8232	29.0775	20.6879	10.8675	3.4488	3.0497
2 (b)	6.068	18.1667	15.9494	19.5843	19.9504	13.9892	6.292
3 (c)	8.0926	24.9524	12.343	19.6031	24.8799	7.0094	3.1196
4 (d)	0.0542	2.9168	6.3244	15.7764	36.3476	30.9035	7.6771
5 (e)	27.7506	53.3692	10.1262	4.2344	1.7365	0.5785	2.2046
6 (f)	27.2424	28.6596	8.092	9.516	13.5634	9.3043	3.6223
7 (g)	2.8551	10.7166	7.8704	14.0894	29.7884	27.1307	7.5494
8 (h)	29.0468	46.1725	9.144	7.239	4.4845	1.6337	2.2795
9 (i)	18.3686	63.895	9.9004	3.4958	1.7086	0.6121	2.0195
10 (j)	3.2503	26.3933	27.8414	26.1595	12.7222	1.6853	1.948

**Tabla 3 Clasificación de las áreas piloto según pendiente en porcentaje (%)**

Las regiones 5 y 9 correspondientes a las figuras (16 e) y (16 i) corresponden a zonas planas en un 70 %, por otra parte las regiones 4 y 7 (figuras 16 d y 16 g) tienen un gran porcentaje de terreno clasificado como empinado y escarpado. Las demás figuras, muestran variación de todo tipo pasando por partes planas, de media pendiente y algunas zonas empinadas. Un análisis estadístico de las áreas piloto, también nos da una idea general de la variabilidad topográfica que se presenta en cada zona; además, nos da un mejor entendimiento de su comportamiento y estructura. Las funciones estadísticas básicas que se calcularon en cada una de las áreas se resumen en la tabla (4):

AREA	NUMERO DATOS	MEDIA	SD	MAX	MEDIANA	MIN	C. VAR
1	79004	749.4	36.17	846.000	746.000	652.000	0.04827
2	97512	720.48	41.47	814.000	720.000	628.000	0.05756
3	98703	888.71	72.25	1028.000	892.000	720.000	0.0813
4	169956	776.33	39.66	890.000	776.000	680.000	0.05108
5	25514	994.19	20.28	1046.000	990.000	960.000	0.02095
6	99779	740.51	28.89	808.000	742.000	680.000	0.03902
7	193150	799.77	31.05	874.000	802.000	712.000	0.0388
8	49156	872.36	18.07	904.000	874.000	816.000	0.02071
9	33364	852.41	14.83	900.000	852.000	820.000	0.01740
10	74709	956.21	30.59	1030.000	956.000	888.000	0.03199

**Tabla 4 Estadísticas básicas de las regiones piloto**

La *tabla (4)* indica que las diferencias de nivel no superan los 86 metros. Las *regiones 3, 4 y 2 (Figura 16 c, d, b)* presentan una gran variabilidad del terreno, con cambios considerables de pendiente; correspondientes a la zona norte y nor-occidental del área metropolitana de Bucaramanga (escarpa). Las *zonas 1, 6, 7 y 8 (figura 16 a, f, h)* corresponden a zonas de pendiente media como el valle de río de Oro y la entrega de éste al río Surata para conformar el río Lebrija.

Este análisis demuestra que la mayoría de las regiones representativas contiene todo tipo de terreno (plano, medio, escarpa). Sin embargo, algunos tienen tendencia a ser escarpados y otros casi planos. Estas alternativas demuestran que la escogencia de las regiones representativas es adecuada, debido a que un mayor porcentaje las regiones contienen todo tipo de topografía y además se analizaran algunas que solo tienen una sola tendencia.

Las 10 regiones descritas anteriormente se analizaran utilizando diferentes algoritmos de interpolación, para determinar cual de todos los procesos es el más eficiente y confiable para la generación completa del DEM. El punto de partida para la generación de la mejor grilla en la construcción del DEM, será un análisis de los datos por medio de la validación cruzada.

La validación cruzada tendrá a cargo la valoración entre cada uno de los métodos interpoladores, y proporcionara la infamación necesaria para descartar o para inspeccionar más a fondo el método interpolador que ajuste mejor los datos.

## **4.2 VALIDACIÓN CRUZADA**

El proceso de la validación cruzada consiste en suprimir un dato muestral en un punto (x,y) e interpolarlo a partir de los datos restantes con base al método de interpolación escogido para la generación del DEM. Usando este valor y el dato

muestral conocido (dato de entrada), el error en la interpolación es calculado como:

$$\text{Error} = \text{Valor interpolado} - \text{Valor muestral} \quad \text{Ec 2}$$

Este procedimiento se repite hasta completar los "n" puntos que se requieran validar. La cantidad de puntos a validar también debe ser una muestra representativa de la cantidad de datos que se tienen para analizar (datos de entrada). Este proceso genera "n" errores de interpolación (uno para cada dato), que pueden ser cuantificados por medio de estadísticos. En este proyecto se utilizo el error medio cuadrático (RMSE, por sus siglas en ingles) como medida para cuantificar los errores y escoger los mejores ajustes.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Ei)^2} \quad \text{Ec 3}$$

*n* = numero de puntos a validar

(surfer V8 tutorial, 1998)

$Ei = Z_{DEM} - Z_{ENTRADA}$

El RMSE es una medida para cuantificar la incertidumbre de los datos en un DEM y mide, la dispersión de la distribución de desviaciones entre la elevación de los datos originales y los datos del DEM. Entre mas grande sea el valor del RMSE, mayor es la discrepancia entre las dos series de datos. La mayor ventaja de este estadístico, radica en su elemental cálculo y sencillo concepto. Sin embargo, el RMSE es esencialmente una medida global de errores para el modelo y no revela su forma de distribución alrededor de la región de estudio.

#### **4.3 ALGORITMOS MATEMÁTICOS DE INTERPOLACIÓN**

Los algoritmos de interpolación para realizar la comparación estadística del RMSE en cada una de las áreas piloto serán:

- Krigeaje ordinario

- Funciones de base radial
- Mínima curvatura
- Inverso de la distancia
- Vecino mas cercano
- Triangulación con interpolación lineal
- Polinomio local

**4.3.1 Krigeaje ordinario.** Es un método de interpolación que pondera el peso de cada punto muestral  $X_i$  en un punto no muestral  $X_o$  según una función estadística de la distancia entre estos puntos.

$$Z_o = \sum_{i=1}^n \lambda_o^i Z_i \quad \text{Ec 4}$$

*(Cressie, statistics for spatial data)*

Se fundamenta en la teoría de las variables regionalizadas, que consiste en la búsqueda de unos interpoladores óptimos que producen residuos insesgados y con mínima varianza.

$$\begin{aligned} \text{Var}[Z^*(x) - Z(x)] &= \text{Minimo} \\ E[Z^*(x) - Z(x)] &= 0 \end{aligned} \quad \text{Ec 5}$$

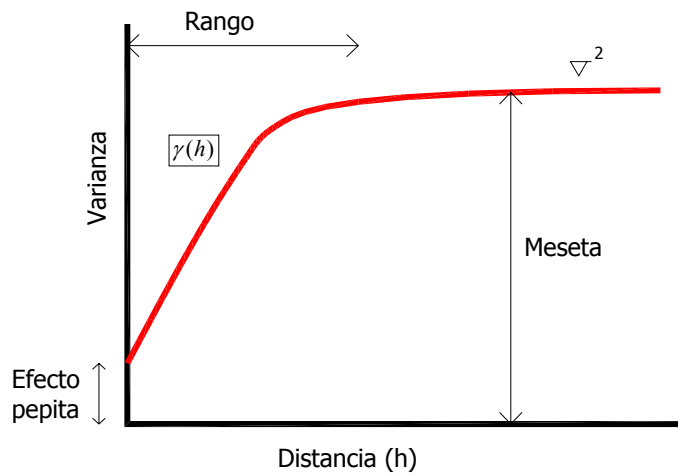
*(Cressie, statistics for spatial data)*

Las principales características que hacen del *krigeado* un método de estimación muy superior a los tradicionales, como el inverso ponderado de la distancia, la triangulación, etc., son las siguientes:

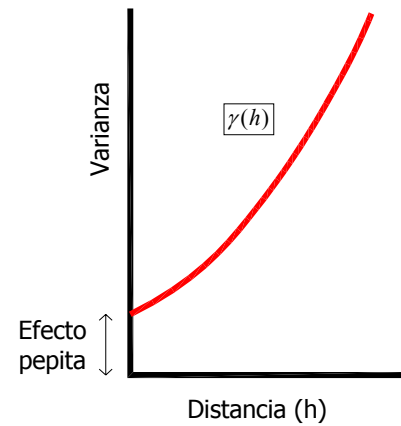
- Mientras que los métodos tradicionales utilizan el concepto euclidiano de la distancia para el cálculo de los pesos que se aplicarán a cada dato muestral, el *krigeado* considera tanto la distancia como la geometría de la localización de las muestras.

- Mediante el *krigeado* se minimiza la varianza del error esperado (diferencia entre el valor real y el estimado). Como el valor real en un punto no muestral es desconocido, el *krigeado* emplea un modelo conceptual con una función aleatoria asociada a los valores reales.
- Los métodos geoestadísticos muestran una gran flexibilidad para la interpolación, pudiéndose estimar valores puntuales o en bloques, así como métodos para incorporar información secundaria que esté relacionada con la variable principal. Todos estos métodos dan lugar a unas superficies muy suaves, además de una estimación de la varianza en todos los puntos, lo cual no puede realizarse con otros métodos de interpolación.

**EL VARIOGRAMA.** El variograma es una función monótona creciente que alcanza un valor límite, denominado meseta, equivalente a la varianza muestral. La meseta se alcanza para un valor de  $h$  conocido como rango o alcance (*Figura 15 (a)*). El rango determina la zona de influencia en torno a un punto, más allá del cual la auto correlación es nula. Sin embargo, no todos los variogramas alcanzan una meseta. Es posible que un variograma no tienda asintóticamente a la varianza, sino que tienda a infinito cuando así lo haga  $h$  (*Figura 17 (b)*). El variograma representa la tasa media de cambio de una propiedad con la distancia. El hecho de que dos observaciones próximas sean más parecidas que si estuvieran más separadas se refleja en el mismo concepto del variograma. La dependencia espacial disminuye a medida que se incrementa la distancia,  $h$ , y finaliza a una cierta distancia, el rango. Más allá del rango, la tasa media de cambio es independiente de la separación entre las observaciones.



**Figura 17 (a)**



**Figura 17 (b)**

Por definición,  $\gamma(0) = 0$ . Pero, con frecuencia,  $\gamma(0)$  muestra un valor positivo, dando lugar al denominado efecto pepita *figura 17 (a)*. Éste se debe a la variabilidad a una distancia más pequeña que el menor  $h$  considerado. Si los variogramas tienden a la horizontalidad, con valores próximos a la varianza muestral, ello indica la ausencia de una estructura espacial. El semivariograma o simplemente variograma, a partir de los datos disponibles, se estima como:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad \text{Ec 6}$$

Para la determinación del variograma experimental o empírico deben cumplirse una serie de etapas:

- El número de observaciones o puntos muestrales requeridos para estimar un variograma debe estar alrededor de 100 en condiciones de isotropía; si existe anisotropía los datos tienen que ser mucho más abundantes.
- El número de pares de datos que se requiere para el cálculo de cada punto del variograma debe ser al menos de 30-50.

Una vez que se han definido los puntos del variograma experimental, será necesario ajustar un modelo a dichos puntos, denominado variograma teórico, debido a la imposibilidad de trabajar con un variograma como el experimental, carente de una función matemática precisa. Los modelos básicos más usados son los denominados esférico, exponencial, gaussiano, potencial, lineal y el efecto pepita puro, pudiéndose combinar linealmente.

**PARÁMETROS DEL VARIOGRAMA.** Los cuatro parámetros fundamentales que componen un variograma experimental se definen como sigue:

**EFFECTO PEPITA (NUGGET).** Describe el efecto de la variabilidad de los datos muestrales en un mismo punto, debido a errores en la toma de la información. Su inclusión dentro del algoritmo interpolador lo convierte en un método inexacto, es decir, cuando un nodo de la grilla coincide exactamente con un dato muestral, se presentara una pequeña diferencia entre el valor muestral y el valor interpolado. Si el valor del efecto nugget es establecido como cero (0), el valor muestral y el valor interpolado en el caso que coincidan sus coordenadas  $(x,y)$  tendran la misma elevación. El valor numérico para el efecto pepita se considera como la suma de los ajustes en la varianza y la microvarianza.

**RANGO (RANGE).** Es la distancia para la cual los datos muestrales dejan de estar correlacionados, es decir, a distancias más grandes que el rango, los datos son independientes.

En el variograma experimental esta distancia se puede observar cuando los datos de varianza se tornan constantes (meseta).

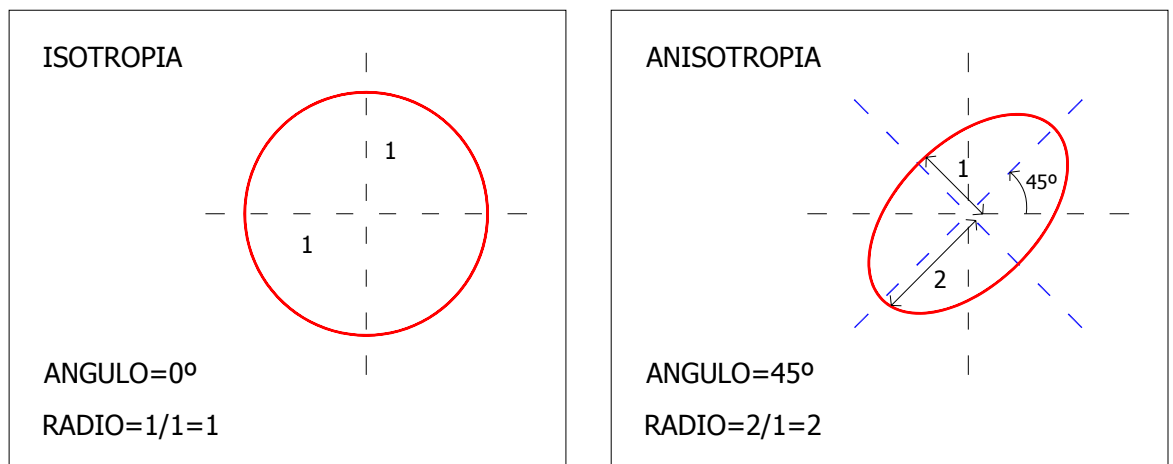
**MESETA (SILL).** Valor para el cual la varianza, representada en el variograma, permanece constante. Para cualquier serie de datos, este valor puede ser

obtenido inspeccionando el variograma experimental a grandes distancias cuando el valor  $\gamma(h)$  tiende a estabilizarse.

**ANISOTROPIA.** En ocasiones, la magnitud de la covarianza entre muestras depende no solo de las distancias entre ellas sino también de su dirección, dando lugar a variogramas experimentales con diferentes parámetros en cada una de las direcciones de análisis.

Este fenómeno direccional se denomina **anisotropía**, que se presenta en dos formas distintas:

- *Anisotropía geométrica:* Ocurre cuando el rango del variograma cambia en diferentes direcciones, permaneciendo constante el valor de la meseta.
- *Anisotropía zonal:* Existe cuando el rango y la meseta del variograma cambian con la dirección.

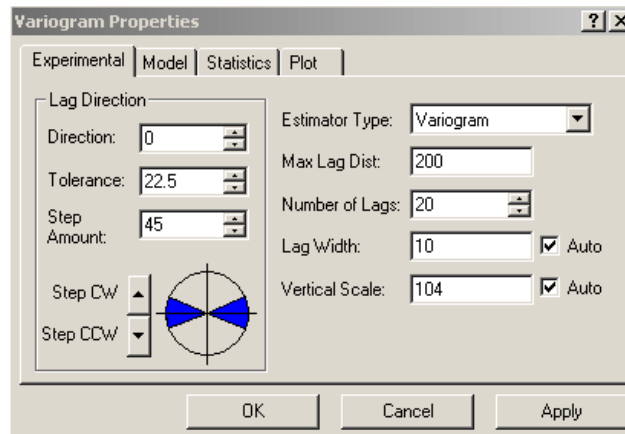


**Figura 18** Relación de anisotropía para búsqueda de datos muestrales

**ANÁLISIS DE LAS ÁREAS PILOTO.** En primer lugar se construirán los variogramas experimentales para cada una de las regiones representativas en cuatro direcciones diferentes. Para ello, utilizaremos SURFER V 8.0 como herramienta para el cálculo y ajuste de dicho variograma. Debido a que el DEM

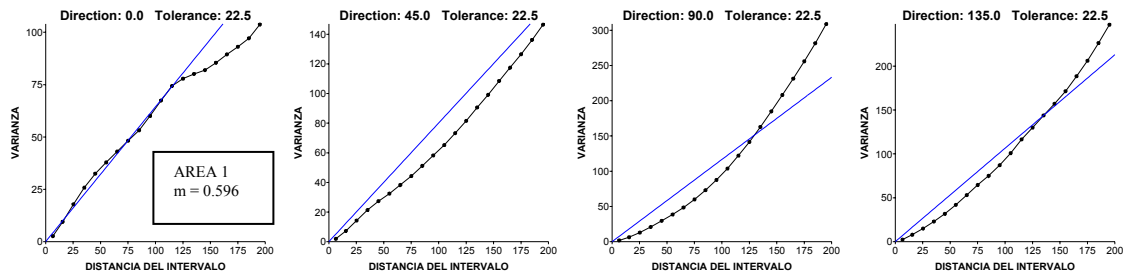
se debe construir necesariamente para toda el área de análisis, se necesita un variograma que represente en su totalidad a dicha región. Por tal razón, se calculará un variograma final teórico como representación topográfica de toda la zona de estudio para posteriormente generar el DEM.

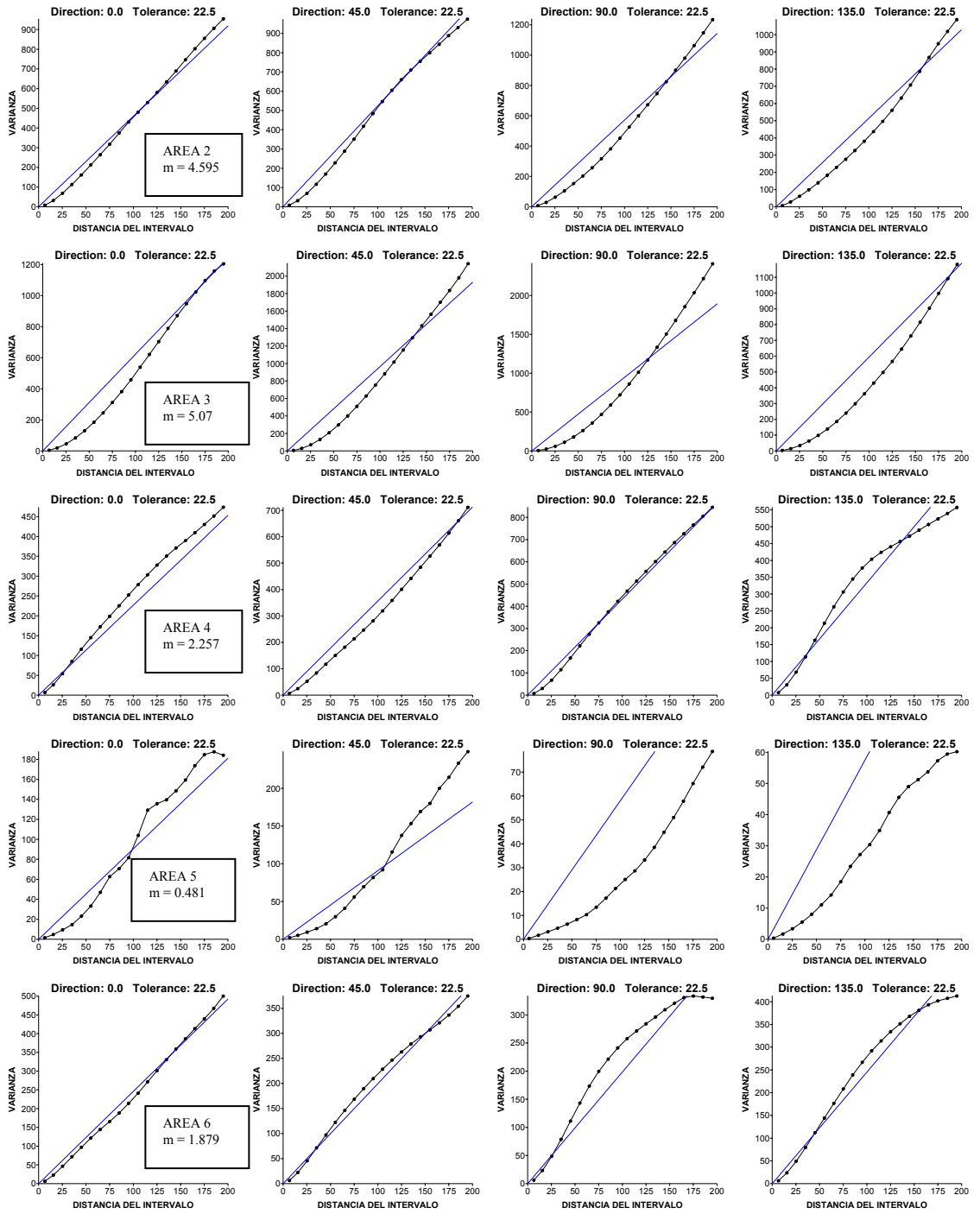
**CONSTRUCCIÓN DEL VARIOGRAMA.** La máxima distancia para el calculo del variograma experimental fue establecida como 200 metros. Este valor se considero debido a la gran densidad de puntos muestrales alrededor del punto a interpolar. Puntos más alejados que esta distancia, no intervendrán en el proceso de interpolación.

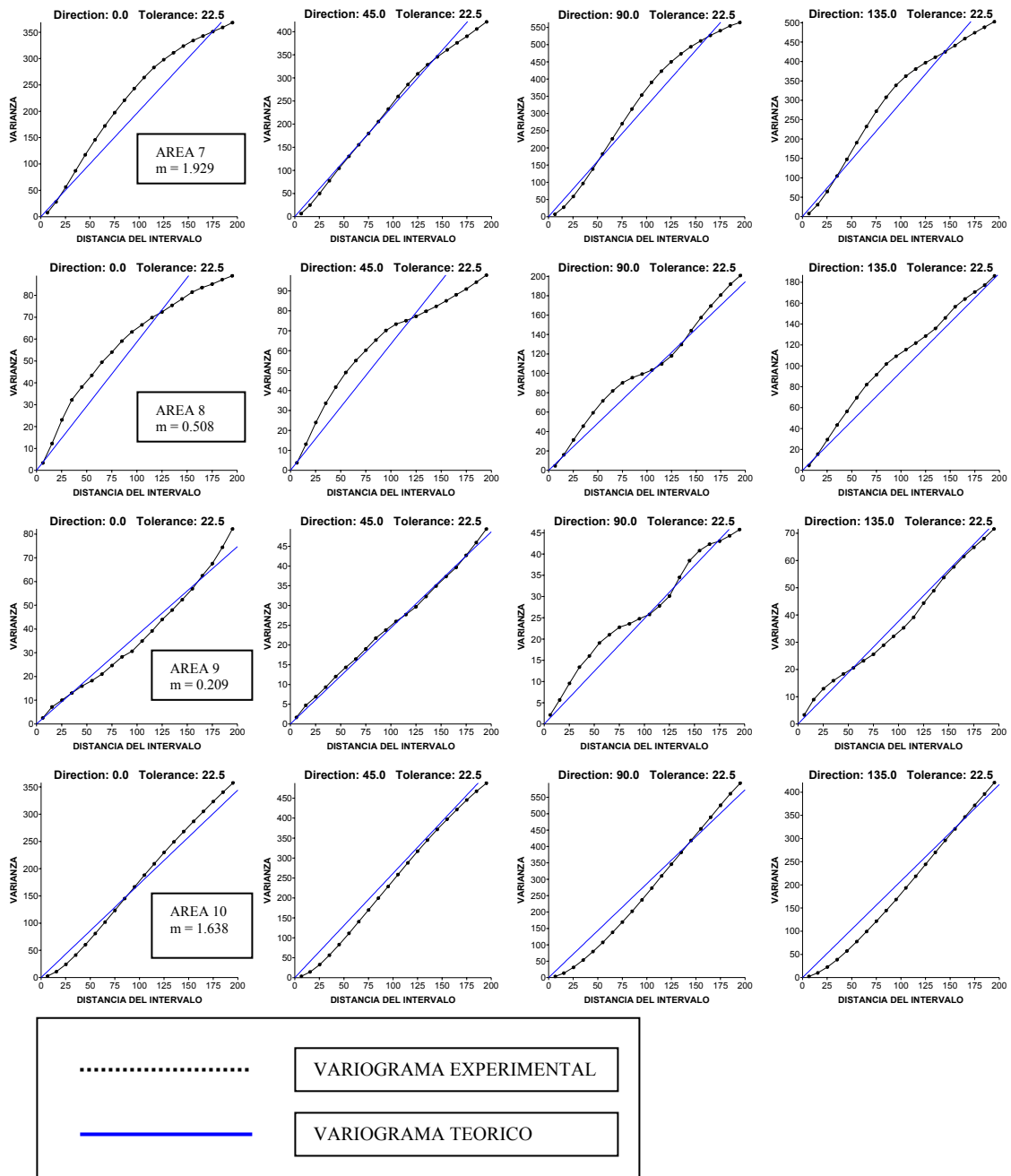


**Figura 19 Cuadro de dialogo (surfer V8.0) para construcción del variograma**

En el cuadro de dialogo anterior (figura 19), se modificarán las direcciones para la construcción de los variogramas experimentales en cada una de las regiones piloto. Las direcciones de análisis serán: 0°, 45°, 90° y 135°:







**Figura 20 Variogramas experimentales y teóricos para las 10 regiones representativas**

En la *figura 20* están representados todos los variogramas experimentales y teóricos lineales para cada una de las 10 regiones representativas.

**ANÁLISIS DE VARIOGRAMAS EXPERIMENTALES.** La representación de los variogramas experimentales (*figura 20*) muestra una tendencia lineal en un 90% de las áreas piloto. Esto nos da una idea general, acerca de cual va a ser la tendencia para el variograma final que se utilizara para generar todo el DEM.

El variograma teórico para la generación de la grilla, es de tipo lineal y su pendiente se estimará como un promedio de todas las pendientes de cada uno de los variogramas teóricos de las áreas piloto. Para el cálculo del promedio se restringirá el uso del *área 5*, debido a que no presenta un ajuste de tipo lineal. De esta manera obtenemos el variograma representativo de toda el área de estudio.

AREA	PENDIENTE
1	0.596
2	4.595
3	5.07
4	2.257
6	1.879
7	1.929
8	0.508
9	0.2091
10	1.638
Prom:	2.0757

**Tabla 5 Pendiente promedio de variogramas teóricos**

Una vez obtenida la pendiente del variograma teórico lineal, se analizara la dirección (anisotropía) en la que el proceso de validación cruzada arroje los menores valores de error medio cuadrático (*tabla 6*).

	AREA	PENDIETE	RMSE		AREA	PENDIETE	RMSE
ANISOTROPIA: 1 ANGULO 0 °	1	2.076	0.096	ANISOTROPIA: 1.5 ANGULO 90°	1	2.076	0.094
	2		0.185		2		0.199

	3		0.08
	4		0.186
	5		0.115
	6		0.124
	7		0.16
	8		0.092
	9		0.139
	10		0.119
		Promedio	0.1296

	3		0.087
	4		0.198
	5		0.105
	6		0.142
	7		0.191
	8		0.096
	9		0.148
	10		0.139
		Promedio	0.1399

AREA	PENDIETE	RMSE
------	----------	------

AREA	PENDIETE	RMSE
------	----------	------

ANISOTROPIA: 1.5 ANGULO 0°	1	2.076	0.122
	2		0.195
	3		0.089
	4		0.201
	5		0.135
	6		0.141
	7		0.175
	8		0.095
	9		0.152
	10		0.126
		Promedio	0.1431

ANISOTROPIA: 2 ANGULO 90°	1	2.076	0.109
	2		0.233
	3		0.107
	4		0.231
	5		0.104
	6		0.168
	7		0.217
	8		0.117
	9		0.16
	10		0.169
		Promedio	0.1615

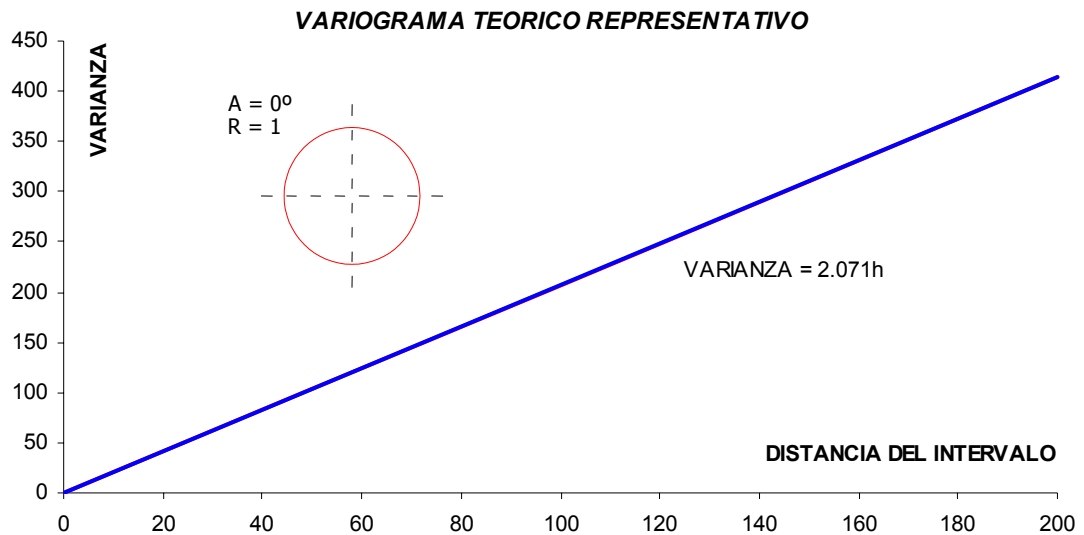
AREA	PENDIETE	RMSE
------	----------	------

ANISOTROPIA: 2 ANGULO 0°	1	2.076	0.143
	2		0.219
	3		0.103
	4		0.227
	5		0.167
	6		0.159
	7		0.201
	8		0.116
	9		0.158
	10		0.145
		Promedio	0.1638

**Tabla 6 Cálculo de la relación de anisotropía y ángulo para el variograma lineal**

De esta manera tenemos todos los parámetros necesarios para construir nuestro DEM desde SURFER V8.0. Tenemos un variograma de tipo lineal con una

pendiente de 2.076, una relación de anisotropía de R=1 y dirección = 0°. Esto se representa en la *figura (21)*:



**Figura 21 Variograma teórico final representativo para construir el modelo por Kriging**

**4.3.2 FUNCIONES DE BASE RADIAL.** Es un interpolador que emplea una función base, dependiente de la distancia entre el punto a interpolar y los datos muestrales vecinos. El valor del punto a interpolar  $Z^*$  se halla de acuerdo a la siguiente ecuación:

$$Z^* = \sum_{i=1}^n \lambda_i \times F(d_i) \quad \text{Ec 7}$$

(PDF.Aguilar Torres, Manuel Ángel, Universidad de Almería)

$Z^*$  es el valor a interpolar,  $\lambda_i$  son los pesos calculados resolviendo un sistema lineal de ecuaciones,  $n$  es la cantidad de puntos muestrales vecinos utilizados en la interpolación y  $F(d_i)$  es la función de base radial dependiente de la distancia entre puntos.

Las funciones  $F(d_i)$  que se consideran en este proyecto son:

**MULTICUADRÁTICA INVERSA.**

$$F(d_i) = \frac{1}{\sqrt{d_i^2 + R^2}} \quad \text{Ec 8}$$

**MULTILOGARITMICA.**

$$F(d_i) = \log(d_i^2 + R^2) \quad \text{Ec 9}$$

**MULTICUADRÁTICA.**

$$F(d_i) = \sqrt{(d_i^2 + R^2)} \quad \text{Ec 10}$$

**NATURAL CUBIC SPLINE.**

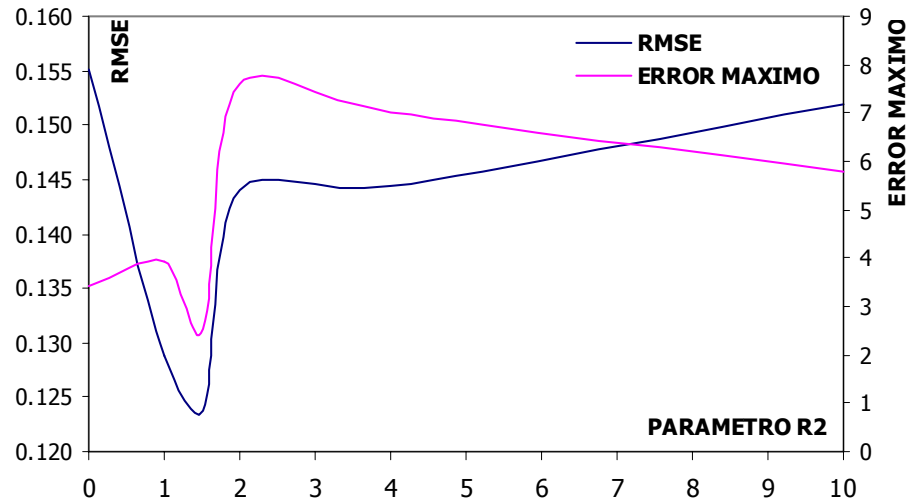
$$F(d_i) = \sqrt{(d_i^2 + R^2)^3} \quad \text{Ec 11}$$

**THIN PLATE SPLINE.**

$$F(d_i) = (d_i^2 + R^2)\log(d_i^2 + R^2) \quad \text{Ec 12}$$

El parámetro  $R^2$  es un factor de suavizado. Estableciendo este factor igual a cero hace que este método se convierta en un interpolador exacto, cualquier punto interpolado que coincida en sus coordenadas  $(x,y)$  con un punto muestral, tomará el valor del dato muestral. Un factor de suavizado elevado producirá una superficie muy suavizada aumentando los errores entre el modelo y la superficie real. Para el análisis de las 10 regiones representativas del área metropolitana se eligió una función de base radial Multicuadrática analizando los errores para los parámetros de  $R^2$  0, 1, 1.5, 2.0 y 10. Haciendo validación cruzada y comparando el error medio cuadrático promedio en las 10 regiones para cada uno de los parámetros,

se concluye que un valor aceptable de  $R^2$  es 1.5, como se muestra en la *figura (22)*:



**Figura 22: RMSE y error máximo para diferentes valores del parámetro  $R^2$**

El valor de este parámetro es utilizado en la validación de las 10 áreas con los resultados que se muestran en *tabla (7)* al final de este capítulo.

**4.3.3 Inverso de la distancia.** Es un método determinístico, que halla el valor de la variable  $Z^*$  basado en la idea de que las observaciones más cercanas deben tener más peso en la determinación del valor interpolado en un punto  $X_0$ .

$$Z(x_0) = \frac{\sum_{i=1}^n w(d_i) \cdot z(x_i)}{\sum_{i=1}^n w(d_i)} \quad \text{Ec 13}$$

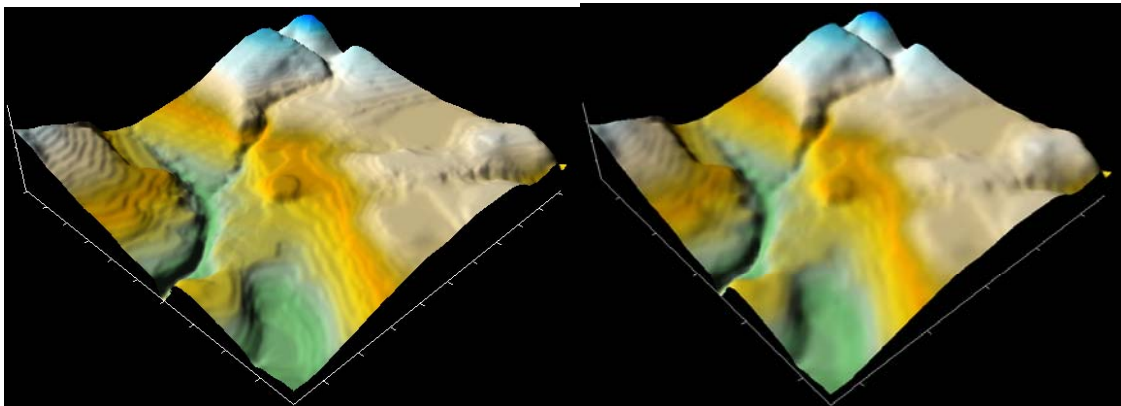
$w(d_i)$  es una función de peso de los  $n$  puntos muestrales que intervienen en el cálculo del valor a interpolar y  $d_i$  la distancia de cada punto a  $x_0$ .

Las funciones de ponderación mas usuales toman la forman de  $w(d_i) = d_i^{-u}$  con valores de  $u$  entre 1 y 2. El uso de exponentes elevados implica dar menos peso relativo a puntos mas alejados de  $x_i$ . Los valores interpolados por este método oscilan siempre entre el rango máximo y mínimo de los datos muestrales, lo que reduce su eficacia para modelar cotas mas altas o mas bajas de una superficie en caso de que estas no pertenezcan al conjunto de puntos muestrales.

La función de ponderación que trabaja surfer V8.0 incluye un parámetro de suavizado  $\delta$  para incorporar una tendencia mas homogénea al conjunto de datos, al igual que lo incluye el Krigeadado con el efecto pepita ò al factor  $R^2$  en la función de base radial;

$$w(d_i, \delta) = (\sqrt{d_i^2 + \delta^2})^{-u} \quad \text{Ec 14}$$

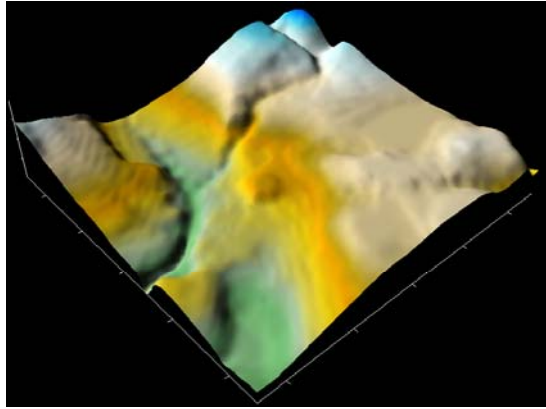
Este efecto se puede apreciar claramente en el siguiente ejemplo; la *figura (23)* muestra la grilla para el área piloto 9, con y sin el parámetro  $\delta$ .



**figura 23 (a) Parámetro de suavizado 0**

**figura 23 (b) Parámetro de suavizado 5**

**5**



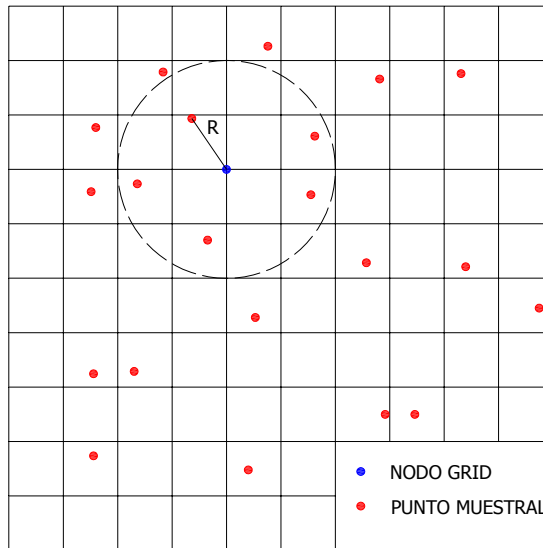
**figura 23 (c) Parámetro de suavizado 10**

En la *figura 23 (a)*, un factor  $\delta = 0$ ; implica una influencia mayor de la distancia en el calculo de la función de ponderación  $w(d_i)$  y los contornos de los datos muestrales son fácilmente identificables (curvas de nivel). Las *figuras 23 (b)* y *23 (c)*, con un factor  $\delta = 5$  y  $\delta = 10$  muestran una superficie mas suavizada donde los contornos no son tan fácilmente identificables. Además, varios de los detalles como depresiones y cambios bruscos de pendiente no se identifican. A medida que el factor  $\delta$  aumente, la influencia de la distancia disminuirá en la *ecuación (14)* y por ende los errores medios cuadráticos aumentaran, generando así un modelo menos preciso.

Los valores correspondientes de RMSE para cada una de las figuras son: 0.466, 0.702 y 0.749 respectivamente.

Los resultados obtenidos con este método de interpolación, se muestran y analizan en la *tabla (7)* al final del capítulo.

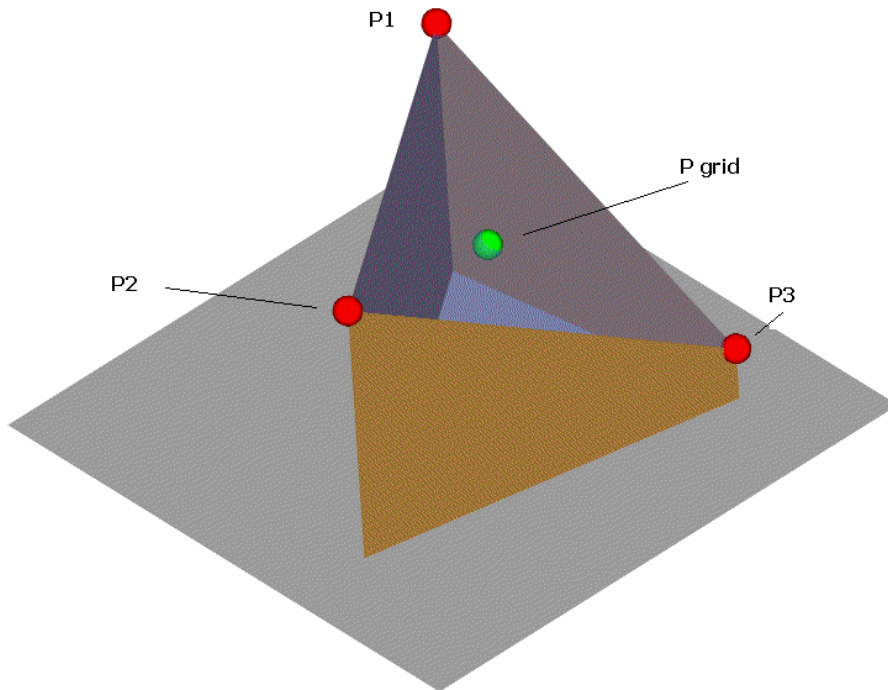
**4.3.4 Vecino más cercano.** El método del vecino más cercano es un interpolador exacto que asigna el valor del punto muestral más cercano al nodo de la grilla que se desea interpolar.



**Figura 24 Radio menor de búsqueda para el método del vecino mas cercano**

Los tiempos de operación de este método interpolador son relativamente cortos comparados con la función de base radial o krigeadó. Debido a la alta densidad de puntos muestrales que contienen los archivos .dxf, un valor de radio de búsqueda de 500 m será más que suficiente para encontrar la muestra más cercana al punto de la grilla.

**4.3.5 Triangulación con interpolación lineal.** Es un método de interpolación exacto basado en la generación de triángulos de Delaunay cuyos vértices coinciden con los puntos muestrales. Para la interpolación de los nodos de la grilla, se supone que estos pertenecen a la superficie plana que se apoya en los vértices de cada triángulo; Así, como se muestra en la *figura (25)*, el proceso de interpolación consiste en calcular la coordenada  $Z$  de cualquier punto  $P(x,y)$  perteneciente al triángulo cuyos vértices son los datos muestrales  $P_1, P_2$  y  $P_3$ .



**Figura 25 Triangulación por el método de Delaunay**

Este método funciona bien cuando los datos se encuentran regularmente distribuidos y por ser un interpolador exacto, los valores de la grilla coinciden exactamente con los datos muestrales cuando ambos yacen en el mismo punto del plano  $(x, y)$ .

**4.3.6 Mínima curvatura.** Con este método se trata de generar una superficie de interpolación tal que pasando por los puntos dato la curvatura es mínima. El procedimiento equivale a adaptar una lámina elástica a cada uno de los puntos muestrales, y de forma que el encorvamiento entre puntos dato sea mínimo.

En resumen, las 10 regiones representativas presentan los siguientes valores de errores utilizando los métodos de interpolación anteriormente descritos:

En todas las áreas, con la función de base radial se obtuvieron los menores errores; seguida por triangulación con interpolación lineal, Krigeado y vecino más cercano.

	<i>KRIGING</i>	<i>RADIAL BASIS</i>	<i>INVERSO DE LA DISTANCIA</i>	<i>MINIMA CURVATURA</i>	<i>VECINO MAS CERCANO</i>	<i>TRIANGULACION I. LINEAL</i>	<i>POLINOMIO LOCAL</i>
<b>AREA</b>	<b>RMSE</b>	<b>RMSE</b>	<b>RMSE</b>	<b>RMSE</b>	<b>RMSE</b>	<b>RMSE</b>	<b>RMSE</b>
1	0.1114	0.0825	0.3795	0.6856	0.1194	0.0972	0.8698
2	0.179	0.142	0.5078	0.932	0.183	0.1739	1.0106
3	0.092	0.0822	0.304	0.5607	0.0916	0.0871	0.6409
4	0.1596	0.1314	0.4983	0.9675	0.3067	0.1447	0.972
5	0.1313	0.1004	0.3781	0.6154	0.1431	0.1043	0.8505
6	0.1265	0.0902	0.4189	0.9905	0.1231	0.1284	0.9456
7	0.1623	0.1182	0.4941	1.1055	0.1642	0.198	0.9983
8	0.0844	0.0536	0.3468	0.7055	0.0675	0.0713	0.8022
9	0.127	0.0998	0.3927	0.63	0.089	0.154	0.7961
10	0.1171	0.0843	0.3824	0.5853	0.1191	0.0769	0.7877
Prom.	0.12906	<b>0.0985</b>	0.41026	0.7778	0.14067	0.12358	0.86737

***Tabla 7 Resumen del valor de RMSE para todos los interpoladores analizados en la construcción del DEM sobre Bucaramanga***

Las diferencias porcentuales entre cada uno de los interpoladores anteriores y la función de base radial son 25%, 31% y 42% respectivamente.

En adelante se desarrollara un análisis exhaustivo para las funciones de base radial, buscando la mejor función y los parámetros de ajuste para encontrar el menor valor de RMSE para todos los datos necesarios en la construcción del DEM.

## 5. ANÁLISIS DEL INTERPOLADOR FINAL PARA EL DEM DE BUCARAMANGA

Como se muestra en la *tabla (7)* del capítulo anterior; el valor promedio del error medio cuadrático para las 10 regiones piloto, es menor cuando se implementa una función de base radial en el proceso de validación cruzada. Sin embargo, se requiere hacer un mejor ajuste de los parámetros para la generación final del modelo, ya que en el análisis de las áreas piloto solo se obtuvo un valor tentativo para el factor  $R^2$  de 1.5. El ajuste completo incluye la determinación de la mejor función de base radial  $F(d_i)$  a emplear en el algoritmo de interpolación, el número total de puntos a validar en el cálculo del error medio cuadrático, el parámetro  $R^2$ , la dirección y la relación de anisotropía.

### 5.1 DETERMINACIÓN DE LA FUNCIÓN DE BASE RADIAL

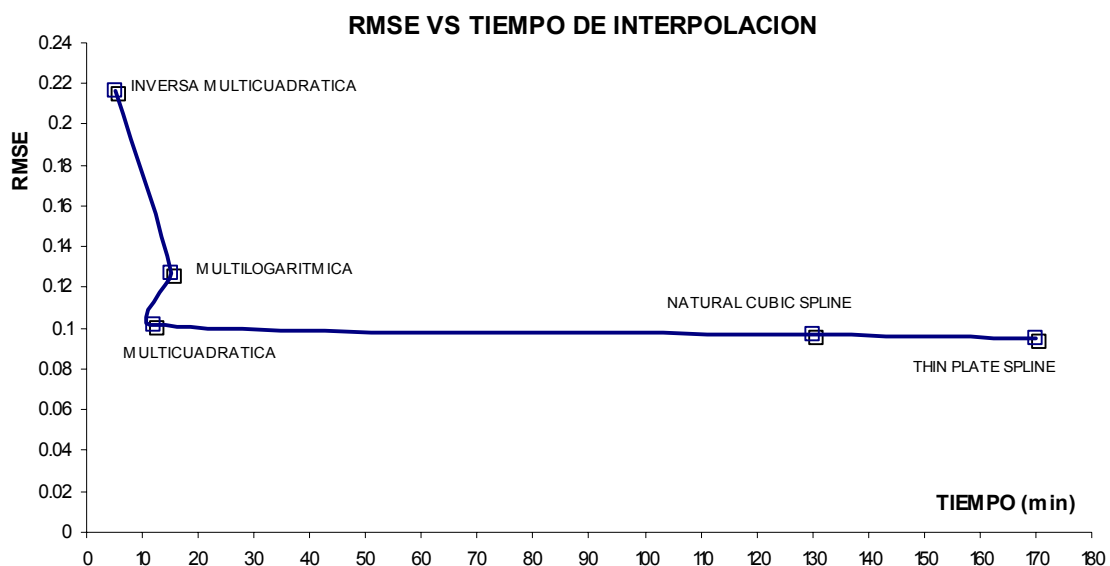
El programa surfer V8 incorpora 5 funciones distintas de base radial: multicuadrática, multicuadrática inversa, multilogarítmica, spline cúbico y spline plano. Estas funciones fueron utilizadas en las 10 áreas piloto para comparar el resultado de sus errores haciendo validación cruzada, tal y como se muestra en la *tabla (8)*.

Analizando los resultados de la *tabla (8)* se concluye que la mejor función de interpolación es el spline plano, con un valor del RMSE de 0.09516, le sigue la función spline cúbico con un RMSE de 0.09735 y la función multicuadrática con un valor de 0.1036. Las funciones multicuadrática inversa y multilogarítmica no proporcionan buenos resultados comparadas con las restantes funciones, generando valores del RMSE de 0.21673 y 0.12715 respectivamente.

<b>FUNCIONES DE BASE RADIAL</b>					
	<b>INV. MULTICUADRATICA</b>	<b>MULTILOGARITMICA</b>	<b>MULTICUADRATICA</b>	<b>CUBIC SPLINE</b>	<b>PLATE SPLINE</b>
AREA	RMSE	RMSE	RMSE	RMSE	RMSE
1	0.2284	0.1126	0.0818	0.0737	0.0743
2	0.2707	0.1697	0.1507	0.144	0.1393
3	0.1643	0.0948	0.0808	0.0848	0.0847
4	0.2323	0.1578	0.1308	0.1374	0.1374
5	0.2564	0.1337	0.0871	0.0829	0.0772
6	0.1735	0.1029	0.094	0.0969	0.0867
7	0.196	0.1481	0.1353	0.1311	0.1261
8	0.1705	0.0863	0.0588	0.058	0.0561
9	0.258	0.1455	0.1067	0.0958	0.0959
10	0.2172	0.1201	0.0876	0.0689	0.0739
Prom	<b>0.21673</b>	<b>0.12715</b>	<b>0.10136</b>	<b>0.09735</b>	<b>0.09516</b>

**Tabla 8. Comparación del RMSE para las funciones de base radial**

Aunque el menor valor del error medio cuadrático se obtuvo con la función spline plano, el tiempo de cómputo para la generación de la grilla utilizando esta función, es aproximadamente 15 veces más grande que para el mismo modelo utilizando la función multicuadrática. Debido a esto, y a que el valor de los errores entre ellas difiere solo en el 6.3% se ha elegido la función multicuadrática para la generación del modelo del área Metropolitana de Bucaramanga (*ver anexo 1*). El valor del RMSE versus el tiempo de cómputo para un modelo se esquematiza en la *figura (26)*:



**Figura 26** Tiempo de cómputo para las funciones de base radial

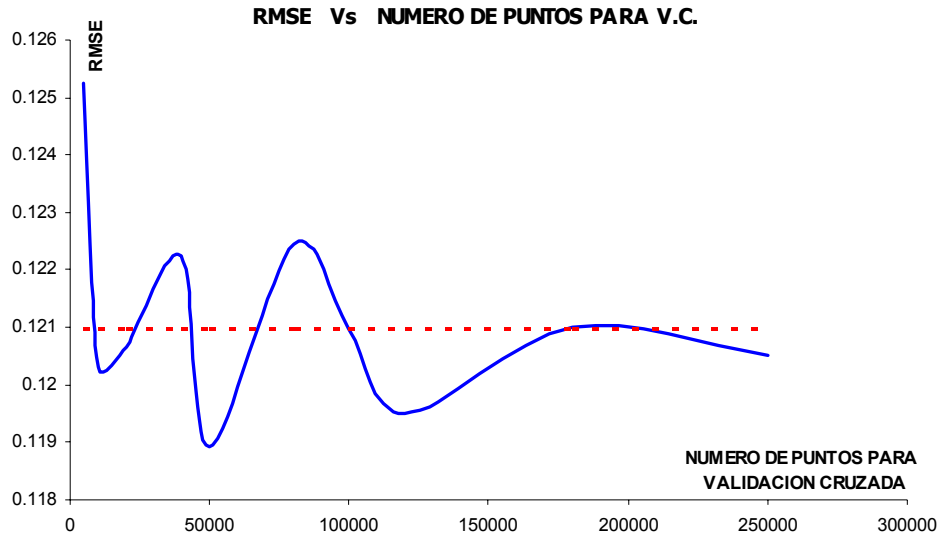
Es estrictamente necesario la utilización de la función multicuadrática; debido a la cantidad de puntos que intervienen en la generación del DEM (12.5 millones de puntos aproximadamente).

En adelante, el ajuste de los parámetros restantes se realizará con base a la totalidad de los datos necesarios para generar el DEM de Bucaramanga; es decir, los datos de entrada para la validación cruzada serán los 12.5 millones de puntos en promedio.

## 5.2 NÚMERO DE PUNTOS NECESARIOS PARA LA VALIDACIÓN CRUZADA

El número de puntos que intervienen en el proceso de la validación cruzada es de vital importancia para determinar la representatividad que tiene la muestra respecto a toda el área que se desea validar. Para estimar la cantidad de puntos

necesaria para validar el modelo, se calculo el valor del RMSE para varias cantidades de datos; desde 5000 hasta 250000 puntos.



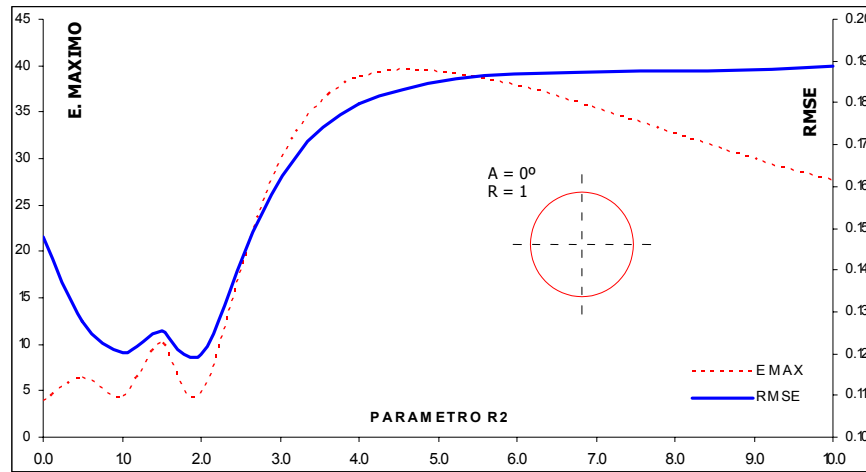
**figura 27** *Número de puntos necesarios para realizar la validación cruzada*

Para valores menores a 20000 puntos, se nota claramente un aumento en el RMSE. La función que describe esta tendencia es de tipo armónica. A pesar que la diferencia en el valor del RMSE entre 300000 y 100000 puntos es de pocos milímetros; estadísticamente se emplearan 100000 puntos para realizar los procesos de validación cruzada. La línea punteada en rojo (*figura 27*) indica la tendencia de los datos hacia un error medio cuadrático de 0.121 m.

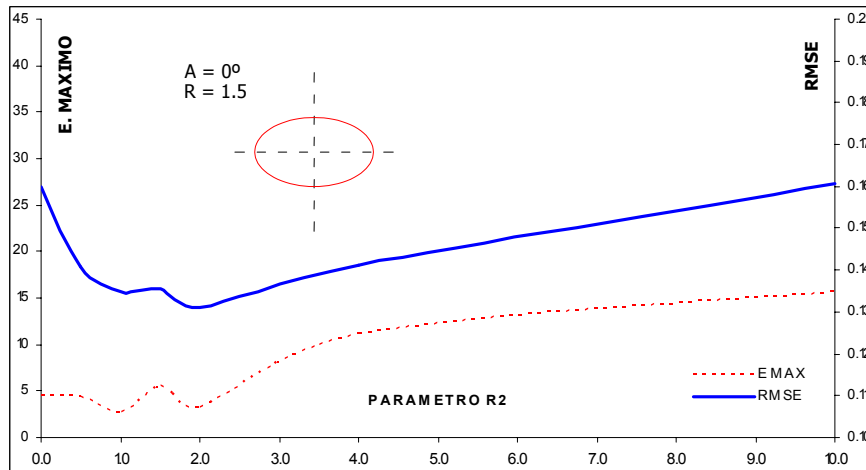
### **5.3 ESTIMACIÓN DEL FACTOR $R^2$ , RELACION DE ANISOTROPÍA Y DIRECCIÓN.**

Una vez obtenido el número de puntos necesario para la validación cruzada y escogida la función de base radial para el análisis, solo queda determinar el factor  $R^2$ , la relación de anisotropía y la dirección para la cual nuevamente los valores del RMSE sean mínimos.

El procedimiento para calcular los parámetros faltantes consiste en fijar la relación de anisotropía y dirección y variar el parámetro  $R^2$  desde 0 hasta 10, para obtener un estimativo del cambio en el valor del RMSE y los errores máximos (positivos o negativos) presentados. Los valores de anisotropía fueron definidos como: 1, 1.5 y 2 y las direcciones para la búsqueda de datos:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  y  $135^\circ$ .



**Figura 28 (a)**



**Figura 28 (b)**

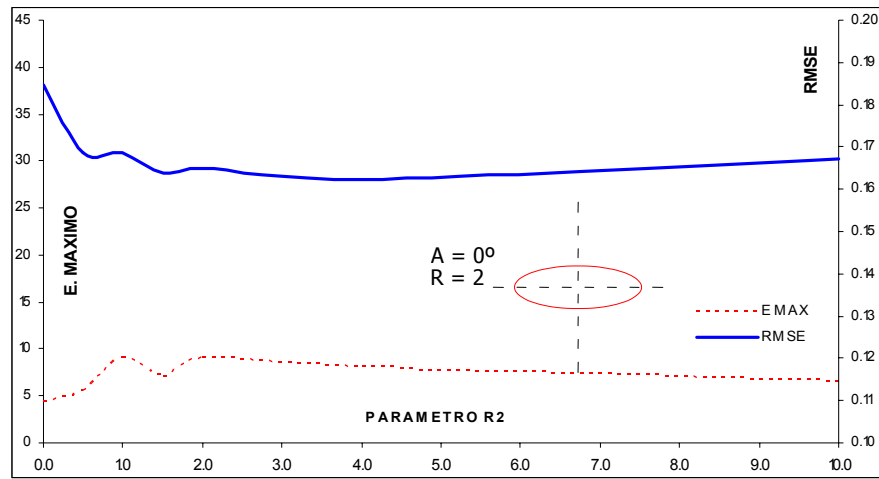


Figura 28 (c)

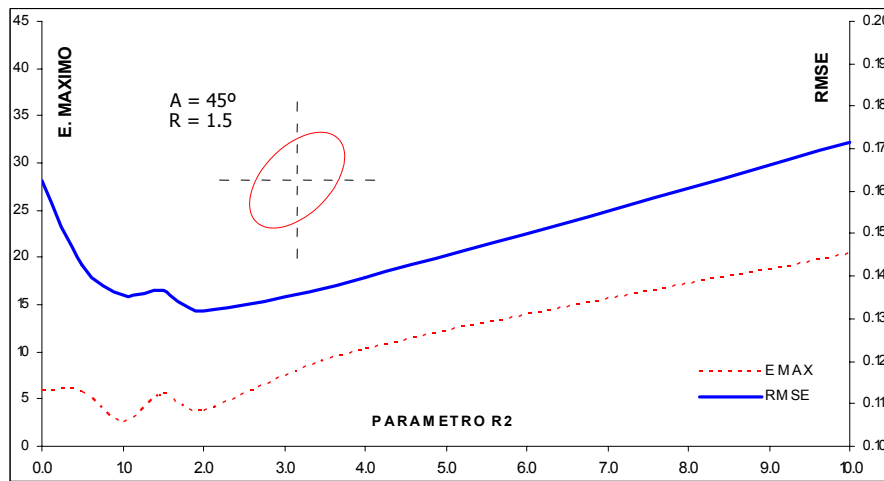


Figura 28 (d)

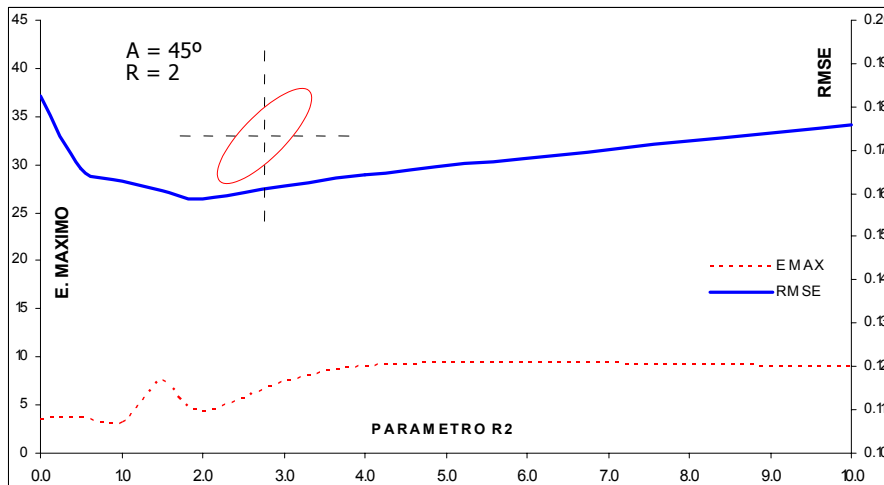
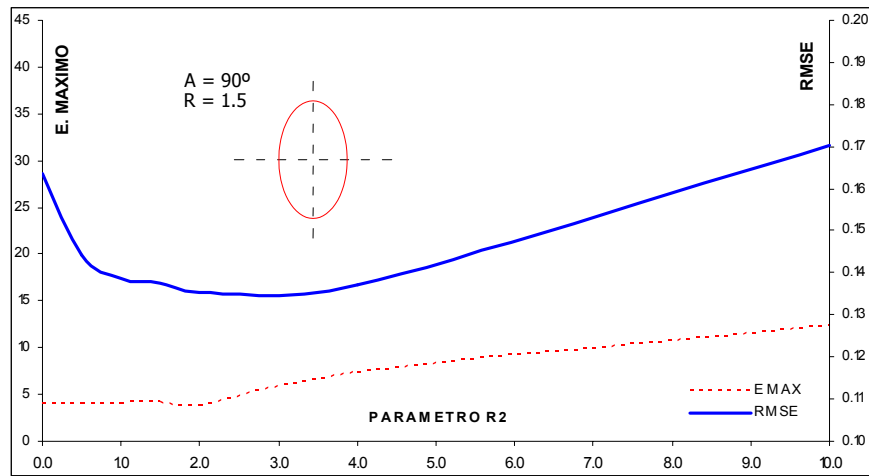
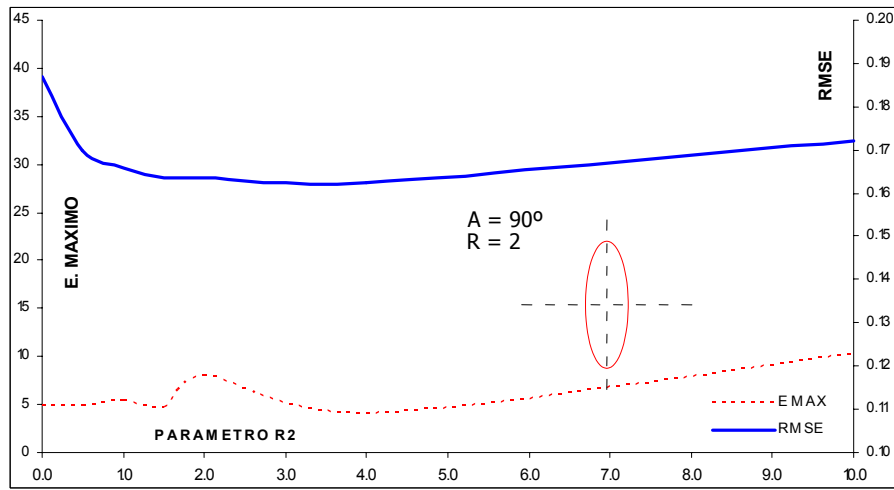


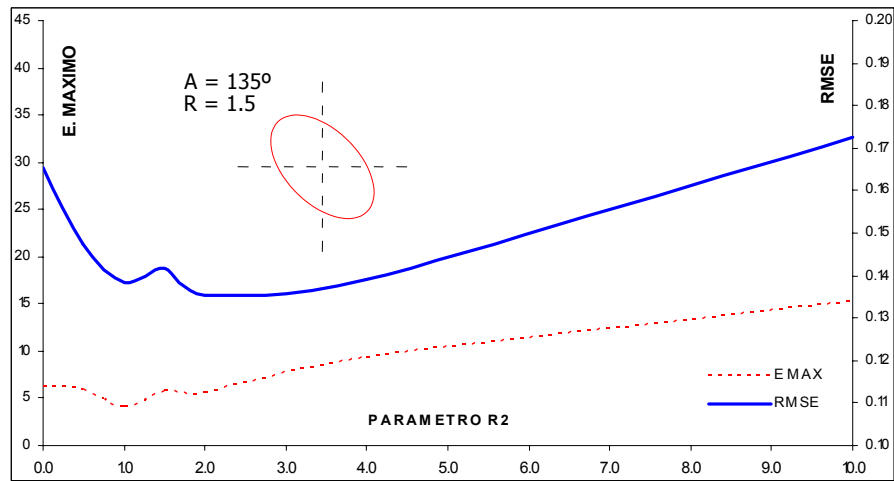
Figura 28(e)



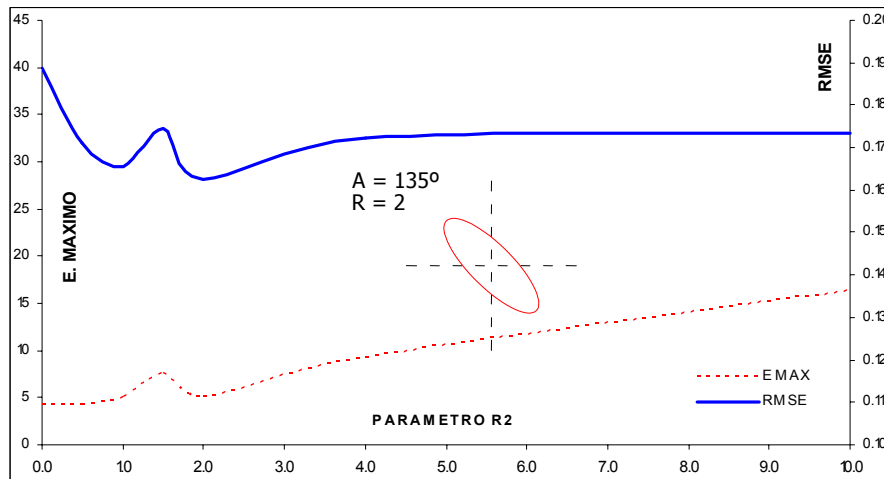
**Figura 28 (f)**



**Figura 28 (g)**



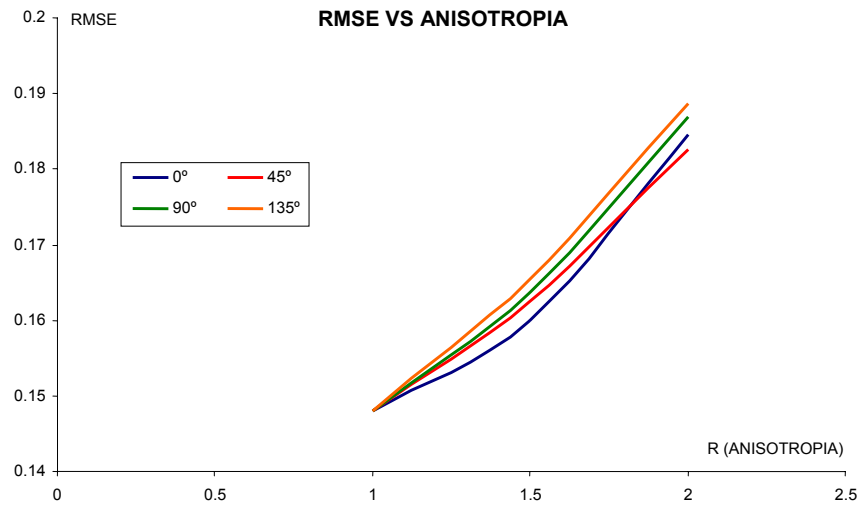
**Figura 28 (h)**



**Figura 28 (i)**

La línea roja representa los errores máximos que se presentaron utilizando validación cruzada para toda el área metropolitana, los cuales varían en la mayoría de las gráficas (*figura 28*) entre 5 y 20 m, excepto para una relación de 1 y un ángulo de 0° donde el error máximo alcanza un valor de 40 m. Las menores magnitudes del error máximo se presentan para el intervalo de R<sup>2</sup> comprendido entre 1 y 2 con relaciones de radio 1 y 1.5.

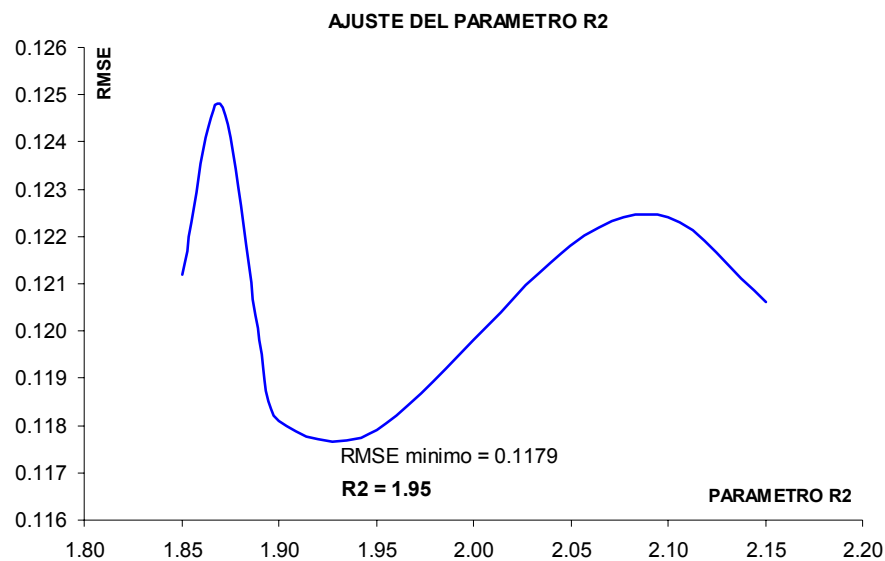
La tendencia general es un aumento en el valor del error a medida que se incrementa el parámetro R<sup>2</sup>. La línea azul modela el error medio cuadrático a lo largo del análisis, presentando mínimos valores para el intervalo de R<sup>2</sup> comprendido entre 1.5 y 2. A partir de 2, el valor del RMSE aumenta progresivamente en cualquiera de las direcciones. Además de la dirección, la relación de anisotropía también influye en el aumento del RMSE a medida que se toma un valor más grande, como se ilustra claramente en la *figura (29)*:



**Figura 29** Cambio del RMSE respecto al radio de anisotropía

#### 5.4 PRUEBA DE SENSIBILIDAD PARA EL FACTOR $R^2$

Un mejor ajuste del parámetro  $R^2$  para la función de base radial multi-cuadrática con relación de anisotropía 1 y dirección  $0^\circ$  se presenta en la *figura (30)*, como resultado de una validación cruzada para parámetros  $R^2$  entre 1.80 y 2.20. El parámetro que arroja el menor valor de RMSE es el de  $R^2 = 1.95$ . (ver anexo 2)



**Figura 30** Prueba de sensibilidad para ajustar el factor  $R^2$  en la función de base radial

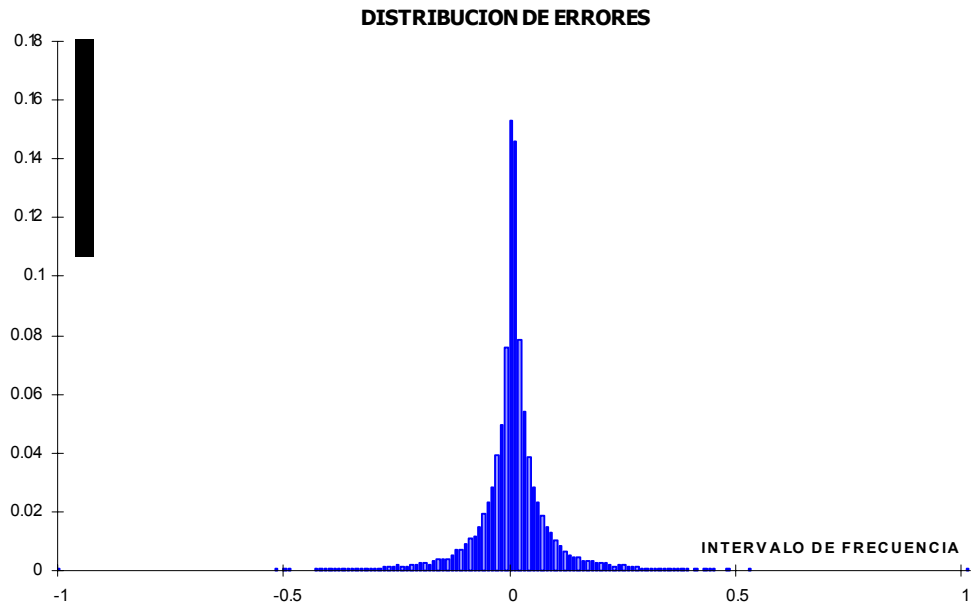
Este parámetro (1.95), será el factor  $R^2$  de la función de base radial para la construcción final del DEM en el área metropolitana de Bucaramanga.

## 5.5 DISTRIBUCIÓN DE ERRORES

La distribución de frecuencias para los errores generados en la validación cruzada con los parámetros anteriormente descritos ( $R^2=1.95$ , función multicuadrática,  $R=1$ , Dirección= $0^\circ$ ) se muestran en la *figura (31)*. Las estadísticas básicas de los RMSE se resumen en la *tabla (9)*:

ESTADISTICAS GENERALES	
NUMERO DE DATOS	100 000
MEDIA ARITMETICA	-0.00179
VARIANZA	0.0139
DESVIACION ESTANDAR	0.1179
CURTOSIS	65.3514
MAXIMO	2.2361
MINIMO	-3.5122

**Tabla 9 Estadísticas de los errores residuales para Bucaramanga**



**figura 31 Distribución de valores residuales**

La mayor cantidad de frecuencias para los valores residuales se encuentran muy cercanas al valor de la media, esto demuestra que el estadístico utilizado para comparar los diferentes métodos interpoladores (RMSE) es una excelente medida para cuantificar la calidad del DEM.

## **5.6 VALIDACIÓN DEL DEM PARA BUCARAMANGA**

En el proceso de validación intervienen básicamente dos variables independientes. Los modelos digitales de elevaciones, construidos a diferentes tamaños de celda y los puntos tomados con equipo GPS. El proceso de validación consistirá en buscar la ubicación espacial (x,y,z) de los puntos tomados con equipo GPS dentro del DEM de Bucaramanga y comparar su valor de elevación como un dato residual. Para esto, surfer incorpora una herramienta que sirve para hallar los valores residuales en cualquier punto (x,y,z) del DEM y compararlo con datos de campo. El valor residual se define como:

$$Z_{RESIDUAL} = Z_{GRID} - Z_{GPS}$$

*Z residual: Diferencia en elevación entre el DEM y los puntos de GPS.*

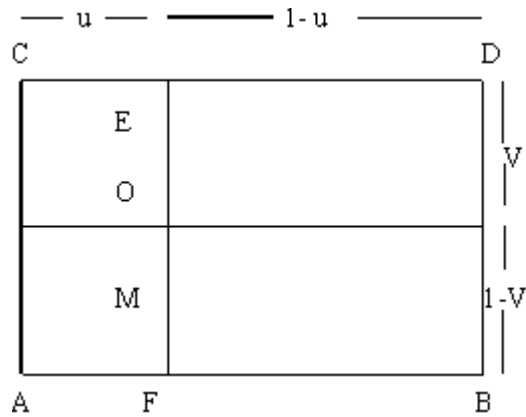
*Z grid: Es el valor de elevación que toma el DEM para una ubicación espacial determinada.*

*Z gps: Es la elevación obtenida con equipo GPS.*

Es preciso aclarar que los puntos tomados con GPS fueron tomados aleatoriamente alrededor de toda el área metropolitana de Bucaramanga. Sin embargo estos puntos no coinciden con los puntos generados por el modelo raster, es decir, las coordenadas halladas con el equipo GPS no coinciden con los nodos del GRID. Por tal razón, surfer introduce una manera especial de hallar los valores de elevación para cualquier coordenada utilizando los 4 puntos del GRID que se encuentran alrededor del objetivo.

Para calcular el valor del Zgrid en el punto específico del Zgps, surfer realiza una interpolación bi-lineal entre los cuatro puntos de la grilla adyacentes al punto del GPS. El proceso de interpolación bi-lineal se describe a continuación:

**5.6.1 Interpolación bilineal.** Suponiendo que los puntos A,B,C,D sean los puntos centrales de la celda en el modelo ráster es decir, cada uno sea un Zgrid, y que la distancia u y v sean las diferencias (este y norte) de los puntos GPS comparados con los GRID; la interpolación bi linear sobre la celda ABCD es realizada en la siguiente secuencia: inicialmente se interpolan los puntos E y F a partir de los puntos C y D, y A y B, respectivamente; y luego se interpola el punto M linealmente a partir de los puntos E y F (*ver figura 32*). Así, el valor de cota Zm calculado en el punto M es:

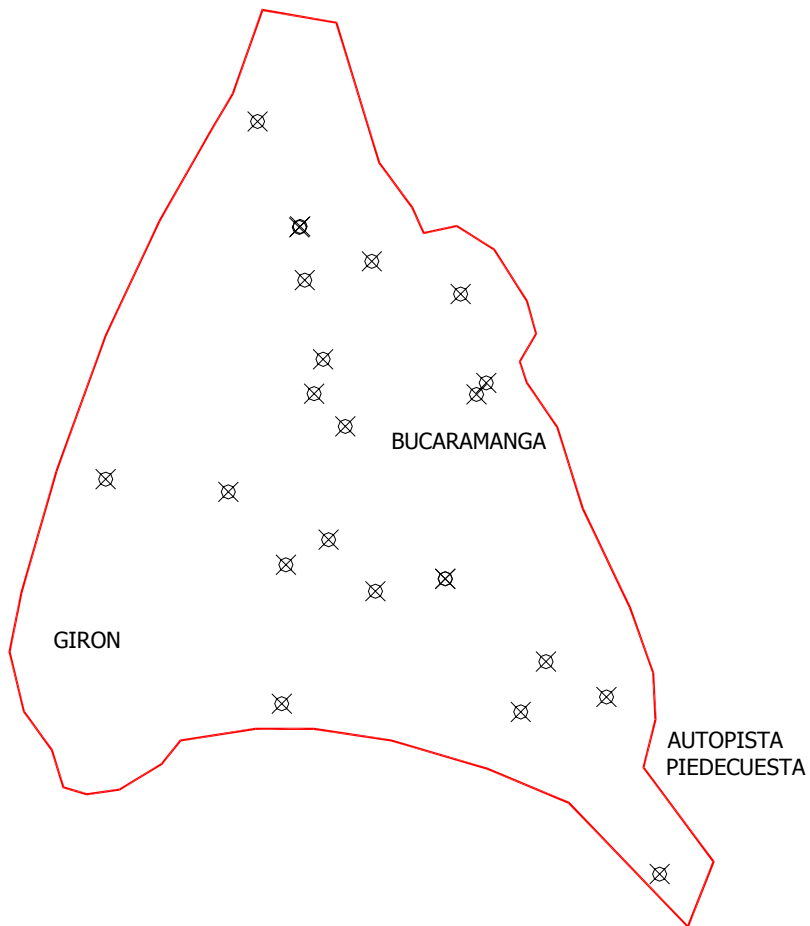


**figura 32 Interpolación bilineal**

- $z = z \times (1-v) + z \times v =$   
 $M-----E-----F$
- $z = (z \times (1-u) + z \times u) \times (1-v) + (z \times (1-u) + z \times u) \times v$   
 $-----C-----D-----A-----B$
- $z = v \times (u \times z + (1-u) \times z) + (1-v) \times (u \times z + (1-u) \times z)$   
 $-----B-----A-----D-----C$

De esta manera se calcula el valor de la elevación en cualquier punto del DEM y se compara con datos reales del GPS.

**5.6.2 Puntos tomados con equipo GPS.** Los puntos fueron tomados con un equipo GPS de precisión sub métrica. Cada punto se corrigió con 100 observaciones del GPS a diferentes horas del día. La ubicación de dichos puntos se muestra en la *figura (33)*.



**figura 33 Puntos tomados con equipo GPS Trimble**

Los puntos tomados con equipo GPS se ordenan en la *tabla 10*:

<b>ESTE</b>	<b>NORTE</b>	<b>ALTURA GPS</b>
111103.777	69922.933	965.118
104145.145	75683.398	780.416
100788.433	77277.904	672.489
104670.043	78869.887	932.743
104837.034	79510.409	944.19
104410.686	81985.658	830.372
103616.931	83939.499	623.568
105744.734	81334.792	992.025
105250.72	78260.919	893.863
107399.066	80726.008	1125.052
107693.9	78856.74	1061.252
107875.452	79070.773	1041.967
104067.604	73095.738	748.43
108518.355	72942.663	868.147
110113.857	73220.665	937.727

107110.202	75422.397	911.4352
104939.91	76150.384	846.4009
108987.362	73880.523	994.1127

**Tabla 10** *Coordenadas de los puntos GPS*



**figura 34** *Punto validado con GPS*

**5.6.3 Tamaño de la celda.** Este es el último parámetro para tener en cuenta en la construcción del DEM. Un valor de celda muy bajo (1 ò 2 m) podría arrojar mejores resultados en términos del RMSE en el modelo. Sin embargo esto no es del todo cierto. El tamaño de la celda estará determinado por un proceso de validación del modelo sobre puntos tomados con equipo GPS alrededor de toda la zona de estudio. Nuevamente el RMSE será el encargado de comparar todos los modelos construidos a diferentes tamaños de celdas (celda del modelo ráster).

Los valores de la *tabla (11)* se introducen en surfer como un archivo .dat para calcular los valores residuales de los DEM construidos con diferentes tamaños de celda. La *tabla (11)* resume algunos de los resultados de estos valores residuales y además, se describen algunas observaciones concernientes a la topografía de la zona de estudio. No todos los puntos tomados con GPS se utilizaran para el cálculo de los errores residuales pues algunos de estos presentan considerables diferencias debido a que los archivos con los que se genera este DEM son del año 2000 y hoy en día existen algunas variaciones topográficas en esos puntos de interés. Sin embargo en la *tabla (10)* se anotan todos, en caso que se quiera actualizar el modelo con cartografía a la fecha.

<b>CELDA 5 m</b>			
<b>ESTE</b>	<b>NORTE</b>	<b>ALTURA</b>	<b>RESIDUO</b>
104145.145	75683.398	780.416	-0.117030431
100788.433	77277.904	672.489	-0.01841894
104670.043	78869.887	932.743	0.594681224
104837.034	79510.409	944.19	0.190018633
104410.686	81985.658	830.372	0.486458043
103616.931	83939.499	623.568	0.230555902
105744.734	81334.792	992.025	0.996028651
107399.066	80726.008	1125.052	-0.162723461
107693.9	78856.74	1061.252	-0.726383866
107875.452	79070.773	1041.967	0.404080226
104067.604	73095.738	748.43	-0.56463336
108518.355	72942.663	868.147	0.173457941
110113.857	73220.665	937.727	0.754229806
107110.202	75422.397	911.4352	-0.028767589
104939.91	76150.384	846.4009	0.745223678
108987.362	73880.523	994.1127	1.076567797
RMSE			<b>0.56102569</b>

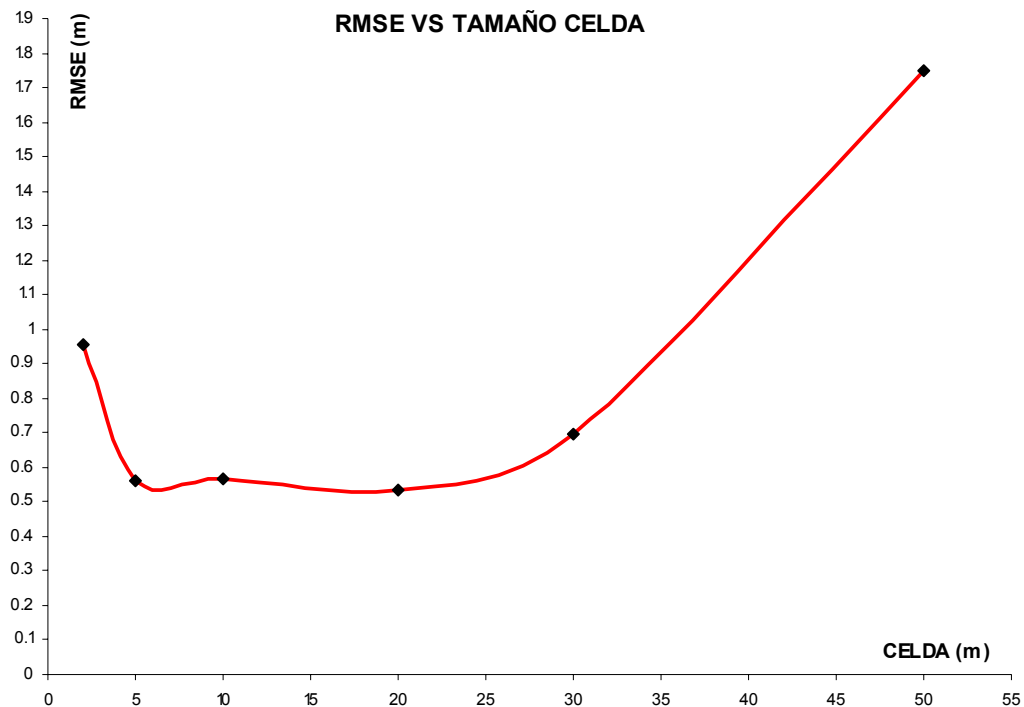
***Tabla 11 Errores residuales para puntos tomados con GPS***

El valor de elevación correspondiente a las coordenadas locales 105250.72 E, 78260.919 N, no se incorporó para el cálculo de los valores residuales debido a cambios de la topografía entre la fecha de rasterización de los archivos .dwg para generar el DEM y los de toma de datos con equipo GPS.

En resumen los valores de RMSE para los diferentes modelos construidos con tamaño de celda distinto se resumen en la *tabla 12*:

Tamaño de celda	RMSE
2	0.955748961
5	0.561025692
10	0.568316083
20	0.536729293
30	0.695028355
50	1.748495204

**Tabla12 RMSE para diferentes tamaño de celda**



**Figura 35 Tamaño de celda óptimo para construcción del DEM en Bucaramanga**

La *figura (35)* muestra que valores de celda entre 20 y 5 m no tienen diferencias significativas en el RMSE dentro del proceso de validación con puntos de GPS. Por tal razón cualquier tamaño de celda que este entre este rango será aceptable para la construcción del DEM en el área metropolitana de Bucaramanga. Valores

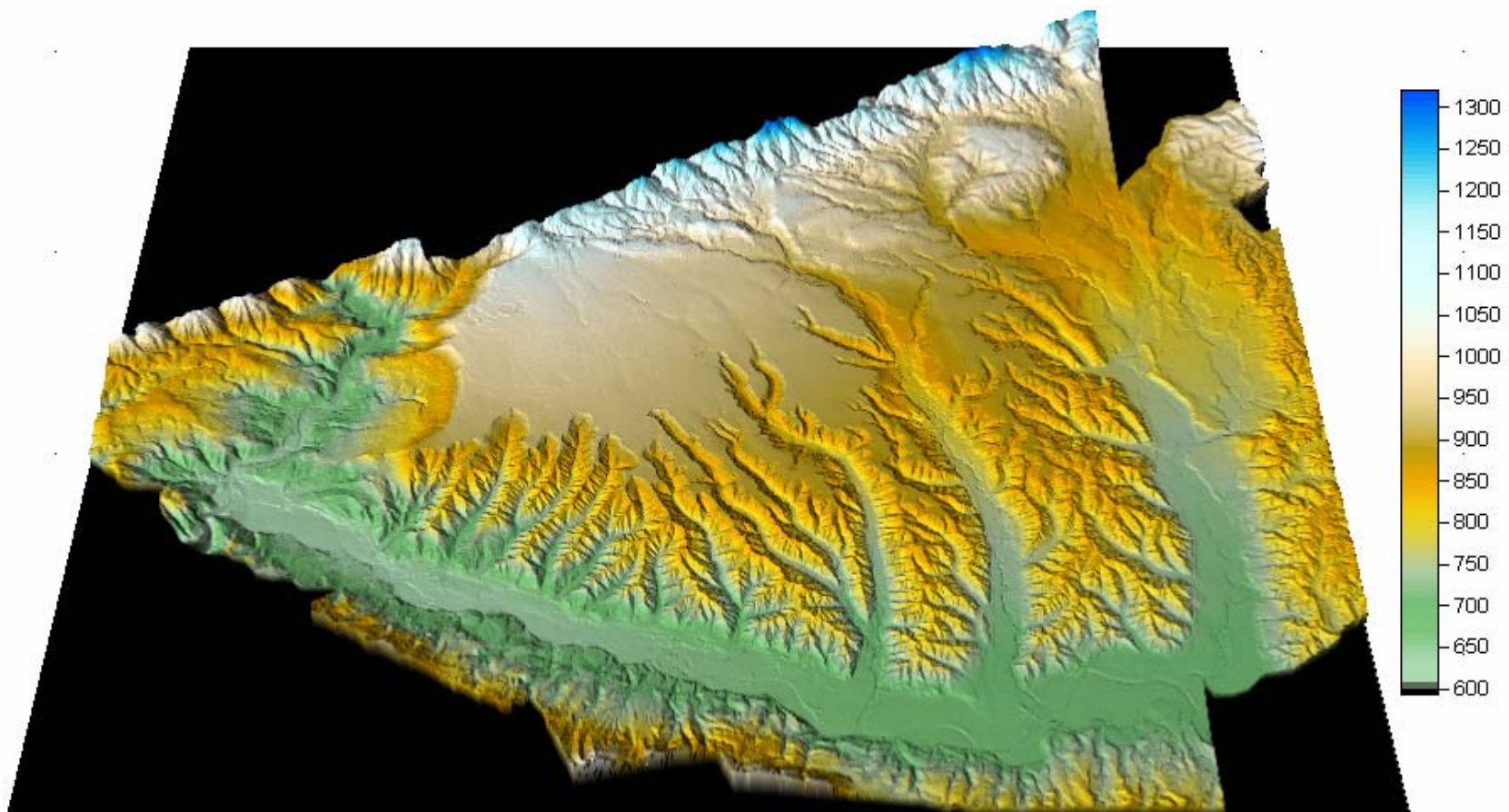
menores a 5 m de celda, disparan el valor del error medio cuadrático considerablemente. Para trabajar estos tamaños de celda, seguramente se necesitaran archivos .dwg que se encuentren a una escala más grande, (1:750, 1:1000 etc).

## **5.8. DEM PARA EL AREA METROPOLITANA DE BUCARAMANGA**

La *figura 36* representa el modelo para el área metropolitana de Bucaramanga utilizando la función de base radial multicuadrática. La variación de elevaciones va desde 600 msnm alrededor de la cuenca de Oro bajo hasta 1300 msnm en los cerros orientales de la ciudad. En la figura se nota claramente la meseta de Bucaramanga, Girón, la escarpa occidental presente en la ciudad y la variabilidad topográfica que existe en la zona norte.

El tiempo de computo para la construcción de este modelo fue de aproximadamente 80 horas.

## DEM PARA EL AREA METROPOLITANA DE BUCARAMANGA



*Figura 36 DEM para Bucaramanga utilizando función de base radial multicuadrática*

## 6. CONSTRUCCIÓN DEL DEM PARA LA CUENCA DE RIO DE ORO

### 6.1 ANALISIS DE INTERPOLADORES

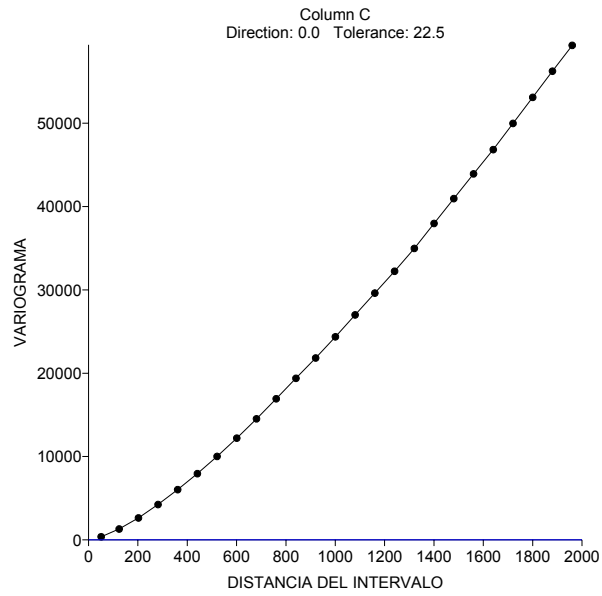
Al igual que como se analizo el área metropolitana de Bucaramanga, la cuenca de río de Oro se someterá a un detallado análisis, implementado diversas técnicas de interpolación y resumiendo solo los resultados obtenidos ya que en capítulos anteriores se describieron los algoritmos matemáticos.

**6.1.1 krigeadado.** El proceso de interpolación por el método del krigeadado ordinario, tendrá como primer punto de análisis el variograma experimental para la cuenca de río de Oro. El variograma se construye teniendo en cuenta solo los puntos que se encuentran dentro de la línea del parte aguas sobre la cuenca.

La distancia del intervalo para el variograma es de 2000 m, es decir, la zona de análisis de los puntos muestrales para la construcción del variograma será de 2 km a la redonda.

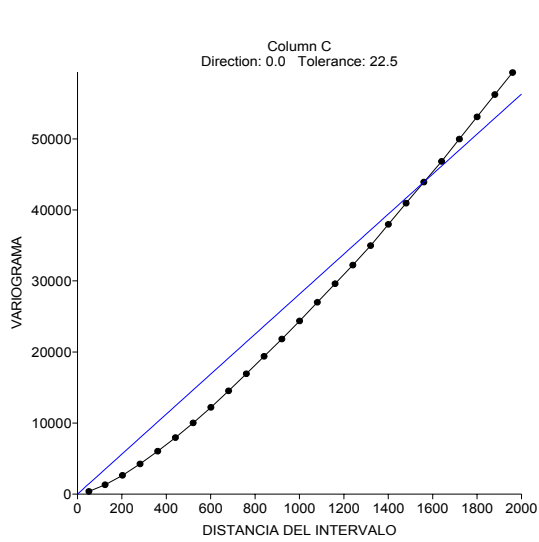
Debido al bajo número de puntos muestrales (150.000 aproximadamente) en la cuenca, se construirá un variograma para toda la zona de estudio. El variograma experimental de la cuenca se muestra en la *figura (37)*:

**ANÁLISIS DE VARIOGRAMAS.** Para llegar al menor valor de RMSE, es conveniente ajustar el variograma experimental a varios tipos de funciones. A simple vista se puede inducir que el variograma teórico podría tener una tendencia lineal o gaussiana. Este ajuste se muestra en la *figuras (38) y (40)*; para las direcciones 0°, 45°, 90° y 135° respectivamente.

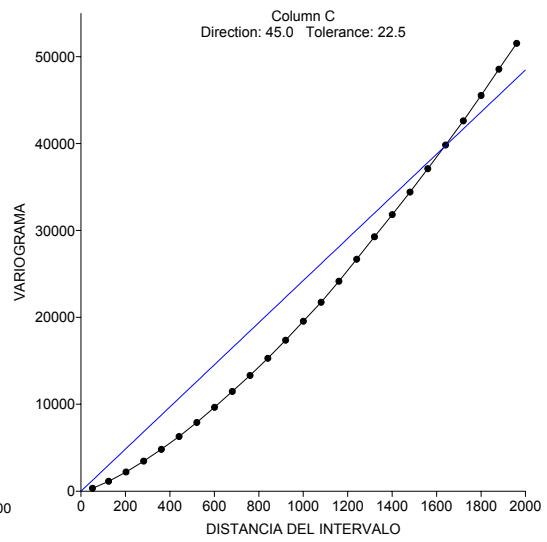


**Figura 37 Variograma experimental para la cuenca de río de Oro**

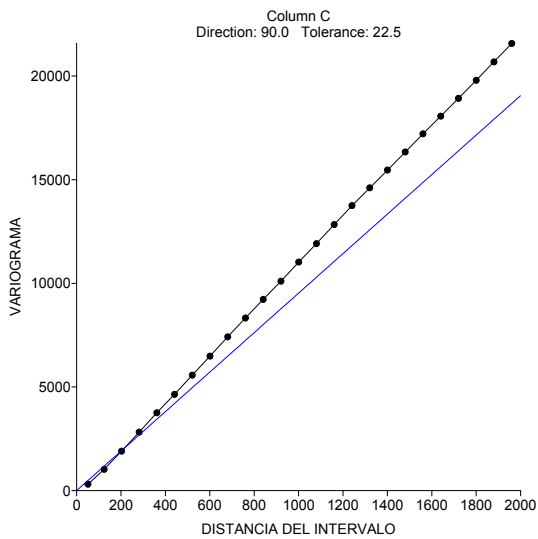
**LINEALES.**



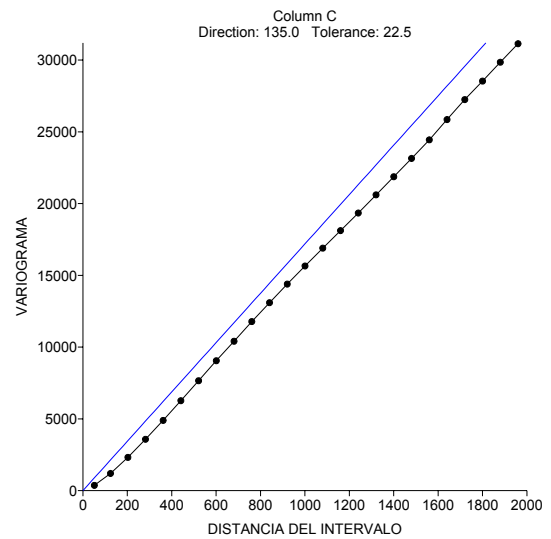
**figura 38 (a) Variograma lineal 0°**



**figura 38 (b) Variograma lineal 45°**



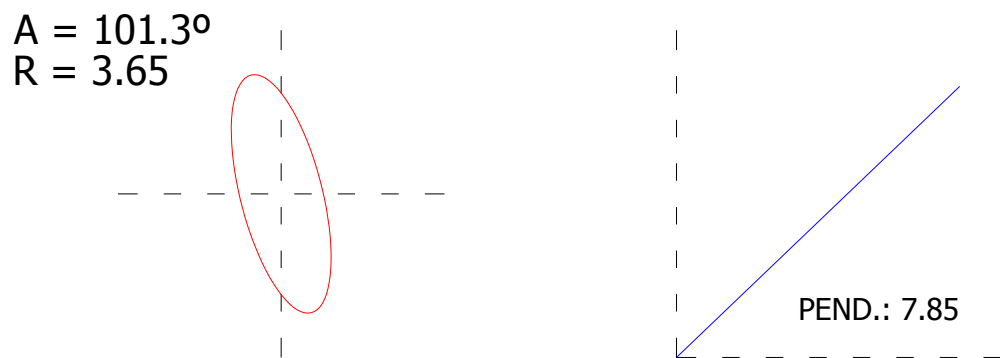
**figura 38 (c) Variograma lineal 90°**



**figura 38 (d) Variograma lineal 135°**

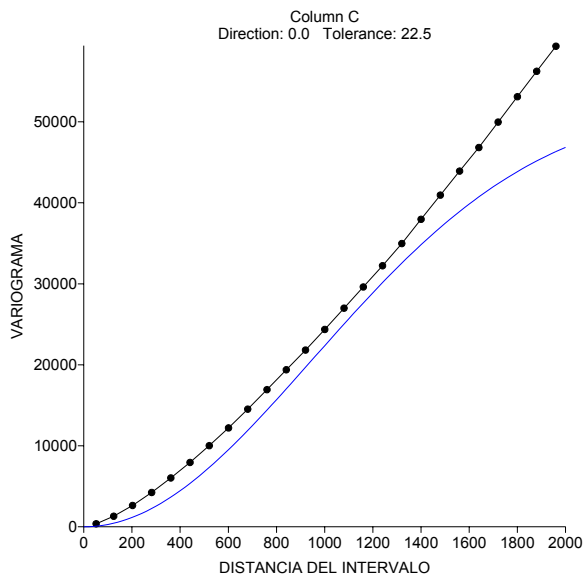
**135°**

Las condiciones de ajuste para el variograma lineal referentes a pendiente y anisotropía se resumen en la *figura (39)*:

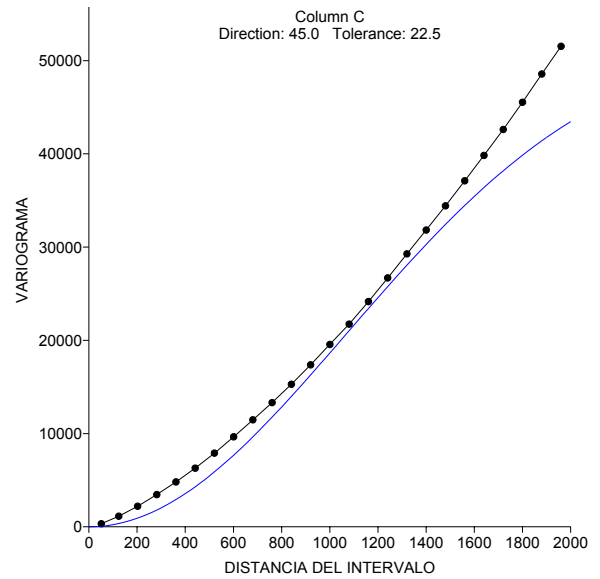


**figura 39 Condiciones de anisotropía para ajuste lineal de variograma experimental**

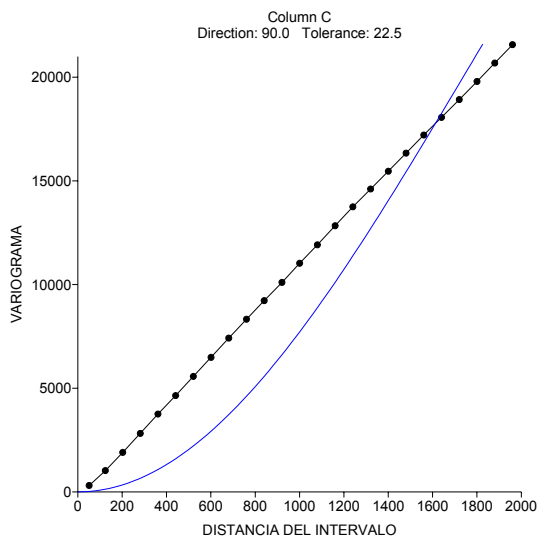
**GAUSSIANOS.**



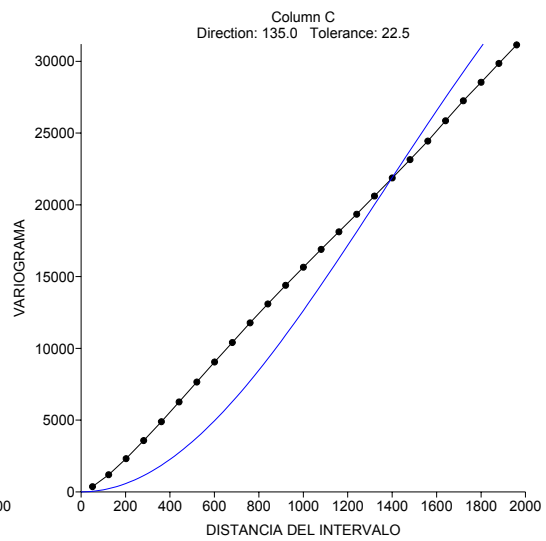
**figura 40 (a) Variograma gaussiano 0°**



**figura 40 (b) Variograma gaussiano**

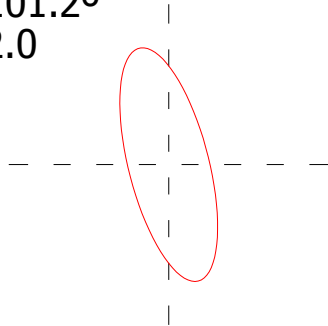


**figura 40 (c) Variograma gaussiano 90°**

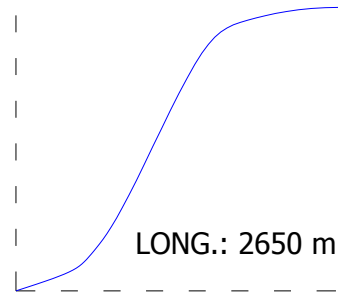


**figura 40 (d) Variograma gaussiano**

A = 101.2°  
R = 2.0



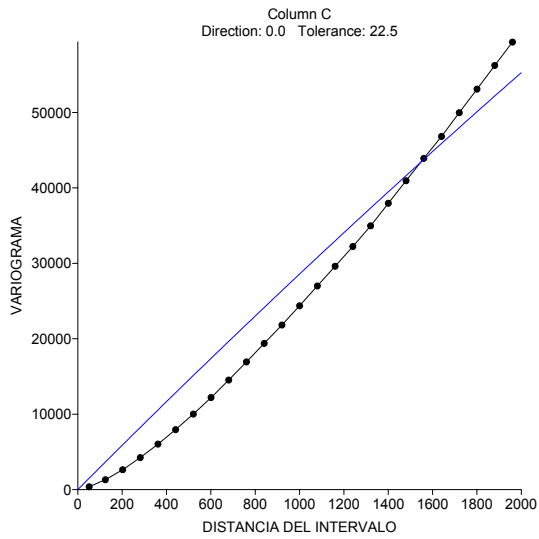
ESCALA: 52600 m



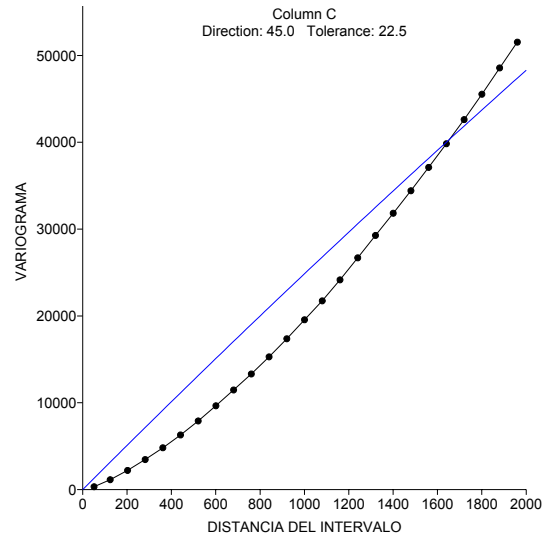
LONG.: 2650 m

**figura 41** Relación de anisotropía, escala y longitud de intervalo para ajuste gaussiano de variograma experimental

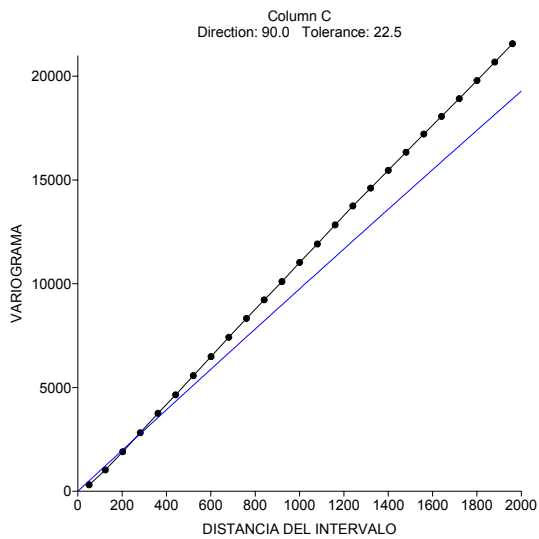
**EXPONENCIALES.**



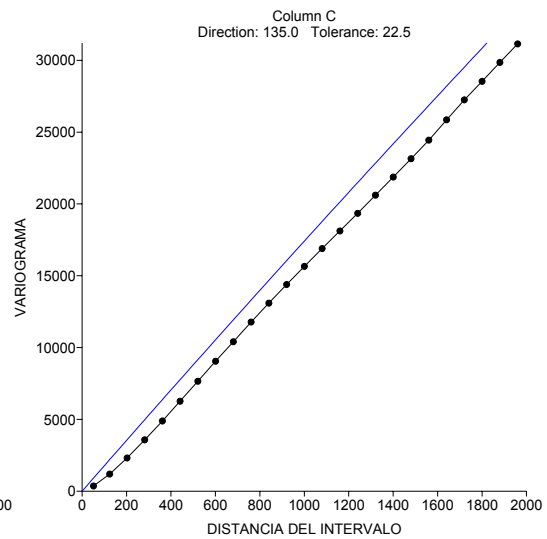
**Figura 42 (a)** Variograma exponencial 0°  
45°



**Figura 42 (b)** Variograma exponencial



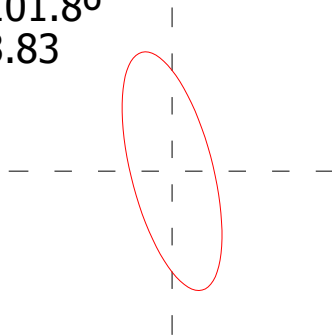
**Figura 42 (c) Variograma exponencial 90° 135°**



**Figura 42 (d) Variograma exponencial**

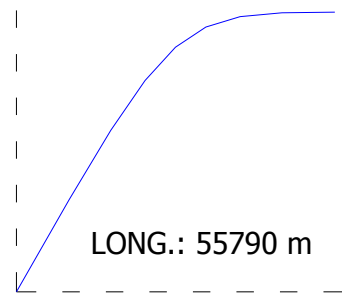
$$A = 101.8^\circ$$

$$R = 3.83$$



ESCALA: 138700

LONG.: 55790 m



**Figura 43 Relación de anisotropía, escala y longitud del intervalo para ajuste exponencial de variograma experimental**

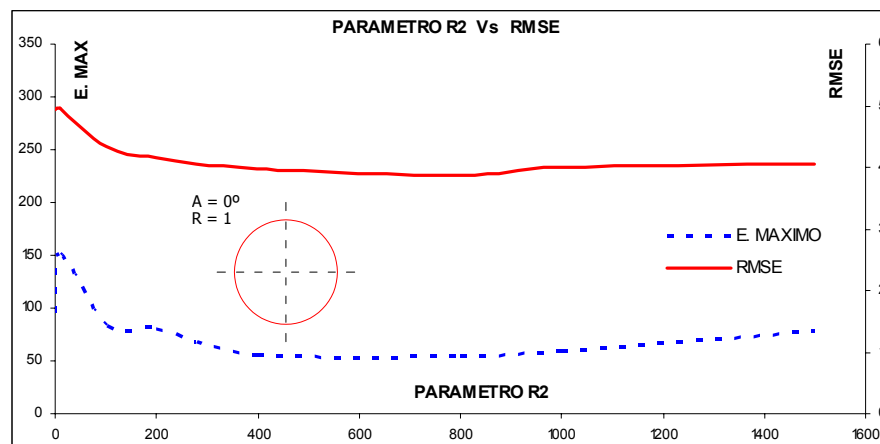
En resumen, por medio de los diferentes ajustes del variograma experimental se obtuvieron valores de RMSE que se consignan en la *tabla (13)*:

Para visualizar los valores del RMSE generados por surfer, ver el *anexo 3*.

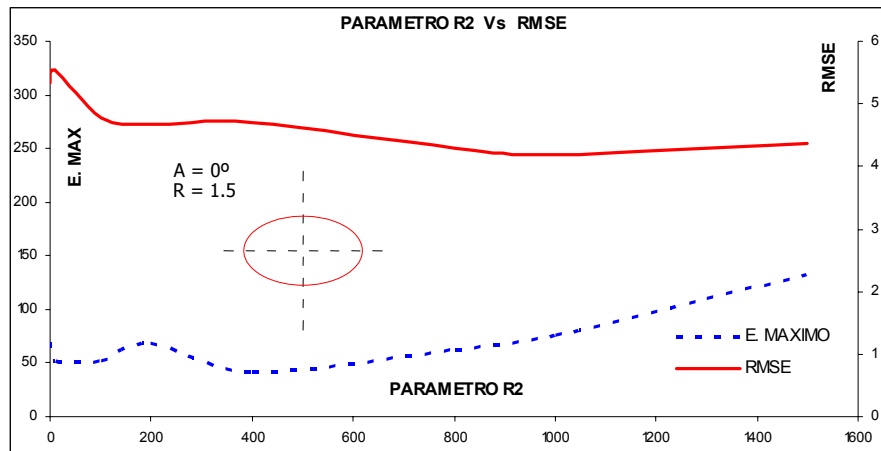
<b>KRIGING</b>			
<b>AJUSTE</b>	<b>PARAMETROS</b>	<b>No Puntos</b>	<b>RMSE</b>
LINEAL	S = 7.85	40000	<b>9.107</b>
	R = 3.65		
	D = 101.3°		
GAUSSIANO	Escala = 52600 m.	30000	<b>11.158</b>
	Longitud = 2650 m.		
	R = 2.0		
	D = 101.2°		
EXPONENCIAL	Escala = 438700 m.	30000	<b>8.628</b>
	Longitud = 55790 m.		
	R = 3.83		
	D = 101.8°		

**Tabla 13 RMSE para ajustes diferentes del variograma experimental**

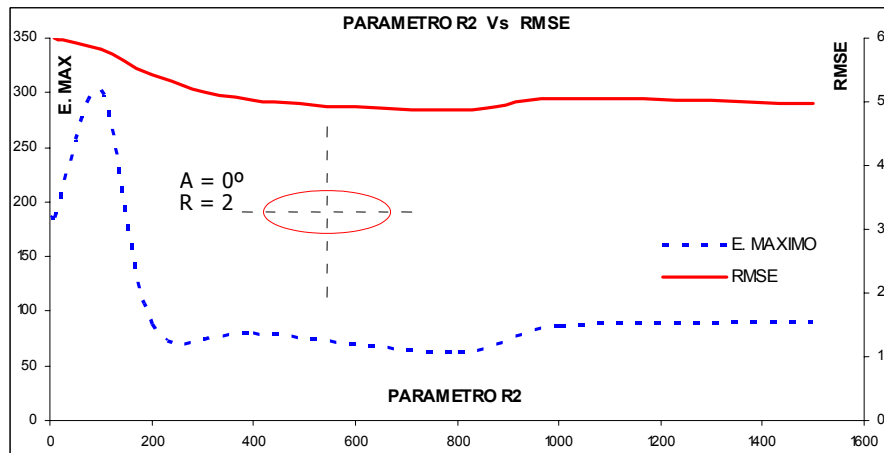
**6.1.2 Función de base radial multicuadrática.** El análisis sobre la cuenca de río Oro referente a la función de base radial se hace de la misma manera que para el área metropolitana de Bucaramanga. Para hallar la menor magnitud en el RMSE, se fija un valor de  $R^2$  que oscila entre 0 y 1500 y se va modificando la anisotropía hasta llegar a 135° en múltiplos de 45°. De esta manera se construyen las graficas de error Vs parámetro  $R^2$ . Las graficas de errores para la función de base radial multicuadrática se resumen en la *figura (44)*:



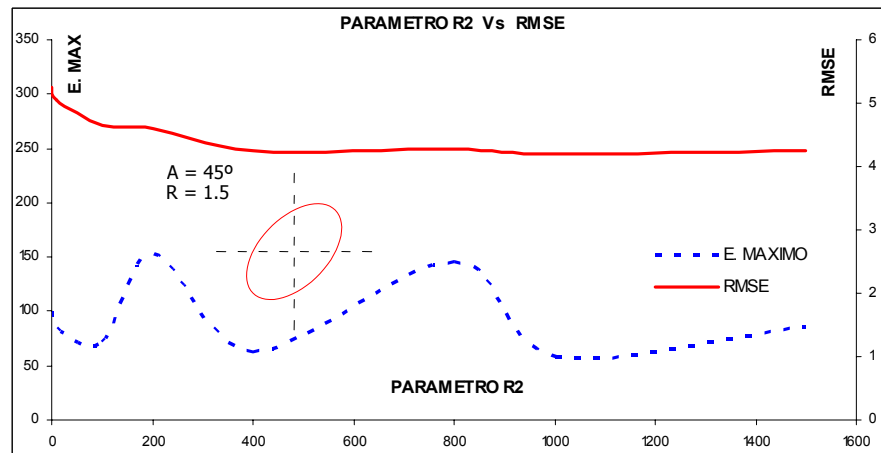
**Figura 44 (a)**



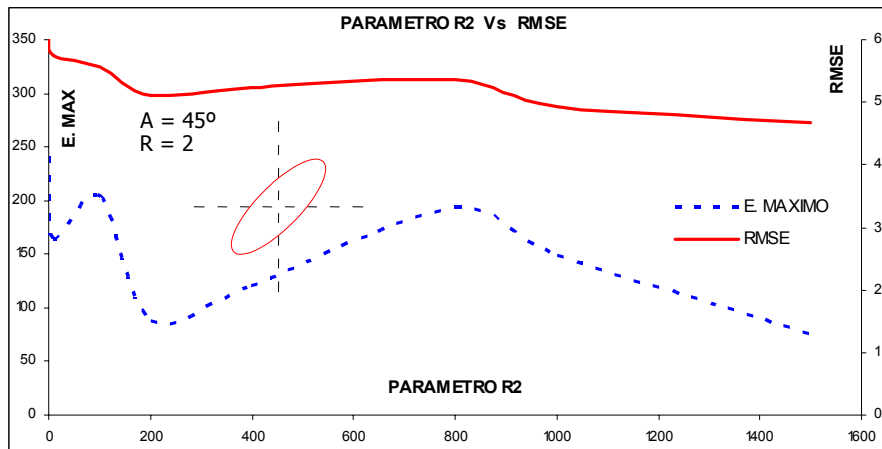
**Figura 44 (b)**



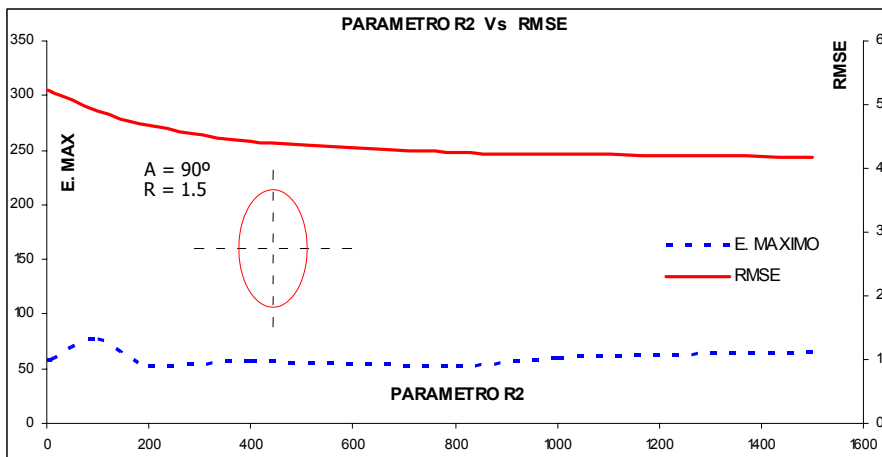
**Figura 44 (c)**



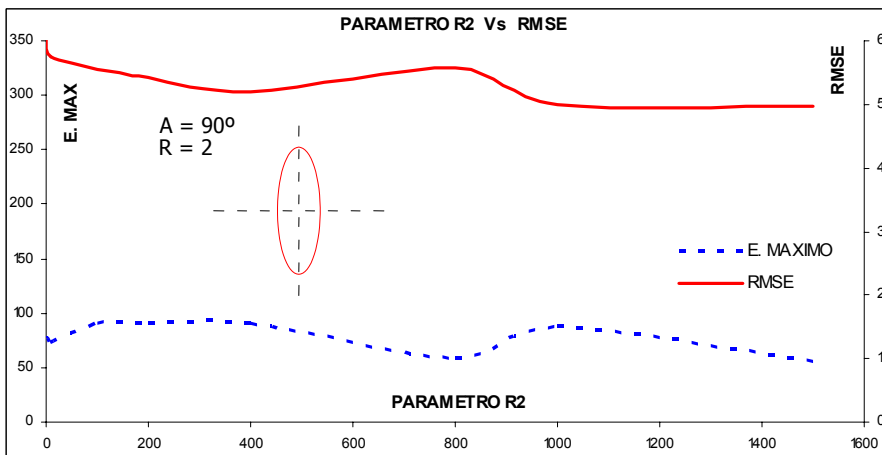
**Figura 44 (d)**



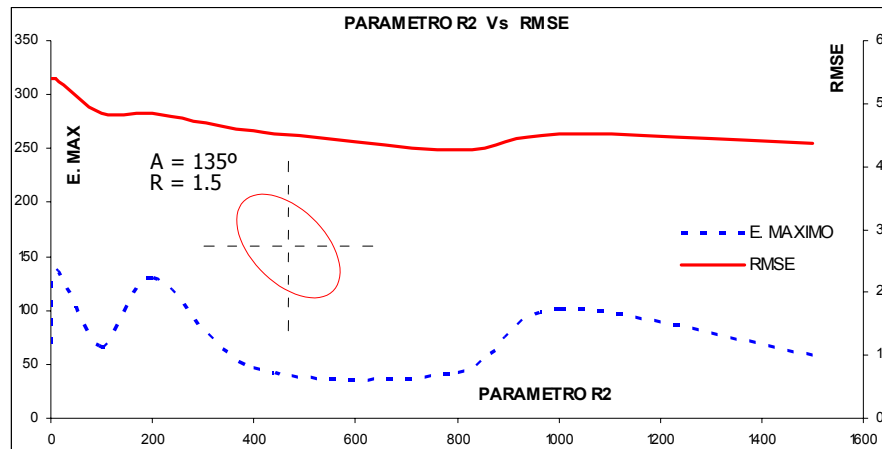
**Figura 44 (e)**



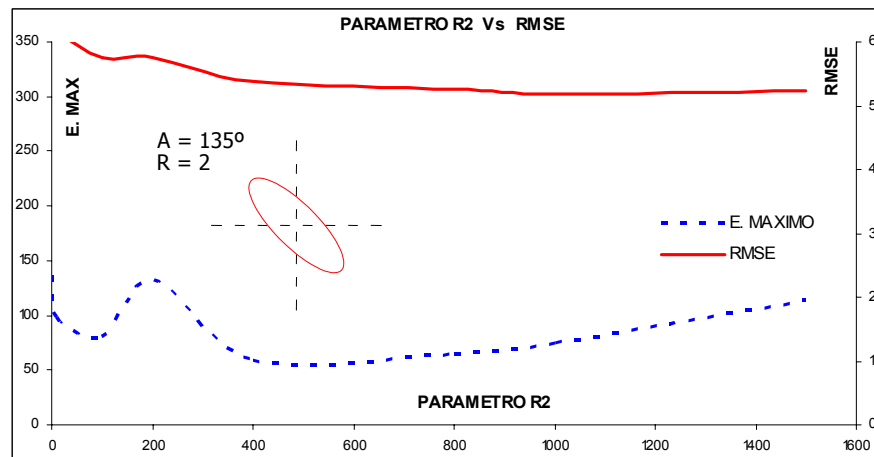
**Figura 44 (f)**



**Figura 44 (g)**



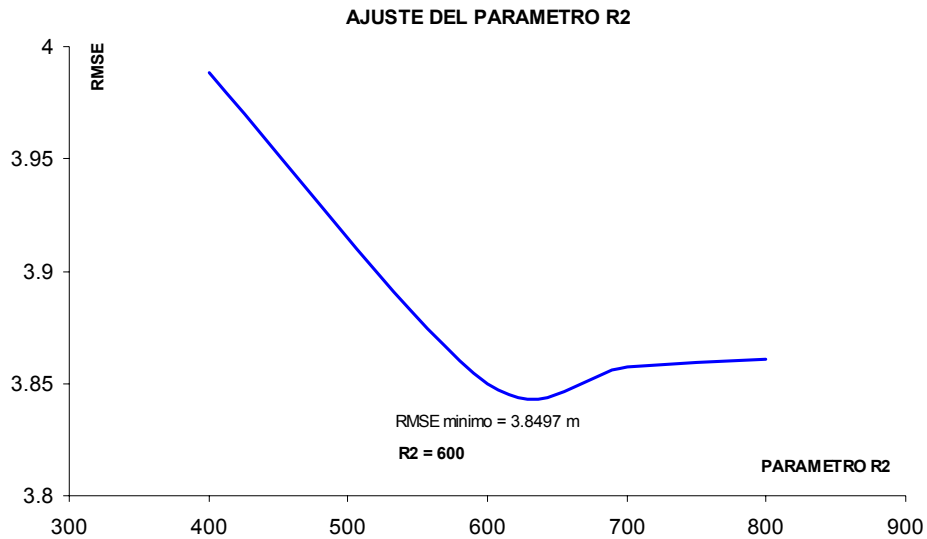
**Figura 44 (h)**



**Figura 44 (i)**

La *figura (44 a)*, muestra que los menores valores del RMSE se encuentran en un amplio rango del parámetro  $R^2$ , desde 400 hasta 800, con valores del error medio cuadrático RMSE cercanos a 3.85 m.

**AJUSTE DEL PARAMETRO  $R^2$ .** Realizando un análisis de sensibilidad se determina que el menor error es para  $R^2=600$ , con un valor del RMSE de 3.84 m. y un error máximo de 55.84 m. La prueba de sensibilidad del parámetro  $R^2$  se muestra en la *figura (45)*:



**figura 45 sensibilidad del parámetro  $R^2$  para río de Oro**

Los reportes de este ajuste se encuentran en el *anexo 4*.

**6.1.3 Inverso de la distancia.** Los valores obtenidos para la función de interpolación IDW, con 20000 puntos utilizados en la validación cruzada, potencia 2 y factor de suavizado 0 se consignan en la *tabla (14)*.

INVERSO DE LA DISTANCIA				
DIRECCION	R	MAXIMO	MINIMO	RMSE
0°	1	820.66	-590.57	14.62
	1.5	122.16	-510.23	13.37
	2	823.28	-537.1	16.06
45°	1.5	70.41	-129.63	12.11
	2	92.08	-148.45	12.74
90°	1.5	88.14	-413.32	12.47
	2	821.26	-93.77	13.93
135°	1.5	127.33	-123.62	12.51
	2	71.01	-116.61	13.21

**Tabla 14 Errores cuadráticos en todas las direcciones para la función IDW**

El mínimo valor para el RMSE utilizando este interpolador es de 12.11 m.

**6.1.4 Vecino más cercano.** El interpolador de vecino más cercano solo presenta variables de anisotropía y las variaciones referentes al RMSE se definen en la tabla 15.

VECINO MAS CERCANO				
DIRECCION	R	MAXIMO	MINIMO	RMSE
0	1	50	-150	9.23
	1.5	150	-150	10.43
	2	150	-150	13.26
45	1.5	50	-150	10.62
	2	50	-150	12.25
90	1.5	100	-150	10.75
	2	100	-150	13.25
135	1.5	150	-50	11.63
	2	50	-150	14.37

**Tabla 15 Valores de RMSE para rio de Oro con vecino natural**

El mínimo valor obtenido de RMSE con este interpolador es de 9.23 m. Sin embargo, los valores residuales máximos y mínimos son considerablemente mayores que otros interpoladores utilizados en el análisis; tal es el caso de la función de base radial multi cuadrática.

**6.1.5 Triangulación con interpolación lineal.** Los valores de errores extremos y RMSE para diferentes condiciones de anisotropía se definen en la tabla 16.

Los valores de RMSE para todos los interpoladores anteriormente descritos y analizados se comparan en la tabla 17 y se determina a la función de base radial multicuadrática con el parámetro  $R^2=600$ , la relación de anisotropía = 1 y la dirección =  $0^\circ$  como el algoritmo interpolador que en promedio arroja los menores valores de errores medios cuadráticos y errores extremos.

TRIANGULACION CON I. LINEAL				
DIRECCION	R	MAXIMO	MINIMO	RMSE
0	1	35.32	-35.56	4.67
	1.5	38.27	-42.41	7.08
	2	32.59	-36.95	6.06
45	1.5	34.8	-36.76	7.15
	2	35.59	-35.06	7.2
90	1.5	26.55	-31.53	6.37
	2	57.38	-37.49	7.57
135	1.5	48.529	-29.587	7.41
	2	39.541	-36.874	6.98

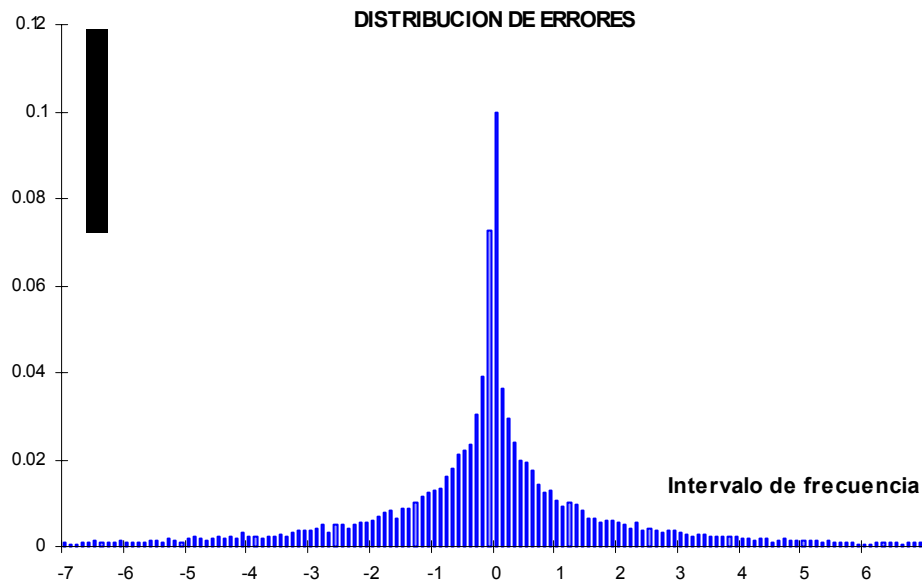
**Tabla 16 RMSE para río de Oro utilizando triangulación de Delaunay**

ERROR MEDIO CUADRATICO RMSE (m)					
RIO DE ORO	KRIGING	RADIAL BASIS	INVERSO DE LA DISTANCIA	VECINO MAS CERCANO	TRIANGULACION I. LINEAL
RMSE	8.628	3.8497	12.115	9.23	4.671

**Tabla 17 Resumen de interpoladores para río de Oro**

## 6.2 DISTRIBUCIÓN DE ERRORES

La distribución de frecuencias para los errores generados en la validación cruzada con los parámetros anteriormente descritos ( $R^2=600$ , función multicuadrática,  $R=1$ , Dirección= $0^\circ$ ) se muestran en la *figura (46)*.



**figura 46 Distribución de valores residuales para la cuenca de río de Oro**

Nuevamente los valores se encuentran distribuidos alrededor de la media garantizando así la eficiencia del estadístico RMSE.

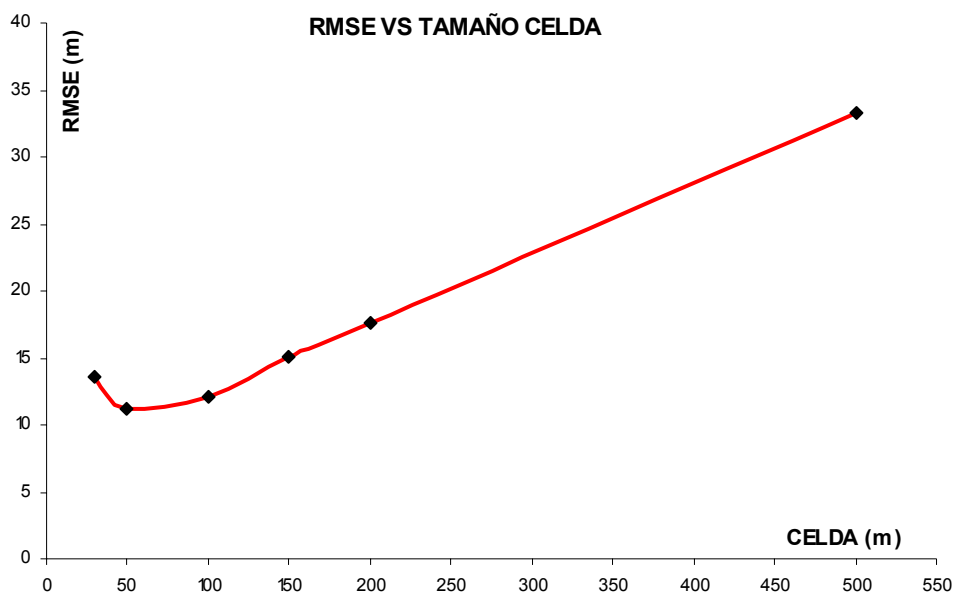
### **6.3 TAMAÑO DE CELDA PARA EL DEM**

Una vez escogido el método interpolador, se analizarán los valores residuales de 34 puntos de GPS que se encuentran dentro de la zona de análisis para compararlos con las elevaciones propias del archivo .grid que se obtuvo para la cuenca de río de Oro utilizando la función de base radial multicuadrática.

Como se hizo para el área metropolitana de Bucaramanga; se construirán modelos a diferentes tamaños de celda, se realizará la interpolación bilineal y se calcularán los valores residuales entre los diferentes .grid y los puntos de GPS.

T. CELDA	RMSE (m)
30	13.56
50	11.25
100	12.13
150	15.07
200	17.56
500	33.29

**Tabla 18** Valores de RMSE para diferentes tamaños de celda en río de Oro



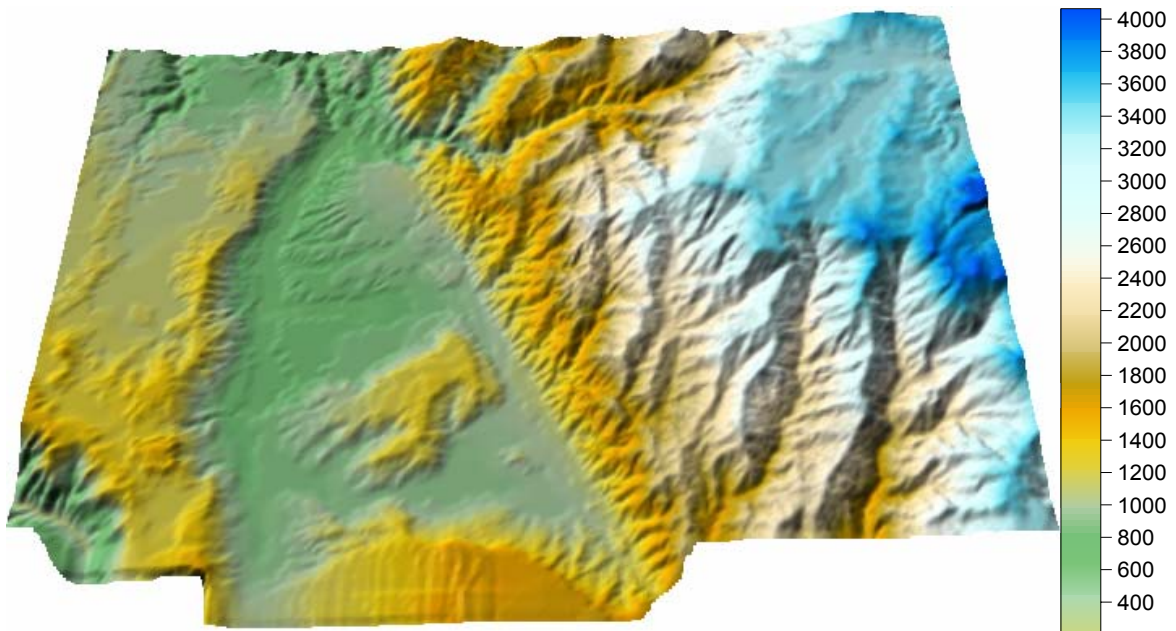
**Figura 47** Rango de celda para la construcción del DEM para la cuenca de río de Oro

La figura 47 muestra claramente que valores menores a 50 disparan considerablemente el valor del RMSE. El tamaño “optimo” de la celda para generar este modelo estará entre 100 y 50 m.

#### 6.4 DEM PARA LA CUENCA DE RIO DE ORO.

Este modelo tiene valores de elevación desde 400 hasta 3800 msnm. El modelo final generado con un tamaño de celda de 100 m se muestra en la *figura 48*. Modelos generados con tamaños de celda menores a 50 m, muestran la tendencia

de las curvas de nivel aumentado los errores en el proceso de validación final del modelo. Los valores recomendados para el tamaño de la grilla varían en un rango de 50 a 100 m. Una vez generado el DEM para la cuenca de río de Oro, mostrado en la *figura (48)* se debe hacer un análisis de pits o ruido para poder parametrizar la cuenca, es decir, hallar las redes de drenaje, pendientes de los cauces, longitud de la divisoria, áreas de la cuenca etc.



***Figura 48 DEM para la cuenca de río de Oro utilizando la función de base radial.***

## **6.5 REMOCIÓN DE PITS**

Las depresiones o pits en un DEM son celdas o grupos de celdas que tienen un valor de elevación más bajo que todas las celdas adyacentes. Los pits en un DEM pueden ser reales, representando características naturales del terreno o en la mayoría de los casos artificiales, causados por errores en los datos de entrada en el procesamiento de los datos.

Varios métodos han sido desarrollados para remover los pits en los DEM's. Una trayectoria de flujo continua es definida desde cada celda (grid) hasta el borde de

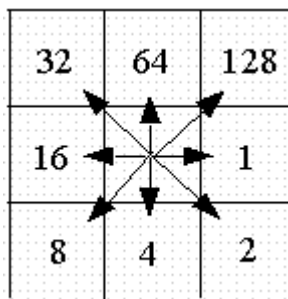
la divisoria de aguas. Se presentan dos maneras prácticas de corregir estas depresiones:

- Elevar la altura de la celda en el pit a la altura del punto adyacente mas bajo. (*Jenson y Domingue (1988)*).
- El pit toma el promedio aritmético de la elevación de sus vecinos.

## 6.6 RED DE DRENAJES

Varios métodos para definir la dirección del flujo en un DEM han sido desarrollados en los últimos años. La más simple y ampliamente usada es la del método D8 el cual asume que una partícula de agua en cada celda del DEM fluye hacia una y solo una de sus celdas vecinas, la cual es la celda que presenta el mayor gradiente (pendiente).

Para asignar un valor de dirección al flujo en una celda, se calcula la diferencia de elevación entre la celda y las celdas vecinas y se divide entre  $\sqrt{2} d$  para celdas que se encuentren diagonales y entre  $d$  para no diagonales; donde  $d$  es el tamaño (arista) de la celda.



**figura 49 ejemplo de modelo D8**

En el ejemplo de la *figura 49*, el valor de la dirección del flujo entre el centro y 2 será, la diferencia de elevación entre ambos sobre  $\sqrt{2}$  X la arista de la celda.

## 7. CONCLUSIONES

- Como conclusión principal de este trabajo podemos considerar al método de interpolación basado en la función de base radial multicuadrática como el que mejor se ajusta al modelo de elevaciones de la región de río de Oro incluida el área metropolitana de Bucaramanga, tanto desde el punto de vista estadístico como visual.
- La función geoestadística de krigeaje arroja resultados de ajuste notablemente inferiores (mayor magnitud de RMSE) debido a que presenta dificultades inherentes al proceso previo de selección del semivariograma. El krigeado puede ser un método muy recomendable cuando se está realmente seguro de haber utilizado el mejor semivariograma posible, circunstancia que ofrece cierta incertidumbre en la mayoría de los casos. Además, la limitante de tener que utilizar un solo semivariograma para representar en buena medida toda el área de interés, es una barrera en la construcción del modelo debido a que la topografía de la zona en estudio es completamente asimétrica y presenta fluctuaciones alrededor de toda su extensión.
- La exactitud de la interpolación, estimada a partir de la técnica de validación cruzada, aumenta con la densidad de puntos muestrales. La validación del modelo y la validación cruzada son procesos que se realizan con objetivos distintos. Los DEM de Bucaramanga y río de Oro, se validaron utilizando puntos tomados con GPS alrededor de la zona de estudio para determinar que tan ajustado estaba el modelo a los datos reales. Esta es la mejor manera de validar ya que si se utilizan los puntos de la validación cruzada, estaríamos comparando el modelo contra los datos muestrales, es decir, con los que se generó el modelo y estaremos induciendo a errores a la hora

de comparar con valores tomados en campo. Es importante aclarar este punto porque el DEM se utilizara para modelar eventos hidrológicos y se requiere que éste sea lo más cercano a la realidad posible.

## BIBLIOGRAFIA

- [3] Aguilar, F.J.; Agüera, F.; Carvajal, F.; “Obtención de modelos digitales de elevación del terreno para su uso en entornos CAD”. Pamplona-Logroño (España): XI Congreso Internacional de Ingeniería Gráfica, Vol. I, p. 707-716, 1999.
- BURROUGH, P. 1986. Principles of Geographical Information Systems for Natural Resources. Clarendon Press, Oxford Inglaterra, 194p.
- CRESSIE NOEL A.C. 1993. statistics for spatial data.
- EASTMAN, R. 1992. IDRISI: A Grid-Based Geographical Analysis System, Version 4 User's Manual and Technical Reference. Clark University Graduate School of Geography, Worcester MA EE UU.
- Felicísimo, A. M., 1994. Curso sobre modelos de elevación digital.
- Garbrecht, J. y Martz, L.W., Digital elevation model issues in water resources modeling.
- Golden Software, Inc.; “Users’ manual of Surfer for Windows (version 6.04)”. Golden (USA): Goleen Software Inc., 1997.
- Hohn Michael Edward. 1988. Geostatistics and petroleum geology. VNB
- Rossiter D., Rodríguez M. Construcción de modelos digitales para la evaluación de tierras.

- USGS. Digital Elevation model Standard. 1997.
- ZINCK, J.A. y C.R. VALENZUELA. 1990. Soil geographic database structure and application examples. ITC Journal 1990-3, 270-294.

SITIOS WEB DE INTERÉS:

<http://www.goldensoftware.com/products/surfer/surfer.shtml>

<http://www.Dem2\USGS.com>

<http://www.csulb.edu/~wechsler/Dissertation/P262/P262.html# Author Information>

[http://www.ncgia.ucsb.edu/conf/SANTA\\_FE\\_CD-](http://www.ncgia.ucsb.edu/conf/SANTA_FE_CD-)

[ROM/sf\\_papers/desawal\\_robert/figure1.gif](http://www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/desawal_robert/figure1.gif)

<http://rockyweb.cr.usgs.gov/nmpstds/acrodcs/dem/3DEM0897.PDF>

<http://grl.ars.usda.gov/topaz/PaperH.html#TOP>

[http://www.sli.unimelb.edu.au/gisweb/GISModule/GIST\\_Raster.htm#rrep](http://www.sli.unimelb.edu.au/gisweb/GISModule/GIST_Raster.htm#rrep)

<http://www.sli.unimelb.edu.au/gisweb/RSAModule/RSAModule.htm>

[http://www.olemiss.edu/depts/geology/courses/ge470/gistop\\_8b.htm#tin](http://www.olemiss.edu/depts/geology/courses/ge470/gistop_8b.htm#tin)

<http://www.olemiss.edu/depts/geology/courses/ge470/rdm.gif>

[http://www.olemiss.edu/depts/geology/courses/ge470/r\\_crd2.gif](http://www.olemiss.edu/depts/geology/courses/ge470/r_crd2.gif)

[http://www.olemiss.edu/depts/geology/courses/ge470/rstr\\_ac.gif](http://www.olemiss.edu/depts/geology/courses/ge470/rstr_ac.gif)

<http://www.olemiss.edu/depts/geology/courses/ge470/rf.gif>

<http://www.geog.ubc.ca/courses/klink/g470/class00/skaleem/fig16.jpg>

<http://www.geog.ubc.ca/courses/klink/g470/class00/skaleem/fig20.jpg>

<http://www.geog.ubc.ca/courses/klink/gis.notes/ncgia/toc.html#UNIT38>

<http://www.geog.ubc.ca/courses/klink/gis.notes/ncgia/u38.html#SEC38.1.1>

<http://edc.usgs.gov/guides/dem.html#background>

## **ANEXOS**

## ANEXO A

### REPORTES DE TIPOS DE FUNCIONES DE BASE RADIAL

#### AREA 1

##### Cross Validation Report

---

Gridding Method:	Radial Basis Function
Basis Kernel Type:	Multiquadric
Shape Factor (R <sup>2</sup> ):	1.5
Anisotropy Ratio:	1
Anisotropy Angle:	0

##### Data Counts at Validation Points

Active Results:	20000
Blanked Results:	0
Attempted Results:	20000
Requested Results:	20000

---

##### Univariate Cross-Validation Statistics

---

	Estimated Z	Residual Z	Data Used
Minimum:	652.04557984144	-1.1507125676759	32
25%-tile:	719.97970383943	-0.020499692516069	63
Median:	742.0315020613	-0.00039743474417264	63
75%-tile:	771.94469175634	0.019285242940782	63
Maximum:	846.01071589021	1.0805374340856	63
Midrange:	749.02814786582	-0.035087566795141	47.5
Range:	193.96513604877	2.2312500017615	31
Interquartile Range:	51.964987916904	0.039784935456851	0
Median Abs. Deviation:	25.934167035791	0.019872231571981	0
Mean:	746.73246828041	-0.0014317195892371	62.8494
Trim Mean (10%):	745.86036900396	-0.00074620214438462	63
Standard Deviation:	35.998052447685	0.081803644022705	1.5541620378846
Variance:	1295.8597800263	0.0066918361753935	2.4154196400016
Coef. of Variation:	0.048207429001422	-1	0.024728351231429
Coef. of Skewness:	0.37335211763843	-0.23752227284551	-11.066332503374
Root Mean Square:	747.59965152759	0.081816171973368	62.868612995675
Mean Square:	558905.23896418	0.0066938859963757	3952.4625

#### AREA 2

##### Cross Validation Report

---

Gridding Method:	Radial Basis Function
Basis Kernel Type:	Multiquadric
Shape Factor (R <sup>2</sup> ):	1.5
Anisotropy Ratio:	1
Anisotropy Angle:	0

##### Data Counts at Validation Points

Active Results:	20000
Blanked Results:	0
Attempted Results:	20000
Requested Results:	20000

---

##### Univariate Cross-Validation Statistics

---

	Estimated Z	Residual Z	Data Used
Minimum:	627.99809810151	-4.8242575637452	31
25%-tile:	691.85942986684	-0.033953995586671	63
Median:	723.97181211134	0.0001767470428149	63
75%-tile:	756.63074892971	0.032571903687085	63
Maximum:	814.21240666186	3.531617348065	63

Midrange:	721.10525238169	-0.6463201078401	47
Range:	186.21430856035	8.3558749118101	32
Interquartile Range:	64.771319062874	0.066525899273756	0
Median Abs. Deviation:	32.240880900819	0.033142091125342	0
Mean:	723.11620281548	-0.0021971845190609	62.89715
Trim Mean (10%):	723.46590392249	-0.00062516469158708	63
Standard Deviation:	42.287293155824	0.15073260542844	1.2849793296
Variance:	1788.2151624466	0.022720318339247	1.6511718774993
Coef. of Variation:	0.058479249934072	-1	0.020429849835804
Coef. of Skewness:	-0.087330169184834	-1.9402659487809	-13.728281383575
Root Mean Square:	724.35161208955	0.15074861843167	62.910274598034
Mean Square:	524685.25793672	0.022725145959058	3957.70265

### AREA 3 Cross Validation Report

#### Gridding Rules

Gridding Method:	Radial Basis Function
Basis Kernel Type:	Multiquadric
Shape Factor (R^2):	1.5
Anisotropy Ratio:	1
Anisotropy Angle:	0

#### Data Counts at Validation Points

Active Results:	20000
Blanked Results:	0
Attempted Results:	20000
Requested Results:	20000

#### Univariate Cross-Validation Statistics

	Estimated Z	Residual Z	Data Used
Minimum:	720.04197909798	-1.619320361766	32
25%-tile:	836.08553563568	-0.022526167074261	63
Median:	901.95395665524	-1.4214806469681E-006	63
75%-tile:	957.65483534771	0.022367146485976	63
Maximum:	1028.0074647793	1.5359244509488	63
Midrange:	874.02472193866	-0.041697955408608	47.5
Range:	307.96548568135	3.1552448127147	31
Interquartile Range:	121.56929971203	0.044893313560237	0
Median Abs. Deviation:	59.941003757825	0.022411797438508	0
Mean:	894.73486123775	0.00036123775049282	62.8879
Trim Mean (10%):	896.1957101324	-0.0001286752209052	63
Standard Deviation:	73.31020973328	0.080825686479331	1.3390793815147
Variance:	5374.3868511376	0.0065327915948552	1.7931335899979
Coef. of Variation:	0.081935121687183	223.74651145697	0.021293116505953
Coef. of Skewness:	-0.23874423466814	0.59130033059209	-13.309990708059
Root Mean Square:	897.733177935	0.080826493723083	62.902154971034
Mean Square:	805924.85876527	0.0065329220875675	3956.6811

### AREA 4 Cross Validation Report

#### Gridding Rules

Gridding Method:	Radial Basis Function
Basis Kernel Type:	Multiquadric
Shape Factor (R^2):	1.5
Anisotropy Ratio:	1
Anisotropy Angle:	0

#### Data Counts at Validation Points

Active Results:	35000
Blanked Results:	0

Attempted Results: 35000  
 Requested Results: 35000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	680.25510810996	-2.2791504610581	31
25%-tile:	743.97587825649	-0.034417516983808	63
Median:	773.99086115329	-0.00073530668748845	63
75%-tile:	801.97452432365	0.030719560417651	63
Maximum:	890.06453532133	1.456276240502	63
Midrange:	785.15982171565	-0.41143711027803	47
Range:	209.80942721137	3.7354267015601	32
Interquartile Range:	57.99864606717	0.065137077401459	0
Median Abs. Deviation:	28.150226336694	0.032569197740941	0
Mean:	773.77461898497	-0.0035524436021027	62.899257142857
Trim Mean (10%):	773.14318257653	-0.0021578058116476	63
Standard Deviation:	40.607823584587	0.13076392638661	1.2679429987825
Variance:	1648.9953362769	0.017099204444043	1.6076794481616
Coef. of Variation:	0.052480170049847	-1	0.020158314364552
Coef. of Skewness:	0.15983988824054	-0.91883105286182	-13.591561250676
Root Mean Square:	774.83943905922	0.1308121718327	62.912035641612
Mean Square:	600376.15632161	0.017111824299589	3957.9242285714

**AREA 5  
 Cross Validation Report**

**Gridding Rules**

Gridding Method: Radial Basis Function  
 Basis Kernel Type: Multiquadric  
 Shape Factor (R^2): 1.5  
 Anisotropy Ratio: 1  
 Anisotropy Angle: 0

**Data Counts at Validation Points**

Active Results: 5000  
 Blanked Results: 0  
 Attempted Results: 5000  
 Requested Results: 5000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	959.90858570117	-1.1378114725624	33
25%-tile:	979.4883511924	-0.0094826020430219	63
Median:	992.00133029568	-0.00015647296140742	63
75%-tile:	1013.9997510156	0.0082272790559728	63
Maximum:	1046.0350683108	1.251418818565	63
Midrange:	1002.971827006	0.056803673001298	48
Range:	86.126482609661	2.3892302911273	30
Interquartile Range:	34.511399823216	0.017709881098995	0
Median Abs. Deviation:	16.002685669822	0.0087991327890222	0
Mean:	996.49459846546	0.0017984654572897	62.7298
Trim Mean (10%):	995.96935798109	-0.00013944104172732	63
Standard Deviation:	21.04697989172	0.08715620800407	1.9780778447774
Variance:	442.97536256247	0.0075962045936487	3.9127919599991
Coef. of Variation:	0.021121017539012	48.461430076847	0.031533303864788
Coef. of Skewness:	0.37979759231378	2.0091710734121	-8.258134000489
Root Mean Square:	996.71684049854	0.087174761666722	62.760979915868
Mean Square:	993444.4601334	0.0075994390716498	3938.9406

**AREA 6  
 Cross Validation Report**

---

**Gridding Rules**

Gridding Method: Radial Basis Function  
Basis Kernel Type: Multiquadric  
Shape Factor (R<sup>2</sup>): 1.5  
Anisotropy Ratio: 1  
Anisotropy Angle: 0

**Data Counts at Validation Points**

Active Results: 20000  
Blanked Results: 0  
Attempted Results: 20000  
Requested Results: 20000

**Univariate Cross-Validation Statistics**

---

	Estimated Z	Residual Z	Data Used
Minimum:	679.9996529995	-4.2638196804767	34
25%-tile:	720.09116277686	-0.022367023686343	63
Median:	741.96064253985	-5.0996395657421E-006	63
75%-tile:	761.98603672127	0.021626805889582	63
Maximum:	807.79324468292	1.6072709570656	63
Midrange:	743.89644884143	-1.3282743617056	48.5
Range:	127.79359168297	5.8710906375422	29
Interquartile Range:	41.894873944414	0.043993829575925	0
Median Abs. Deviation:	20.037646439371	0.022021732059102	0
Mean:	740.31962634928	-0.0017736507188445	62.90345
Trim Mean (10%):	740.47461302896	-0.00056963027886942	63
Standard Deviation:	28.653412694545	0.093988455830332	1.2092262391712
Variance:	821.01805904391	0.0088338298293703	1.4622280975002
Coef. of Variation:	0.038704110595909	-1	0.019223528108096
Coef. of Skewness:	-0.12429086294949	-6.8674303519384	-13.71219389469
Root Mean Square:	740.87392126932	0.094005189570804	62.915071723713
Mean Square:	548894.16721698	0.0088369756662429	3958.30625

---

**AREA 7****Cross Validation Report**

---

**Gridding Rules**

Gridding Method: Radial Basis Function  
Basis Kernel Type: Multiquadric  
Shape Factor (R<sup>2</sup>): 1.5  
Anisotropy Ratio: 1  
Anisotropy Angle: 0

**Data Counts at Validation Points**

Active Results: 40000  
Blanked Results: 0  
Attempted Results: 40000  
Requested Results: 40000

**Univariate Cross-Validation Statistics**

---

	Estimated Z	Residual Z	Data Used
Minimum:	711.9673796484	-7.4349237240669	31
25%-tile:	773.9792405317	-0.033856516738069	63
Median:	798.01640734938	-1.478984734149E-005	63
75%-tile:	819.93332428857	0.032627658805268	63
Maximum:	874.1462852123	5.9544147186097	63
Midrange:	793.05683243035	-0.74025450272865	47
Range:	162.1789055639	13.389338442677	32
Interquartile Range:	45.954083756877	0.066484175543337	0
Median Abs. Deviation:	23.847561169564	0.033223147078502	0
Mean:	796.46445848398	-0.0011915160177642	62.939575
Trim Mean (10%):	796.6721716153	-0.00076737889660287	63

---

Standard Deviation:	32.471172818615	0.13538942037366	0.96903241399591
Variance:	1054.3770642163	0.018330295149115	0.93902381937473
Coef. of Variation:	0.040769142267091	-1	0.015396233832146
Coef. of Skewness:	-0.13542459133102	-1.6787653774676	-18.013716772254
Root Mean Square:	797.1260946001	0.13539466333477	62.9470342828
Mean Square:	635410.0106924	0.018331714859536	3962.329125

**AREA 8**  
**Cross Validation Report**

**Gridding Rules**

Gridding Method:	Radial Basis Function
Basis Kernel Type:	Multiquadric
Shape Factor (R^2):	1.5
Anisotropy Ratio:	1
Anisotropy Angle:	0

**Data Counts at Validation Points**

Active Results:	10000
Blanked Results:	0
Attempted Results:	10000
Requested Results:	10000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	816.00026461474	-0.88891778837069	31
25%-tile:	859.97666226338	-0.010807919609078	63
Median:	873.99450896681	-0.00014039104883068	63
75%-tile:	886.74598030586	0.0091725601085955	63
Maximum:	904.003167631	0.7476031971039	63
Midrange:	860.00171612287	-0.070657295633396	47
Range:	88.002903016265	1.6365209854746	32
Interquartile Range:	26.769318042482	0.019980479717674	0
Median Abs. Deviation:	13.992466511105	0.010066213786331	0
Mean:	872.12696570004	-0.00023429996785266	62.8127
Trim Mean (10%):	872.73666999147	-0.0004767914503611	63
Standard Deviation:	17.745914857147	0.058848473609371	1.7215163984115
Variance:	314.91749411712	0.0034631428461528	2.9636187099997
Coef. of Variation:	0.020347857083978	-1	0.027407138976855
Coef. of Skewness:	-0.39665935435692	-0.1863092698033	-10.54965401795
Root Mean Square:	872.3074926855	0.058848940029773	62.836286491167
Mean Square:	760920.36179527	0.0034631977426278	3948.3989

**AREA 9**  
**Cross Validation Report**

**Gridding Rules**

Gridding Method:	Radial Basis Function
Basis Kernel Type:	Multiquadric
Shape Factor (R^2):	1.5
Anisotropy Ratio:	1
Anisotropy Angle:	0

**Data Counts at Validation Points**

Active Results:	6000
Blanked Results:	0
Attempted Results:	6000
Requested Results:	6000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	819.99166266135	-2.2142979073718	31

25%-tile:	841.99849123409	-0.0064167085521376	63
Median:	852.027848498	-0.00014824117806711	63
75%-tile:	863.98484103862	0.0046015521107847	63
Maximum:	899.99417678917	1.7109002766871	63
Midrange:	859.99291972526	-0.25169881534237	47
Range:	80.002514127827	3.9251981840589	32
Interquartile Range:	21.98634980453	0.011018260662922	0
Median Abs. Deviation:	10.048832216923	0.0054894397516136	0
Mean:	853.85730991903	0.0023099190309624	62.836333333333
Trim Mean (10%):	853.44421698078	-0.00060290084109374	63
Standard Deviation:	15.042532445515	0.10671574803961	1.5971474224033
Variance:	226.27778237437	0.011388250879654	2.5508798888894
Coef. of Variation:	0.017617150161707	46.198912866288	0.02541757829711
Coef. of Skewness:	0.37664543929833	0.29752937830034	-11.080350112929
Root Mean Square:	853.98980291602	0.10674074482401	62.856627865856
Mean Square:	729298.58348454	0.011393586605584	3950.9556666667

**AREA 10**  
**Cross Validation Report**

**Gridding Rules**

Gridding Method:	Radial Basis Function
Basis Kernel Type:	Multiquadric
Shape Factor (R <sup>2</sup> ):	1.5
Anisotropy Ratio:	1
Anisotropy Angle:	0

**Data Counts at Validation Points**

Active Results:	15000
Blanked Results:	0
Attempted Results:	15000
Requested Results:	15000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	887.94860436641	-1.3719892422827	35
25%-tile:	932.06365337705	-0.021277746680084	63
Median:	955.99545482308	-0.00029924685566129	63
75%-tile:	981.96528686969	0.01861988730343	63
Maximum:	1030.0014304919	3.287938110755	63
Midrange:	958.97501742913	0.95797443423618	49
Range:	142.05282612545	4.6599273530377	28
Interquartile Range:	49.901633492634	0.039897633983514	0
Median Abs. Deviation:	24.01065857771	0.020100868290911	0
Mean:	956.72071646645	-0.0014168668855248	62.884133333333
Trim Mean (10%):	956.63283851281	-0.0011859742774896	63
Standard Deviation:	30.730395549621	0.087687256569381	1.3085519664461
Variance:	944.35721063614	0.0076890549646645	1.7123082488901
Coef. of Variation:	0.032120549937625	-1	0.020808936962045
Coef. of Skewness:	0.037938508814133	3.0929890746125	-12.279821691145
Root Mean Square:	957.21412783489	0.087698702820713	62.897746647502
Mean Square:	916258.88652671	0.0076910624764358	3956.1265333333

## ANEXO B

### PRUEBA DE SENSIBILIDAD PARA FACTOR $R^2$ EN DEM BUCARAMANGA

**$R^2 = 1.85$**

#### Cross Validation Report

Gridding Method:	Radial Basis Function
Basis Kernel Type:	Multiquadric
Shape Factor ( $R^2$ ):	1.85
Anisotropy Ratio:	1
Anisotropy Angle:	0

#### Data Counts at Validation Points

Active Results:	50000
Blanked Results:	0
Attempted Results:	50000
Requested Results:	50000

#### Univariate Cross-Validation Statistics

	Estimated Z	Residual Z	Data Used
Minimum:	601.99228440303	-4.223107158296	34
25%-tile:	760.05184379321	-0.025575209541557	63
Median:	812.01887256073	-0.00012556484682591	63
75%-tile:	884.04652618789	0.023696463156512	63
Maximum:	1317.9594026818	3.2606461849792	63
Midrange:	959.97584354242	-0.48123048665838	48.5
Range:	715.96711827879	7.4837533432752	29
Interquartile Range:	123.99468239468	0.049271672698069	0
Median Abs. Deviation:	60.011750607315	0.024675084625528	0
Mean:	838.52386490357	-0.0018150964255947	62.97364
Trim Mean (10%):	830.77120471042	-0.0011825297232051	63
Standard Deviation:	117.74972007517	0.11911197721683	0.58311675537664
Variance:	13864.996577781	0.014187663116503	0.34002515040098
Coef. of Variation:	0.14042500756816	-1	0.009259695888258
Coef. of Skewness:	1.0785568252613	-1.0381979256535	-24.649139881348
Root Mean Square:	846.75100743407	0.12122580615273	62.976339684043
Mean Square:	716987.26859061	0.014190957691537	3966.01936

**$R^2 = 1.87$**

#### Cross Validation Report

Gridding Method:	Radial Basis Function
Basis Kernel Type:	Multiquadric
Shape Factor ( $R^2$ ):	1.87
Anisotropy Ratio:	1
Anisotropy Angle:	0

#### Data Counts at Validation Points

Active Results:	50000
Blanked Results:	0
Attempted Results:	50000
Requested Results:	50000

#### Univariate Cross-Validation Statistics

	Estimated Z	Residual Z	Data Used
Minimum:	601.99355826854	-5.3761890010949	34
25%-tile:	769.96447272788	-0.024957656648667	63
Median:	822.04377130047	7.9825175589576E-006	63
75%-tile:	884.01916519443	0.024296365479472	63
Maximum:	1316.7350567524	4.5471936684578	63

Midrange:	959.36430751047	-0.41449766631854	48.5
Range:	714.74149848387	9.9233826695527	29
Interquartile Range:	114.05469246655	0.049254022128139	0
Median Abs. Deviation:	57.489998884978	0.024632991065346	0
Mean:	843.06076387554	-0.0009161244705954	62.97868
Trim Mean (10%):	835.99438695891	-0.00049369767472766	63
Standard Deviation:	114.99260451132	0.12480968994858	0.52750872751064
Variance:	13223.299092297	0.01557745870506	0.2782654575999
Coef. of Variation:	0.13639895181778	-1	0.0083759889459519
Coef. of Skewness:	1.0353877262385	-1.9480143002339	-28.735646245243
Root Mean Square:	850.86705816989	0.12481305215844	62.980889164889
Mean Square:	723974.75067869	0.015578297989106	3966.5924

**R<sup>2</sup> = 1.89**

**Cross Validation Report**

Gridding Method: Radial Basis Function  
 Basis Kernel Type: Multiquadric  
 Shape Factor (R<sup>2</sup>): 1.89  
 Anisotropy Ratio: 1  
 Anisotropy Angle: 0

**Data Counts at Validation Points**

Active Results: 50000  
 Blanked Results: 0  
 Attempted Results: 50000  
 Requested Results: 50000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	601.98171482538	-4.267087713489	47
25%-tile:	768.01374413595	-0.025105615147481	63
Median:	822.00324197462	-0.00014274353691235	63
75%-tile:	885.54883924659	0.024110937129535	63
Maximum:	1318.0317933487	2.8098166399725	63
Midrange:	960.00675408703	-0.72863553675825	55
Range:	716.0500785233	7.0769043534615	16
Interquartile Range:	117.53509511064	0.049216552277016	0
Median Abs. Deviation:	57.983657292254	0.024581600175225	0
Mean:	842.88107401263	-0.0011659873642338	62.97846
Trim Mean (10%):	835.84074003946	-0.0006896975070346	63
Standard Deviation:	115.38007307174	0.11982149247241	0.51208986359888
Variance:	13312.56126204	0.014357190058317	0.26223602840071
Coef. of Variation:	0.13688772548001	-1	0.0081311906261105
Coef. of Skewness:	1.0246444210599	-1.6203096569224	-26.003020376273
Root Mean Square:	850.741480234	0.11982716547115	62.980541915738
Mean Square:	723761.06619073	0.01435854958485	3966.54866

**R<sup>2</sup> = 1.90**

**Cross Validation Report**

Gridding Method: Radial Basis Function  
 Basis Kernel Type: Multiquadric  
 Shape Factor (R<sup>2</sup>): 1.9  
 Anisotropy Ratio: 1  
 Anisotropy Angle: 0

**Data Counts at Validation Points**

Active Results: 50000  
 Blanked Results: 0  
 Attempted Results: 50000  
 Requested Results: 50000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	600.18819198213	-3.3529280426663	38
25%-tile:	769.95405156434	-0.025889724007357	63
Median:	822.00633935205	-0.00029077049305215	63
75%-tile:	883.99383447895	0.023466153591926	63
Maximum:	1312.0864389198	2.5963862283136	63
Midrange:	956.13731545095	-0.37827090717639	50.5
Range:	711.89824693764	5.9493142709799	25
Interquartile Range:	114.03978291461	0.049355877599282	0
Median Abs. Deviation:	56.058898134003	0.024694483791791	0
Mean:	841.98075992911	-0.001840070889882	62.977
Trim Mean (10%):	835.11264671434	-0.0013231555618284	63
Standard Deviation:	113.3791668433	0.11817837197675	0.53493083665159
Variance:	12854.83547408	0.013966127603074	0.28615100000077
Coef. of Variation:	0.13465766943753	-1	0.008494066669603
Coef. of Skewness:	1.0193276815193	-0.62072566958787	-26.173759727526
Root Mean Square:	849.58015252528	0.11819269632238	62.979271828118
Mean Square:	721786.43556488	0.013969513463954	3966.38868

**R<sup>2</sup> = 1.95****Cross Validation Report**

Gridding Method: Radial Basis Function  
Basis Kernel Type: Multiquadric  
Shape Factor (R<sup>2</sup>): 1.95  
Anisotropy Ratio: 1

**Data Counts at Validation Points**

Active Results: 50000  
Blanked Results: 0  
Attempted Results: 50000  
Requested Results: 50000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	602.03770604497	-4.4816800714393	47
25%-tile:	768.01403306311	-0.024858468915227	63
Median:	821.99783802682	-0.00014495853656626	63
75%-tile:	884.04030952151	0.024250148497686	63
Maximum:	1315.9957406171	3.1961997780816	63
Midrange:	959.01672333104	-0.64274014667888	55
Range:	713.95803457215	7.6778798495209	16
Interquartile Range:	116.0262764584	0.049108617412912	0
Median Abs. Deviation:	57.986372039916	0.024527347240678	0
Mean:	842.11053857064	-0.00034142936208302	62.9808
Trim Mean (10%):	835.10540869216	-0.00036821069702726	63
Standard Deviation:	114.34286291358	0.12322261454884	0.48040749369634
Variance:	13074.290299274	0.015183812736252	0.2307913599996
Coef. of Variation:	0.13578129910079	-1	0.0076278404481419
Coef. of Skewness:	1.0233478979813	-0.2845783038033	-27.661869676373
Root Mean Square:	849.83789599606	0.11792308756991	62.982632209205
Mean Square:	722224.449471	0.015183929310261	3966.81196

**R<sup>2</sup> = 2.05****Cross Validation Report**

Gridding Method: Radial Basis Function  
Basis Kernel Type: Multiquadric  
Shape Factor (R<sup>2</sup>): 2.05  
Anisotropy Ratio: 1  
Anisotropy Angle: 0

**Data Counts at Validation Points**

Active Results: 50000  
 Blanked Results: 0  
 Attempted Results: 50000  
 Requested Results: 50000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	599.95036756921	-5.4663974692223	47
25%-tile:	769.9718156657	-0.024616619910148	63
Median:	822.00393236829	-8.34488023429E-005	63
75%-tile:	885.96488778745	0.024037351828383	63
Maximum:	1317.9476042918	2.8531782549758	63
Midrange:	958.94898593053	-1.3066096071233	55
Range:	717.99723672263	8.3195757241981	16
Interquartile Range:	115.99307212175	0.048653971738531	0
Median Abs. Deviation:	57.974552365288	0.024355562212349	0
Mean:	843.17391342785	-0.00080657214915639	62.97768
Trim Mean (10%):	836.17716680549	-0.00034728400601901	63
Standard Deviation:	115.17415033916	0.12180270290145	0.53533337052755
Variance:	13265.084906348	0.0148358984341	0.28658181760039
Coef. of Variation:	0.13659596022241	-1	0.0085003666462079
Coef. of Skewness:	1.0213705555604	-2.029668270086	-25.898070446421
Root Mean Square:	851.003721021	0.12180537341485	62.979955223865
Mean Square:	724207.33319158	0.014836548992731	3966.47476

**R<sup>2</sup> = 2.1****Cross Validation Report**

Gridding Method: Radial Basis Function  
 Basis Kernel Type: Multiquadric  
 Shape Factor (R<sup>2</sup>): 2.1  
 Anisotropy Ratio: 1  
 Anisotropy Angle: 0

**Data Counts at Validation Points**

Active Results: 50000  
 Blanked Results: 0  
 Attempted Results: 50000  
 Requested Results: 50000

**Univariate Cross-Validation Statistics**

	Estimated Z	Residual Z	Data Used
Minimum:	599.9951555384	-4.9104659001323	41
25%-tile:	769.92813455475	-0.024550021915502	63
Median:	822.03220503899	2.5806912162807E-011	63
75%-tile:	884.00586273448	0.02382891002469	63
Maximum:	1317.9321275354	5.793913512445	63
Midrange:	958.96364153691	0.44172380615635	52
Range:	717.93697199702	10.704379412577	22
Interquartile Range:	114.07772817973	0.048378931940192	0
Median Abs. Deviation:	56.902737781462	0.02422192232018	0
Mean:	842.42038663491	-5.336509311242E-005	62.97662
Trim Mean (10%):	835.454853826	-0.00041307317050763	63
Standard Deviation:	114.08770594267	0.12242635518867	0.5298994013963
Variance:	13016.004647262	0.014988212444782	0.28079337560015
Coef. of Variation:	0.13542847223629	-1	0.008414224221565
Coef. of Skewness:	1.025116104578	1.3728593566775	-24.875976943429
Root Mean Square:	850.11064718974	0.12242636681947	62.978849306732
Mean Square:	722688.11246537	0.014988215292615	3966.33546

## ANEXO C

### VALIDACION CRUZADA CON VARIOGRAMA EXPONENCIAL

#### Cross Validation Report

##### Gridding Rules

Gridding Method: Kriging  
Kriging Type: Point

##### Semi-Variogram Model

Component Type: Exponential  
Anisotropy Angle: 101.8  
Anisotropy Length: 55790  
Anisotropy Ratio: 3.833  
Variogram Scale: 438700

##### Data Counts at Validation Points

Active Results: 30000  
Blanked Results: 0  
Attempted Results: 30000  
Requested Results: 30000

#### Univariate Cross-Validation Statistics

	Estimated Z	Residual Z	Data Used
Minimum:	520.87206806334	-113.02419578568	31
25%-tile:	950.65478060151	-2.5692856943729	63
Median:	1227.7318716189	0	63
75%-tile:	1797.1365048336	2.2200463318222	63
Maximum:	3550.8865007645	68.998080772849	63
Midrange:	2035.8792844139	-22.013057506417	47
Range:	3030.0144327012	182.02227655853	32
Interquartile Range:	846.48172423214	4.7893320261951	0
Median Abs. Deviation:	328.63778812563	2.3943892265307	0
Mean:	1417.7594432766	-0.21305672335791	62.761566666667
Trim Mean (10%):	1373.2849436316	-0.17667970034662	63
Standard Deviation:	595.31937371917	8.6257173532593	1.7608472048686
Variance:	354405.15672539	74.402999858318	3.1005828788937
Coef. of Variation:	0.41990154009717	-1	0.028056138468
Coef. of Skewness:	1.0441534830147	-0.2168417057625	-8.1332703435781
Root Mean Square:	1537.675842213	8.6283482211654	62.786263094194
Mean Square:	2364446.9957255	74.448393025688	3942.1148333333

## ANEXO D

### SENSIBILIDAD DEL PARAMETRO $R^2$ PARA RIO DE ORO

#### Cross Validation Report

##### Gridding Rules

Gridding Method: Radial Basis Function  
Basis Kernel Type: Multiquadric  
Shape Factor ( $R^2$ ): 600  
Anisotropy Ratio: 1  
Anisotropy Angle: 0

##### Data Counts at Validation Points

Active Results: 20000  
Blanked Results: 0  
Attempted Results: 20000  
Requested Results: 20000

#### Univariate Cross-Validation Statistics

	Estimated Z	Residual Z	Data Used
Minimum:	648.70044969507	-70.843171914211	32
25%-tile:	950.10138321388	-0.87346570186924	63
Median:	1207.9397906864	-0.001694133609135	63
75%-tile:	1798.2218445445	0.70550322579527	63
Maximum:	3550.0000938757	36.257717652342	63
Midrange:	2099.3502717854	-17.292727130934	47.5
Range:	2901.2996441806	107.10088956655	31
Interquartile Range:	848.12046133059	1.5789689276645	0
Median Abs. Deviation:	340.61867494953	0.78739017152338	0
Mean:	1415.0019782182	-0.11802178184508	62.764
Trim Mean (10%):	1369.6303434519	-0.087457726261442	63
Standard Deviation:	594.32893001888	3.8479475618542	1.7621305286508
Variance:	353226.87705738	14.80670043878	3.105104000003
Coef. of Variation:	0.42001985804097	-1	0.028075497556733
Coef. of Skewness:	1.0596898991794	-0.74132524181592	-8.1973561615018
Root Mean Square:	1534.7499716301	3.8497570806182	62.788731473092
Mean Square:	2355457.4754187	14.82062957977	3942.4248