

CAPÍTULO I

1. INTRODUCCIÓN

Un mapeo batimétrico detallado es fundamental para una correcta cuantificación del vital líquido elemento sobre todo en lugares donde escasea. Sin embargo, un levantamiento detallado implica un alto precio siendo algunas veces incluso innecesario el nivel de detalle para la obtención de resultados similares. En este contexto, Actualmente, existen diferentes métodos de interpolación espacial descritos en la literatura que se utilizan para generar mapas batimétricos.

En la actualidad existen más de 40 métodos de interpolación espacial descritos en la literatura, que se clasifican en tipos deterministas, geoestadísticos y combinados. Varios de estos métodos se han utilizado para la modelación de superficies batimétricas. Sin embargo, existen muchos factores (por ejemplo, el tamaño de la muestra y la naturaleza de los datos) que afectan la estimación de un interpolador espacial universal y hasta ahora no hay hallazgos consistentes con respecto a cuál es el mejor método de interpolación.

La distancia de separación entre secciones de levantamientos hidrográficos es quizás la variable más importante que afecta de manera directa a los métodos de interpolación y por ende a los resultados obtenidos. La determinación de la distancia adecuada ha sido tema de debate por años tanto por investigadores como instituciones hidrográficas.

El objetivo de este estudio es analizar modelos digitales de elevación obtenidos por tres de las metodologías de interpolación más utilizadas en estudios hidrográficos a distintas distancias de separación de líneas de sondeo esto para obtener la distancia óptima. En el capítulo I se presenta la introducción y los antecedentes de investigaciones similares y los objetivos de la presente investigación, en el capítulo II se explican los conceptos relacionados al análisis de variables espacialmente distribuidas, las características del fenómeno en estudio y los métodos de interpolación que se utilizaron, el capítulo III se encuentra el marco metodológico el conjunto de pasos y los procesos que se desarrollaron y la metodología empleada, el capítulo IV se exponen y analizan los resultados obtenidos de los tres métodos de interpolación.

1.1. Antecedentes

Son varios los investigadores e instituciones que contribuyeron a tener una mejor comprensión sobre los métodos de interpolación sin embargo existen varias opiniones, Azpurua, Dos Ramos, (2010) afirma que la ponderación de distancia inversa (IDW), un método determinista, es más probable que produzca la mejor estimación de una superficie continua de la magnitud promedio de la intensidad del campo eléctrico. Por otro lado, el artículo “Evaluación de kriging de regresión para interpolación espacial comparaciones de siete métodos de interpolación GIS.” Meng et al, (2013). declaran que, para algunos casos, los métodos deterministas, como IDW y la Función de Base Radial (RBF), mostraron mejores resultados que los métodos geoestadísticos, como Ordinary Kriging (OK), y en otras situaciones, los métodos geoestadísticos arrojaron los mejores resultados. Este hecho refleja la importancia de evaluar el método de interpolación para cada conjunto de datos y caso específico.

La interpolación de datos batimétricos en particular, no hay consenso sobre qué método proporciona los mejores resultados. Bello – Pineda y Hernández – Stefanoni, (2006) plantean que el método kriging era el más adecuado para mapear la batimetría de la plataforma sumergida de Yucatán. En contra posición Merwade et al. (2006) sugieren que los métodos deterministas, como el IDW elíptico, pueden proporcionar mejores resultados en comparación con los métodos geoestadísticos, como kriging Ordinario.

De igual manera existen varias opiniones sobre la distancia de líneas de sondeo que corresponden a la precisión que cada institución o investigador requiere. la agencia United States Bureau of Reclamation de EE.UU. Departamento del Interior, Denver, Colorado asegura que para las encuestas con el equipo monohaz (haz individual), la separación entre secciones debe establecerse teniendo en cuenta el propósito de la encuesta a mejores puntos de densidad variando en función de los cambios. una forma diferente para determinar secciones batimétricas se define por Yuqian (1989). Tomando como referencia el mapa batimétrico anterior para llenar el depósito y su volumen, las secciones se asignan en este mapa, de modo que el volumen no supera el 5% del valor de referencia. Las secciones en este caso no tienen la necesidad de ser equidistantes.

El Servicio Geológico de Estados Unidos indica que el intervalo mínimo entre transectos es de 10 metros debido a la dificultad de mantener, durante la recolección, la posición del barco alineado con las líneas planificadas (Wilson; Richards, 2006). La Junta de Desarrollo del Agua de Texas – TWDB Inició un programa de estudios batimétricos en 1991 y ha realizado más de 100 batimetrías de embalses en Texas-USA. Este departamento utiliza la separación entre las secciones transversales paralelas a una distancia que oscila generalmente de 75 a 150 m (Furnans; Austin, 2008).

El cuerpo de ingenieros de E.E.U.U. también da su contribución en la distancia entre las secciones. Se establece que cuando el objetivo es calcular el volumen de un depósito de fondo irregular, existe una separación estricta entre los transectos. La distancia de las secciones es por lo general entre 60 y 120 metros con un límite especificado para el que no debe excederse. También podría estar espaciada más si el fondo del depósito es uniforme. Especificar un espaciado mucho menor puede ser económicamente inviable. En cuanto a la precisión horizontal, esto no es crítico para estudios de sedimentos (Usace, 2002)

1.2. Problema de investigación

1.2.1. Planteamiento del problema

Actualmente no se conoce el método de interpolación más adecuado para levantamientos batimétricos y la distancia de separación de líneas de sondeo óptima para realizar levantamientos batimétricos en embalse “El Tranque” con niveles de precisión aceptables que permitan la reducción de los trabajos de recopilación y que generen superficies confiables de monitoreo del embalse sin perder la representatividad del fondo sub acuático. Esto afecta generalmente a instituciones encargadas de realizar trabajos de levantamiento hidrográficos las cuales muchas veces realizan trabajos con distancias demasiado extensas las cuales no representan correctamente la superficie sub acuática o sumamente detallados que no conllevan a una mejora en la precisión, el presente trabajo propone la generación de varios modelos variando las distancias de líneas de sondeo y los métodos de interpolación para la obtención de volúmenes y analizar su variación, los resultados obtenidos serian de gran apoyo tanto para instituciones hidrográficas como para la el laboratorio de hidráulica de la U.A.J.M.S.

1.2.2. Formulación del problema

¿Cuál es el método de interpolación más adecuado para realizar levantamiento hidrográfico en el embalse “El tranque” y a que distancia de separación de líneas de sondeo?

1.3. Objetivo de proyecto

1.3.1. Objetivo general

Analizar modelos digitales de elevación obtenidos por tres metodologías de interpolación a distintas distancias de separación de líneas de sondeo por medio de validación cruzada y el (RMSE) para determinar el método de interpolación más adecuado en el embalse “El Tranque” y la distancia de separación de líneas de sondeo.

1.3.2. Objetivos específicos

- ✓ Realizar la red geodésica del embalse “El Tranque” alrededor del embalse dirigido a tener puntos que permitan el correcto levantamiento del margen del embalse.
- ✓ Realizar el levantamiento topográfico del margen del embalse “El tranque” a nivel superior de vertedero cota 2023.9 m.s.n.m. dirigido para la obtención del volumen útil.
- ✓ Realizar el levantamiento batimétrico del embalse “El tranque” con distancias de líneas de sondeo entre 10 y 20 metros dirigido a obtener un levantamiento detallado del embalse.
- ✓ Generar modelos digitales batimétricos con tres metodologías de interpolación variando la separación de líneas de sondeo para obtener diferentes resultados.
- ✓ Comparar los modelos generados por medio de la validación cruzada y el RMSE para conocer el modelo más confiable y la distancia más óptima.

1.4. Justificaciones del proyecto

Las razones por la cual se realiza este proyecto son:

1.4.1. Justificación académica

Este estudio contribuye al avance de investigación de modelos digitales de superficies sub acuáticas mejorando la precisión del post proceso de datos brutos de batimetrías en embalses.

1.4.2. Justificación técnica

Mediante el análisis a realizar, se brindará información que sirva para generar una planificación sobre la realización de levantamientos en menor tiempo.

1.4.3. Justificación social

Con el producto de este estudio será posible determinar el volumen de almacenamiento con mayor precisión lo que contribuye a mejorar la planificación y operación de embalses.

1.5. Alcance

“El tranque” es una represa de tierra, de 195 m. de longitud de coronamiento y una altura de presa es de 40 m., cuyo vaso de almacenamiento es de aproximadamente 62 Has. de área y un volumen total de almacenamiento de 5,00 hm³., por lo cual los resultados obtenidos del presente estudio están sujetos a la morfología del vaso de almacenamiento y al periodo que se realizaron los trabajos de batimetría correspondientes al mes de junio del 2019.

El estudio está limitado para levantamientos batimétricos en el embalse “El tranque” con sonda electroacústica de haz simple, no incluye el uso de los métodos de “Perfilador Acústico Doppler de Corriente (ADCP)” o sondas de haz multiple, para los cuales se deben realizar otro tipo de análisis.

La determinación de la superficie subacuática de un embalse requiere únicamente la obtención de coordenadas y profundidades de cada punto, por lo tanto, en el presente trabajo se detalla solamente el procedimiento realizado para la obtención y procesado de los mismos.

CAPÍTULO II

2. FUNDAMENTO TEÓRICO

2.1. Análisis de variables espacialmente distribuidas

2.1.1. Influencia Espacial de Datos

La estadística espacial es una rama de la estadística y de los SIG, que a menudo se pasa por alto (ESRI 2001) (Lynch y Krause 2014). La capacidad de analizar los datos de variabilidad y la integridad debe exigirse como parte de la validación y verificación de la aplicabilidad de un conjunto de datos.

Los puntos utilizados en la interpolación, las características de la superficie, dependen de la distribución espacial estadística, modifica la superficie resultante (Kravchenko 2003). Los métodos de interpolación se pueden aplicar ya sea con escasos conjuntos de datos o densos, aunque cuantos más puntos contenidos en un área, es más probable que la superficie pueda imitar las condiciones verdaderas. Los conjuntos de datos deben ser analizados antes de la interpolación con el fin de determinar la idoneidad de la algoritmo o método. Factores tales como la distribución normal, la Estacionariedad, tendencias y autocorrelación espacial puede influir en la decisión de un algoritmo (Chiles y Delfiner 1999, ESRI 2001, Bohling 2005, Kalkhan 2011).

2.1.2. Distribución normal

En estadística, distribución normal se describe como un caso en el que los valores de los puntos del conjunto de datos se agrupan mayoritariamente cerca del valor medio. Esta distribución normal es también conocido como una curva en forma de campana, cuando se visualizan en un histograma (Berry n.d.). Los histogramas no deben contener valores demasiado distantes de los valores de media y la mediana. Algunos los casos en que los datos no se distribuyen normalmente pueden ser también vistos en un Quantile - normal Cuantil (QQ normal) trazar la existencia de valores atípicos, clusters o agrupación de los datos puede aumentar un problema dentro de algunos de los algoritmos de modelado mediante la creación de la asimetría en la distribución.

El factor de asimetría debe ser cercano a cero, lo que significa la media y la mediana son casi iguales. Si los datos no se distribuyen normalmente, a continuación, aplicar una transformación podría resolver algunos de las irregularidades. Transformaciones modifican el histograma mediante el uso de ya sea empírico o Logaritmo (log) empírica con el fin de reducir la asimetría de los datos (ND Berry, Kalkhan 2011, Lynch y Krause 2014).

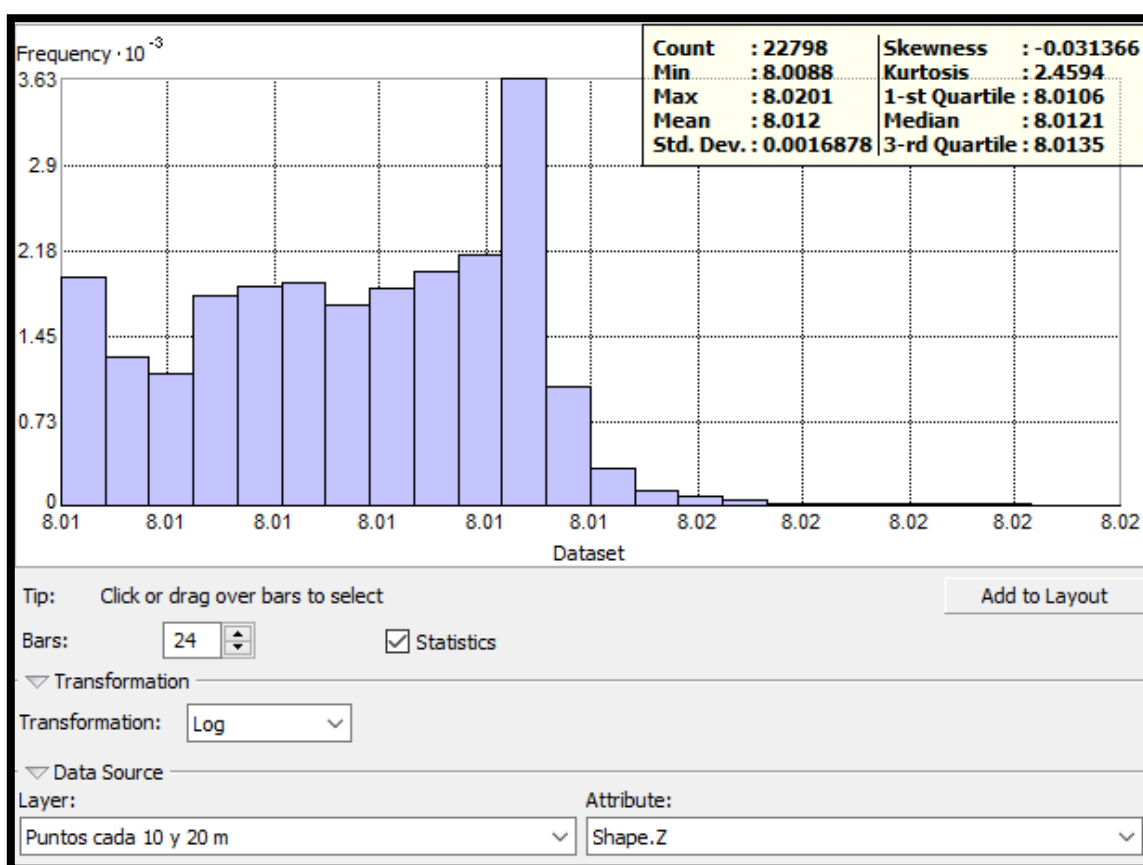


Figura 2. 1 Histograma distribución de frecuencias de elevación embalse “El Tranque”

2.1.3. Estacionariedad de los datos

Si los datos son estacionarios la relación estadística entre dos puntos depende de la distancia entre ellos. Si los datos no son estacionarios, transformaciones podrían estabilizar las variaciones o la utilización de ciertos algoritmos de interpolación, que no requieren estacionariedad. Para Visualizar estacionariedad se utiliza el mapa de Voronoi, simbolizado por la entropía o la desviación estándar. polígonos de entropía se clasifican utilizando cinco

clases en base a una agrupación natural de los valores de datos (cuantiles inteligentes). El valor asignado a un polígono es la entropía que se calculan a partir del polígono y sus vecinos- donde la entropía es la proporción de polígonos que están asignados a cada clase (Kalkhan 2011, Lynch y Krause 2014).

El empleo de técnicas geoestadísticas requiere de estacionaridad de segundo orden, es decir, al menos la varianza debe ser igual en las diferentes zonas del área de estudio. La falta de estacionariedad puede deberse bien a la existencia de anomalías en el espacio, bien a la existencia de una tendencia o gradiente espacial cuya dimensión es mayor que el área de estudio. La estacionariedad puede ser un problema a la hora de la interpolación de puntos en el espacio, pero no justifica el abandono de la geoestadística a favor de otras técnicas de interpolación (como la técnica del inverso de la distancia) ya que son igualmente sensibles a la falta de estacionariedad (Isaaks y Srivastava, 1989)

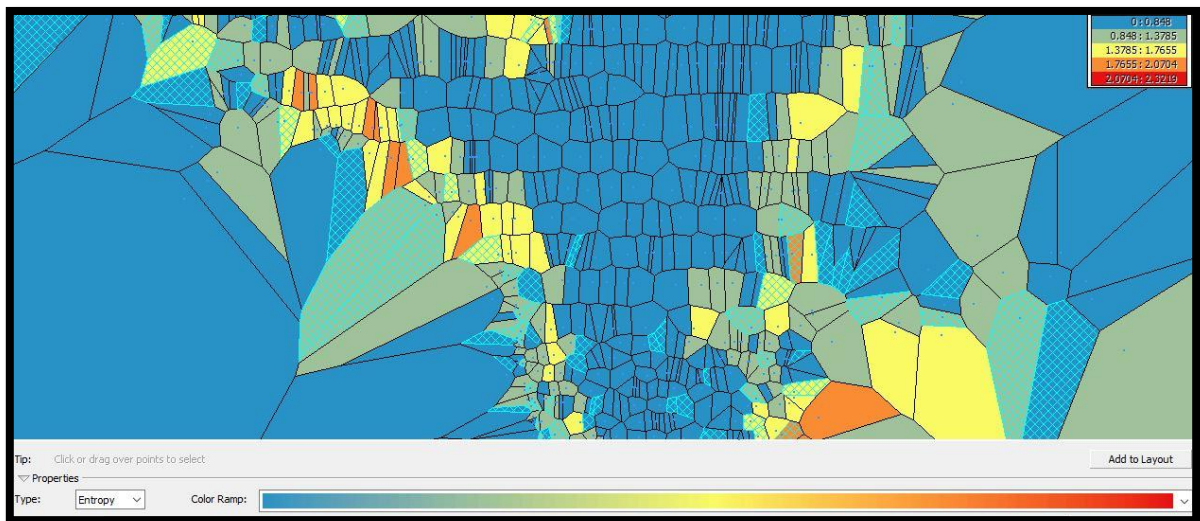


Figura 2. 2 Mapa de Voronoi embalse “El Tranque” entropía de los datos

2.1.4. Tendencias

La identificación de tendencias contenidas en un conjunto de datos es esencial para entender los cambios de los valores de los datos a través de la zona de estudio. El análisis de tendencias dentro de ArcGIS puede ser completado utilizando herramienta de exploración espacial Análisis de Datos (ESDA) se puede utilizar para encontrar las tendencias globales en el conjunto de datos. Tendencias en la X, Y, o eje Z pueden ser visualizados por manipular el

gráfico tridimensional (3D). Si en la pantalla de datos de tendencias de signos, hay podría haber problemas con el conjunto de datos, pero que podrían ser indistinguible de auto correlación y anisotropía. herramientas de eliminación de tendencia o un método de interpolación determinista (Por ejemplo, polinómica global o local), puede reducir esta influencia global tendencia.

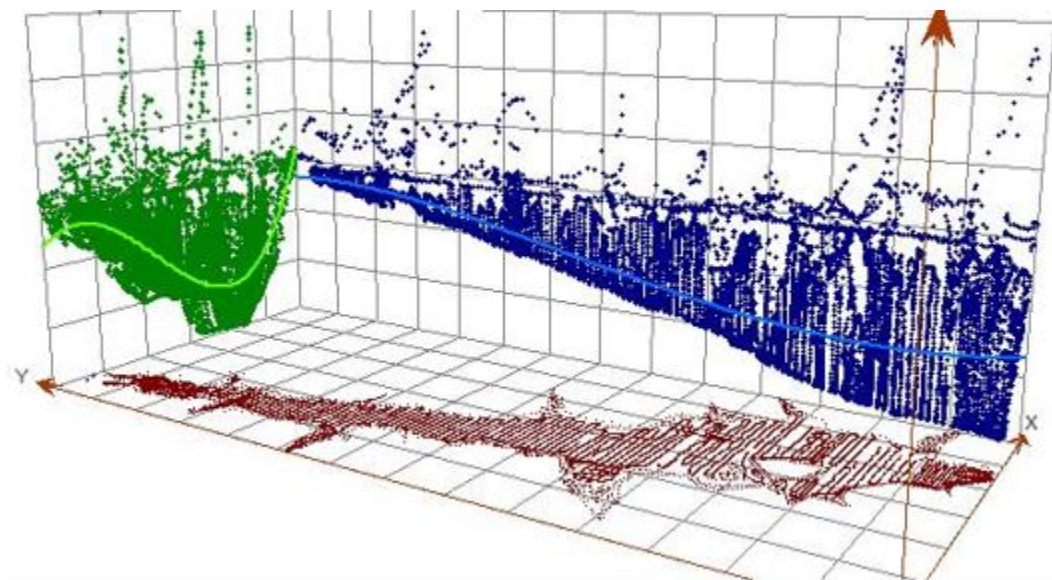


Figura 2. 3 análisis de tendencias de datos embalse “El Tranque”

2.1.5. Autocorrelación espacial

Espacial de autocorrelación, también conocida como primera ley de la geografía, es visualizado mediante la herramienta de covarianza nube semivariogram / en ArcMap. Esta herramienta examina las características locales de autocorrelación espacial contenidos en un conjunto de datos. En el semivariograma el valor es la diferencia al cuadrado entre los valores de dos puntos de datos que constituyen un par que representa el eje y, y la distancia entre los dos puntos del eje x. Con el semivariograma la variación de los valores debe aumentar con la distancia. Los gráficos Covarianza ilustra una versión a escala de correlación por que ilustra la correlación entre la dirección y la distancia de los pares de valores y sus ubicaciones, donde el eje Y está representando la similitud de los valores y el eje X representa la distancia relación. Con el gráfico de covarianza, la función debe disminuir con la distancia.

Estos dos gráficos se utilizan en la comprensión de la relación de los puntos tienen dentro de su espacio geográfico (Kalkhan 2011, Lynch y Krause 2014). Para este análisis, la utilización de los dos conjuntos de datos recolectados a través de dos métodos diferentes y por lo tanto dos diferentes normas de precisión, los requisitos de procesamiento, y las limitaciones espaciales. La exactitud se asume desde la producción de datos. Si el método de interpolación seleccionado no hace generalmente conforme a los parámetros del modelo de interpolación, los datos no serán lo más preciso. Al analizar los diversos conjuntos de datos de puntos, la identificación de las tendencias espaciales, la correlación espacial, la distribución de puntos de los valores y ubicaciones, y la Estacionariedad del conjunto de datos debe ser tomado en consideración para su uso con los diversos métodos de interpolación.

2.2. Características del fenómeno en estudio

2.2.1. Superficie terrestre

La superficie terrestre es un fenómeno continuo que puede ser medido y descrito en cualquier punto sobre la misma con base en un conjunto de variables (Longley et al., 2005). Entre dichas variables, la elevación es una de las variables más estudiadas, ya que es utilizada en un amplio rango de aplicaciones científicas y civiles (INEGI, 2002; Fisher y Tate, 2006), y se define como la distancia vertical desde un punto sobre la superficie terrestre hasta un nivel de referencia arbitrario, como puede ser el nivel medio del mar (definido como datum vertical), (Anderson y Mikhail, 1998).

Debido a la naturaleza de la superficie terrestre, la elevación es continua y debe ser representada digitalmente como tal. Desafortunadamente, debido a limitantes técnicas y económicas, no es factible medir y almacenar la elevación de cada punto sobre la superficie terrestre (Kumler, 1994). Aun considerando nuevas tecnologías como Lidar (Light Detection and Ranging), que brinda la posibilidad de medir la elevación de un considerable número de puntos sobre el terreno en un corto periodo de tiempo, no es viable representar la superficie terrestre utilizando todas las muestras de elevación (puntos sobre el terreno) disponibles y es necesario reducir el volumen de datos (Anderson et al., 2006). Por tanto, la elevación del terreno se representa digitalmente como una superficie continua mediante Modelos Digitales de Elevación (MDE) generados con base en un número finito de muestras de elevación (Li et al., 2005).

2.2.2. Embalses de montaña

Todos los grandes ríos del mundo tienen sus cabeceras en tierras montañosas y más de la mitad de la humanidad depende del agua dulce que se acumula en las zonas montañosas. Aunque constituyen una proporción relativamente pequeña de cuencas fluviales, la mayor parte del caudal se origina en las montañas, y esta proporción depende de las estaciones. Estos «depósitos de agua» son esenciales para el sustento de la humanidad. Al crecer la demanda, aumentan las posibilidades de conflicto por el uso del agua de las montañas. La gestión cuidadosa de los recursos hídricos debe ser, por lo tanto, una prioridad absoluta en un mundo que está destinado a una crisis de agua en el próximo siglo. Muchas razones impulsan a centrar la atención en las montañas. Las más importantes son:

- **Abundancia de las precipitaciones.** Las montañas forman barreras en la circulación de las masas de aire. Al tener que elevarse, el aire se enfría, lo que da lugar a las precipitaciones. En las regiones semiáridas y áridas, sólo las tierras altas tienen pluviosidad suficiente para producir escorrentía y recargar las capas de agua subterránea.

- **Almacenamiento y distribución del agua en las tierras bajas.** Las aguas captadas a altitudes elevadas fluyen por gravedad por la red fluvial o las faldas acuíferas subterráneas hacia las tierras bajas, donde hay una fuerte demanda de los centros urbanos, la agricultura y la industria.

En las zonas húmedas del mundo, la proporción de agua generada en las montañas puede llegar hasta el 60 por ciento del total de agua dulce disponible en la cuenca, mientras que en las zonas semiáridas y áridas esta proporción suele superar el 90 por ciento.

- **El agua como sustentadora de vida.** El suministro de agua limpia y potable es fundamental para la existencia y la salud humanas. Desde 1940, las extracciones mundiales de agua dulce de todas las fuentes (es decir, el uso de agua de superficie o subterráneas) han aumentado en más del cuádruple. El 70 por ciento del agua se utiliza para riego. La relación entre el agua de las montañas y la producción mundial de alimentos es evidente, en particular en los climas áridos y semiáridos de las zonas tropicales y subtropicales donde se encuentran la mayoría de los países en desarrollo y más de la mitad de la población mundial. Además, el agua almacenada en los lagos y embalses de las montañas tiene un valor económico

adicional como fuente potencial de energía hidroeléctrica. El agua dulce de las montañas mantiene también muchos hábitats naturales, en tierras altas o bajas, contribuyendo así a la conservación de la biodiversidad.(H. Liniger y R. Weingartner, 1998).

La vigilancia de los recursos naturales y de su uso y la evaluación de los efectos causados por los cambios en el uso de las tierras altas sobre la disponibilidad y la calidad del agua en los valles son las primeras medidas que se precisan para una buena ordenación (Liniger y Gichuki, 1994).

2.2.3. Modelo Digital de Terreno (M.D.T.)

Un MDT es una estructura de celdas numéricas, que almacenan datos espaciales de una variable cuantitativa y continúa tales como: Temperatura, presión o altitud. En el caso que represente a la última variable mencionada, se denomina MDE.

Un MDT a partir de coordenadas tridimensionales representa numéricamente las características del terreno. Tanto el MDT como el MDE están basados en el modelo de datos raster. (Priego y Porres, 2002).

La forma más simple de ejemplificar lo anterior es tomando como base las curvas de nivel, que son el insumo principal. A través de estas curvas de nivel se puede representar el relieve mediante un MDT, que hoy es la forma de representarlo en formato digital como lo fue anteriormente con los mapas topográficos.



Figura 2. 4 MDT en formato TIN con curvas de nivel en formato vectorial

Fuente: (Anaya, 2008)

2.2.4. Modelo Digital de Elevación (M.D.E.)

Las muestras de elevación o puntos sobre el terreno no representan la superficie terrestre de manera continua por sí mismas, sin importar si su distribución espacial es regular o irregular (Kumler,1994), y, en consecuencia, es necesario utilizar un formato que permita modelar la elevación como una superficie continua (Burrough y McDonell, 1998). De los diferentes formatos disponibles (curvas de nivel, redes irregulares de triángulos (TIN), celdas (Li et al., 2005)), el formato de celdas, o raster, es el más común entre los usuarios y es normalmente conocido como Modelo Digital de Elevación (MDE), (Fisher y Tate, 2006). Sin embargo, las curvas de nivel y las redes irregulares de triángulos pueden ser también consideradas como MDE si se utiliza la definición tradicional de MDE de Miller y Laflamme (1958): “representación numérica digital de la superficie terrestre”. En este proyecto se utiliza el término MDE para referirse únicamente al formato de celdas, como es común en la literatura científica.

Los MDE en formato de celdas representan el terreno como una superficie continua mediante la definición de una cuadrícula o malla regular sobre el área en estudio y el almacenamiento

de un valor de elevación para cada celda de la cuadrícula (Burrough y McDonell, 1998). La definición de la cuadrícula depende directamente de la distribución espacial de las muestras de elevación. Si los puntos sobre el terreno disponibles se distribuyen espacialmente en forma de matriz, entonces es recomendable definir la malla con base en las muestras de elevación disponibles y almacenar la elevación directamente en cada celda. En este caso el tamaño de celda o resolución espacial está determinado por la distancia entre cada muestra de elevación (Li et al., 2005).

En caso contrario, si las muestras de elevación se distribuyen irregularmente sobre el área de estudio, es necesario definir el tamaño de celda de manera arbitraria, empírica (Hutchinson, 1996) o estadística (Hengl, 2006), y seleccionar un método de interpolación para generar un valor de elevación para cada celda de la malla tomando como entrada los puntos sobre el terreno disponibles (Li et al., 2005). Desafortunadamente, diferentes métodos de interpolación calculan diferentes valores para una misma celda aun utilizando las mismas muestras de elevación como entrada (Lloyd y Atkinson, 2002), por lo que la selección del algoritmo de interpolación puede ser considerada como un factor que introduce errores en los MDE.

Un MDE es considerado una representación matemática de los valores de las alturas del terreno, se considera modelo de elevación porque a través de algoritmos matemáticos, los datos altimétricos pueden modelarse automáticamente mediante un ordenador, lo cual permite el estudio de análisis de la superficie terrestre de forma tridimensional.

La palabra elevación, enfatiza el concepto de medición de altura con respecto a un datum y en la generación de valores absolutos de altitud por parte del modelo.

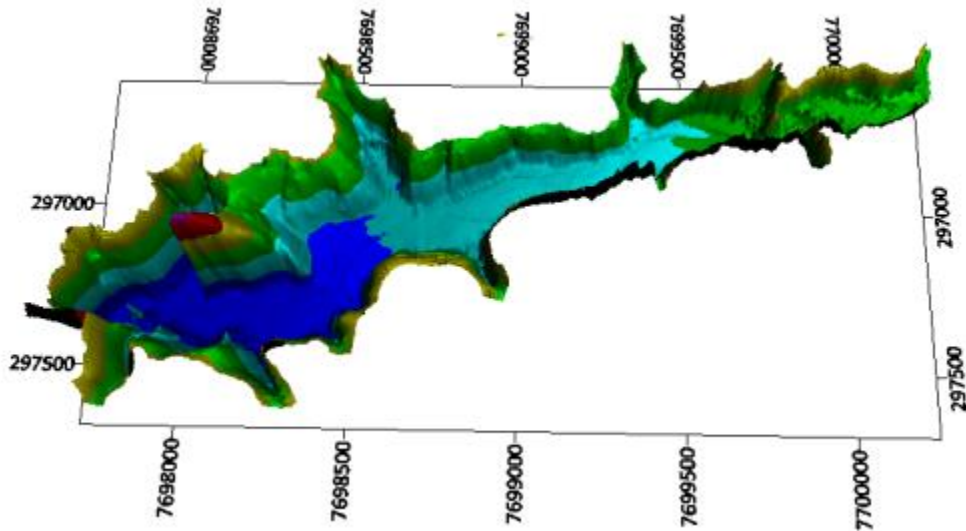


Figura 2. 5 Modelo Digital de Elevación

Resumidamente, es una representación digital de las elevaciones que presenta la superficie topográfica (generalmente sobre el nivel del mar). Aquí se encuentran excluidos la vegetación y las construcciones antropogénicas.

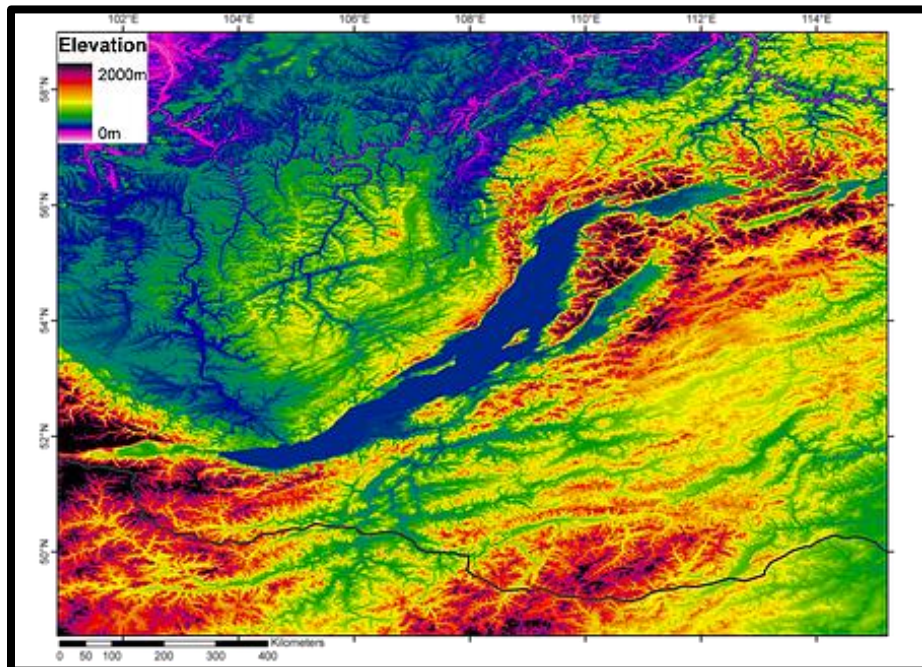


Figura 2. 6 representación del Modelo Digital de Elevación

Fuente: (Anaya, 2008)

2.2.5. Aplicación de los modelos digitales de elevación (MDE's)

La aplicación de los MDE se extiende a campos tan diversos como la geografía, la hidrología, la ecología, los estudios de impacto ambiental y la cartografía en general.

- ✓ Generación de curvas de nivel.
- ✓ Creación de mapas de relieve.
- ✓ Planificación de vuelo en tres dimensiones.
- ✓ Rectificación geométrica de fotografías aéreas o de imágenes satelitales.
- ✓ Correcciones topográficas.
- ✓ Proyectos de grandes obras de ingeniería.
- ✓ Trazados de perfiles topográficos.
- ✓ Cálculos de volúmenes.
- ✓ Análisis de riesgos ambientales.

A través de los Modelos Digitales de Elevación (MDE) se generan las cuencas hidrográficas. (Ruíz y Torrez, 2015)

2.3. Métodos de Interpolación

2.3.1. Métodos de Interpolación determinísticos

La interpolación espacial es “un procedimiento que permite calcular el valor de una variable en una posición del espacio, conociendo los valores de esa variable en otras posiciones del espacio” (Bosque, 2000). Los diferentes métodos de interpolación espacial se clasifican en dos grandes categorías: determinísticos y probabilísticos (Maune et al., 2001). Los métodos de interpolación determinísticos calculan un valor para celda de la cuadrícula usando únicamente las propiedades físicas de las muestras de elevación (Goovaerts, 1997).

2.3.2. Método IDW

El método de interpolación determinístico más comúnmente empleado es la ponderación en función inversa de la distancia o IDW (Inverse Distance Weighting; Wise, 2000). IDW se encuentra disponible en la mayoría de los programas de sistemas

de información geográfica (SIG; Pérez y Mas, 2009) y su definición formal es (O'Sullivan y Unwin, 003:228):

$$\hat{e}_c = \frac{1}{n} \sum_{m=1}^n p_{mc} e_m$$

Ecuación 1

donde \hat{e}_c es la elevación calculada para la celda c utilizando n muestras de elevación; e_m es el valor de las muestras de elevación; y p_{mc} es el peso (o influencia) entre 0 y 1 asignado a cada muestra de elevación, calculado como:

$$p_{mc} = \frac{1}{d_{mc}}$$

Ecuación 2

donde d_{mc} es la distancia euclidiana entre la muestra de elevación m y el centro de la celda c de la malla. Opcionalmente, con el fin de mejorar resultados, es común utilizar un exponente a para disminuir la influencia de las muestras conforme se incrementa la distancia (Bosque, 2000).

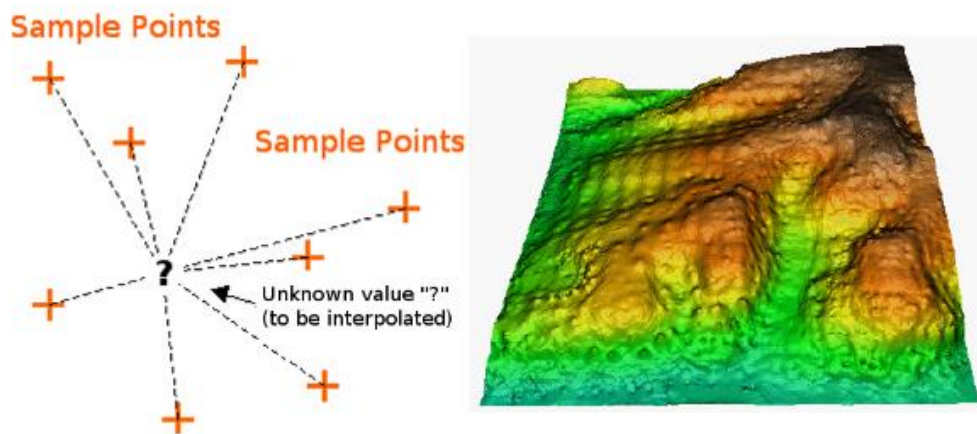


Figura 2. 7 Interpolación Distancia Inversa Ponderada basada en la distancia del punto de muestra ponderada (izquierda). Superficie IDW interpolada de puntos vectoriales de elevación(derecha).

Fuente: Mitas, L., Mitasova, H. (1999).

2.3.3. Método TIN

El modelado de redes irregulares triangulares (TIN) es el método más recomendado para modelar características, como las profundidades de valles y embalses (Felgueiras y Goodchild 1995; Kidner y col. 2000), y el método de triangulación de Delaunay se usa comúnmente en la generación de un TIN. El método de triangulación de Delaunay forma triángulos a partir de todos los puntos dentro de un polígono. Esto requiere que un círculo formado por los tres nodos del triángulo no tenga otro punto, por lo que los puntos están conectados a los puntos. vecinos más cercanos, conservando todos los datos recogidos. Después de eso, se interpolan las curvas de nivel (Tsai 1993; Ferrari 2006; Ferrari y Collins 2006).

En la generación del modelo TIN por el método de triangulación de Delaunay, se generan curvas de nivel artificialmente curvas que se extienden hacia el interior del yacimiento (Furnans y Austin 2008), creando características irregulares cerca del borde del reservorio, expresadas por grandes superficies planas y discontinuidades poco profundas que ocurren por interpolación de algunos datos de la línea de costa del embalse con los datos de RL, lo que provoca una pérdida de volumen en el cálculo final.

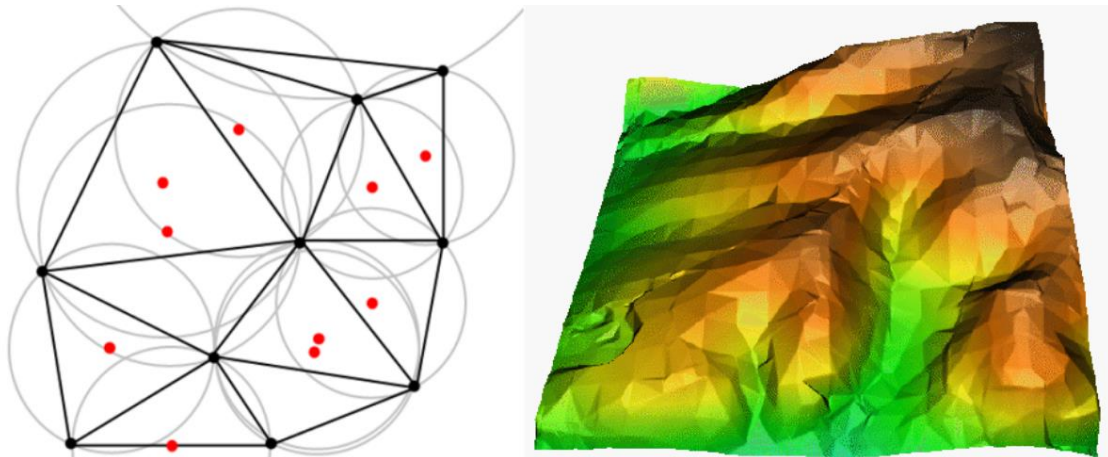


Figura 2. 8 Triangulación de Delaunay con circunferencias alrededor de los datos de muestra rojos. La superficie TIN interpolada resultante creada a partir de puntos de vector de elevación se muestra a la derecha.

Fuente: Mitas, L., Mitsova, H. (1999).

2.3.4. Métodos de Interpolación geo estadísticos

Los métodos de interpolación probabilísticos, en contraste con los métodos determinísticos, infieren una gran cantidad de información utilizando las muestras disponibles (Maune et al., 2001). Los métodos de interpolación probabilísticos más comunes son los métodos geo estadísticos derivados de la forma genérica de Kriging (Goovaerts, 1997:126)

2.3.5. Método Kriging

Este método de interpolación define la altitud como una variable regionalizada, suponiendo que funciones de correlación espacial pueden explicar en su totalidad o parcialmente la variación espacial de esta, que es la variable estimada que se va a representar a través de funciones homogéneas que se deducen de toda el área de análisis, es decir, los valores próximos de z (Giraldo-Henao, 2003).

Estas funciones homogéneas son derivadas teniendo en cuenta la distancia d que existe entre los datos medidos a través de la semivarianza, analizando así la correlación espacial entre ellos en función de d y estimando una media ponderada:

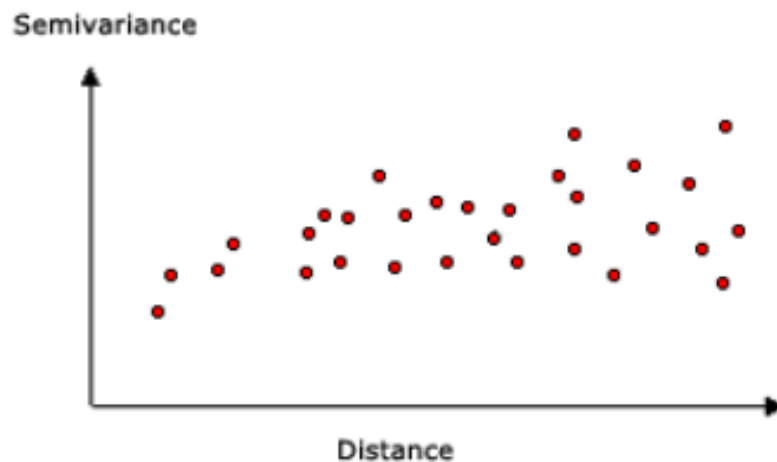


Figura 2. 9 semivariograma

Fuente: arcgis.com

$$\widehat{Z}_{x,y} = \sum_{i=1}^k \lambda_i \cdot Z_j$$

Ecuación 3

Donde un conjunto de k datos se sitúan en torno del punto interpolado y λ representa los pesos asignados a cada uno de ellos. La ventaja del método es el criterio de asignar a cada dato su valor ponderado junto a la asignación de un valor de error a cada punto interpolado, lo que permite aplicar criterios estadísticos brindando una buena definición del área de análisis.

Los supuestos estimados en este tipo de métodos suponen que la altitud en un punto esté determinada por los valores de los puntos más próximos distribuidos a distancias variables, y que en función de ellas habrá una mayor o menor inferencia.

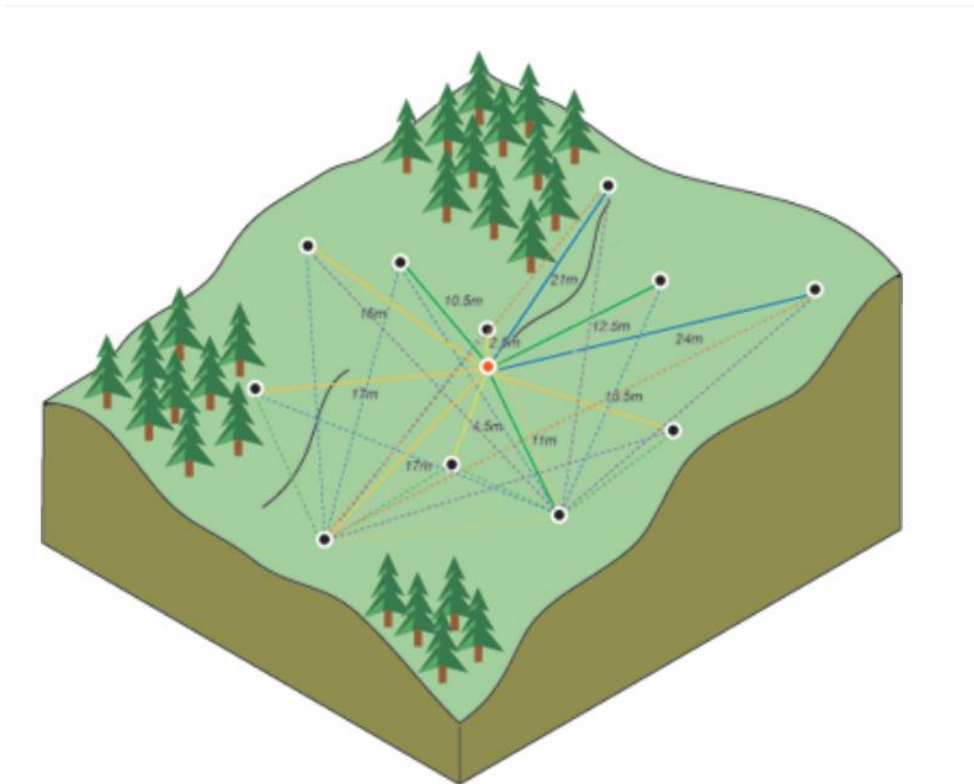


Figura 2. 10 cálculo de la diferencia cuadrada entre las ubicaciones asociadas

Fuente: arctgis.com

2.3.6. Errores por métodos de interpolación

Los errores introducidos por el método de interpolación seleccionado pueden ser aleatorios, debido por ejemplo al redondeo numérico (USGS, 1997); sistemáticos, manifestados como patrones en los MDE (Wise, 2000); o inclusive garrafales (blunders), causados por alta variabilidad del terreno y la aplicación de métodos de interpolación sin restricciones (Smith et al., 2005). Los errores sistemáticos y garrafales pueden ser eliminados o reducidos mediante la aplicación de filtros estadísticos apropiados (López, 2000), por lo cual no es común observarlos en MDE comerciales u oficiales (Fisher y Tate, 2006). Sin embargo, los errores aleatorios no pueden ser filtrados, debido al poco entendimiento de su naturaleza y localización (Li et al., 2005; Pérez y Mas, 2009), y en algunos casos la aplicación de filtros para la reducción/eliminación de errores sistemáticos y garrafales resulta en la adición de nuevos errores aleatorios en los MDE (Wise, 2000).

Estudios previos han sugerido que los principales factores que influyen en la naturaleza y magnitud de los errores introducidos en los MDE por el método de interpolación son: a) la cantidad de información (ej. número de muestras, correlación espacial) empleada por el método de interpolación para calcular la elevación en cada celda de la malla (Lloyd y Atkinson, 2006) y b) los rasgos topográficos del área geográfica representada por el MDE (Fisher y Tate, 2006). No obstante, no existe un consenso en cuanto a ninguno de estos dos factores. Por ejemplo, para la generación de MDE en áreas con relieve moderado, Lloyd y Atkinson (2002), Reuter et al. (2007) y Erdogan (2009) concluyeron que métodos probabilísticos que emplean gran cantidad de información para calcular la elevación de cada celda son más convenientes que métodos determinísticos. Sin embargo, Bishop y McBratney (2002), Su y Bork (2006) y Bater y Coops (2009) sugieren que métodos determinísticos producen MDE más precisos en áreas de relieve similar cuando una gran cantidad de muestras de elevación es utilizada. En áreas con relieve predominantemente montañoso, los resultados de Reuter et al. (2007) sugieren un mejor desempeño de métodos determinísticos en este tipo de terreno. En contraste, Morillo et al. (2002) no encontraron diferencias significativas en la aplicación de métodos determinísticos y probabilísticos para la generación de MDE y lo atribuyeron a la alta densidad de datos disponibles para su área en estudio. De forma similar, Guo et al. (2010) sugieren que diferentes métodos de interpolación

generan resultados similares, aunque establecen que métodos probabilísticos son más confiables en la generación de MDE en áreas con relieve montañoso. Por lo tanto, es difícil establecer la superioridad de ciertos métodos de interpolación sobre otros y es necesario ampliar el nivel de entendimiento de los mismos con base en nuevos estudios (Fisher y Tate, 2006; Wise, 2011).

2.3.7. Métodos de cálculo de volúmenes y modelamiento

Hasta mediados de la década de 1990 la literatura presenta varios métodos tradicionales desarrollados para el cálculo del volumen de levantamientos batimétricos. Tales métodos se han desarrollado a principios de siglo y la mayor parte de ellas se encuentran por Eakin en 1939, que todavía son citados por muchos autores. Con la evolución de las computadoras, su capacidad de procesamiento y el desarrollo de herramientas de geoprocésamiento, los métodos tradicionales se están sustituyendo por métodos automatizados (Usace, 2002).

El grupo sedimentos Bureau of Reclamation (USBR) utiliza el método TIN para modelar y generar las curvas de nivel del depósito utilizando el software ArcGIS. El modelo TIN genera un conjunto de triángulos no solapados adyacentes de puntos espaciados con coordenadas x, y, z. Este método está diseñado para trabajar con los datos continuos como elevaciones. ArcGIS utiliza el criterio Delaunay TIN para el método de triangulación donde los triángulos se forman desde todos los puntos dentro de un polígono. Esto requiere que un círculo formado por los tres nodos del triángulo contiene ningún otro punto, por lo que los puntos están conectados a los puntos vecinos más cercanos. Este método preserva todos los puntos recogidos. Posteriormente líneas de contorno se interpolan (Ferrari, Collins, 2006).

El TWDB (Texas Water Development Board) también utiliza Arcgis y el método TIN para la generación de MDT y afirma que esta metodología es la más apropiada para representar la superficie del depósito cuando los datos recogidos son perpendiculares al canal sumergido. Sin embargo, un inconveniente del método de triangulación Delaunay es la generación que produce en el MDT, que son generados artificialmente líneas de contorno curvados que se extienden en el depósito (Furnans; Austin, 2008).

Para resolver este problema TWDB (Texas Water Development Board) desarrolló un método, propiedad de la empresa, lo que crea secciones intermedias entre un transecto y otra llamada

La interpolación autónomos similares y después de la introducción de estas nuevas secciones artificiales se genera de nuevo un MDT. En la Figura 4 (A) se pueden ver los bancos en los próximos pasos de margen. Estos no representan la realidad y aparecen en la generación de TIN cuando son transectos muy separados. En la Figura 4 (B) se muestran las secciones intermedias generado artificialmente y la Figura 4 (C) generada por el método MDT La interpolación autónomos similares (Furnans; Austin, 2008).

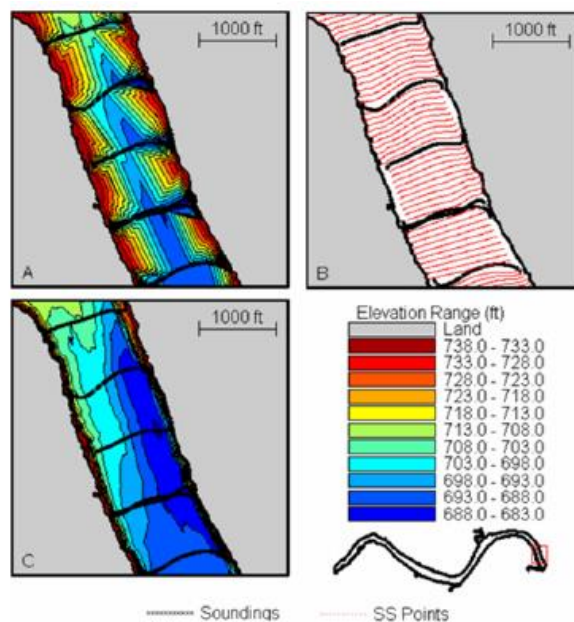


Figura N° 2.11: Auto similares Interpolación - (A) con el TIN generada solamente transectos; (B) Puntos secciones recogidos (negro) y interpolados (rojo); (C) generado con secciones TIN interpolada.

Fuente: (Furnans, Austin, 2008)

TWDB solamente expone los resultados de usar este método, que no muestran procesos o algoritmos derivados de los datos finales.

Los estudios realizados por la batimetría núcleo de CESE-USP también indican las generaciones de características irregulares, cerca del borde del depósito, causadas por las reglas método de triangulación Delaunay. Tales irregularidades se expresan en superficies planas grandes y discontinuidades que ocurren en la interpolación de algunos puntos de los transectos puntos del contorno del depósito (Matos et al., 2011).

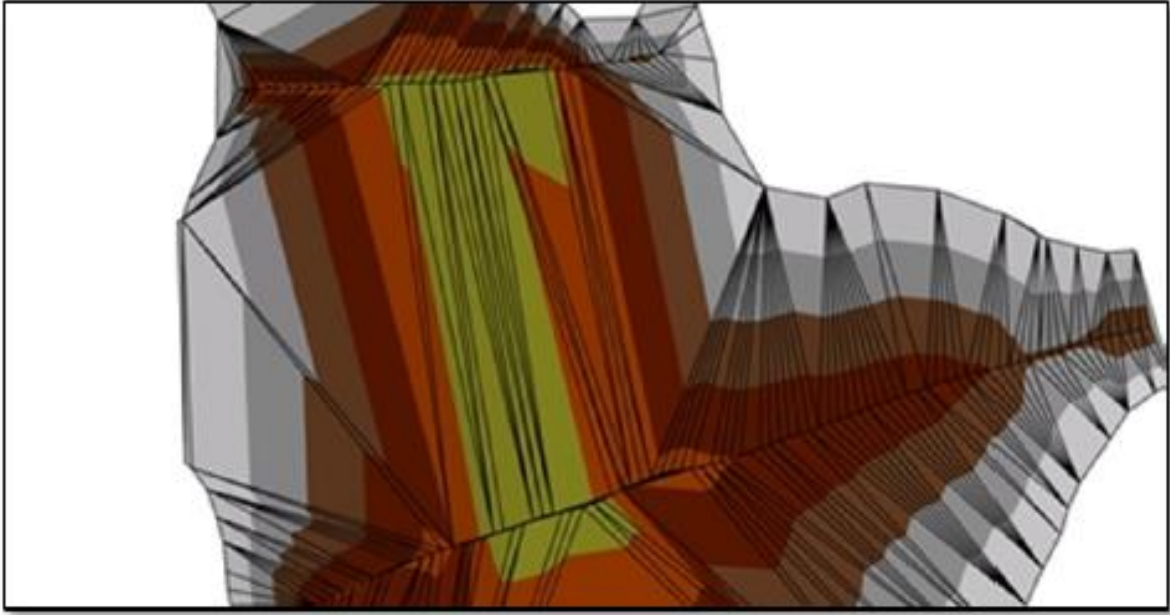


Figura N° 2.12: Detalles de modelo digital del terreno y TIN generados para el depósito Bariri entre dos secciones batimétricas

Fuente: (Matos et al., 2011).

2.3.8. Validación cruzada

La validación cruzada permite evaluar y comparar el rendimiento de los métodos de interpolación dividiendo el conjunto de datos en 2, un conjunto se utiliza para construir el modelo y el otro para validar el modelo (Castellarin et al., 2009; Erdogan 2010; Bailly y col., 2010).

Badea (2013) llega a la conclusión de que no existe una metodología práctica para la exactitud y evaluación, pero la utilización de la Root Mean Square (RMS / RMSE) dato de error de predicción aumenta la comprensión de precisión utilizando la validación cruzada. RMSE es una medida de validación cruzada en la que se realizan las comparaciones de los valores previstos.

2.3.9. Raíz del error medio cuadrático (RMSE)

La raíz del error medio cuadrático (RMSE), se refiere a elevar al cuadrado el error medio de cada dato, lo que provoca que todos los valores se vuelvan positivos y de esta manera la suma no es cero sino generan una suma de cuadrados que si los divides entre el número de datos

generan la varianza que es un indicador de dispersión, pero refleja que tan separados están los datos entre sí o que tan juntos u homogéneos estos son. (Sandoval, 2013) Para su cálculo se utiliza la siguiente fórmula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - y_j)^2}{N}}$$

Ecuación 4

Dónde:

y_i = Es el valor de la cota generada del modelo interpolado

y_j = Es el valor de la cota medida

N = Es el número de pares de valores modelados.

CAPÍTULO III

3. PROCESO GENERACIÓN DE SUPERFICIES TOPOBATIMÉTRICAS

Balestrini (2000) señala que el marco metodológico “es el conjunto de procedimientos a seguir con la finalidad de lograr objetivos de la información válida y con una alta precisión”. En otras palabras, es la estructura sistemática para la recolección, ordenamiento y análisis de la información, que permite la interpretación de los resultados en función del problema que se investiga.

Dado que el objetivo de la investigación será determinar la distancia optima de líneas de sondaje en función del método de interpolación y variando las distancias de separación de líneas se recurre a un diseño experimental que se aplicará de manera transversal, considerando que el tema de investigación tiene un sustento teórico suficiente, se procedió a realizar una investigación de tipo descriptivo para conocer a detalle los parámetros que influyen en la elección de la distancia optima de sondaje.

De acuerdo con Hernandez, Fernandez y Baptista (2003) la investigación experimental “es aquella que se realiza manipulando deliberadamente las variables; para observar las respuestas del fenómeno y después analizarlo”. Estos mismos autores señalan que los diseños de investigación transversales “recolectan datos en un solo momento, en un tiempo único. Su propósito es describir variables y analizar su incidencia en un momento dado”

El presente trabajo será diseñado bajo el planteamiento metodológico del enfoque cuantitativo puesto que éste es el que mejor se adapta a las características y necesidades de la investigación.

El enfoque cuantitativo utiliza la recolección y el análisis de datos para contestar preguntas de investigación y probar hipótesis establecidas previamente, y confía en “la medición numérica, el conteo y frecuentemente en el uso de la estadística para establecer con exactitud patrones de comportamientos en una población (Hernandez, Fernandez & Baptista, 2003)

3.1. Instrumentos de investigación

Los instrumentos empleados en el presente trabajo de investigación que fueron adquiridos por el laboratorio de Hidráulica e Hidrología de la Universidad Autónoma Juan Misael Saracho el año 2016 fueron los siguientes a excepción de los GPS estacionarios:

- Estación total Sokkia Set 5X
- Z-Boat Ecosonda
- 2 equipos GPS de doble frecuencia marca Geofennel

Para la aplicación de estos equipos utilizados para levantamientos batimétricos, se realizó la recopilación de información y herramientas necesarias, las cuales fueron:

- Software Hypack
- Software Odom eChart
- Manual de Usuario Z-Boat
- Manual de Usuario Ecosonda CV 100
- Programa Google Earth Pro
- Software Civil 3D 2018
- Arc Gis 10.5

Estos instrumentos forman parte del desarrollo tecnológico de los métodos para levantamientos batimétricos, los cuales nos permiten realizar mediciones de forma más eficiente, mejorando principalmente el tiempo de medición e incrementando la seguridad del personal encargado de la actividad.

3.2. Pre - proceso: Planificación del levantamiento batimétrico.

3.2.1. Trazado de líneas de sondeo

Obtención de mapas del embalse mediante Google Earth para obtener una referencia aproximada de la forma, tamaño del embalse y sus vías de acceso estas deben coincidir o aproximarse con el nivel actual del embalse al momento del levantamiento.



Figura 3. 1 Imagen satelital embalse “El Tranque”

Para un levantamiento batimétrico con el equipo Z-Boat se traza un recorrido previo con la finalidad de definir los puntos de inicio y fin de las correspondientes trayectorias de las secciones a levantar. La separación entre líneas de sondaje se la realiza en función a la magnitud del embalse, condición climática, tiempo de trabajo proyectado y precisión del proyecto.

Para este proyecto la separación entre líneas de sondaje es de aproximadamente 10-20 m. las cuáles son dibujadas mediante hypack para referencia al momento del sondeo.

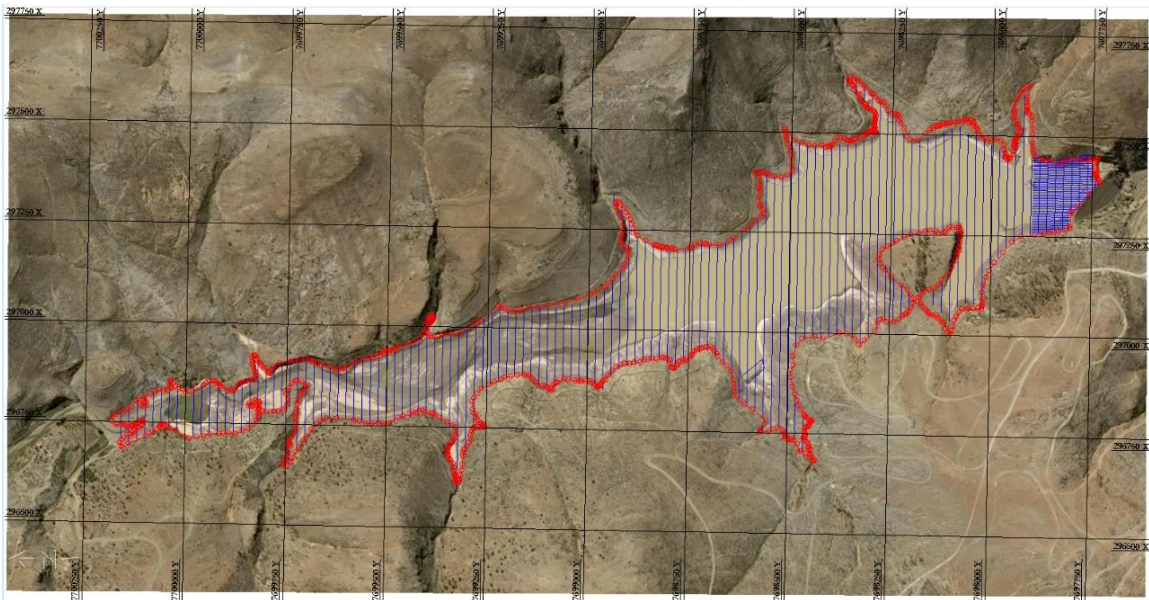


Figura 3. 2 Imagen satelital con líneas planeadas embalse “El Tranque”

3.2.2. Ubicación de la red geodésica en el embalse El Tranque

La presa “El Tranque” cuenta con seis puntos geodésicos colocados y bien monumentados alrededor del embalse los cuales son identificados y ubicados mediante una imagen satelital. Estos puntos nos ayudan con la georreferenciación para el levantamiento batimétrico a realizarse.



Figura 3. 3 Ubicación de pares geodésicos embalse “El Tranque”

3.2.3. Estudio e inspección preliminar del lugar

La inspección tiene como objetivo evaluar en campo lo proyectado en los mapas y ver la accesibilidad a la zona para no tener contratiempos cuando se esté transportando al equipo y personal competente, a la zona de estudio.

El reconocimiento del terreno se realiza con el fin de posicionar físicamente los puntos de inicio y final de los recorridos sobre el vaso de almacenamiento de la presa, con una separación entre transectos de aproximadamente entre 10 - 20 m. En el levantamiento con el equipo Z-Boat no es necesario mucha precisión en esta señalización, ya que el equipo nos brinda directamente las coordenadas de cada punto de sondeo.

3.3. Proceso: Recolección de datos.

3.3.1. Monumentación de pares Geodésicos.

Se realizó la materialización de seis puntos GPS, debidamente encofrados de acuerdo a la normativa emitida en Navstar Global Positioning, con placa de aluminio e inscripción en ellas, las mismas con foto identificables.



Figura 3. 4 Monumentación de pares geodésicos embalse “El Tranque”

3.3.2. Posicionamiento de pares geodésicos.

La observación de los seis puntos GPS se realizó con equipos geodésicos de doble frecuencia, cumpliendo con los tiempos de observación según la longitud de la línea base fueron ajustadas con estaciones IRTF 2014.

Tabla 3. 1 Resumen de coordenadas red geodésica embalse “El Tranque”

COORDENADAS AJUSTADAS CON ITRF 2014						
Punto	ESTE	NORTE	ZONA	ALTURA ELIPSOIDAL	ALTURA ORTOMETRICA	ONDULACION GEOIDAL
LAB-H-28	296843,602	7700120,450	20	3080,597	3041,540	39,057
LAB-H-27	297246,313	7698096,056	20	3078,149	3039,195	38,954
LAB-H-24	296838,540	7699711,616	20	3070,570	3031,528	39,042
LAB-H-26	297237,387	7697791,920	20	3071,677	3032,733	38,944
LAB-H-25	296935,830	7698693,939	20	3065,880	3026,883	38,997
LAB-H-29	297425,722	7698538,539	20	3068,812	3029,853	38,959

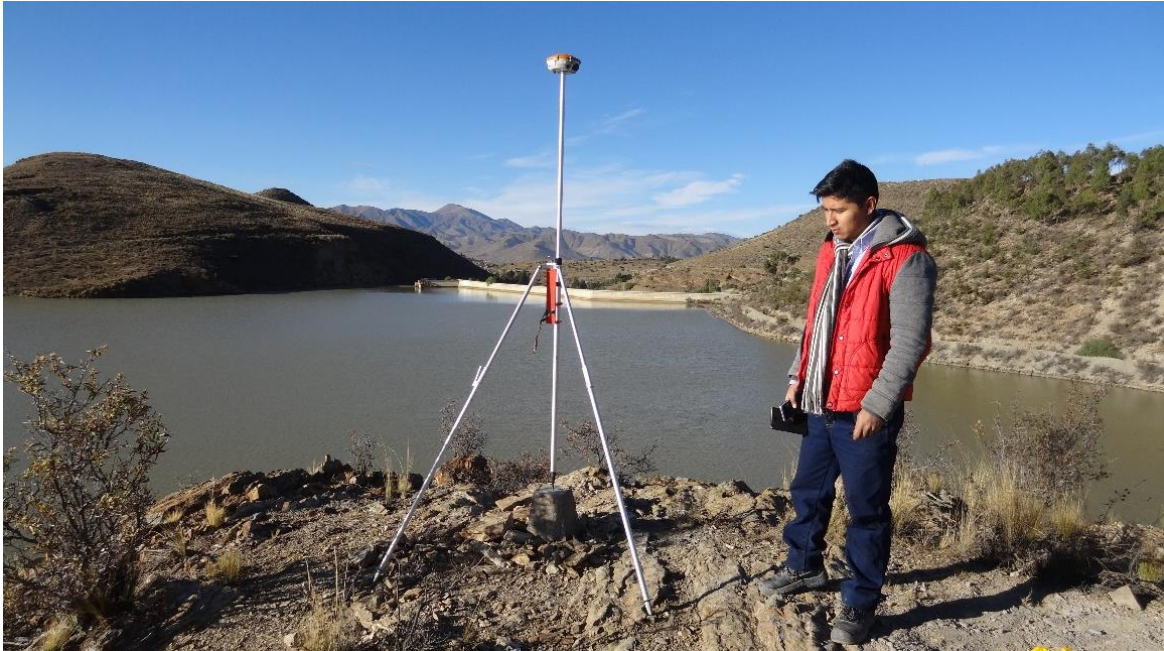


Figura 3. 5 Posicionamiento de pares geodésicos embalse “El Tranque”

3.3.3. Levantamiento Batimétrico.

Para el proceso del levantamiento batimétrico se utilizó el equipo Z-Boat, compuesto por una ecosonda Echotrac CV100 de doble frecuencia con una precisión de 0,10 m (Anexo F)



Figura 3. 6 Levantamiento Batimétrico embalse “El Tranque”

Previo al levantamiento batimétrico se realiza la calibración de profundidad de la ecosonda, por medio de la prueba de barra o bar check que consiste en medir y comparar las profundidades medidas por la sonda y compararlas con la distancia patrón, la prueba de plancha deberá presentar un valor medio que no difiera a +/- 15 centímetros, con respecto a lecturas parciales.

Otro parámetro importante a medir es la velocidad del sonido en el agua esto se logra a través de un perfilador de sonido el utilizado en este proyecto fue un mini SVP, con rango de precisión de velocidad del sonido: rango 1375-1900m/s, resolución: 0,001m/s.



Figura 3. 7 Medición de la velocidad de sonido en el embalse “El Tranque”

3.3.4. Levantamiento Topográfico.

A partir de los puntos de control de partida (materializados por los mojones) y empleados también en la batimetría, se estacionó la Estación Total, se hizo origen en el punto de control correspondiente y se procedió a realizar la taquimetría o registro de punto empleado prisma de reflexión portados por el personal de alarifes.

La taquimetría de ribera se realizó con medición de nubes de puntos en los lugares más accidentados, esta nube fue densificada de acuerdo al criterio técnico y considerando la escala del producto final.

De acuerdo a la topografía e inter-visibilidad de la ribera, se generaron puntos de control (X, Y y Z), auxiliares a objeto de cubrir la taquimetría de todo el perímetro del lago y no dejar áreas de sombra o sin levantar (considerando siempre la escala del producto final).



Figura 3. 8 Levantamiento Topográfico embalse “El Tranque”

3.4. Post - proceso: Procesamiento de datos.

3.4.1. Generación de datos y reducción de la muestra total.

Una vez descargados los datos de topo-batimetría y realizada la corrección de errores brutos se procede a generarlos y reducirlos de manera manual para obtener muestras de datos reducidos que simulan una topo-batimetría con un mayor distanciamiento de líneas de sondeaje esto por medio del software Civil 3D.

Para este paso se utilizó como el levantamiento en el embalse el tranque en 2019 que consta de 22801 puntos levantados en total de los cuales se fue reduciendo la distancia de líneas de sondeaje.

Aumentando la distancia de líneas de sondeo de forma manual al eliminar las líneas intermedias se obtuvieron nubes de puntos correspondientes a diferentes distancias de sondeo.

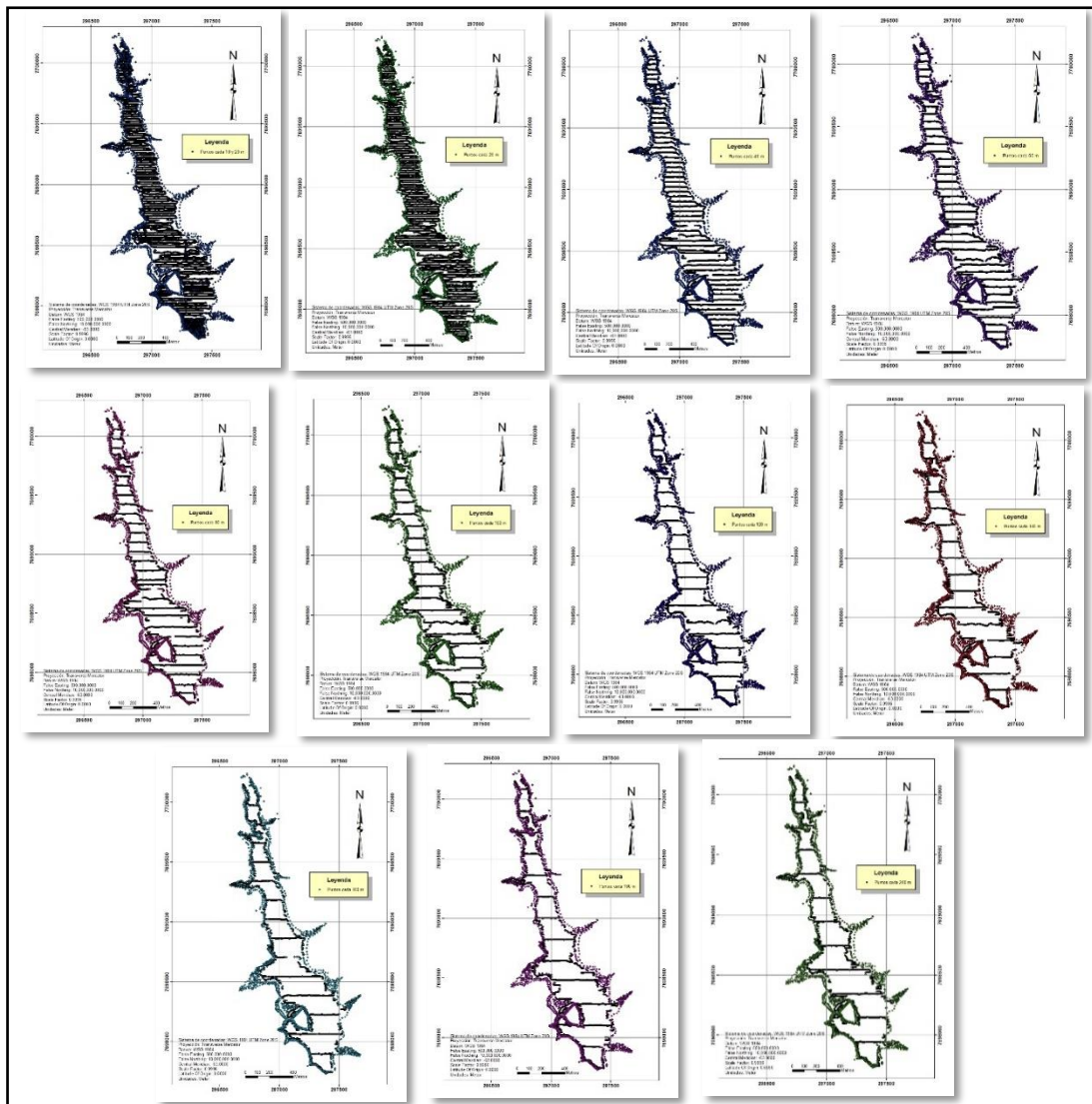


Figura N° 3.9: Nube de todos los puntos cada 10-20, 20, 40, 60, 80, 100, 120, 140, 160, 180 y 200 metros de distancia

3.4.2. Análisis de distribución de frecuencias

Inicialmente y con el objetivo de explorar y comprobar la bondad de los datos de elevación, analizados estadísticamente, normalizadas y promediados; la comprobación realizada fue en el Geostatistical Analyst del software ArcGIS. 10.3, cuyos valores obtenidos en la desviación estándar se encuentran en el anexo C lo cual se encuentra dentro del valor ($\sigma = 0,695$) de acuerdo a J. Martínez-Frutos, P. Martí en su libro “Rev. Internacional de Diseño Óptimo Robusto de modelos Interpolación Kriging 2014 pág. 101 indica “.conjunto de diseños óptimos obtenidos tras 3 actualizaciones del modelo Kriging permite obtener una solución de compromiso entre el diseño adecuado y óptimo a una mínima media ($\sigma = 0,915$) y el diseño de mínima desviación estándar ($\sigma = 0,695$)”.

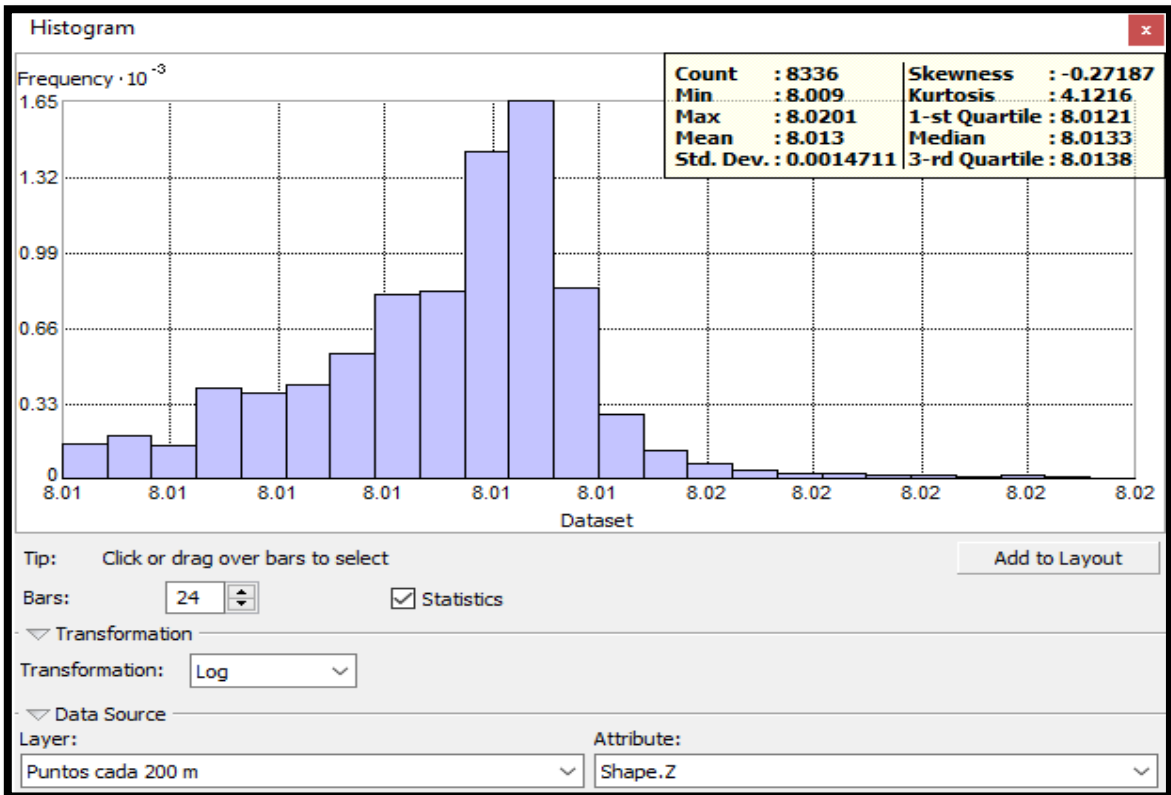


Figura 3. 10 Análisis de distribución de frecuencias de alturas

3.4.3. Análisis de tendencias

Se utiliza para la observación y eliminación de tendencias en el uso de métodos probabilísticos y para tener una comprensión de la tendencia de la distribución espacial de los datos.

En este trabajo se removieron tendencias de segundo grado para la generación de modelos con el método de interpolación Kriging.

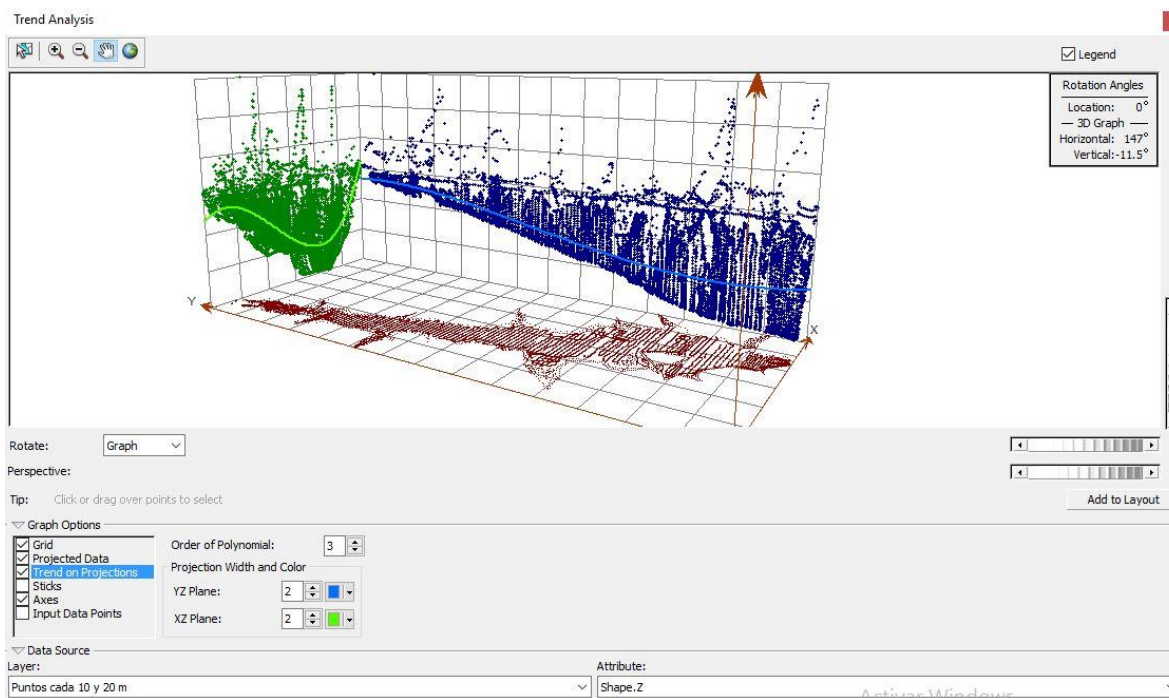


Figura 3. 11 Análisis de Tendencias de distribución espacial

3.4.4. Modelación de superficies batimétricas

Se generaron 33 modelos batimétricos con tres métodos de interpolación 11 modelos con TIN, 11 con IDW y 11 con Kriging.

Para los modelos interpolados por medio de IDW se utilizó una potencia variable resultado de un proceso iterativo que se ajustó por medio de la validación cruzada para la obtención de la superficie más precisa, mismo caso para el ajuste del semivariograma del método kriging, los detalles de ajustes de cada semivariograma se encuentran en anexos. Así como los modelos generados, se realizó la aplicación la transformación logarítmica del kriging,

obteniendo los siguientes resultados que según S. Hara y Tokunaga et al. Es $\pm 0,5$, estos estarían dentro de lo establecido.

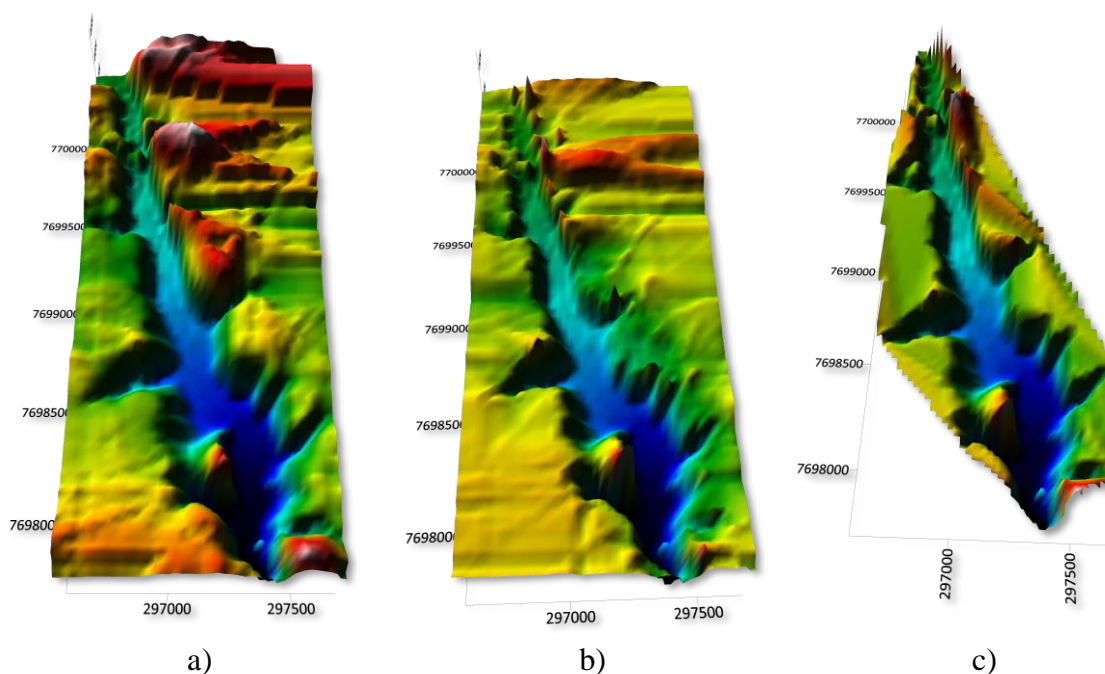


Figura 3. 12 Modelos digitales del embalse “El Tranque” interpolados por diferentes métodos de interpolación a) Kriging b) IDW c) TIN

3.4.5. Estimación de errores

Apartando un conjunto de datos aleatorios distribuidos por toda la superficie batimétrica se realiza la validación cruzada de los modelos generados conociendo su distribución de error además de la ubicación de los mismos esto sirve para observar los sectores en que se distribuye el mismo, como se puede ver en la figura 3.13 donde se puede observar que los errores métricos se ubican en el contorno del embalse en la transición de puntos topográficos y batimétricos.

La mayoría de los errores como se observa en el grafico se encuentran cercanos a cero lo cual indica una buena representatividad del modelo siendo pocos los valores que varían en gran magnitud.

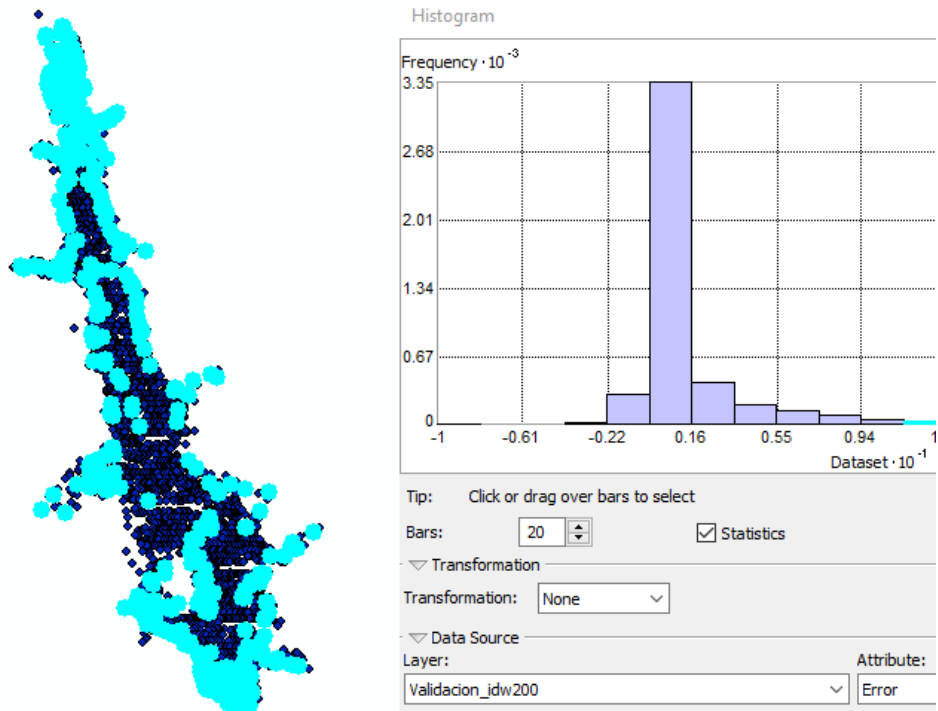


Figura 3. 13 Estimación de errores y ubicación de errores métricos

3.4.6. Cálculo de volúmenes

La estimación de volúmenes se realizó en Arc Gis 10.5 hasta la cota 3023,9 m.s.n.m. correspondiente al nivel del vertedero del embalse El tranque.

Una superficie de ráster se evalúa utilizando la extensión del punto central de cada celda, en lugar de la extensión del área completa de la celda. El análisis resultante reducirá el área de datos del ráster a la mitad de una celda en relación con el área de datos que se muestra para el ráster.



Figura 3. 14 Conversión de MDE a ráster para el cálculo de volumen

3.4.7. Validación cruzada y RMSE

Estimar la distribución del error de cada modelo y del RMSE son fundamentales para validar y estimar la calidad de cada modelo. La figura 3.12 muestra la distribución de frecuencia de errores donde la mayoría de los puntos evaluados presentan errores mayores a -0.4 y menores a 0.8 teniendo una desviación estándar de 0.31 m y un error medio de -0.012435 m.

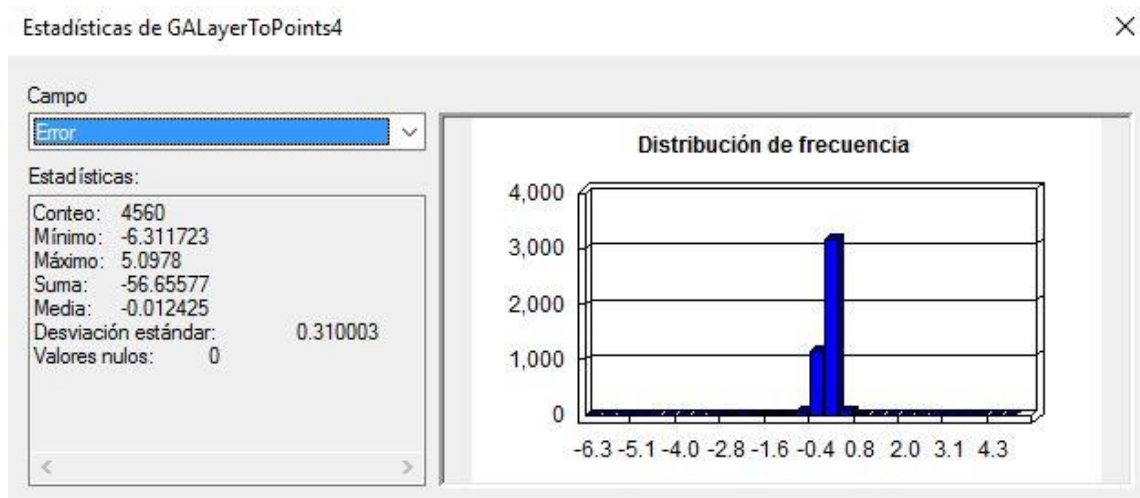


Figura 3. 15 Distribución de frecuencia de error

El proceso de validación cruzada se realizó tanto para la generación de modelos y ajuste de los parámetros de cada método de interpolación así como para validar los modelos.

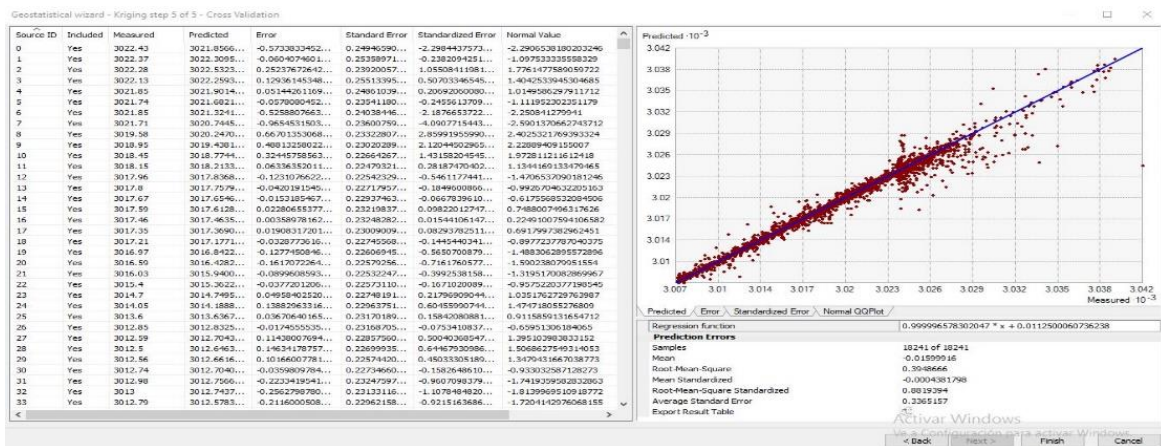


Figura 3. 16 Validación cruzada utilizada para la optimización de modelos

El calculo de la raiz del error medio cuadratico se realizo para todos los modelos generados por los tres metodos de interpolacion por medio de una hoja de calculo Excel a una cantidad de 4560 puntos extraidos que no forman parte de la generacion del modelo, estos son reservados para la validacion del modelo.

Validacion modelo 10-20										
OBJECTID *	Shape *	ID	NORTE	ESTE	Z	F5	Included	Predicted	Error	Standard Error
2894	Punto ZM	142	7698236.31	297273.29	3009.57		Yes	3009.628755	0.058755	0.228074
4084	Punto ZM	202	7699965.82	296800.2	3021.35		Yes	3021.408503	0.058503	0.218154
189	Punto ZM	990	7697795.03	297400.35	3009.35		Yes	3009.408356	0.058356	0.2367
3308	Punto ZM	163	7698001.88	297175.98	3019.89		Yes	3019.948349	0.058349	0.228928
1863	Punto ZM	912	7698722.53	297199.17	3016.44		Yes	3016.498243	0.058243	0.225348
4169	Punto ZM	207	7699897.11	296799.71	3021.17		Yes	3021.228168	0.058168	0.221822
1013	Punto ZM	499	7699306.32	296699.81	3022.01		Yes	3022.068102	0.058102	0.226118
1278	Punto ZM	636	7699059.82	297038.63	3016.12		Yes	3016.178077	0.058077	0.226752
3485	Punto ZM	172	7698543.72	296925.39	3017.47		Yes	3017.527766	0.057766	0.22176
1353	Punto ZM	671	7699019.8	296895.43	3018.93		Yes	3018.987714	0.057714	0.217646
3517	Punto ZM	174	7698526.57	296938.21	3020.19		Yes	3020.247674	0.057674	0.235408
3866	Punto ZM	191	7700065.07	296740.03	3023.68		Yes	3023.737563	0.057563	0.286262
689	Punto ZM	348	7699541.47	296804.54	3019.78		Yes	3019.837511	0.057511	0.232047
3745	Punto ZM	186	7699714.07	296762.5	3021.73		Yes	3021.787493	0.057493	0.224812
4075	Punto ZM	202	7699967.62	296746.03	3021.84		Yes	3021.897377	0.057377	0.229907
3464	Punto ZM	171	7698542.55	296961.76	3019.35		Yes	3019.407227	0.057227	0.21711
3673	Punto ZM	183	7699766.06	296837.35	3021.42		Yes	3021.477201	0.057201	0.244105
3878	Punto ZM	192	7700144.59	296759.63	3022.66		Yes	3022.717118	0.057118	0.236178
3763	Punto ZM	187	7699696.78	296767.85	3020.69		Yes	3020.746745	0.056745	0.240478
4213	Punto ZM	209	7699845.93	296830.52	3020.99		Yes	3021.048397	0.056397	0.224704
3792	Punto ZM	188	7699695.74	296828.81	3021.47		Yes	3021.526311	0.056311	0.230986
4202	Punto ZM	208	7699857.8	296768.91	3022.6		Yes	3022.656244	0.056244	0.226428
2092	Punto ZM	102	7698580.49	297131.9	3012.69		Yes	3012.746207	0.056207	0.225369
127	Punto ZM	666	7697835.62	297390.44	3009.71		Yes	3009.76567	0.05567	0.225887



Figura 3. 16 Cálculo del RSME en ArcGis

CAPÍTULO IV

4. ANÁLISIS DE MÉTODOS DE INTERPOLACIÓN

4.1. Análisis de tamaño de celda

El tamaño de celda juega un papel fundamental al momento de la generación de modelos digitales es tan importante e incluso más que el método de interpolación en este trabajo se utilizaron dos criterios empíricos para la determinación del tamaño de celda a utilizar en la generación de los modelos batimétricos.

4.1.1. Criterio de variación de volumen

Generando diez modelos digitales con el mismo método de interpolación (TIN) y la misma cantidad de muestra (todos los puntos) se generaron modelos variando el tamaño de celda o grilla de manera que esta no influya de manera significativa en el cálculo del volumen el tamaño seleccionado fue de 1 x 1 m.

Distancia de línea de sondeo	Grilla	Volumen	Volumen	Porcentaje
(m)	(m)	(m ³)	(hm ³)	(%)
10 - 20	50 x 50	4817814,85	4,818	95,85
10 - 20	30 x 30	4918390,15	4,918	97,85
10 - 20	25 x 25	4940947,11	4,941	98,30
10 - 20	20 x 20	4973729,20	4,974	98,95
10 - 20	15 x 15	498773,31	4,988	99,23
10 - 20	10 x 10	5008294,01	5,008	99,64
10 - 20	5 x 5	5021305,02	5,021	99,90
10 - 20	1 x 1	5026150,66	5,026	99,99
10 - 20	0,5 x 0,5	5026345,80	5,026	100,00
10 - 20	0,1 x 0,1	5026424,23	5,026	100,00

Tabla 4. 1 Variación de volúmenes de modelos digitales del embalse “El Tranque” en función del tamaño de grilla.

El tamaño de grilla utilizado para este trabajo fue de 1x1 m. con el cual solo hay una variación de volumen del 0,01 % todos los modelos topo batimétricos se crearon con este tamaño de celda o grilla.

4.1.2. Criterio de velocidad de procesamiento

Al trabajar con cuadrículas de grilla más pequeñas el tiempo de procesamiento aumenta en gran medida llegando incluso a colapsar la memoria de los ordenadores y no generan una mejora significativa en la precisión de los modelos digitales.

4.2. Análisis del método de interpolación TIN (Triangulated Irregular Network)

Asumiendo que el volumen estimado con el total de puntos representa el 100% que se obtendría de un sondeo detallado y considerando una pérdida aceptable del 5% del volumen como límite para la determinación de distancia óptima entre líneas de sondeo para el método TIN podría trabajarse con distancias de hasta 100 metros y se perdería un 2,96 % de volumen. Para una variación menor al 2% se podría trabajar hasta con distancias de 60 metros la cual varía solo 1,79 % con respecto.

Distancia de línea de sondeo		Grilla			Volumen	Volumen	Porcentaje
N°	(m)		(m)		(m ³)	(hm ³)	(%)
1	15	1	x	1	5027676,95	5,028	100,00
2	20	1	x	1	5019933,00	5,020	99,85
3	40	1	x	1	4978081,84	4,978	99,01
4	60	1	x	1	4937760,41	4,938	98,21
5	80	1	x	1	4871383,42	4,871	96,89
6	100	1	x	1	4878709,21	4,879	97,04
7	120	1	x	1	4711839,46	4,712	93,72
8	140	1	x	1	4690451,75	4,690	93,29
9	160	1	x	1	4601619,45	4,602	91,53
10	180	1	x	1	4585499,95	4,585	91,21
11	200	1	x	1	4391240,65	4,391	87,34

Tabla 4. 2 Variación de volúmenes en función de la separación de líneas de sondeo para los modelos digitales generados por el método (TIN).

Para una distancia de 80 metros se evidencia una pérdida de volumen del 3,11 % incluso mayor que para 100 metros. Aun así, no supera la variación del 5 % puesta como límite.

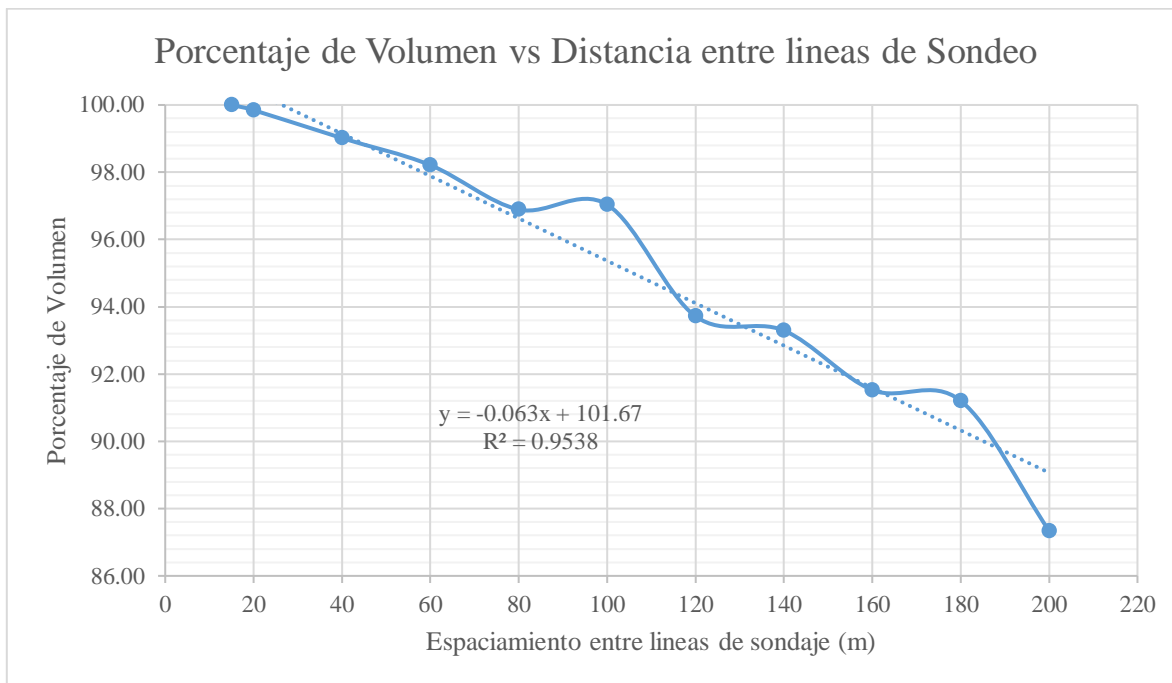


Figura 4. 1 Variación de volúmenes en función de la separación de líneas de sondeo para los modelos digitales generados por el método (TIN).

Como puede verse en la gráfica porcentaje de volumen vs distancia a una distancia de 100 m existe un mayor aumento en el porcentaje de volumen generado, sin embargo, para una distancia de 120 m se pierde de manera significativa el porcentaje de volumen.

4.3. Análisis del método de interpolación IDW (Inverse Distance Weighting)

Asumiendo que el volumen estimado con el total de puntos representa el 100% que se obtendría de un sondeo detallado y considerando una pérdida aceptable del 5% del volumen como límite para la determinación de distancia óptima entre líneas de sondeo para el método IDW podría trabajarse con distancias de hasta 100 metros y se perdería un 4,30 % de volumen. Considerando un límite del 2% la distancia máxima sería de 40 metros con una variación del 1,27%.

Distancia de línea de sondeo		Grilla			Volumen	Volumen	Porcentaje
N°	(m)	(m)			(m³)	(hm³)	(%)
1	15	1	x	1	5085348,03	5,085	100,00
2	20	1	x	1	5101139,54	5,101	100,31
3	40	1	x	1	5020817,46	5,021	98,73
4	60	1	x	1	4964607,24	4,965	97,63
5	80	1	x	1	4882141,43	4,882	96,00
6	100	1	x	1	4866826,19	4,867	95,70
7	120	1	x	1	4677563,61	4,678	91,98
8	140	1	x	1	4650026,55	4,650	91,44
9	160	1	x	1	4517558,65	4,518	88,83
10	180	1	x	1	4522401,35	4,522	88,93
11	200	1	x	1	4317964,56	4,318	84,91

Tabla 4. 3 Variación de volúmenes en función de la separación de líneas de sondeo para los modelos digitales generados por el método (IDW).

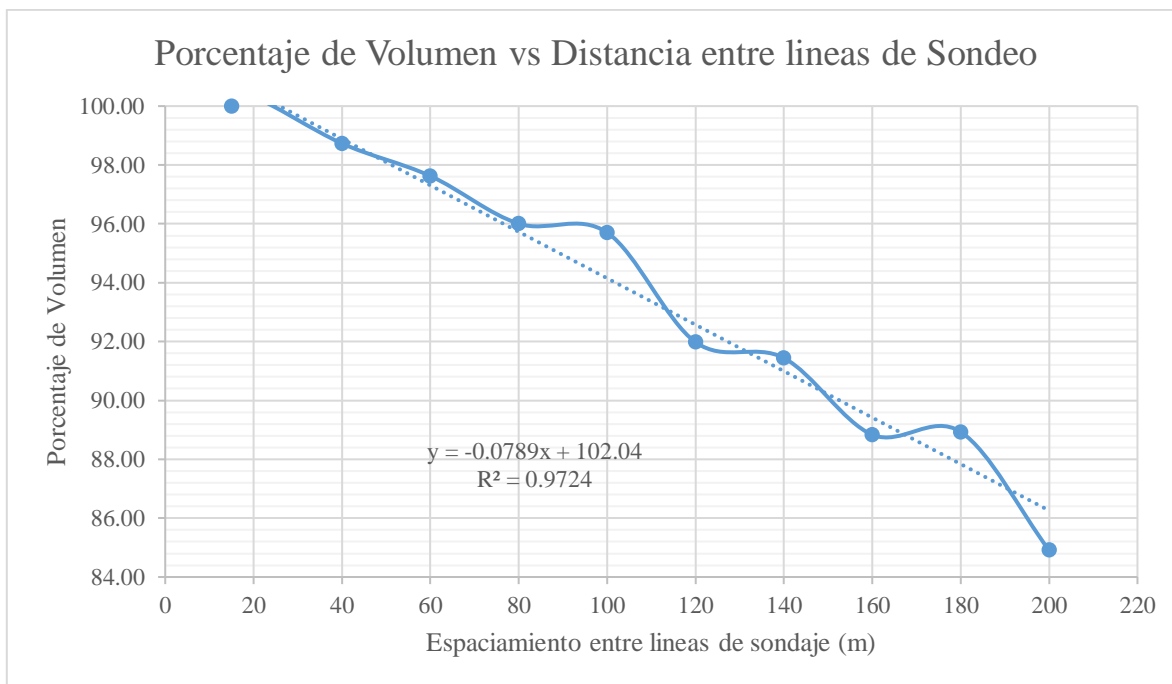


Figura 4. 2 Variación de volúmenes en función de la separación de líneas de sondeo para los modelos digitales generados por el método (IDW).

Se observa un aumento mayor al 100% para el modelo generado con una distancia de 20 m. lo que muestra que la reducción de datos genero un aumento de volumen en este método de interpolación, ese aumento corresponde a una variación del 0,31 %.

4.4. Análisis del método de interpolación Kriging Ordinario (K.O.)

Asumiendo que el volumen estimado con el total de puntos representa el 100% que se obtendría de un sondeo detallado y considerando una perdida aceptable del 5% del volumen como límite para la determinación de distancia optima entre líneas de sondeo para el método Kriging Ordinario podría trabajarse con distancias de hasta 100 metros y se perdería un 3,54 % de volumen. Considerando un límite del 2% la distancia máxima seria de 60 metros con una variación del 1,17%.

Distancia de línea de sondeo		Grilla	Volumen	Volumen	Porcentaje
N°	(m)	(m)	(m ³)	(hm ³)	(%)
1	15	1 x 1	5079537,98	5,080	100,00
2	20	1 x 1	5037975,60	5,038	99,18
3	40	1 x 1	5010350,11	5,010	98,64
4	60	1 x 1	5020334,27	5,020	98,83
5	80	1 x 1	4949459,88	4,949	97,44
6	100	1 x 1	4899686,83	4,900	96,46
7	120	1 x 1	4718007,73	4,718	92,88
8	140	1 x 1	4687057,03	4,687	92,27
9	160	1 x 1	4630504,73	4,631	91,16
10	180	1 x 1	4573429,56	4,573	90,04
11	200	1 x 1	4338730,07	4,339	85,42

Tabla 4. 4 Variación de volúmenes en función de la separación de líneas de sondeo para los modelos digitales generados por el método (K.O.).

Se puede observar que no existe una gran diferencia en entre los valores de porcentaje entre los modelos generados cada 40 y cada 60 m.

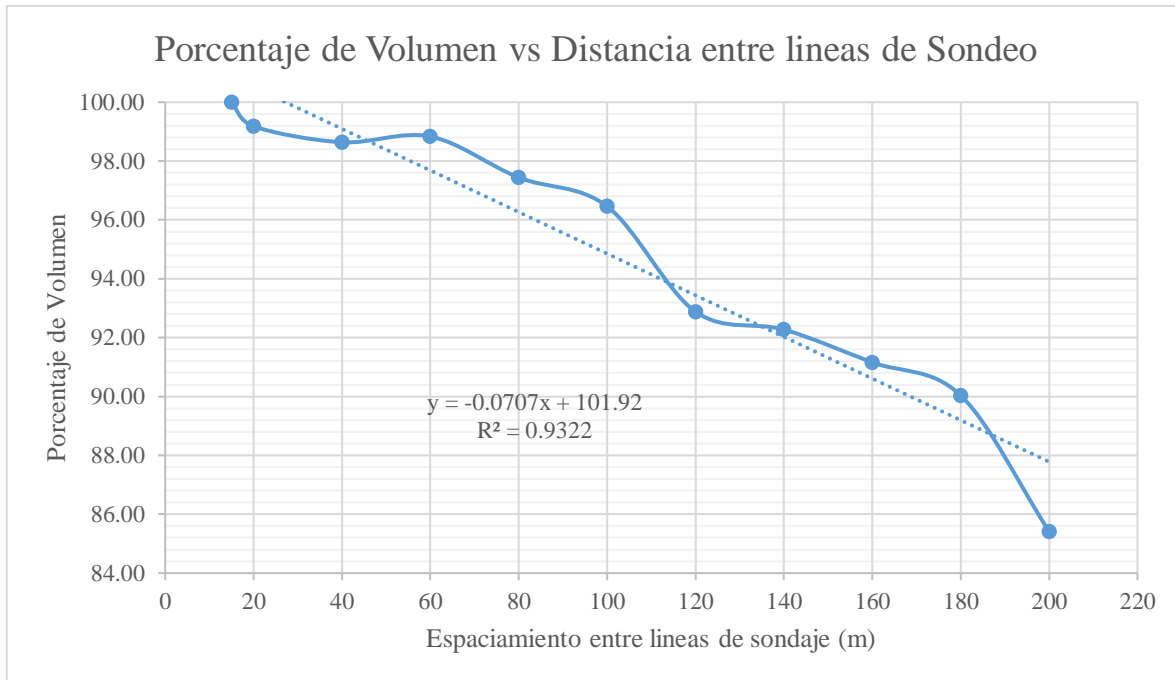


Figura 4. 3 Variación de volúmenes en función de la separación de líneas de sondeo para los modelos digitales generados por el método (K.O.).

4.5. Comparación de volúmenes

Analizando los volúmenes obtenidos por los tres tipos de interpolación en las distancias comprendidas entre distancias de entre 0 a 100 m. se observa que el método de interpolación TIN genero los volúmenes más bajos en sus modelos generados, en contraste el método Kriging genera valores más altos que varían poco incluso disminuyendo la cantidad de datos y el método IDW se mantiene en medio con algunos valores atípicos.

Se observa también una tendencia de reducción del volumen mientras aumenta la distancia de sondeo.

Kriging para una distancia de 60 m de separación de línea de sondeo genera un mayor volumen que los otros dos métodos 5020334,27 m³ valor muy cercano al generado por los otros métodos trabajando con distancias de 10 y 20 metros (totalidad de datos).

Distancia de línea de sondeo		Grilla	Volumen TIN	Volumen IDW	Volumen Kriging
N°	(m)	(m)	(m ³)	(m ³)	(m ³)
1	15	1 x 1	5027676,95	5085348,03	5079537,98
2	20	1 x 1	5019933,00	5101139,54	5037975,60
3	40	1 x 1	4978081,84	5020817,46	5010350,11
4	60	1 x 1	4937760,41	4964607,24	5020334,27
5	80	1 x 1	4871383,42	4882141,43	4949459,88
6	100	1 x 1	4878709,21	4866826,19	4899686,83
7	120	1 x 1	4711839,46	4677563,61	4718007,73
8	140	1 x 1	4690451,75	4650026,55	4687057,03
9	160	1 x 1	4601619,45	4517558,65	4630504,73
10	180	1 x 1	4585499,95	4522401,35	4573429,56
11	200	1 x 1	4391240,65	4317964,56	4338730,07

Tabla 4. 5 Variación de volúmenes en función de la separación de líneas de sondeo para los modelos digitales generados por TIN, IDW y K.O.

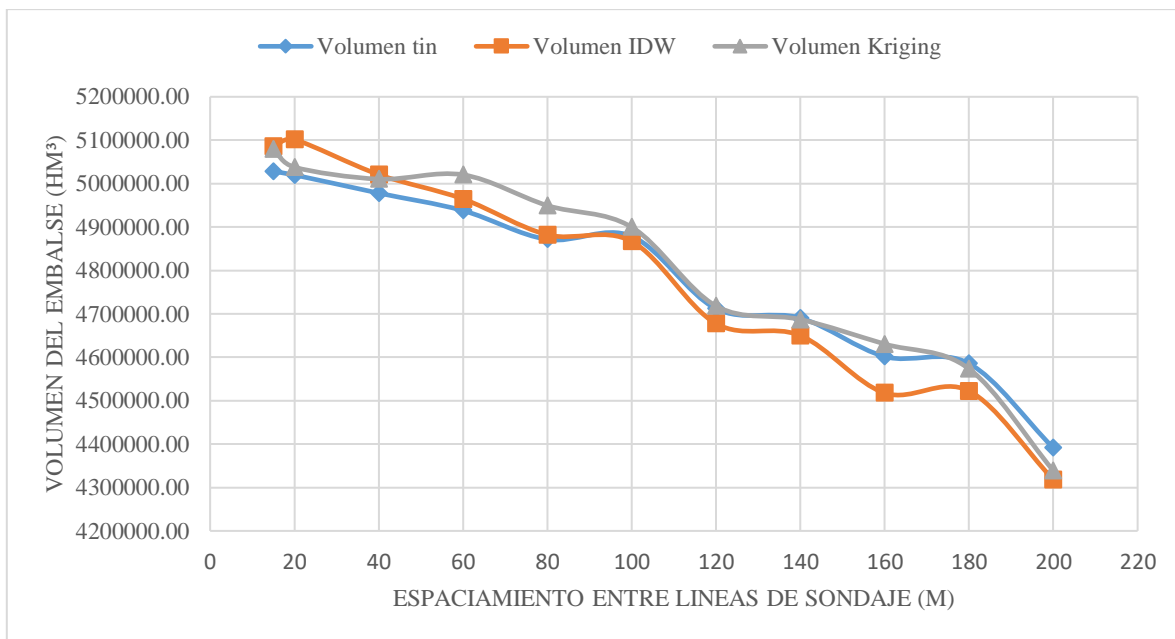


Figura 4. 4 Variación de volúmenes en función de la separación de líneas de sondeo para los modelos digitales generados por TIN, IDW y K.O.

4.6. Comparación de RSME

Mientras más Bajo sea el RSME el modelo se podría considerar que tiene mayor representatividad y es más preciso. para las distancias comprendidas entre 20 a 160 se puede observar que kriging tiene valores más bajos de RSME **1,2494** siendo el valor más bajo en promedio con respecto a los otros dos métodos de interpolación.

Distancia de línea de sondeo		Grilla	Kriging	IDW	TIN
N°	(m)	(m)	RSME	RSME	RSME
1	15	1 x 1	0,3103	0,3234	0,6547
2	20	1 x 1	0,3061	0,3387	0,6940
3	40	1 x 1	0,4788	0,7635	0,7823
4	60	1 x 1	0,8929	0,9284	1,0718
5	80	1 x 1	0,8735	1,3866	1,1337
6	100	1 x 1	0,9774	1,4166	1,2593
7	120	1 x 1	1,7366	2,4224	1,7768
8	140	1 x 1	1,7303	2,2733	1,7071
9	160	1 x 1	1,6292	2,4114	1,7538
10	180	1 x 1	2,2391	2,6859	2,1044
11	200	1 x 1	2,5701	2,9455	2,3072
Promedio			1,2494	1,6268	1,3859

Tabla 4. 6 Variación del RSME en función de la separación de líneas de sondeo para los modelos digitales generados por TIN, IDW y K.O.

El método de interpolación con mayor RSME es IDW en sus modelos generados con distancias superiores a 80 metros teniendo valores mayores de 1,3866 un valor demasiado alto para considerar modelos generados con distancias superiores.

Para la distancia de 100 y 60 metros por el método kriging se obtienen valores de 0,9774 y 0,8929 respectivamente, los cuales no superan la unidad indicando que estos modelos predicen un modelo más próximo a los datos observados.

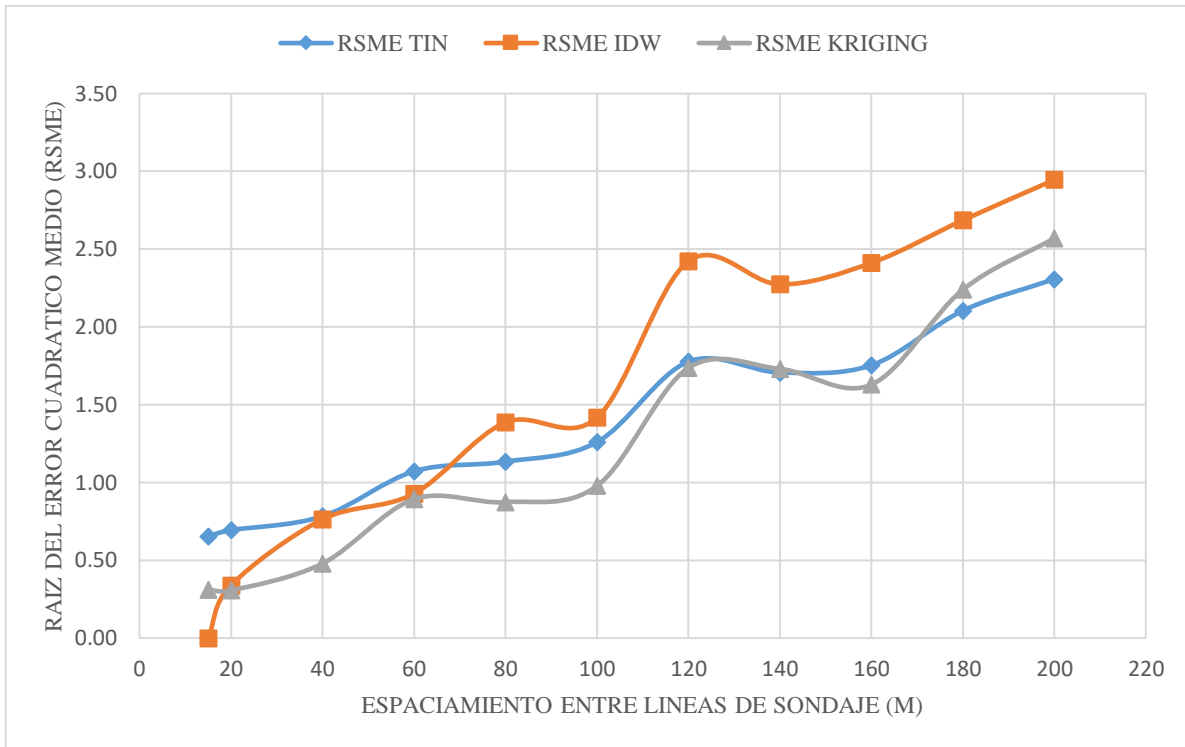


Figura 4. 5 Variación del RSME en función de la separación de líneas de sondeo para los modelos digitales generados por TIN, IDW y K.O.

CAPÍTULO V

5. CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

- ✓ De acuerdo al análisis realizado la distancia óptima entre líneas de sondeo para el embalse “El Tranque” es de **100 m.** esto considerando una pérdida de precisión no mayor al **5%** del volumen total que se obtiene de un muestreo más detallado, la finalidad de esta distancia es exclusiva para la determinación de la capacidad de almacenamiento.
- ✓ Para trabajos donde se requiera una mayor precisión se puede reducir la distancia incluso a **60 m.** donde la variación será del **2%** del volumen total que se obtiene de un muestreo más detallado, la finalidad de esta distancia es exclusiva para la determinación de la capacidad de almacenamiento.
- ✓ El método de interpolación que mejor se adapta al embalse “El Tranque” es el método Kriging al presentar en promedio **1,2494** de RMSE más bajo que el promedio del método IDW de **1,6268**.
- ✓ Las ubicaciones de errores métricos en los modelos generados se distribuyen generalmente entre la transición de datos batimétricos y topográficos.
- ✓ El método TIN genera errores debido a concentración de datos obtenidos por las líneas de sondeo batimétricos en las cuales se encuentran la mayor cantidad de datos dejando un gran margen datos no medidos, teóricamente se sabe que el comportamiento de la topografía no es lineal entre dos puntos, por lo tanto, la red de triángulos no es la adecuada para trabajos de precisión.
- ✓ El método IDW es susceptible al parámetro de la potencia la cual es una limitante para adoptar largas distancia debidamente por la distribución de datos que se concentra en las líneas batimétricas que no son lo ideal para este método.
- ✓ No existe un método de interpolación universal que se adapte a todos los embalses debido a que cada embalse es un caso de estudio particular, la elección del mejor depende de los datos con que se disponga y de cómo se comporten dichos datos.

- ✓ El levantamiento del margen del embalse es una de las zonas más influyentes en la generación de modelos batimétricos debido a que esta zona afecta en el cálculo del volumen total.
- ✓ El análisis de datos revela que todas las nubes de puntos no se ajustan a una distribución normal debido a que hay una mayor cantidad de puntos de batimetría que de topografía, esto no es recomendable al momento de usar métodos de interpolación ya sean determinísticos o geoestadísticos.

5.2. Recomendaciones

- ✓ Utilizar distancias de 100 m. de líneas de sondaje para embalses cuya área de superficie de inundación sea mayor a 80 hectáreas.
- ✓ Normalizar los histogramas de distribución de frecuencia de elevaciones con una transformación logarítmica ayuda a mejorar la interpolación con el método Kriging.
- ✓ Realizar el contorno del embalse con fotogrametría con dron debido a que este sector es de los que más influencia tienen en el cálculo del volumen así mismo es donde se presentan lo mayores errores de interpolación.
- ✓ Realizar las campañas batimétricas en dos periodos de tiempo la batimetría cuando el embalse este lleno y la fotogrametría o topografía cuando el embalse baje su nivel para al momento de combinar las superficies se pueda generar Modelos Digitales más Precisos.
- ✓ Utilizar distancias de líneas de sonde de 60 metros para el embalse “El Tranque en para la determinación de la capacidad de almacenamiento.
- ✓ Realizar el mismo estudio a embalses mayores debido a que cada embalse es un caso particular sin embargo a través de varios estudios se podría obtener una conclusión general con respecto a cuál es la distancia optima y al método de interpolación para todo tipo de embalses esto ayudaría a realizar un monitoreo más constante y óptimo de las fuentes de agua.
- ✓ Analizar la estacionariedad de la nube de puntos de evidenciarse una variación regionalizada de los datos se recomienda interpolar por sectores.