



UNIVERSIDAD MAYOR DE SAN ANDRÉS

FACULTAD DE TECNOLOGÍA

CARRERA DE GEODESIA, TOPOGRAFÍA Y GEOMÁTICA
CENTRO DE INVESTIGACIONES Y APLICACIONES GEOMÁTICAS – CIAG



DOCUMENTO DE INVESTIGACION TECNICA

DETERMINACIÓN DE UN MODELO PSEUDO GEOIDAL LOCAL PARA EL MUNICIPIO DE LA CIUDAD DE PAZ

M.SC. HUBER AUGUSTO MAMANI GUTIERREZ
Investigador – CIAG

<https://orcid.org/0000-0003-3937-3966>



LA PAZ – BOLIVIA
2020

*“La parte más difícil de la investigación es siempre encontrar una pregunta que sea lo suficientemente grande como para que valga la pena responderla, pero lo suficientemente poco como para que realmente puedas responderla
(Edward Witten)*

Mientras la investigación sea considerada simplemente un complemento y no la esencia del conocimiento, el desarrollo de la ciencia y tecnología no alcanzará su esplendor.

ÍNDICE

	Página
RESUMEN.....	1
CAPÍTULO I.....	3
PLANTEAMIENTO GENERAL.....	3
1.1. ANTECEDENTES	3
1.2. JUSTIFICACIÓN	10
1.3. PLANTEAMIENTO DEL PROBLEMA	14
1.4. FORMULACIÓN DEL PROBLEMA.....	14
1.5. HIPÓTESIS	15
1.6. OBJETIVOS	15
1.6.1. Objetivo General.....	15
1.6.2. Objetivos Específicos	15
CAPÍTULO II.....	16
MATERIALES Y METODOLOGÍA	16
2.1. MATERIALES Y EQUIPOS.....	16
2.2. METODOLOGÍA.....	17
2.2.1. Tipo de investigación.....	17
2.2.2. Método experimental.....	18
CAPÍTULO III.....	19
APLICACIÓN PRÁCTICA	19
3.1. DELIMITACIÓN DEL ÁREA DE ESTUDIO	19
3.2. RECOPIACIÓN DE INFORMACIÓN.....	20
3.3. DEPURACIÓN DE LA INFORMACIÓN RED GEODÉSICA 2005 - 2006.....	21
3.3.1. Prueba de los valores atípicos	21
3.3.2. Determinación de las estaciones de validación y del modelo.....	24
3.4. GENERACIÓN DEL MODELO A PARTIR DEL ANÁLISIS ESTADÍSTICO	26
3.4.1. Aplicación de la Regresión Lineal Múltiple.....	26
3.4.1.1. Prueba de hipótesis en la regresión lineal múltiple.....	27
3.4.1.2. Procedimiento realizado en Microsoft Excel mediante el Análisis de Datos.....	29
3.4.2. Aplicación de la Regresión Polinomial Múltiple.....	39
3.5. DETERMINACIÓN DE VALORES DE ONDULACIÓN GEOIDAL A PARTIR DE MODELOS GLOBALES	55
3.5.1. Cálculos de ondulación geoidal mediante puntos definidos por el usuario	57

	Página
CAPÍTULO IV	61
RESULTADOS	61
4.1. COMPARACIÓN DE VALORES DE ONDULACIÓN PSEUDO GEOIDAL CON LAS ESTACIONES DE VALIDACIÓN	61
4.2. COMPARACIÓN DE VALORES DE ONDULACIÓN PSEUDO GEOIDAL CON MODELOS GLOBALES	62
4.2.1. Comparación con el modelo global EGM96.....	62
4.2.2. Comparación con el modelo global EGM08.....	63
4.2.3. Comparación con el modelo global EIGEN-6C4.....	64
CAPÍTULO V	65
CONCLUSIONES Y RECOMENDACIONES	65
5.1. CONCLUSIONES.....	65
5.2. RECOMENDACIONES	67
CAPÍTULO VI	68
BIBLIOGRAFÍA	68

ÍNDICE DE FIGURAS

	Página
Figura 1: Discrepancias en las alturas niveladas entre países vecinos en América del Sur.....	4
Figura 2: Altura nivelada referida al geoide local definida por el datum vertical “Arica”.....	12
Figura 3: Portadas del Informe del proyecto Red Geodésica Municipal de la Ciudad de La Paz 2005	16
Figura 4: Programas utilizados para la investigación.....	17
Figura 5: Diseño del método experimental.	18
Figura 6: Ubicación geográfica del área de estudio.....	19
Figura 7: Gráfica de valores atípicos de N – Primer análisis.	23
Figura 8: Distribución de las estaciones de validación y del modelo en el área de estudio.	25
Figura 10: Submenú de Análisis de datos de Excel.....	31
Figura 11: Elección de valores de entrada y opciones de salida para la regresión.....	32
Figura 12: Representación de los valores de F y $F_{Crítica}$ del análisis de varianza.....	36
Figura 13: Herramienta de Análisis de datos, Regresión de Excel.....	48
Figura 14: Valores de entrada y opciones de salida para la regresión polinomial múltiple.....	49
Figura 15: Valores de F y $F_{Crítica}$ del análisis de varianza – Regresión polinomial.	52
Figura 16: Lista de funciones de campo de gravedad disponibles en el servicio de cálculo ICGEM...	56
Figura 17: Valores de latitud y longitud en grados de las estaciones utilizadas en el modelo.....	57
Figura 18: Instantánea del servicio de cálculo del valor de puntos en ICGEM.	58
Figura 19: Instantánea del archivo devuelto por el servicio de cálculo del ICGEM.....	58
Figura 20: Instantánea del mapeo de la ubicación de los puntos definidos por el usuario.	59

ÍNDICE DE CUADROS

	Página
Cuadro 1: Tipos básicos de regresión.	14
Cuadro 2: Tamaño de muestra y relación recomendada por Dixon.	21
Cuadro 3: Hipótesis nula y alterna para la prueba de valores atípicos.....	22
Cuadro 4: Prueba de relación r^2 de Dixon – Primer análisis.	22
Cuadro 5: Valor atípico determinado – Primer análisis.....	22
Cuadro 6: Estaciones identificadas con valores atípicos.	24
Cuadro 7: Estaciones para validación del modelo.	25
Cuadro 8: Valores de las variables independientes y dependientes en el análisis de datos.	30
Cuadro 9: Valores de los coeficientes determinados en la regresión lineal.	32
Cuadro 10: Valores de N estimados y residuales por estación.	34
Cuadro 11: Análisis de varianza para la regresión lineal.	35
Cuadro 12: Valores de correlación de la regresión lineal.	37
Cuadro 13: Valores de correlación determinados para polinomios del tercer al sexto grado.	41
Cuadro 14: Valores de variables definidas para el proyecto.	43
Cuadro 15: Valores de promedio de las coordenadas Este y Norte.	44
Cuadro 16: Valores nuevos determinados para las coordenadas Este y Norte.	45
Cuadro 17: Valores de las variables independientes utilizadas en la regresión polinomial múltiple.	48
Cuadro 18: Valores de los coeficientes determinados en la regresión polinomial múltiple.	49
Cuadro 19: Valores de N estimados mediante regresión polinomial y residuales por estación.	51
Cuadro 20: Valores de promedio, valor máximo y mínimo de los residuos de regresión polinomial. ...	51
Cuadro 21: Análisis de varianza para la regresión polinomial.	52
Cuadro 22: Valores de correlación de la regresión polinomial	53
Cuadro 23: Comparación de valores de ondulación pseudo geoidal de las estaciones de validación.	61
Cuadro 24: Comparación de valores de ondulación pseudo geoidal con el modelo EGM96.	62
Cuadro 25: Comparación de valores de ondulación pseudo geoidal con el modelo EGM08.	63
Cuadro 26: Comparación de valores de ondulación pseudo geoidal con el modelo EIGEN-6C4.....	64

RESUMEN

La presente investigación determina un Modelo Pseudo Geoidal Local para el Municipio de la Ciudad de La Paz, aplicando análisis estadístico comprendido en regresión polinomial múltiple de quinto grado, utilizando datos de las estaciones de la Red Geodésica Municipal 2005 – 2006.

Se establece como espacio de estudio el área que comprende la parte urbana del Municipio de La Paz, donde se tiene perfectamente implantada la red geodésica municipal con 113 estaciones con información de coordenadas geodésicas, coordenadas UTM y alturas niveladas que corresponden a alturas geométricas.

Para el desarrollo de la investigación, se establece las diferencias de alturas elipsoidales (h) y las alturas niveladas (H_n), que representan la distancia entre el elipsoide de referencia y el pseudo geoide local que pasa por el datum vertical “Arica”, denominado en la investigación como ondulación pseudo geoidal (N), que permite generar el modelo pseudo geoidal local del área de estudio.

Así también, antes de generar el modelo se realizó la depuración de datos, mediante análisis estadístico de valores atípicos, aplicando el programa Minitab 18 y las pruebas de Dixon.

Para generar el modelo, se aplica en primera instancia la regresión lineal múltiple para comprender el comportamiento e influencia de las variables independientes (coordenadas Este y Norte de las estaciones utilizadas), respecto a la variable dependiente (N – Ondulación pseudo geoidal), para finalmente utilizar la regresión polinomial múltiple mediante un polinomio de quinto grado que mejor se ajusta a los valores utilizados en la investigación, de tal manera que los resultados y valores estadísticos interpretados, establecen la calidad de la regresión así como la fiabilidad para la predicción de valores de N , por un alto grado de relación entre variables dependientes e independiente que alcanza a un porcentaje del 88%.

El modelo finalmente es verificado mediante un grupo de estaciones de validación seleccionados del total de estaciones depuradas de la Red Geodésica Municipal de La Paz 2005 – 2006.

Así también, se realizan comparación de los resultados del modelo con valores de ondulaciones geoidales determinadas mediante modelos geoidales globales como el EGM96, EGM08 y EIGEN-6C4, cuyas diferencias oscilan entre 2 a 3 metros, explicando que no se pueden comparar alturas ortométricas con alturas niveladas.

CAPÍTULO I

PLANTEAMIENTO GENERAL

1.1. ANTECEDENTES

La Asociación Internacional de Geodesia (IAG), y la Asamblea General Internacional de Geodesia y Geofísica (IUGG), publicaron en julio de 2015 una resolución bajo el consentimiento del Comité de Expertos de las Naciones Unidas en Gestión Global de la información Geoespacial (United Nations Initiative on Global Geospatial Information Management UN-GGIM), para la definición y realización de un Sistema Internacional de Referencia de Alturas (IHRs), promoviendo la mejora y la compatibilidad de los sistemas espaciales de navegación y determinación de la posición, incluidos los Sistemas Mundiales de Navegación por Satélite (GNSS), y el acceso universal a estos (Leal Vélez Anderson, 2017).

El Sistema de alturas o Sistemas Verticales de Referencia, se componen básicamente de: una superficie de referencia (datum vertical), y una coordenada vertical (altura). Si la superficie de referencia y la altura dependen directamente del campo de gravedad se habla de un sistema de alturas físicas (Altura ortométrica y el Geoide, o altura normal y el Cuasigeoide). Si se habla de alturas que no dependen del campo de gravedad se refiere a sistema de alturas geométricas (por ejemplo, alturas elipsoidales y su elipsoide de referencia, también se podría considerar dentro de estas a las alturas niveladas).

Sin duda, esta iniciativa bastante ambiciosa, refleja la importancia de estandarizar a través de un Sistema Internacional de Referencia de Alturas, la componente vertical o control vertical que se va realizando en una serie de trabajos topográficos, geodésicos y demás aplicaciones, en todas las regiones de nuestro planeta. En el entendido, que por ejemplo en América Latina, los datum verticales se refieren a diferentes mareógrafos, por tanto, a diferentes niveles del mar y sin duda a diferentes épocas de observación.

Esto significa que no se toma en cuenta las variaciones de las alturas y del nivel de referencia con el tiempo, y en general, la extensión del control vertical mediante redes de nivelación no incluyen las reducciones por los efectos de gravedad, en Bolivia simplemente

se determinan los delta h (alturas), para el tema de nivelación geométrica, suponiendo que las superficies equipotenciales son paralelas¹.

Asimismo, el nivel medio del mar también varía a través del tiempo, es decir, el nivel registrado en un mareógrafo de referencia durante el período de definición, definitivamente no es el mismo que se registraría si se tomara otra época para definir la altura cero en ese mismo mareógrafo. Por ejemplo, en el mareógrafo Mar del Plata (Argentina), se ha registrado un descenso del nivel medio del mar de 1.3 mm por año, es decir que el nivel medio del mar registrado en la década de los años 50 estaba 7 cm “más arriba” que el nivel actual (Laura Sánchez, 2006).

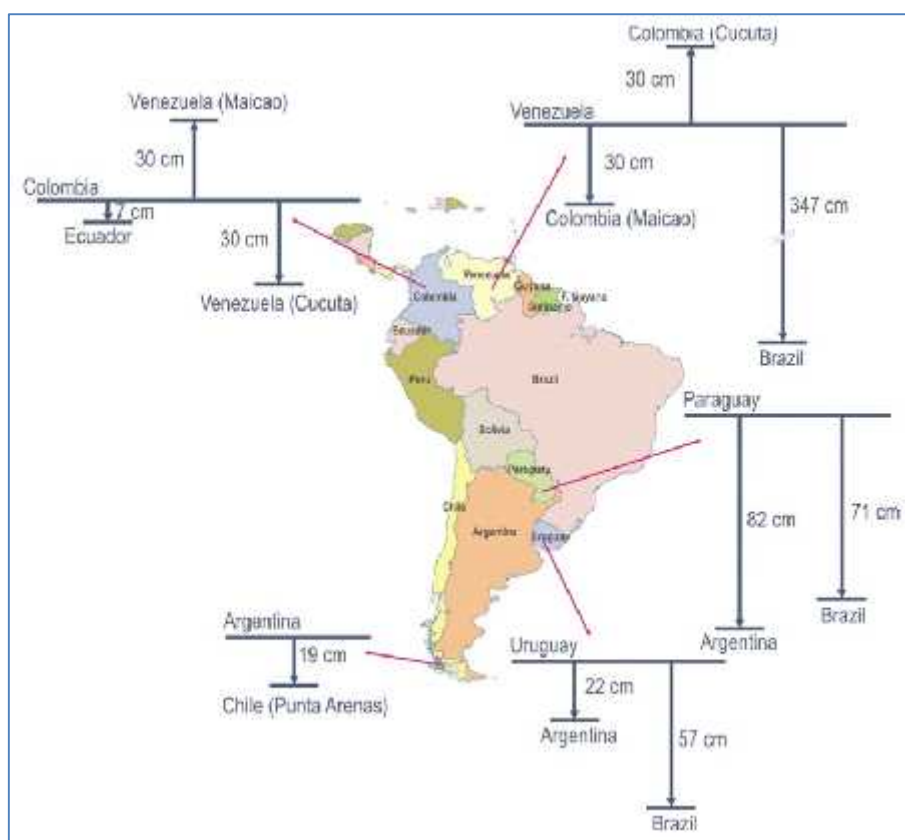


Figura 1: Discrepancias en las alturas niveladas entre países vecinos en América del Sur.
Fuente: Laura Sánchez, 2006.

¹ D. Flores Vargas, Seminario virtual sobre Sistemas de Altitudes, UMSA, La Paz Bolivia, 2020.

En consecuencia, las alturas presentan discrepancias considerables entre países vecinos, no permiten el intercambio de información vertical ni a escala continental ni a escala global y no están en capacidad de soportar la determinación práctica de alturas a partir de las técnicas GNSS. Por tanto, no se cumple la relación que la ondulación geoidal (N), es resultado de la diferencia entre la altura elipsoidal (h), y la altura ortométrica (H).

Por otro lado, estas últimas décadas han surgido una serie de versiones de modelos geoidales globales como el EGM96, EGM2008, EIGEN-6C4, XGM2019e_2159 entre otros, que son de acceso abierto y gratuito, incluyendo descargas de modelos y servicios de cálculo y visualización tal el caso del Centro Internacional de Modelos Globales de la Tierra (ICGEM)².

Los modelos geopotenciales o geoidales se basan, en la determinación de los coeficientes del desarrollo en serie de funciones armónicas esféricas del potencial gravitacional terrestre, donde el procedimiento se expresa en coeficientes determinados en función de la variación de los seis elementos orbitales. En un campo gravitacional ideal no perturbado, la órbita satelital sería una elipse kepleriana. En el campo gravitatorio real el satélite es perturbado por el campo gravitacional de la Tierra, pero esa misma perturbación que sufre el satélite se utiliza para derivar información sobre dicho campo (Pacino, 1999). Asimismo, las observaciones son las variaciones de los elementos orbitales y las incógnitas son los coeficientes armónicos, surgiendo de esta manera los modelos geopotenciales donde las ondulaciones geoidales son expresadas en función de armónicos esféricos (Tierra, 2015).

Existen diferentes recursos de datos complementarios que se utilizan para el desarrollo de modelos de campos de gravedad globales de alta calidad. Por ejemplo, las mediciones

² El Centro Internacional de Modelos Globales de la Tierra (ICGEM), alojado en el GFZ German Research Center for Geosciences (GFZ), es uno de los cinco servicios coordinados por el International Gravity Field Service (IGFS), de la International Association of Geodesy (IAG). El objetivo del servicio ICGEM es proporcionar a la comunidad científica un archivo de última generación de modelos de campos gravitatorios globales estáticos y temporales de la Tierra, y desarrollar y operar servicios interactivos de cálculo y visualización de funciones del campo gravitatorio en el usuario. ICGEM cuenta con modelos globales de campos de gravedad, incluidos los de los años 60 a los 90, así como los más recientes, que se han desarrollado utilizando datos de misiones de gravedad satelitales dedicadas, CHAMP, GRACE, GOCE, metodologías de procesamiento avanzadas, y fuentes de datos adicionales como altimetría satelital y gravedad terrestre.

satelitales avanzadas o las cantidades derivadas son una de ellas y pueden ser en forma de perturbaciones orbitales de los satélites derivadas de las mediciones GNSS, mediciones de la frecuencia de alcance del láser y microondas entre dos satélites, observaciones del alcance del láser por satélite (SLR), desde la superficie terrestre a los satélites cercanos a la Tierra y, finalmente, los gradientes de gravedad y las aceleraciones no gravitacionales medidas a bordo de naves espaciales. Los satélites recientes que contribuyen enormemente a las mejoras en el modelado global del campo de gravedad son las misiones de gravedad dedicadas CHAMP, GRACE, GRACE Follow-On, GOCE, y satélites SLR como LAGEOS 1 y LAGEOS 2, así como la flota de satélites altimétricos como Topex / Poseidon y Jason 1 y 2. Otros conjuntos de datos fundamentales utilizados en el desarrollo de modelos de campo de gravedad global son las mediciones de gravedad terrestre, incluidas las recopiladas en plataformas móviles. Además de las mediciones de gravedad, los modelos digitales de elevación (DEM) de alta resolución complementan los modelos de campo de gravedad global para mapear características detalladas del campo de gravedad y en las áreas donde faltan mediciones de gravedad real, como la Antártida (Ince et al., 2019).

El modelo geoidal permite calcular las ondulaciones geoidales y en consecuencia determinar el valor de alturas de carácter físico (alturas ortométricas), que son necesarias tanto para la investigación científica como para trabajos de ingeniería. Asimismo, estos modelos globales, son utilizados en combinación al uso de equipos GNSS, que nos proporcionan alturas geométricas (elipsoidales), a las cuales aplicando los valores de ondulación geoidal, permite obtener la altura ortométrica. Surgiendo la incertidumbre sobre la precisión que tienen estos modelos en relación a la región donde son utilizadas.

Según investigaciones sobre comparaciones de modelos geoidales globales, hacen referencia, por ejemplo, en Argentina que del análisis estadístico de los residuos de ondulación y anomalías de gravedad de los modelos globales, los que mejor se ajustan a la región de Argentina son aquellos modelos combinados que incluyen en su cálculo datos satelitales, de altimetría satelital y de campañas terrestres. Los modelos basados en datos satelitales solamente acusan una gran diferencia respecto de los valores terrestres, con rangos de valores en desviación estándar de 0.27 hasta 0.61 metros, recomendando el uso del modelo EGM2008, siendo que es el que presenta un mayor desarrollo en serie de los

coeficientes armónicos esféricos y cuenta con el respaldo de un procesamiento llevado a cabo por instituciones de reconocido valor científico (Cornero Cecilia et al., 2015).

En el caso de Ecuador la diferencia entre la ondulación geoidal calculada aplicando observaciones GNSS y alturas ortométricas en el estudio y la obtenida por el modelo EGM08 se encuentra entre un máximo de $\pm 0,95$ m y un mínimo de $\pm 0,09$ m. Sin embargo es importante tomar en cuenta que el EGM08 es un modelo global y que su adaptación a la zona de estudio puede resultar no muy próxima a la realidad (Canizares et al., 2015).

En nuestro país, es difícil de poder determinar las diferencias entre modelos geoidales globales y alturas ortométricas, pues la red de control vertical de Bolivia se encuentra definida por alturas niveladas y no existen puntos específicos con alturas físicas como las ortométricas. Asimismo, al no existir un modelo geoidal local, para obtener ondulaciones geoidales con buena precisión, lo que generalmente realizan los profesionales del área, son las aplicaciones de modelos geopotenciales que tienen cobertura mundial, como es el caso del EGM96 (Earth Geopotential Model), cuyo error medio es de ± 1 metro (Lemoine et. al, 1998), aplicados generalmente a la mayoría de los programas comerciales para el procesamiento y ajuste de datos GPS y así transformar alturas elipsoidales en alturas ortométricas.

Es indudable que el desarrollo del Sistema Global de Navegación por Satélite (Global Navigation Satellite System, GNSS), para la determinación de puntos de control horizontal sobre la superficie terrestre, con altas precisiones en posición, favorecen el trabajo topográfico geodésico en cuanto a tiempos, costos y procedimientos en su determinación, sin embargo, es importante considerar que el método ofrece una altura elipsoidal referida a una superficie matemática.

El posicionamiento satelital, proporciona la posibilidad de disponer de coordenadas geodésicas de alta precisión, que tienen como referencia un datum geocéntrico y un modelo geométrico o elipsoide de referencia para su representación en valores angulares, consiguiendo de esta forma aproximar la forma terrestre. Cada punto sobre la superficie del planeta cuenta con una posición definida por su latitud, longitud y altura elipsoidal; los dos primeros valores proveen la ubicación horizontal con respecto a referencias convencionales

que están definidas sobre el elipsoide de referencia. La altura elipsoidal por su parte representa la distancia medida sobre la normal al elipsoide desde el punto posicionado hasta dicha superficie de referencia, es decir esta altura es una expresión geométrica de la distancia existente entre el punto posicionado y la superficie de referencia (elipsoide); esto se convierte en un inconveniente al tomar en consideración que para ser utilizadas en proyectos de ingeniería las alturas medidas sobre la superficie terrestre deben tener un significado físico, esto implica tener en cuenta la distribución de masas del planeta y su campo gravitatorio. Siendo indispensable disponer de una superficie de referencia para las alturas que considere estas variables físicas (Carrión Sánchez, 2013).

Para una mejor comprensión, podemos explicar que un punto sobre la superficie topográfica necesita de tres coordenadas, para así estar totalmente definida sobre un espacio tridimensional. Las dos primeras coordenadas se refieren a una posición planimétrica (por ejemplo latitud y longitud), mientras que la última coordenada será la altura del punto.

La altura de un punto sobre la superficie topográfica, puede ser definida como la distancia existente entre éste y una superficie de referencia, medida a lo largo de la línea normal a dicha superficie que pasa por el punto de trabajo. En un proceso convencional de determinación de alturas mediante nivelación geométrica, el telescopio del instrumento es tangente a la superficie equipotencial local y la línea de la plomada coincide con el vector de la fuerza de gravedad, el cual es perpendicular a aquellas superficies. Generalmente, el geoide es aceptado como la superficie equipotencial de referencia para los sistemas de alturas, sin embargo las distintas superficies equipotenciales no son paralelas entre sí, debido principalmente a que la distribución de masas en el interior de la Tierra no es homogénea. Este no paralelismo tiene como consecuencia que el desnivel entre dos puntos, obtenido únicamente por nivelación geométrica, dependa del camino seguido por la línea de nivelación. Es por esto que el conjunto de operaciones de nivelación geométrica exigen el conocimiento de valores de la gravedad para la corrección de los desniveles observados (Canizares et al., 2015).

Dentro de la Geodesia consideramos dos tipos de alturas: geométricas y físicas. Las alturas niveladas y las alturas elipsoidales son alturas de tipo geométrico, siendo que fueron obtenidas mediante mediciones geométricas. Las alturas niveladas son obtenidas bajo

proceso de nivelación geométrica con métodos ópticos de medición, en el cual se observa los desniveles o distancias existentes entre las superficies equipotenciales del campo de gravedad terrestre. Las alturas elipsoidales (h), por su parte, representan la separación entre la superficie topográfica y el elipsoide, calculada sobre la línea perpendicular a este último, son obtenidas a partir de las coordenadas geocéntricas cartesianas (X, Y, Z), definidos sobre un elipsoide de referencia y como se había mencionado estas alturas son obtenidas directamente a partir de observaciones GNSS.

Las alturas consideradas físicas, son las alturas dinámicas que utilizan una gravedad teórica constante, normalmente a una latitud de 45° , careciendo de significado geométrico; las alturas normales que toman el valor medio de la gravedad teórica a lo largo de la normal al elipsoide de referencia, entre el elipsoide y un punto P cuyo potencial normal es igual al potencial real de un punto P situado en la superficie terrestre (originan el Cuasigeoide), y finalmente tenemos a las alturas ortométricas, que toman el valor medio de la gravedad observada a lo largo de la línea de fuerza entre el geoide y un punto P situado en la superficie terrestre (originan el geoide).

De ahí la importancia de la búsqueda de una superficie de referencia para alturas con significado físico (Geoide), continúa siendo una necesidad básica y una tarea fundamental en la mayoría de países, con el objetivo de integrarlo con otras tecnologías. De esta manera, se hace indispensable y necesario el conocimiento de un modelo geoidal con la finalidad de transformar la altura elipsoidal a una altura de carácter físico y de esta manera poder explotar de una forma más eficiente la tecnología de posicionamiento por satélites artificiales.

En los últimos años el cálculo del Geoide ha sido, es y seguirá siendo una tarea fundamental de muchos investigadores, por lo cual han sido desarrollados modelos geoidales en casi todos los países del mundo. Diferentes métodos se han desarrollado a lo largo de la historia para posibilitar el cálculo de modelos del geoide en diferentes regiones del mundo. Todos los métodos involucran de una u otra manera el conocimiento del campo potencial de gravedad. Actualmente, los métodos empleados en la determinación de las ondulaciones geoidales son los modelos geopotenciales, modelos gravimétricos, modelos por satélites artificiales y últimamente se han incorporado los modelos estadísticos.

1.2. JUSTIFICACIÓN

Con la necesidad de establecer claramente las precisiones y sobre todo una referencia vertical, en la región se han ido generando una serie de investigaciones apoyados en determinar nuevos modelos geoidales de carácter local, así como la evaluación de los modelos globales. Para Ecuador tenemos por ejemplo la Evaluación del EGM08 y EGM96 a partir de datos GPS y Nivelación Geométrica, que concluye que el modelo EGM08 presenta una mejoría respecto al EGM96 para el cálculo de las ondulación geoidales, así también, que para el área de Ecuador, ambos modelos no serían aconsejables utilizar para transformaciones de alturas elipsoidales a alturas niveladas para proyectos aplicativos cuyo error sea menor a 4 metros (Tierra, 2015). Otra investigación similar corresponde al Modelo neuronal para la predicción de la altura geoidal local del Ecuador, publicada en la Revista Geoespacial de la Universidad de las Fuerzas Armadas del Ecuador, del Departamento de Ciencias de la Tierra y Construcción, cuyo objetivo fue desarrollar un modelo geoidal local mediante la utilización del método geométrico, con datos provenientes del GPS conjuntamente con la nivelación geométrica y el uso de la técnica de redes neuronales artificiales para modelar el geoide local, cuyo resultados muestran que se obtuvieron errores menores a 40 cm, con error medio de 0.2 cm y una desviación estándar de 15 cm (Tierra & Acurio, 2016).

En el Perú, se tiene los proyectos de determinación del Modelo Geoidal con receptores GNSS en método cinemático y nivelación geométrica para la estimación de alturas ortométricas en la Ciudad Universitaria de UNA – Puno, desarrollado a partir de la aplicación de la ecuación de regresión múltiple alcanzando una desviación estándar de 0.00459 metros con respecto a la nivelación geométrica (Acero Velo & Serruto Flores, 2018).

Como se observa en las investigaciones de la región, la necesidad de contar con un modelo “geoidal” local, que permita la combinación de levantamientos realizados con equipos GNSS y estos modelos, a fin de obtener alturas niveladas o pseudo ortométricas como algunos autores lo denominan, permite tener clara que la realidad de los países vecinos, es similar a la nuestra, primero, que la mayoría determinó su Datum Vertical como origen para sus alturas desde un nivel medio del mar a través de un mareógrafo (conociendo además que cambia en función del tiempo), por ejemplo, para Bolivia el mareógrafo de Arica – Chile, para

el Ecuador el mareógrafo La Libertad, situación que como señala L. Sánchez (2006), las alturas entre países vecinos presentan discrepancias considerables, no permiten el intercambio de la información vertical a nivel continental y mecho menos a nivel global.

En este sentido, al determinar los países su origen vertical referido a un determinado mareógrafo, se lo puede denominar como datum vertical local, y a partir de estos, mediante nivelación geométrica y observaciones gravimétricas (aunque algunos muchos no hicieron gravimetría), materializaron su red de control vertical con la finalidad de determinar la altura de un punto. Ahora bien, algunos países sin hacer correcciones gravimétricas, haciendo o no solamente correcciones de paralelismo de las superficies equipotenciales, denominaron a estas alturas (erróneamente), como “alturas ortométricas” (Tenzer et al., 2005) (Tierra & Acurio, 2016).

Por todo lo citado, podemos concluir que el datum local no coincide con el geoide global, por tal razón, en muchos países, a estas alturas “ortométricas”, pasaron a llamarlas alturas niveladas referidas a su propio datum vertical. Para Bolivia se definió como datum vertical “Arica” que se encuentra en Chile, que representa el nivel medio de los mares (para esa ubicación geográfica y una determinada época), pudiendo ser considerada este nivel como un geoide local que pasa por ese mareógrafo. A partir de este datum, se extendió la red control vertical a lo largo y ancho de nuestro país, mediante nivelación geométrica (sin realizar correcciones gravimétricas y considerando que las superficies equipotenciales eran paralelas). Por tanto, implica que las alturas o cotas definidas y existentes en nuestro territorio sean consideradas de carácter geométrico, referidos al nivel medio de los mares definido por el datum vertical. También resaltar, que estas alturas han sido y siguen siendo utilizadas para fines prácticos en los distintos trabajos de topografía, geodesia e ingenieriles.

En la práctica generalmente, se hacen uso de los diferentes modelos geoidales globales, en el procesamiento y ajuste de datos GPS (destinados a densificación de puntos e incluso al establecimiento de redes geodésicas locales, como las municipales o para proyectos específicos), calculando de esta manera la altura ortométrica (que conceptualmente estaría correctamente determinada, por la definición que la ondulación geoidal es resultado de la diferencia de la altura elipsoidal y la altura ortométrica), pero no olvidemos que la red de control vertical de Bolivia corresponde a alturas niveladas (alturas geométricas), asimismo,

como se había mencionado, al no existir puntos específicos con alturas ortométricas, es imposible verificar la diferencia o precisión de los modelos geoidales globales, solo por comparación con un estudio realizado en Ecuador, considerando una altura de unos 3602.388 metros sobre el nivel del mar similar al de la ciudad de La Paz, la diferencia entre una altura nivelada y una altura ortométrica alcanza los -1.071 metros³.

Esto significa que desde el aspecto práctico se debería seguir utilizando la nivelación clásica, que si bien representa un método preciso para obtener alturas niveladas y los correspondientes desniveles, también conlleva mucho tiempo y sobre todo representa un ítem de costo muy elevado para los proyectos, mucho más si en las áreas próximas no se encuentra un BM para realizar el denominado “arrastre de cotas”.

Por todo lo expuesto, por la disponibilidad y distribución de datos de la Red Geodésica de la Ciudad de La Paz (2005-2006), se plantea la alternativa de generar un Modelo Geoidal Local utilizando datos resultantes del levantamiento con equipos GNSS y de nivelación geométrica.

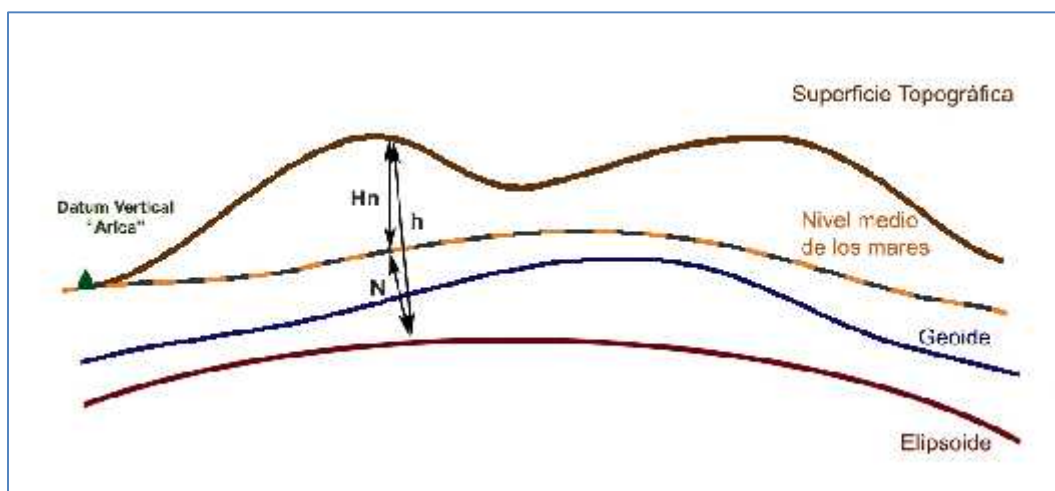


Figura 2: Altura nivelada referida al geoide local definida por el datum vertical “Arica”.

Fuente: Figura modificada de Tierra & Acurio, 2016.

La diferencia entre las alturas elipsoidales (h) y las alturas niveladas (H_n), darán como resultado la distancia entre el elipsoide de referencia WGS84 y el pseudo geoide local que pasa por el Datum Vertical “Arica”, y es denominado como ondulación pseudo geoidal (N),

³ D. Flores Vargas, Seminario virtual sobre Sistemas de Altitudes, UMSA, La Paz Bolivia, 2020.

que permitirá generar el Modelo Pseudo Geoidal para la Ciudad de La Paz, y poder ser aplicado por los profesionales del área, para transformar alturas elipsoidales obtenidas mediante equipos GNSS a alturas geométricas utilizadas en nuestra ciudad y también en nuestro país.

La metodología utilizada corresponde a un modelo generado a partir de la regresión polinomial múltiple, la cual permitió ajustar la tendencia de regresión considerando todos los valores de N.

La regresión lineal múltiple permite generar un modelo lineal en el que el valor de la variable dependiente o respuesta (Y) se determina a partir de un conjunto de variables independientes llamadas predictores (X_1, X_2, X_3, \dots). Es una extensión de la regresión lineal simple, por lo que es fundamental comprender esta última. Los modelos de regresión múltiple pueden emplearse para predecir el valor de la variable dependiente o para evaluar la influencia que tienen los predictores sobre ella (Joaquín Amat Rodrigo, 2016).

El análisis de regresión engloba a un conjunto de métodos estadísticos que usamos cuando tanto la variable de respuesta como la(s) variable(s) predictiva(s) son continuas y queremos predecir valores de la primera en función de valores observados de las segundas. En esencia, el análisis de regresión consiste en ajustar un modelo a los datos, estimando coeficientes a partir de las observaciones, con el fin de predecir valores de la variable de respuesta a partir de una (regresión simple) o más variables (regresión múltiple) predictivas o explicativas.

El análisis de regresión juega un papel central en la estadística moderna y se usa para:

-) Identificar a las variables predictivas relacionadas con una variable de respuesta.
-) Describir la forma de la relación entre estas variables y para derivar una función matemática óptima que modele esta relación.
-) Predecir la variable de respuesta a partir de las explicativas o predictoras.

El término regresión puede ser confuso porque existen muchas variantes especializadas de regresión, muchas funciones para ajustar una gran gama de modelos de regresión (Vinuesa, 2016)

Tipo de regresión	Uso típico
Lineal simple	Predicción de una variable de respuesta cuantitativa a partir de una variable predictora cuantitativa
Polinomial	Predicción de una variable de respuesta cuantitativa a partir de una variable predictora cuantitativa, donde la relación se modela como una función polinomial de orden n
Lineal múltiple	Predicción de una variable de respuesta cuantitativa a partir de dos o más variables predictoras cuantitativas
Multivariada	Predicción de más de una variable de respuesta cuantitativa a partir de una o más variables predictoras cuantitativas
Logística	Predicción de una variable categórica a partir de una o más predictoras
De Poisson	Predicción de una variable de respuesta que representa un conteo a partir de una o más predictoras
No lineal	Predicción de una variable de respuesta cuantitativa a partir de una o más predictoras, donde el modelo no es lineal
Robusta	Predicción de una variable de respuesta cuantitativa a partir de una o más predictoras, usando una aproximación resistente al efecto de observaciones influyentes

Cuadro 1: Tipos básicos de regresión.

Fuente: Vinuesa, 2016.

1.3. PLANTEAMIENTO DEL PROBLEMA

Existe la necesidad de contar con un Modelo Pseudo Geoidal de la Ciudad de La Paz, aplicando análisis estadístico de regresión polinomial múltiple, que permita llevar alturas elipsoidales obtenidas mediante el empleo de equipos GNSS, a alturas niveladas (geométricas), considerando que la red de nivelación nacional y municipal corresponden a alturas niveladas.

1.4. FORMULACIÓN DEL PROBLEMA

¿La aplicación de la regresión polinomial múltiple, en la generación del modelo pseudo geoidal de la Ciudad de La Paz, permitirá la determinación de un modelo que alcance precisiones para el uso en trabajos topográficos y geodésicos?

1.5. HIPÓTESIS

Hipótesis Alternativa (Ha): La aplicación de la regresión polinomial múltiple, si permite generar el modelo pseudo geoidal de la Ciudad de La Paz.

Hipótesis Nula (Ho): La aplicación de la regresión polinomial múltiple, no permite generar el modelo pseudo geoidal de la Ciudad de La Paz.

1.6. OBJETIVOS

1.6.1. Objetivo General

Determinar el Modelo Pseudo Geoidal del Municipio de la Ciudad de La Paz, a través del análisis estadístico de regresión polinomial múltiple aplicado a datos de alturas elipsoidales y alturas niveladas de la Red Geodésica Municipal de La Paz 2005 – 2006.

1.6.2. Objetivos Específicos

-) Delimitar el área de estudio de la investigación.
-) Definir la información correspondiente a la red geodésica del municipio de La Paz para la generación del modelo.
-) Aplicar la regresión lineal y polinomial múltiple para generar el modelo pseudo geoidal de la Ciudad de La Paz.
-) Determinar valores de ondulación geoidal a partir de modelos globales.
-) Realizar la validación del modelo pseudo geoidal determinado mediante regresión polinomial múltiple.

CAPÍTULO II

MATERIALES Y METODOLOGÍA

2.1. MATERIALES Y EQUIPOS

a) Información recopilada

- Red Geodésica del Municipio de La Paz (2005 – 2006).
- Red Geodésica del Municipio de La Paz (2012 - 2013).
- Macrodistrictos de la Ciudad de La Paz – Gobierno Municipal Autónomo de La Paz.

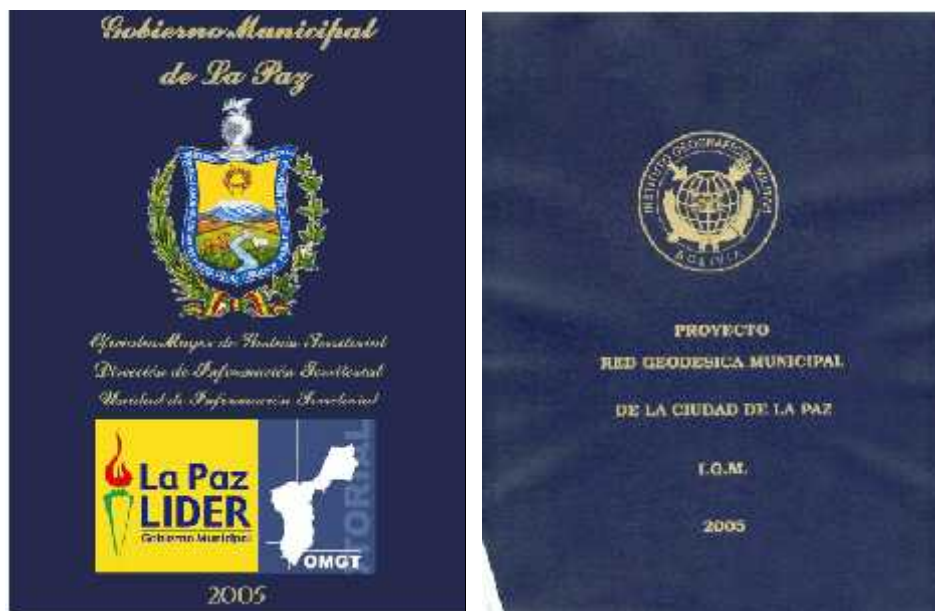


Figura 3: Portadas del Informe del proyecto Red Geodésica Municipal de la Ciudad de La Paz 2005
Fuente: Propia.

b) Programas (Software)

- Microsoft Excel 2010
- Microsoft Word 2010
- Minitab 18 - Versión 18.1
- Google Earth Pro, Versión 7.3.2.5776 (64-bit)
- ArcGIS 10.0 – ArcMap (ESRI)



*Figura 4: Programas utilizados para la investigación.
Fuente: Propia.*

c) Equipo de computación

- Computadora Core i5
- Red WiFi de internet
- Impresora

2.2. METODOLOGÍA

2.2.1. Tipo de investigación

La investigación corresponde a un análisis cuantitativo (determinación de valores de ondulación pseudo geoidal (N), en función de valores de alturas elipsoidales y alturas niveladas), del tipo correlacional – transversal, puesto que mediante el análisis estadístico tiene el propósito de relacionar el comportamiento de las ondulaciones pseudo geoidales con respecto a una posición definida mediante coordenadas proyectadas UTM.

Asimismo, se realizó comparaciones con modelos geoidales globales más conocidos y utilizados en el país, así también, se vio como alternativa la generación de un modelo geoidal a partir de los datos de la investigación y modelos globales aplicados a Geoestadística mediante el programa ArcGIS.

2.2.2. Método experimental

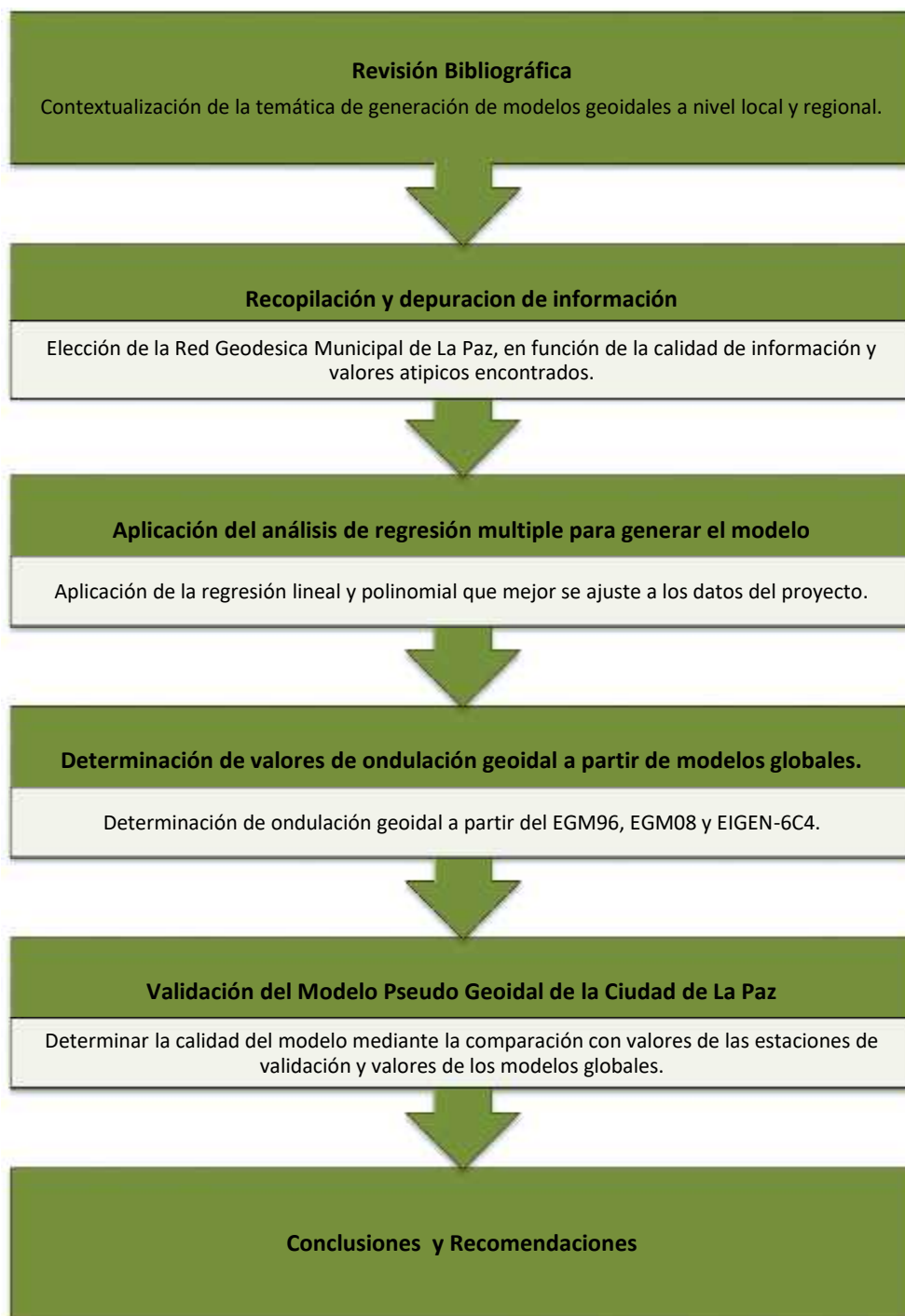


Figura 5: Diseño del método experimental.
Fuente: Elaboración Propia.

CAPÍTULO III

APLICACIÓN PRÁCTICA

3.1. DELIMITACIÓN DEL ÁREA DE ESTUDIO

Para la investigación se delimitó espacialmente como área de estudio la parte urbana de la Ciudad de La Paz, dentro del Municipio de Nuestra Señora de La Paz, distribuida administrativamente por nueve macrodistritos.

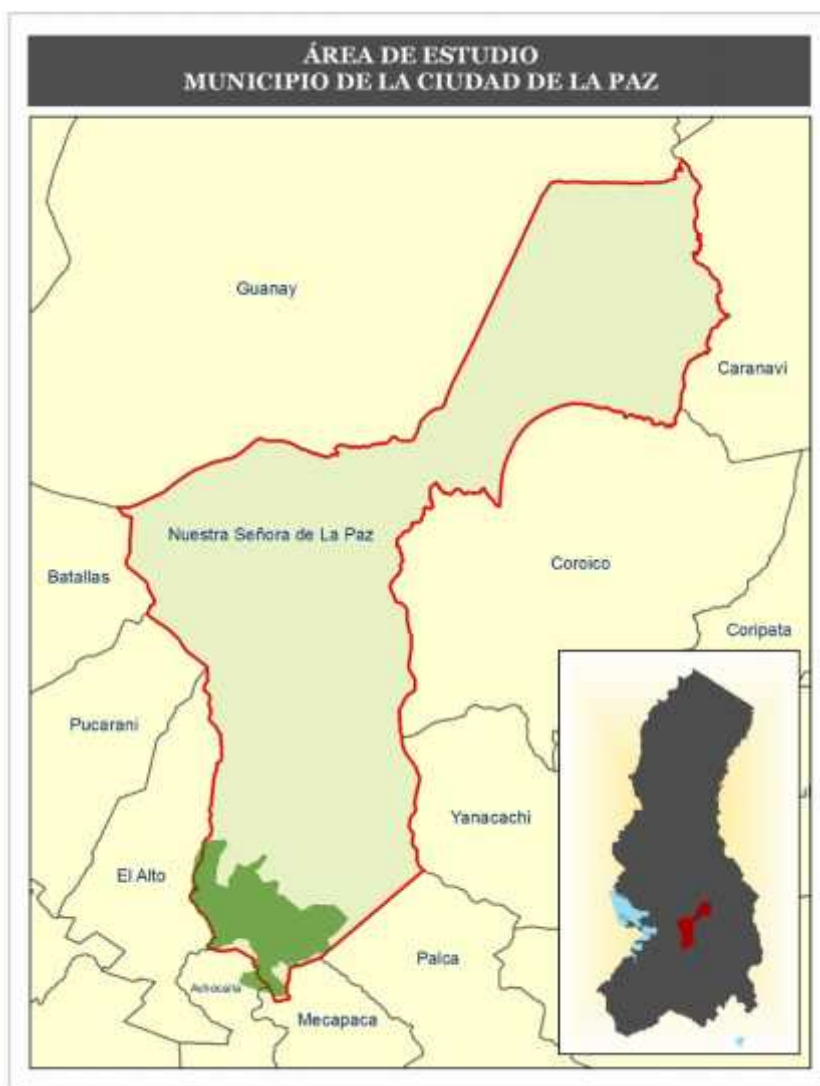


Figura 6: Ubicación geográfica del área de estudio.
Fuente: Elaboración propia.

3.2. RECOPIACIÓN DE INFORMACIÓN

El Gobierno Autónomo Municipal de La Paz, cuenta con dos campañas realizadas para el establecimiento de la Red Geodésica Municipal, las cuales fueron consideradas para la investigación.

La primera se ejecutó a través del Instituto Geográfico Militar (IGM), durante los años 2005 y 2006, estableciendo 114 puntos vinculados a la Red Geodésica MARGEN (Marco de Referencia Geodésico Nacional), administrado por el IGM, cuyos resultados se obtuvieron en coordenadas de la proyección Universal Transversa de Mercator (UTM), y Geodésicas en el sistema de referencia WGS 84.

De estos puntos establecidos, 113 cuentan con nivelación geométrica determinando en cada punto bancos de nivel referidos al nivel medio del mar (Datum Vertical Arica – República de Chile), cuyo orden de precisión corresponde al tercer orden geodésico, con una tolerancia de $12 \text{ mm } k$, donde k , es la distancia en kilómetros, y la metodología empleada corresponde a una nivelación geométrica de doble corrida (ida y regreso) (Instituto Geográfico Militar, 2006).

La segunda campaña también fue ejecutada por el IGM, en los años 2012 y 2013, realizando la reposición de monumentos en una cantidad de 44 puntos y 312 vértices nuevos monumentados con medición GPS y nivelación geométrica de segundo orden las cuales iniciaron corridas de nivelación a partir de los BM (Red Geodésicas de La Paz 2005), más cercanos a cada vértice, realizando corridas de ida y regreso en cada tramo y verificando el cumplimiento de la tolerancia mediante la fórmula $8 \text{ mm } k$. Y donde todas las estaciones se encuentran referenciadas a la Red MARGEN-ROC para la época 2011.2, ITRF 2005 realización IGS 2005 (Instituto Geográfico Militar, 2013).

Revisando la información de ambas campañas, se vio conveniente hacer uso para la investigación la campaña de los años 2005 – 2006, toda vez que para la campaña 2013, los puntos establecidos como BM, fueron la base para realizar la nivelación geométrica, además que la red geodésica 2013 presenta una serie de inconsistencias en cuanto a valores de las coordenadas horizontales y verticales.

3.3. DEPURACIÓN DE LA INFORMACIÓN RED GEODÉSICA 2005 - 2006

3.3.1. Prueba de los valores atípicos

Determinada como información confiable la Red Geodésica de la Ciudad de La Paz (Campaña 2005 – 2006), se procedió a realizar la depuración de la información apoyada en una serie de análisis estadísticos.

Para identificar aquellos valores de N (Ondulación pseudo geoidal), que puedan generar inconsistencias en el momento de generar el modelo, se utilizó el programa Minitab 18.

Todas las pruebas de valores atípicos de Minitab están diseñadas para detectar un único valor atípico en una muestra. Por lo general, la prueba de Grubbs funciona adecuadamente. Sin embargo, si una muestra contiene más de un posible valor atípico, entonces la prueba de Grubbs y la relación Q de Dixon podrían no ser efectivas. Si no sabe si los datos incluyen valores atípicos, es recomendable utilizar la prueba de Grubbs. Si se conoce que los datos incluyen uno o más valores atípicos, se debe utilizar una de las pruebas de Dixon. Las pruebas de Dixon están diseñadas para superar el efecto de enmascaramiento que pueden causar múltiples valores atípicos posibles.⁴

Las muestras más grandes de una población normal tienen mayor probabilidad de incluir valores extremos. Dixon propuso las siguientes directrices generales para las relaciones.

Tamaño de la muestra (n)	Relación recomendada
$3 \leq n \leq 7$	r10 (también denominada relación Q de Dixon)
$8 \leq n \leq 10$	r11
$11 \leq n \leq 13$	r21
$n \geq 14$	r22

Cuadro 2: Tamaño de muestra y relación recomendada por Dixon.
Fuente: Soporte Minitab 18.

⁴ Soporte Minitab 18, Prueba de valores atípicos.

Considerando que la Red Geodésica de la Ciudad de La Paz, presenta 113 estaciones, se utilizó la Prueba de valores atípicos “Relación r22 de Dixon”, considerando el Nivel de significancia igual a 0.05.

Un nivel de significancia (denotado como α o alfa) de 0.05 funciona adecuadamente. Un nivel de significancia de 0.05 indica un riesgo de 5% de concluir que existe una diferencia cuando no hay una diferencia real.

Para realizar el análisis estadístico, es importante considerar que los datos presentan una distribución normal, aspecto que puede corroborarse mediante el “valor p”. Este valor nos permite ser comparado con el nivel de significancia para decidir si puede rechazar o no la hipótesis nula (H_0). Entenderemos para la prueba, las siguientes hipótesis:

Hipótesis nula	Todos los valores de los datos provienen de la misma población normal
Hipótesis alterna	El valor más pequeño o más grande de los datos es un valor atípico
Nivel de significancia	$\alpha = 0.05$

Cuadro 3: Hipótesis nula y alterna para la prueba de valores atípicos.
Fuente: Minitab 18.

La interpretación de la prueba corresponde a: Si el valor p es menor que el nivel de significancia, la interpretación es que los resultados son estadísticamente significativos y por tanto se rechaza H_0 (Hipótesis nula).

Variable	N	Mín.	x[2]	x[3]	x[N-2]	x[N-1]	Máx.	r22	P
N	113	16.370	44.708	45.445	52.744	86.435	86.953	0.82	0.000

x[i] denota la observación *i*-ésima más pequeña.

Cuadro 4: Prueba de relación r22 de Dixon – Primer análisis.
Fuente: Minitab 18.

Variable	Fila	Valor atípico
N	3	86.9533

Cuadro 5: Valor atípico determinado – Primer análisis.
Fuente: Minitab 18.

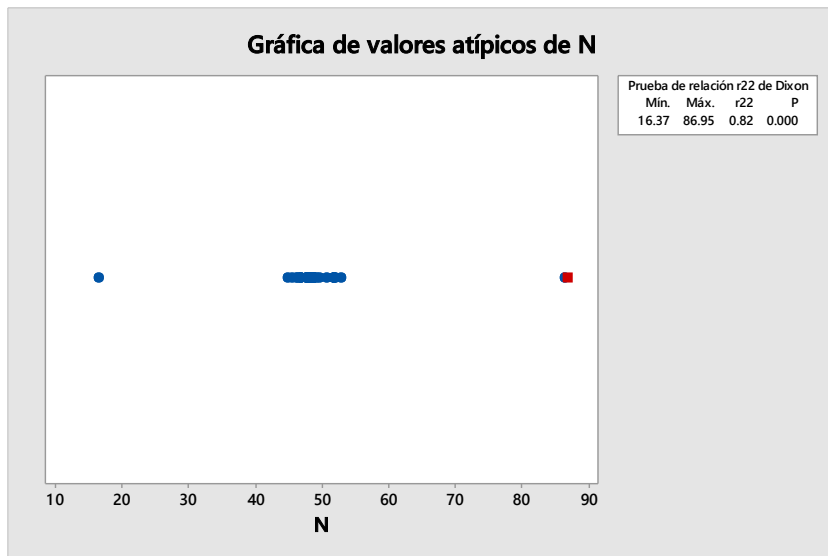


Figura 7: Gráfica de valores atípicos de N – Primer análisis.
Fuente: Minitab 18.

Como se puede observar en los cuadros y figuras anteriores, la prueba de valores atípicos identifica que valores de N, corresponden a valores irregulares en relación a la muestra total de la Red Geodésica de la Ciudad de La Paz. Por lo cual, se hizo el análisis estadístico hasta que la prueba evidencie que no existe más valores atípicos, identificándose 21 estaciones con valores irregulares considerado atípicos.

SISTEMA DE REFERENCIA WGS-84, COORDENADAS UTM ZONA 19S				
No.	ESTACIÓN	ESTE	NORTE	N
1	AFLO	597672.300	8170253.273	86.953
2	CHAJ	601420.008	8174133.052	46.838
3	OVEJ	603359.657	8171188.736	51.814
4	PC15	602483.745	8177389.930	49.633
5	PC16	605056.067	8171364.647	52.744
6	PC17	603840.388	8169020.771	51.849
7	PC18	601938.774	8167817.436	44.709
8	PC22	601138.937	8167030.134	45.445
9	PC50	600983.083	8176697.296	46.190
10	PC56	601606.780	8173906.752	46.814
11	PC57	603579.310	8174205.521	47.764

12	PC58	604015.177	8173332.713	50.673
13	PC59	602386.251	8173148.936	51.828
14	PC60	602961.078	8172223.323	51.789
15	PC62	605001.677	8170369.006	51.794
16	PC53	602067.649	8175171.580	49.129
17	PC72	597187.761	8164734.981	46.719
18	PC80	600638.043	8163568.259	46.375
19	PFM1	592177.913	8171395.289	47.736
20	PC74	598081.063	8168836.775	86.435
21	PC81	598038.011	8163945.405	16.371

Cuadro 6: Estaciones identificadas con valores atípicos.
Fuente: Elaboración propia.

Excluyendo de la lista general las estaciones con valores atípicos, se quedó con un total de 92 estaciones (Anexo 1), que fueron utilizadas para realizar los análisis estadísticos posteriores, que permitieron generar el modelo pseudo geoidal de la Ciudad de La Paz.

3.3.2. Determinación de las estaciones de validación y del modelo

Para realizar la validación del modelo pseudo geoidal de la Ciudad de La Paz, se determinó un conjunto de estaciones de validación, las cuales permitieron establecer el grado de precisión a las que puede llegar el modelo.

Para tal efecto, se procedió a definir estaciones de validación de la Red Geodésica de La Paz, que se encuentren homogéneamente distribuidas en toda el área de estudio. Estas estaciones de validación fueron seleccionadas de las 92 estaciones anteriormente depuradas mediante el análisis estadístico de prueba de valores atípicos. Se determinaron un 10% de vértices de validación del total de vértices depurados, definiéndose 8 vértices distribuidos en toda el área de estudio.

SISTEMA DE REFERENCIA WGS-84, COORDENADAS UTM ZONA 19S				
No.	ESTACIÓN	NORTE	ESTE	N
1	085X	8169107.842	600085.966	47.820
2	ACHA	8178692.407	591213.178	48.394

3	BAIR	8173101.400	591909.120	47.895
4	JUPA	8165614.003	599074.723	47.828
5	PC12	8174169.317	596977.025	47.973
6	PC24	8167209.210	595282.368	47.971
7	PC31	8181723.486	591057.266	48.363
8	PC42	8180386.566	595624.980	48.561

Cuadro 7: Estaciones para validación del modelo.
Fuente: Elaboración propia.

Excluyendo las estaciones que fueron utilizadas para validar el modelo, el listado de estaciones consideradas para el análisis estadístico corresponden a 84 estaciones, las mismas que observando su ubicación en el área de estudio, se encuentran distribuidas de manera homogénea.

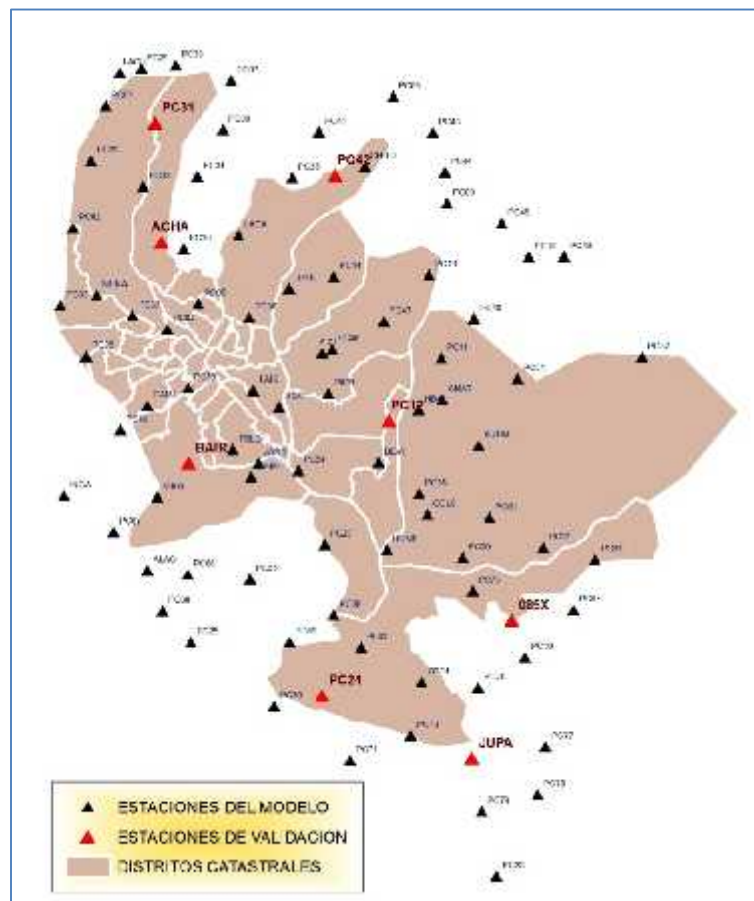


Figura 8: Distribución de las estaciones de validación y del modelo en el área de estudio.
Fuente: Elaboración propia.

3.4. GENERACIÓN DEL MODELO A PARTIR DEL ANÁLISIS ESTADÍSTICO

Para una mejor comprensión del análisis estadístico realizado para generar el modelo pseudo geoidal de la Ciudad de La Paz, en función de los valores de ondulación pseudo geoidal (N), y valores de coordenadas Este y Norte en proyección UTM. La investigación se desarrolló partiendo del modelo determinado mediante la regresión lineal múltiple para finalmente aplicar la regresión polinomial múltiple que permitió en función al grado del polinomio, ajustar a la tendencia de distribución de los valores de N respecto a las coordenadas proyectadas horizontales.

3.4.1. Aplicación de la Regresión Lineal Múltiple

Para generar el modelo pseudo geoidal de la Ciudad de La Paz, se consideró como primera alternativa apoyarse en la Regresión Lineal Múltiple, considerando que este análisis estadístico puede utilizar más de una variable independiente (X_1, X_2, \dots, X_n), (para la investigación corresponden como variables independientes a las coordenadas UTM de cada estación), que permiten relacionar con una variable respuesta o dependiente Y (los valores de la ondulación pseudo geoidal - N), asimismo, la regresión lineal múltiple permite entender o predecir el comportamiento de la variable respuesta (N), con respecto a sus variables independientes.

La regresión lineal múltiple tiene la ventaja de trabajar con más de una variable independiente (en el caso del análisis estadístico realizado, las variables independientes corresponderán a los valores de coordenadas; Coordenada Este (X_1) y Coordenada Norte (X_2)), que permiten determinar la variable dependiente Y (que en el caso de la investigación llegan a ser los valores de N).

Para mejor entendimiento, las variables consideradas en la regresión lineal múltiple corresponden a:

Variables Independientes : Coordenada Este, Coordenada Norte (UTM Zona 19 Sur)
Variables Dependientes : Valor de N (Ondulación pseudo geoidal)

El modelo de regresión lineal múltiple con más de una variable independiente, es el polinomio de primer orden, cuya expresión se representa como:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \epsilon$$

Dónde:

- Y = Variable dependiente
- X_1, X_2, \dots, X_p = Variables independientes
- $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ = Coeficientes de regresión
- ϵ = Error

La solución se realizó mediante el método de mínimos cuadrados mediante el programa Excel, de manera secuencial y explicativa, a fin de entender todos los procedimientos necesarios y sobre todo las interpretaciones de los valores estadísticos obtenidos.

3.4.1.1. Prueba de hipótesis en la regresión lineal múltiple

La hipótesis global más importante sobre un modelo de regresión múltiple es ver si la regresión es significativa. Esto se logra probando las siguientes hipótesis.

- Hipótesis Nula (H_0) : $\beta_1 = \beta_2 = \dots = \beta_k = 0$
- Hipótesis Alterna (H_A) : $\beta_j \neq 0$ (para al menos un $j = 1, 2, \dots, k$)

Aceptar la hipótesis nula (H_0), significa que ningún término o variable independiente en el modelo tiene una contribución significativa al explicar la variable de respuesta Y . Mientras que rechazar (H_0), implica que por lo menos un término en el modelo contribuye de manera significativa al ajuste. Para hacer la prueba de estas hipótesis el procedimiento es similar al de la regresión simple. Primeramente se obtiene la tabla para el análisis de varianza y con los valores de F (valor crítico) y F_0 (valor calculado), se comparan y se obtiene una conclusión. Se rechaza H_0 si $F_0 > F_{(\alpha, k, n-k-1)}$. Donde k es el igual al número de variables independientes.

Para la investigación interpretamos que, aceptar la hipótesis nula (H_0), representaría que los valores de las coordenadas Este y Norte de cada estación, no tienen ninguna relación con

los valores determinados para N, o por lo menos no contribuyen significativamente al ajuste y determinación de N.

El procedimiento de estimación por mínimos cuadrados es más sencillo si se representa en notación matricial, donde el modelo en término de datos y ecuación puede describirse como:

$$Y = X \beta + \epsilon$$

La representación de forma matricial resulta ser:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

Para dar solución y determinar el modelo por regresión lineal múltiple, es necesario determinar los valores de los coeficientes beta (β), que representan el vector de los estimadores de mínimos cuadrados, para lo cual de la ecuación anterior es necesario despejar β , donde el resultado se reduce a la solución:

$$\beta = A^{-1}g = (X'X)^{-1}X'y$$

Matricialmente podemos representarlas como:

$$A = X'X = \begin{bmatrix} n & \sum_{i=1}^n X_{i1} & \sum_{i=1}^n X_{i2} & \cdots & \sum_{i=1}^n X_{ip} \\ \sum_{i=1}^n X_{i1} & \sum_{i=1}^n X_{i1}^2 & \sum_{i=1}^n X_{i1}X_{i2} & \cdots & \sum_{i=1}^n X_{i1}X_{ip} \\ \sum_{i=1}^n X_{i2} & \sum_{i=1}^n X_{i1}X_{i2} & \sum_{i=1}^n X_{i2}^2 & \cdots & \sum_{i=1}^n X_{i2}X_{ip} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^n X_{ip} & \sum_{i=1}^n X_{i1}X_{ip} & \sum_{i=1}^n X_{i2}X_{ip} & \cdots & \sum_{i=1}^n X_{ip}^2 \end{bmatrix}$$

El vector g resulta ser:

$$g = X'y = \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_{i1}Y_i \\ \sum_{i=1}^n X_{i2}Y_i \end{pmatrix}$$

Para la solución del modelo se procedió a resolver mediante la hoja electrónica Excel, que cuenta con la herramienta de Análisis de datos (Complemento de Microsoft Excel que permite analizar datos científicos y financieros).

3.4.1.2. Procedimiento realizado en Microsoft Excel mediante el Análisis de Datos.

Considerando que tenemos dos variables independientes (Coordenada Este y Coordenada Norte) y una variable independiente de predicción (N – Ondulación pseudo geoidal), el modelo de regresión lineal múltiple para este caso se define como:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

Dónde:

- Y : *N – Ondulación pseudo geoidal*
- X_1 : *Coordenada Este*
- X_2 : *Coordenada Norte*
- ϵ : *Error*

- Estructura de datos

Es recomendable establecer una estructura de los datos que nos permita entender que valores corresponden a las variables independientes y cuales a las dependientes, en tal sentido los valores fueron ordenados de la siguiente manera:

No.	ESTACIÓN	X ₁ : ESTE	X ₂ : NORTE	Y: N
1	ALAC	590859.784	8170386.780	48.312
2	ARUN	599246.172	8173542.594	47.734
3	BEVI	596706.068	8173120.005	48.103
4	BPUL	594444.113	8177535.495	48.309
43	PC28	595579.583	8169244.231	47.950
44	PC29	590705.303	8183091.160	48.547
45	PC30	591579.587	8183176.162	48.499
46	PC32	589427.012	8180756.494	48.473

5	CALV	590855.881	8174560.981	48.345	47	PC33	590747.664	8180107.898	48.514
6	CHUQ	596359.121	8180612.610	48.374	48	PC34	592121.462	8180356.428	48.575
7	CMAT	598329.107	8174700.545	47.753	49	PC35	594521.284	8180334.556	48.686
8	COL6	597948.742	8171789.800	47.843	50	PC36	591781.731	8178536.647	48.450
9	FELC	593020.519	8173433.945	48.306	51	PC37	590487.067	8176840.537	48.206
10	HB83	597722.255	8174452.041	47.732	52	PC38	593434.804	8176784.820	48.123
11	HUMB	596931.142	8170900.732	47.991	53	PC39	591901.952	8175003.477	47.852
12	IGM1	594196.788	8174522.075	48.173	54	PC40	590177.611	8173946.269	48.347
13	INGA	588751.181	8172270.466	48.390	55	PC41	595208.321	8181493.306	48.497
14	JARD	593673.237	8173091.328	48.235	56	PC43	598090.267	8181460.296	48.172
15	LACA	593170.421	8178864.897	48.320	57	PC44	598387.636	8180477.034	48.435
16	LAGU	590169.314	8182988.746	48.598	58	PC45	599836.307	8179174.627	47.872
17	LAIK	593544.602	8174947.552	48.124	59	PC46	595532.826	8175988.994	48.069
18	MAME	593482.947	8172736.320	48.190	60	PC47	596842.752	8176695.979	48.010
19	MUNA	589571.325	8177342.266	48.608	61	PC48	599132.324	8176761.243	47.863
20	P028	595442.053	8174868.639	48.155	62	PC49	601412.767	8178330.774	48.123
21	P043	596271.576	8168418.133	47.928	63	PC51	600238.521	8175232.091	48.183
22	P-27	595288.695	8175890.002	48.142	64	PC52	603379.291	8175788.311	48.555
23	PC01	589804.984	8182138.836	48.464	65	PC54	594680.366	8172920.575	48.153
24	PC02	588982.341	8179056.176	48.531	66	PC55	597734.783	8172319.347	47.916
25	PC03	588655.830	8177087.146	48.606	67	PC61	599527.060	8171718.407	48.048
26	PC04	591361.556	8176483.866	48.200	68	PC63	602194.356	8170653.206	47.836
27	PC05	592151.440	8177132.830	48.308	69	PC64	601652.291	8169377.663	47.701
28	PC06	592771.371	8181538.531	48.569	70	PC65	593454.029	8170162.507	48.189
29	PC07	592976.097	8182789.885	48.582	71	PC66	591885.034	8170269.057	48.285
30	PC08	597081.058	8182386.140	48.804	72	PC67	590007.686	8171354.938	48.336
31	PC09	598443.415	8179673.100	48.150	73	PC68	591246.873	8169355.277	48.223
32	PC10	600514.072	8178320.897	48.105	74	PC69	594453.414	8168569.568	47.987
33	PC11	598299.396	8175753.484	48.088	75	PC70	594067.491	8166942.902	47.982
34	PC13	597993.173	8177865.794	48.094	76	PC71	595995.273	8165567.617	48.106
35	PC14	595577.062	8177826.392	47.784	77	PC73	597524.531	8166207.127	47.781
36	PC19	600424.857	8168176.664	47.935	78	PC75	599102.323	8169865.026	47.878
37	PC20	598846.822	8170695.924	47.739	79	PC76	599229.319	8167400.540	48.089
38	PC21	600886.937	8170954.256	47.776	80	PC77	600939.907	8165907.848	47.693
39	PC23	599702.764	8162633.956	47.765	81	PC78	600738.293	8164698.951	47.802
40	PC25	591961.420	8168548.034	48.069	82	PC79	599323.422	8164274.580	47.684
41	PC26	589312.774	8175815.962	48.483	83	VIRG	591104.498	8172237.546	48.334
42	PC27	595355.933	8171030.182	47.835	84	ZOOL	597810.816	8167541.373	47.929

Cuadro 8: Valores de las variables independientes y dependientes en el análisis de datos.

Fuente: Elaboración propia.

- Aplicación de la herramienta Análisis de Datos

Esta herramienta resulta ser bastante sencilla en su uso, dentro del submenú que cuenta se define la opción “Regresión”, con la cual se procederá a dar solución al modelo de Regresión lineal múltiple planteada.

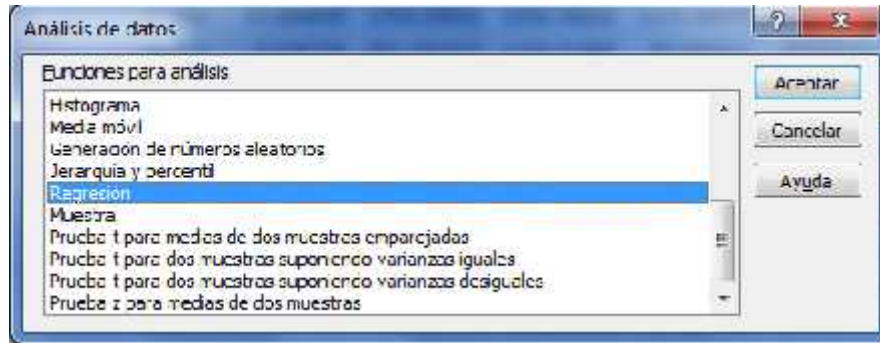


Figura 10: Submenú de Análisis de datos de Excel.
Fuente: Elaboración propia.

Definida la opción de regresión, se procede a establecer los valores para el análisis y solución del modelo.

Se considera como valores de entrada:

- Rango Y de entrada : Los valores (Y: N – Ondulación pseudo geoidal)
- Rango X de entrada : Los valores (X_1 : Coordenada ESTE y X_2 : Coordenada NORTE)
- Nivel de confianza : 95%
- Residuales : Todas las determinadas

Esta selección permite que la herramienta solucione el modelo aplicando mínimos cuadrados, además de obtener una serie de valores estadísticos que deben ser interpretados a fin de verificar la calidad del ajuste de regresión lineal múltiple que garanticen la predicción de la variable dependiente.

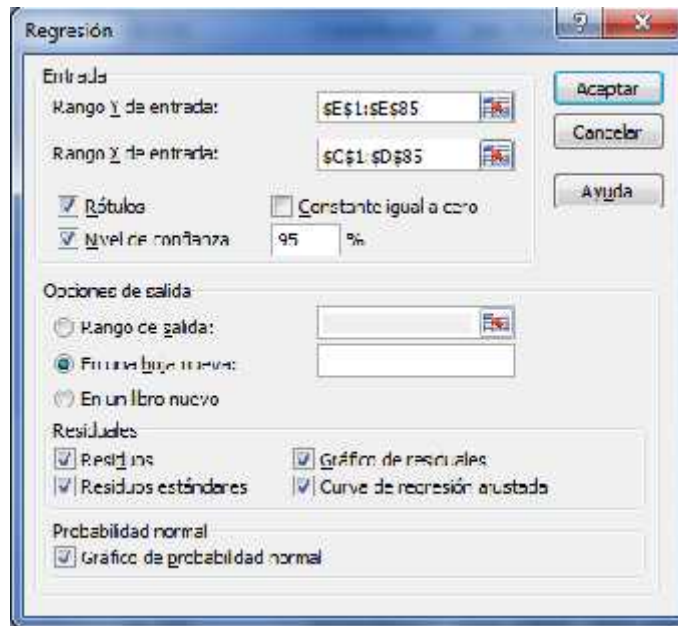


Figura 11: Elección de valores de entrada y opciones de salida para la regresión.
Fuente: Elaboración propia.

Como resultado se genera una serie de valores estadísticos que a continuación se detallan y explican.

- Coeficientes determinados

Coeficientes	
Intercepción	-167.570165338
X1: ESTE	-0.000034073
X2: NORTE	0.000028872

Cuadro 9: Valores de los coeficientes determinados en la regresión lineal.
Fuente: Elaboración propia.

Los resultados de los coeficientes permiten establecer la ecuación de predicción del valor de la ondulación pseudo geoidal (N), en función de la siguiente interpretación:

$$Y = 0 + {}_1X_1 + {}_2X_2$$

$$\begin{aligned}
 0 &= -167.570165338 \\
 1 &= -0.000034073 \\
 2 &= 0.000028872
 \end{aligned}$$

Por tanto, resultando ser el modelo igual a:

$$\mathbf{N} = -167.570165338 + -0.000034073 (\mathbf{ESTE}) + 0.000028872 (\mathbf{NORTE})$$

Esta ecuación permite determinar el valor de N, simplemente introduciendo los valores de las coordenadas Este y Norte de un punto cualquiera que se encuentre dentro del área de estudio. Sin embargo, antes de su aplicación es importante verificar la calidad del ajuste de la regresión lineal con valores estadísticos que resuelve la herramienta de Excel.

- Ondulaciones pseudo geoidales estimadas y residuales

Observación	Estación	N (m)	Pronóstico Y: N (m)	Residuos (m)	Observación	Estación	N (m)	Pronóstico Y: N (m)	Residuos (m)
1	ALAC	48.312	48.193	0.119	43	PC28	47.950	47.999	-0.049
2	ARUN	47.734	47.998	-0.264	44	PC29	48.547	48.565	-0.018
3	BEVI	48.103	48.072	0.031	45	PC30	48.499	48.537	-0.038
4	BPUL	48.309	48.277	0.032	46	PC32	48.473	48.541	-0.068
5	CALV	48.345	48.313	0.032	47	PC33	48.514	48.477	0.037
6	CHUQ	48.374	48.301	0.073	48	PC34	48.575	48.438	0.137
7	CMAT	47.753	48.063	-0.310	49	PC35	48.686	48.355	0.331
8	COL6	47.843	47.992	-0.149	50	PC36	48.450	48.397	0.053
9	FELC	48.306	48.207	0.099	51	PC37	48.206	48.392	-0.186
10	HB83	47.732	48.076	-0.344	52	PC38	48.123	48.290	-0.167
11	HUMB	47.991	48.001	-0.010	53	PC39	47.852	48.291	-0.439
12	IGM1	48.173	48.198	-0.025	54	PC40	48.347	48.319	0.028
13	INGA	48.390	48.319	0.071	55	PC41	48.497	48.365	0.132
14	JARD	48.235	48.175	0.060	56	PC43	48.172	48.266	-0.094
15	LACA	48.320	48.359	-0.039	57	PC44	48.435	48.228	0.207

16	LAGU	48.598	48.580	0.018	58	PC45	47.872	48.141	-0.269
17	LAIK	48.124	48.233	-0.109	59	PC46	48.069	48.195	-0.126
18	MAME	48.190	48.171	0.019	60	PC47	48.010	48.171	-0.161
19	MUNA	48.608	48.437	0.171	61	PC48	47.863	48.095	-0.232
20	P028	48.155	48.166	-0.011	62	PC49	48.123	48.063	0.060
21	P043	47.928	47.952	-0.024	63	PC51	48.183	48.013	0.170
22	P-27	48.142	48.201	-0.059	64	PC52	48.555	47.922	0.633
23	PC01	48.464	48.568	-0.104	65	PC54	48.153	48.136	0.017
24	PC02	48.531	48.507	0.024	66	PC55	47.916	48.014	-0.098
25	PC03	48.606	48.461	0.145	67	PC61	48.048	47.936	0.112
26	PC04	48.200	48.352	-0.152	68	PC63	47.836	47.814	0.022
27	PC05	48.308	48.344	-0.036	69	PC64	47.701	47.796	-0.095
28	PC06	48.569	48.450	0.119	70	PC65	48.189	48.098	0.091
29	PC07	48.582	48.479	0.103	71	PC66	48.285	48.154	0.131
30	PC08	48.804	48.327	0.477	72	PC67	48.336	48.250	0.086
31	PC09	48.150	48.202	-0.052	73	PC68	48.223	48.150	0.073
32	PC10	48.105	48.093	0.012	74	PC69	47.987	48.018	-0.031
33	PC11	48.088	48.094	-0.006	75	PC70	47.982	47.984	-0.002
34	PC13	48.094	48.166	-0.072	76	PC71	48.106	47.879	0.227
35	PC14	47.784	48.247	-0.463	77	PC73	47.781	47.845	-0.064
36	PC19	47.935	47.803	0.132	78	PC75	47.878	47.897	-0.019
37	PC20	47.739	47.930	-0.191	79	PC76	48.089	47.821	0.268
38	PC21	47.776	47.867	-0.091	80	PC77	47.693	47.720	-0.027
39	PC23	47.765	47.668	0.097	81	PC78	47.802	47.692	0.110
40	PC25	48.069	48.102	-0.033	82	PC79	47.684	47.728	-0.044
41	PC26	48.483	48.402	0.081	83	VIRG	48.334	48.238	0.096
42	PC27	47.835	48.058	-0.223	84	ZOOL	47.929	47.874	0.055

Cuadro 10: Valores de N estimados y residuales por estación.
Fuente: Microsoft Excel, Análisis de datos - Regresión

Este cuadro permite apreciar la diferencia (residual), que existe entre el valor de N determinado con valores de la Red Geodésica de la Ciudad de La Paz 2005-2006 y los valores de N estimados a partir de la ecuación generada con los valores de los coeficientes beta ().

- Análisis de Varianza

El análisis de varianza nos permitirá realizar la prueba de hipótesis, para determinar si existe una relación entre las variables independientes, con la variable dependiente, recordando que para la investigación es importante rechazar la hipótesis nula.

	<i>Grados de libertad</i>	<i>Suma de cuadrados</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Valor crítico de F (Alfa 5%)</i>
Regresión	2	4.246316	2.123158	73.234687	3.109311
Residuos	81	2.348283	0.028991		
Total	83	6.594599			

Cuadro 11: Análisis de varianza para la regresión lineal.
Fuente: Microsoft Excel, Análisis de datos - Regresión

Un valor crítico ($F_{\text{Crítico}}$), es un punto en la distribución del estadístico de prueba bajo la hipótesis nula que define un conjunto de valores que apoyan el rechazo de la hipótesis nula, este conjunto se denomina región crítica o de rechazo. Por lo general, las pruebas unilaterales tienen un valor crítico y las pruebas bilaterales tienen dos valores críticos. Los valores críticos se determinan de manera que la probabilidad de que el estadístico de prueba tenga un valor en la región de rechazo de la prueba (cuando la hipótesis nula sea verdadera), sea igual al nivel de significancia (denotado como α o alfa).⁵

La interpretación del análisis de varianza corresponde a comparar los valores de F y el Valor crítico de F (denominada también como $F_{\text{Crítica}}$). Si el valor de F es menor a $F_{\text{Crítica}}$ entonces se acepta la hipótesis nula (H_0), lo cual significa que entre las variables independientes (Coordenada Este y Coordenada Norte) y la variable dependiente (N – Ondulación pseudo geoidal), no existe ninguna relación, vale decir, que en el modelo ningún término variable contribuye significativamente a la variable dependiente, en resumen podemos interpretarla que las coordenadas Este y Norte no influyen en la determinación del valor de N. En caso de

⁵ Minitab 18, Soporte de Minitab, 2020.

que el valor de F es mayor a $F_{Crítica}$ entonces se rechaza la hipótesis nula (H_0), por tanto, si existirá relación entre variables independientes y dependientes.

-) Si $F < F_{Crítica}$, entonces se acepta la hipótesis nula (H_0)
-) Si $F > F_{Crítica}$, entonces se rechaza la hipótesis nula (H_0)

Comparando los valores obtenidos tenemos:

$$F (73.234687) > F_{Crítica} (3.109311)$$

Por tanto, rechazamos la hipótesis nula (H_0), determinando que si existe relación entre las variables independientes (Coordenada Este y Coordenada Norte), con la variable dependiente (N – Ondulación pseudo geoidal).

Una manera sencilla de interpretar y entender los valores del análisis de varianza corresponde a una representación gráfica que se muestra a continuación.



Figura 12: Representación de los valores de F y $F_{Crítica}$ del análisis de varianza.
Fuente: Elaboración propia.

- Correlación

Los valores determinados para la correlación son muy importantes siendo que nos permiten evaluar la calidad del ajuste de la regresión lineal múltiple, para lo cual es imprescindible interpretar los valores mostrados a continuación.

<i>Estadísticas de la regresión</i>	
Coefficiente de determinación R ²	0.643908
R ² ajustado	0.635116
Coefficiente de correlación múltiple	0.802439
Error típico	0.170268
Observaciones	84

Cuadro 12: Valores de correlación de la regresión lineal.
Fuente: Microsoft Excel, Análisis de datos - Regresión

La interpretación de estos valores corresponden a:

) Coeficiente de determinación R²

En estadística, el coeficiente de determinación, denominado R² y pronunciado R cuadrado, es un estadístico usado en el contexto de un modelo estadístico cuyo principal propósito es predecir futuros resultados o probar una hipótesis. El coeficiente determina la calidad del modelo para replicar los resultados, y la proporción de variación de los resultados que puede explicarse por el modelo.⁶

El valor obtenido oscila entre $0 < R^2 < 1$, interpretándose como la proporción de las variables en los datos de la predicción de (N), que es explicada por el modelo. Mientras más cerca se encuentre el valor de R² a 1 será mejor. Si el valor es 1 significaría que el modelo esta explicado al 100%, en el caso de los resultados analizados observamos que corresponde a un 64.4 %, lo cual implica que la ecuación definida como (**N** = -167.570165338 + -0.000034073 (**ESTE**) + 0.000028872 (**NORTE**)), no establece una predicción confiable para los valores de N.

) Coeficiente de determinación R² ajustado

Denominado también coeficiente de determinación corregido, mide el porcentaje de variación de la variable dependiente (al igual que el coeficiente de determinación), pero teniendo en cuenta el número de variables incluidas en el modelo. Sabemos que a medida que vamos incluyendo variables en el modelo, el coeficiente de determinación aumenta aunque las variables que incluyamos no sean significativas.

⁶ Wikipedia la enciclopedia libre, Coeficiente de determinación, 2020.

Esto supone un problema, ya que no debemos olvidar que la inclusión de nuevas variables supone un aumento en el número de parámetros a estimar para el modelo. El coeficiente de determinación corregido viene a resolver este problema del coeficiente de determinación.⁷

El R cuadrado ajustado (o coeficiente de determinación ajustado), se utiliza en la regresión múltiple para ver el grado de intensidad o efectividad que tienen las variables independientes en explicar la variable dependiente. En palabras más simples, el R cuadrado ajustado nos dice qué porcentaje de variación de la variable dependiente es explicado colectivamente por todas las variables independientes.⁸

Para fines de predicción o pronósticos se recomienda que el coeficiente de determinación ajustado deba estar al menos en un 0.7.

Revisando los valores determinados en la regresión, vemos que el R cuadrado ajustado se encuentra con un valor de 0.63, un valor por debajo del recomendado, lo cual significa que el modelo determinado no sirve para realizar predicciones para el valor de N (variable dependiente).

) Coeficiente de correlación múltiple

En estadística y probabilidades, el coeficiente de correlación múltiple es una medida de qué tan bien se puede predecir una variable dada, usando una función lineal de un conjunto de otras variables independientes. Es la correlación entre los valores de la variable y las mejores predicciones que se pueden calcular linealmente a partir de las variables predictivas. El coeficiente de correlación múltiple toma valores entre 0 y 1; un valor más alto indica una alta predictibilidad de la variable dependiente a partir de las variables independientes, con un valor de 1 que indica que las predicciones son exactamente correctas y un valor de 0 que indica que ninguna combinación lineal de las variables independientes es un mejor predictor que la fija⁹.

⁷ Wikipedia la enciclopedia libre, Coeficiente de determinación, 2020.

⁸ Economipedia, R cuadrado ajustado (Coeficiente de determinación ajustado), 2020.

⁹ Wikipedia la enciclopedia libre, Correlación múltiple, 2020.

En conclusión podemos señalar que es una medida de la intensidad de la relación entre la variable dependiente (N), y el conjunto de las variables independientes o regresoras (Este y Norte). Entre más cerca se encuentre al valor de 1 resulta mejor.

-) Revisando el valor calculado en la regresión, vemos que corresponde a 0.80, por tanto la intensidad de relación entre variables independientes y dependientes de nuestro modelo podría considerarse buena, sin embargo, el coeficiente de determinación R^2 ajustado no cumple las condiciones para el modelo.

3.4.2. Aplicación de la Regresión Polinomial Múltiple

Con las interpretaciones realizadas a los resultados y valores estadísticos determinados en la regresión lineal múltiple, prácticamente se descarta el modelo determinado.

Considerando que la ecuación generada, no permite realizar estimaciones y predicciones confiables de N (Ondulación pseudo geoidal), se realizó la modificación a una Regresión polinomial múltiple, la cual en función al grado del polinomio permite ajustar la tendencia de la distribución de los valores de N.

La Regresión Polinomial es un caso especial de la Regresión Lineal, extiende el modelo lineal al agregar predictores adicionales, obtenidos al elevar cada uno de los predictores originales a una potencia. Por ejemplo, una regresión cúbica utiliza tres variables, como predictores. Este enfoque proporciona una forma sencilla de proporcionar un ajuste no lineal a los datos. El método estándar para extender la Regresión Lineal a una relación no lineal entre las variables dependientes e independientes, consiste en reemplazar el modelo lineal con una función polinomial.¹⁰

En ese sentido la ecuación del modelo determinado anteriormente con la regresión lineal múltiple, viene a ser modificada a una ecuación polinomial, cuya ecuación puede ser determinada mediante:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \beta_4 X_1^2 + \beta_5 X_2^2 + \beta_6 X_1^3 + \beta_7 X_2^3 + \dots + \beta_i X_1^i + \beta_{i+1} X_2^{i+1} +$$

¹⁰ AprendeIA, Regresión Polinomial – Teoría, Curso Machine Learning con Python, 2020.

Dónde:

- Y : N – Ondulación pseudo geoidal
 X_1 : Coordenada Este
 X_2 : Coordenada Norte
: Error
 n : Grado del polinomio
 i : Número de coeficiente en función del grado del polinomio

Algo importante para aplicar la regresión polinomial múltiple en la generación del modelo pseudo geoidal, corresponde en determinar el modelo de la ecuación que mejor se ajuste a la distribución de los valores de N , en función al grado del polinomio, pudiendo ser de tercer, cuarto, quinto, sexto o más grados. Estadísticamente, como se vio en la regresión lineal múltiple los valores del coeficiente de determinación R^2 , coeficiente de determinación R^2 ajustado y el coeficiente de correlación múltiple, permiten establecer la correlación de las variables independientes y dependiente, que denoten la calidad del ajuste de la regresión (recordando que valores más cercanos a 1 vale decir a 100%, corresponden a ser mejores).

En este sentido, preliminarmente se realizó regresiones de tercer, cuarto, quinto y sexto grado, a fin de determinar el mejor modelo de ajuste para la serie de datos definimos en la investigación, obteniendo como resumen los siguientes valores en cuanto a la correlación:

Estadísticas de la regresión - Tercer Grado		Porcentaje (%)
Coeficiente de determinación R^2	0.736965	74
R^2 ajustado	0.712738	71
Coeficiente de correlación múltiple	0.858467	86
Error típico	0.151076	
Observaciones	84	

Estadísticas de la regresión - Cuarto grado		Porcentaje (%)
Coeficiente de determinación R^2	0.762863	76
R^2 ajustado	0.734022	73
Coeficiente de correlación múltiple	0.873420	87
Error típico	0.145371	
Observaciones	84	

<i>Estadísticas de la regresión - Quinto grado</i>		<i>Porcentaje (%)</i>
Coeficiente de determinación R ²	0.767005	77
R ² ajustado	0.731409	73
Coeficiente de correlación múltiple	0.875788	88
Error típico	0.146083	
Observaciones	84	

<i>Estadísticas de la regresión - Sexto grado</i>		<i>Porcentaje (%)</i>
Coeficiente de determinación R ²	0.776413	77
R ² ajustado	0.734890	73
Coeficiente de correlación múltiple	0.881143	88
Error típico	0.145134	
Observaciones	84	

Cuadro 13: Valores de correlación determinados para polinomios del tercer al sexto grado.
Fuente: Elaboración propia.

Revisando los valores expresados en porcentajes, vemos que a medida que sube el grado del polinomio va mejorando el ajuste, llegando a mejores resultados para el quinto y sexto grado, con valores idénticos expresado porcentualmente. Por tanto, ya no corresponde seguir subiendo el grado del polinomio.

En este sentido podemos concluir, que para los datos utilizados en la investigación, basta con determinar los valores de los coeficientes para un polinomio de quinto grado, que permitirá generar la ecuación del modelo para determinar predicciones del valor de N.

En consecuencia el modelo de regresión polinomial múltiple, resulta ser:

$$Y = a_0 + a_1X_1 + a_2X_2 + a_3X_1X_2 + a_4X_1^2 + a_5X_2^2 + a_6X_1^3 + a_7X_2^3 + a_8X_1^4 + a_9X_2^4 + a_{10}X_1^5 + a_{11}X_2^5 +$$

Dónde:

- Y : N – Ondulación pseudo geoidal
- X₁ : Coordenada Este
- X₂ : Coordenada Norte
- : Error

n : 5 (Polinomio de quinto grado)

i : 11

La solución también se determinó mediante el método de mínimos cuadrados mediante el programa Excel, que a continuación se detalla de manera secuencial a fin de puntualizar algunas consideraciones que se realizaron durante el proceso y además sustentan la investigación.

- Estructura de datos

Los valores definidos como variables independientes y dependientes utilizados para la regresión lineal múltiple se mantienen, en vista de que representan los datos definidos y depurados para el área de estudio.

No.	ESTACIÓN	X ₁ : ESTE	X ₂ : NORTE	Y: N	No.	ESTACIÓN	X ₁ : ESTE	X ₂ : NORTE	Y: N
1	ALAC	590859.784	8170386.780	48.312	43	PC28	595579.583	8169244.231	47.950
2	ARUN	599246.172	8173542.594	47.734	44	PC29	590705.303	8183091.160	48.547
3	BEVI	596706.068	8173120.005	48.103	45	PC30	591579.587	8183176.162	48.499
4	BPUL	594444.113	8177535.495	48.309	46	PC32	589427.012	8180756.494	48.473
5	CALV	590855.881	8174560.981	48.345	47	PC33	590747.664	8180107.898	48.514
6	CHUQ	596359.121	8180612.610	48.374	48	PC34	592121.462	8180356.428	48.575
7	CMAT	598329.107	8174700.545	47.753	49	PC35	594521.284	8180334.556	48.686
8	COL6	597948.742	8171789.800	47.843	50	PC36	591781.731	8178536.647	48.450
9	FELC	593020.519	8173433.945	48.306	51	PC37	590487.067	8176840.537	48.206
10	HB83	597722.255	8174452.041	47.732	52	PC38	593434.804	8176784.820	48.123
11	HUMB	596931.142	8170900.732	47.991	53	PC39	591901.952	8175003.477	47.852
12	IGM1	594196.788	8174522.075	48.173	54	PC40	590177.611	8173946.269	48.347
13	INGA	588751.181	8172270.466	48.390	55	PC41	595208.321	8181493.306	48.497
14	JARD	593673.237	8173091.328	48.235	56	PC43	598090.267	8181460.296	48.172
15	LACA	593170.421	8178864.897	48.320	57	PC44	598387.636	8180477.034	48.435
16	LAGU	590169.314	8182988.746	48.598	58	PC45	599836.307	8179174.627	47.872
17	LAIK	593544.602	8174947.552	48.124	59	PC46	595532.826	8175988.994	48.069
18	MAME	593482.947	8172736.320	48.190	60	PC47	596842.752	8176695.979	48.010
19	MUNA	589571.325	8177342.266	48.608	61	PC48	599132.324	8176761.243	47.863
20	P028	595442.053	8174868.639	48.155	62	PC49	601412.767	8178330.774	48.123
21	P043	596271.576	8168418.133	47.928	63	PC51	600238.521	8175232.091	48.183
22	P-27	595288.695	8175890.002	48.142	64	PC52	603379.291	8175788.311	48.555

23	PC01	589804.984	8182138.836	48.464	65	PC54	594680.366	8172920.575	48.153
24	PC02	588982.341	8179056.176	48.531	66	PC55	597734.783	8172319.347	47.916
25	PC03	588655.830	8177087.146	48.606	67	PC61	599527.060	8171718.407	48.048
26	PC04	591361.556	8176483.866	48.200	68	PC63	602194.356	8170653.206	47.836
27	PC05	592151.440	8177132.830	48.308	69	PC64	601652.291	8169377.663	47.701
28	PC06	592771.371	8181538.531	48.569	70	PC65	593454.029	8170162.507	48.189
29	PC07	592976.097	8182789.885	48.582	71	PC66	591885.034	8170269.057	48.285
30	PC08	597081.058	8182386.140	48.804	72	PC67	590007.686	8171354.938	48.336
31	PC09	598443.415	8179673.100	48.150	73	PC68	591246.873	8169355.277	48.223
32	PC10	600514.072	8178320.897	48.105	74	PC69	594453.414	8168569.568	47.987
33	PC11	598299.396	8175753.484	48.088	75	PC70	594067.491	8166942.902	47.982
34	PC13	597993.173	8177865.794	48.094	76	PC71	595995.273	8165567.617	48.106
35	PC14	595577.062	8177826.392	47.784	77	PC73	597524.531	8166207.127	47.781
36	PC19	600424.857	8168176.664	47.935	78	PC75	599102.323	8169865.026	47.878
37	PC20	598846.822	8170695.924	47.739	79	PC76	599229.319	8167400.540	48.089
38	PC21	600886.937	8170954.256	47.776	80	PC77	600939.907	8165907.848	47.693
39	PC23	599702.764	8162633.956	47.765	81	PC78	600738.293	8164698.951	47.802
40	PC25	591961.420	8168548.034	48.069	82	PC79	599323.422	8164274.580	47.684
41	PC26	589312.774	8175815.962	48.483	83	VIRG	591104.498	8172237.546	48.334
42	PC27	595355.933	8171030.182	47.835	84	ZOOL	597810.816	8167541.373	47.929

Cuadro 14: Valores de variables definidas para el proyecto.
Fuente: Elaboración propia.

- Adecuación de los valores de las variables independientes

Las variables independientes de la investigación que corresponden a las coordenadas Este (X_1) y Norte (X_2), de las estaciones de la Red Geodésica de la Ciudad de La Paz, (Sistema de Referencia WGS-84, Proyección UTM Zona 19 Sur), presentan como características numéricas valores que van desde quinientos mil hasta ocho millones respectivamente. Este aspecto debe ser considerado para la solución del polinomio de quinto grado, en el sentido que estas variables serán elevadas hasta la quinta potencia, que en valores, es de entender representan cifras muy grandes, las cuales dificultan el análisis estadístico.

En este sentido a fin de reducir sus valores sin afectar su correspondencia e influencia respecto al valor de la pseudo ondulación (N), se procedió a calcularlos con el siguiente criterio.

Se determinó el promedio de estas dos variables independientes, en función del número de estaciones consideradas para la investigación.

PROMEDIO X ₁	PROMEDIO X ₂
595241.502	8174450.112

Cuadro 15: Valores de promedio de las coordenadas Este y Norte.
Fuente: Elaboración propia.

Se determinó nuevos valores para las variables (X₁) y (X₂), realizando el siguiente procedimiento:

$$(X_1) = \frac{(P \quad x_1) - x_1}{1}$$

$$(X_2) = \frac{(P \quad x_2) - x_2}{1}$$

Obteniendo de esta manera los nuevos valores para las variables Este y Norte de las estaciones utilizadas en el modelo, que a continuación se detallan:

No.	ESTACIÓN	X ₁ : ESTE	X ₂ : NORTE	(X ₁)	(X ₂)	No.	ESTACIÓN	X ₁ : ESTE	X ₂ : NORTE	(X ₁)	(X ₂)
1	ALAC	590859.784	8170386.780	43.81718	40.63332	43	PC28	595579.583	8169244.231	-3.38081	52.05881
2	ARUN	599246.172	8173542.594	-40.04670	9.07518	44	PC29	590705.303	8183091.160	45.36199	-86.41048
3	BEVI	596706.068	8173120.005	-14.64566	13.30107	45	PC30	591579.587	8183176.162	36.61915	-87.26050
4	BPUL	594444.113	8177535.495	7.97389	-30.85383	46	PC32	589427.012	8180756.494	58.14490	-63.06382
5	CALV	590855.881	8174560.981	43.85621	-1.10869	47	PC33	590747.664	8180107.898	44.93838	-56.57786
6	CHUQ	596359.121	8180612.610	-11.17619	-61.62498	48	PC34	592121.462	8180356.428	31.20040	-59.06316
7	CMAT	598329.107	8174700.545	-30.87605	-2.50433	49	PC35	594521.284	8180334.556	7.20218	-58.84444
8	COL6	597948.742	8171789.800	-27.07240	26.60312	50	PC36	591781.731	8178536.647	34.59771	-40.86535
9	FELC	593020.519	8173433.945	22.20983	10.16167	51	PC37	590487.067	8176840.537	47.54435	-23.90425
10	HB83	597722.255	8174452.041	-24.80753	-0.01929	52	PC38	593434.804	8176784.820	18.06698	-23.34708
11	HUMB	596931.142	8170900.732	-16.89640	35.49380	53	PC39	591901.952	8175003.477	33.39550	-5.53365
12	IGM1	594196.788	8174522.075	10.44714	-0.71963	54	PC40	590177.611	8173946.269	50.63891	5.03843
13	INGA	588751.181	8172270.466	64.90321	21.79646	55	PC41	595208.321	8181493.306	0.33181	-70.43194
14	JARD	593673.237	8173091.328	15.68265	13.58784	56	PC43	598090.267	8181460.296	-28.48765	-70.10184

15	LACA	593170.421	8178864.897	20.71081	-44.14785	57	PC44	598387.636	8180477.034	-31.46134	-60.26922
16	LAGU	590169.314	8182988.746	50.72188	-85.38634	58	PC45	599836.307	8179174.627	-45.94805	-47.24515
17	LAIK	593544.602	8174947.552	16.96900	-4.97440	59	PC46	595532.826	8175988.994	-2.91324	-15.38882
18	MAME	593482.947	8172736.320	17.58555	17.13792	60	PC47	596842.752	8176695.979	-16.01250	-22.45867
19	MUNA	589571.325	8177342.266	56.70177	-28.92154	61	PC48	599132.324	8176761.243	-38.90822	-23.11131
20	P028	595442.053	8174868.639	-2.00551	-4.18527	62	PC49	601412.767	8178330.774	-61.71265	-38.80662
21	P043	596271.576	8168418.133	-10.30074	60.31979	63	PC51	600238.521	8175232.091	-49.97019	-7.81979
22	P-27	595288.695	8175890.002	-0.47193	-14.39890	64	PC52	603379.291	8175788.311	-81.37789	-13.38199
23	PC01	589804.984	8182138.836	54.36518	-76.88724	65	PC54	594680.366	8172920.575	5.61136	15.29537
24	PC02	588982.341	8179056.176	62.59161	-46.06064	66	PC55	597734.783	8172319.347	-24.93281	21.30765
25	PC03	588655.830	8177087.146	65.85672	-26.37034	67	PC61	599527.060	8171718.407	-42.85558	27.31705
26	PC04	591361.556	8176483.866	38.79946	-20.33754	68	PC63	602194.356	8170653.206	-69.52854	37.96906
27	PC05	592151.440	8177132.830	30.90062	-26.82718	69	PC64	601652.291	8169377.663	-64.10789	50.72449
28	PC06	592771.371	8181538.531	24.70131	-70.88419	70	PC65	593454.029	8170162.507	17.87473	42.87605
29	PC07	592976.097	8182789.885	22.65405	-83.39773	71	PC66	591885.034	8170269.057	33.56468	41.81055
30	PC08	597081.058	8182386.140	-18.39556	-79.36028	72	PC67	590007.686	8171354.938	52.33816	30.95174
31	PC09	598443.415	8179673.100	-32.01913	-52.22988	73	PC68	591246.873	8169355.277	39.94629	50.94835
32	PC10	600514.072	8178320.897	-52.72570	-38.70785	74	PC69	594453.414	8168569.568	7.88088	58.80544
33	PC11	598299.396	8175753.484	-30.57894	-13.03372	75	PC70	594067.491	8166942.902	11.74011	75.07210
34	PC13	597993.173	8177865.794	-27.51671	-34.15682	76	PC71	595995.273	8165567.617	-7.53771	88.82495
35	PC14	595577.062	8177826.392	-3.35560	-33.76280	77	PC73	597524.531	8166207.127	-22.83029	82.42985
36	PC19	600424.857	8168176.664	-51.83355	62.73448	78	PC75	599102.323	8169865.026	-38.60821	45.85086
37	PC20	598846.822	8170695.924	-36.05320	37.54188	79	PC76	599229.319	8167400.540	-39.87817	70.49572
38	PC21	600886.937	8170954.256	-56.45435	34.95856	80	PC77	600939.907	8165907.848	-56.98405	85.42264
39	PC23	599702.764	8162633.956	-44.61262	118.16156	81	PC78	600738.293	8164698.951	-54.96791	97.51161
40	PC25	591961.420	8168548.034	32.80082	59.02078	82	PC79	599323.422	8164274.580	-40.81920	101.75532
41	PC26	589312.774	8175815.962	59.28728	-13.65850	83	VIRG	591104.498	8172237.546	41.37004	22.12566
42	PC27	595355.933	8171030.182	-1.14431	34.19930	84	ZOOL	597810.816	8167541.373	-25.69314	69.08739

Cuadro 16: Valores nuevos determinados para las coordenadas Este y Norte.
Fuente: Elaboración propia.

Fueron estas variables independientes (X_1) y (X_2), utilizadas para determinar los coeficientes beta (), de la regresión polinomial múltiple, a través de la ecuación polinómica de quinto grado, donde es necesario verificar que los valores obtenidos para la ondulación pseudo geoidal presenten residuos mínimos, así también el coeficiente de determinación se

aproxime a 1. Aspectos que representarán que la función matemática se adaptó a los valores de las ondulaciones (N), generando un alto nivel de calidad de la regresión.

No	ESTACIÓN	(X ₁)	(X ₂)	(X ₁)(X ₂)	(X ₁) ²	(X ₂) ²	(X ₁) ³	(X ₂) ³	(X ₁) ⁴	(X ₂) ⁴	(X ₁) ⁵	(X ₂) ⁵
1	ALAC	43.81718	40.63332	1780.4375	1919.94541	1651.06658	84126.5968	67088.3143	3686190.37	2726020.85	161518473	110767273
2	ARUN	-40.04670	9.07518	-363.430939	1603.73805	82.3588661	-64224.4138	747.421417	2571975.72	6782.98283	-102999136	61556.7804
3	BEVI	-14.64566	13.30107	-194.802906	214.495308	176.918425	-3141.425	2353.2041	46008.2372	31300.1292	-673820.922	416325.164
4	BPUL	7.97389	-30.85383	-246.025109	63.5829483	951.958914	507.003542	-29371.5799	4042.79132	906225.774	32236.78	-27960537.3
5	CALV	43.85621	-1.10869	-48.623006	1923.3673	1.22919668	84351.6035	-1.36279983	3699341.78	1.51092449	162239116	-1.67514903
6	CHUQ	-11.17619	-61.62498	688.732398	124.907186	3797.63834	-1395.98623	-234029.392	15601.805	14422056.9	-174368.711	-888758991
7	CMAT	-30.87605	-2.50433	77.3238582	953.330361	6.2716759	-29435.0743	-15.7063551	908838.777	39.3339186	-28061350	-98.5051687
8	COL6	-27.07240	26.60312	-720.210223	732.914752	707.725918	-19841.7601	18827.7165	537164.033	500875.975	-14542318.7	13324862.9
9	FELC	22.20983	10.16167	225.688948	493.276623	103.259508	10955.5908	1049.2889	243321.826	10662.526	5404136.81	108349.056
10	HB83	-24.80753	-0.01929	0.47857266	615.413462	0.00037216	-15266.8869	-7.1795E-06	378733.729	1.385E-07	-9395447.72	-2.6719E-09
11	HUMB	-16.89640	35.49380	-599.717359	285.488277	1259.80974	-4823.72364	44715.433	81503.5561	1587120.57	-1377116.55	56332937.9
12	IGM1	10.44714	-0.71963	-7.51809148	109.142769	0.51786939	1140.22997	-0.37267509	11912.144	0.26818871	124447.856	-0.19299702
13	INGA	64.90321	21.79646	1414.66016	4212.42688	475.085606	273400.034	10355.1837	17744540.3	225706.333	1151677652	4919598.74
14	JARD	15.68265	13.58784	213.093339	245.945563	184.629357	3857.0786	2508.7139	60489.2201	34087.9995	948631.369	463182.234
15	LACA	20.71081	-44.14785	-914.337836	428.93772	1949.03279	8883.64833	-86045.6099	183987.568	3798728.8	3810531.86	-167705715
16	LAGU	50.72188	-85.38634	-4330.95591	2572.70928	7290.8273	130492.656	-622537.069	6618833.04	53156162.8	335719666	-4538810262
17	LAIK	16.96900	-4.97440	-84.4106261	287.947018	24.7446696	4886.17342	-123.08992	82913.4849	612.298672	1406959.06	-3045.81939
18	MAME	17.58555	17.13792	301.379752	309.251627	293.708253	5438.36047	5033.54812	95636.5691	86264.5379	1681821.83	1478394.63
19	MUNA	56.70177	-28.92154	-1639.90264	3215.09091	836.455559	182301.351	-24191.5841	10336809.6	699657.902	586115415	-20235185
20	P028	-2.00551	-4.18527	8.39359673	4.02206368	17.5164969	-8.06628222	-73.3112941	16.1769962	306.827665	-32.4431007	-1284.15706
21	P043	-10.30074	60.31979	-621.338358	106.10521	3638.47689	-1092.96201	219472.157	11258.3156	13238514.1	-115968.963	798544372
22	P-27	-0.47193	-14.39890	6.79524955	0.22271635	207.328362	-0.10510616	-2985.30065	0.04960257	42985.0498	-0.02340886	-618937.495
23	PC01	54.36518	-76.88724	-4179.98885	2955.57298	5911.64789	160680.262	-454530.299	8735411.63	34947580.8	474902240	-2687023085
24	PC02	62.59161	-46.06064	-2883.00978	3917.70985	2121.58269	245215.774	-97721.4595	15348450.5	4501113.11	960684252	-207324157
25	PC03	65.85672	-26.37034	-1736.66424	4337.10779	695.394907	285627.7	-18337.8011	18810504	483574.077	1238798124	-12752013.5
26	PC04	38.79946	-20.33754	-789.085659	1505.39823	413.615591	58408.6407	-8411.92422	2266223.82	171077.857	87928264.1	-3479303.01
27	PC05	30.90062	-26.82718	-828.976584	954.848419	719.697663	29505.4098	-19307.4598	911735.504	517964.727	28173193.9	-13895533.7
28	PC06	24.70131	-70.88419	-1750.9325	610.154798	5024.56859	15071.6238	-356162.482	372288.878	25246289.6	9196023.6	-1789562822
29	PC07	22.65405	-83.39773	-1889.29652	513.206057	6955.18161	11626.1965	-580046.368	263380.457	48374551.2	5966634.48	-4034327828
30	PC08	-18.39556	-79.36028	1459.87669	338.396566	6298.05427	-6224.99378	-499815.359	114512.236	39665487.6	-2106516.52	-3147864256
31	PC09	-32.01913	-52.22988	1672.35528	1025.22458	2727.96051	-32826.7974	-142481.054	1051085.44	7441768.57	-33654839.5	-388682690
32	PC10	-52.72570	-38.70785	2040.8985	2779.99926	1498.29776	-146577.403	-57995.8872	7728395.91	2244896.18	-407485071	-86895108
33	PC11	-30.57894	-13.03372	398.557364	935.07147	169.877894	-28593.4928	-2214.14115	874358.653	28858.499	-26736959.3	-376133.636
34	PC13	-27.51671	-34.15682	939.883293	757.169238	1166.68845	-20834.8051	-39850.3691	573305.254	1361161.94	-15775473.5	-46492965.3
35	PC14	-3.35560	-33.76280	113.2944	11.2600402	1139.92676	-37.784172	-38487.1209	126.788505	1299433.02	-425.451295	-43872499

36	PC19	-51.83355	62.73448	-3251.75063	2686.71673	3935.6148	-139262.062	246898.742	7218446.8	15489063.9	-374157711	971698345
37	PC20	-36.05320	37.54188	-1353.50479	1299.83311	1409.39265	-46863.1409	52911.2476	1689566.11	1986387.63	-60914262.2	74572723.3
38	PC21	-56.45435	34.95856	-1973.56264	3187.09345	1222.10082	-179925.284	42722.883	10157564.6	1493530.41	-573438692	52211670.2
39	PC23	-44.61262	118.1615	-5271.49651	1990.28571	13962.1539	-88791.857	1649789.87	3961237.23	194941742	-176721164	2.3035E+10
40	PC25	32.80082	59.02078	1935.93003	1075.8939	3483.4523	35290.204	205596.067	1157547.69	12134439.9	37968515.3	716184093
41	PC26	59.28728	-13.65850	-809.775421	3514.98177	186.554661	208393.714	-2548.05711	12355096.8	34802.6416	732500106	-475351.931
42	PC27	-1.14431	34.19930	-39.1345423	1.30944156	1169.59202	-1.49840489	39999.2268	1.7146372	1367945.5	-1.96207364	46782776.6
43	PC28	-3.38081	52.05881	-176.000854	11.429865	2710.11955	-38.6421828	141085.595	130.641814	7344747.97	-441.674932	382358829
44	PC29	45.36199	-86.41048	-3919.75154	2057.71029	7466.7713	93341.8369	-645207.303	4234171.63	55752673.7	192070458	-4817615372
45	PC30	36.61915	-87.26050	-3195.40554	1340.96227	7614.39511	49104.9007	-664435.935	1798179.81	57979012.9	65847819.1	-5059277737
46	PC32	58.14490	-63.06382	-3666.8397	3380.82959	3977.04557	196578.004	-250807.692	11430008.7	15816891.5	664596733	-997473621
47	PC33	44.93838	-56.57786	-2542.51753	2019.45815	3201.0544	90751.181	-181108.812	4078211.21	10246749.3	183268212	-579739162
48	PC34	31.20040	-59.06316	-1842.79436	973.465064	3488.45704	30372.501	-206039.301	947634.231	12169332.5	29566568.6	-718759250
49	PC35	7.20218	-58.84444	-423.808357	51.8714208	3462.66829	373.587396	-203758.781	2690.64429	11990071.7	19378.509	-705549070
50	PC36	34.59771	-40.86535	-1413.84765	1197.00165	1669.97695	41413.518	-68244.1948	1432812.96	2788823	49572049.5	-113966232
51	PC37	47.54435	-23.90425	-1136.51214	2260.46538	571.413236	107472.361	-13659.2057	5109703.71	326513.087	242937550	-7805050.92
52	PC38	18.06698	-23.34708	-421.811292	326.415827	545.086211	5897.34875	-12726.1722	106547.292	297118.978	1924987.97	-6936860.97
53	PC39	33.39550	-5.53365	-184.799066	1115.25953	30.6212981	37244.6515	-169.44759	1243803.82	937.663899	41537452.6	-5188.70518
54	PC40	50.63891	5.03843	255.140539	2564.29937	25.3857625	129853.33	127.904351	6575631.28	644.436936	332982812	3246.94947
55	PC41	0.33181	-70.43194	-23.3701399	0.11009898	4960.65837	0.03653213	-349388.8	0.01212179	24608131.5	0.00402215	-1733198476
56	PC43	-28.48765	-70.10184	1997.03661	811.546108	4914.26817	-23119.0401	-344499.248	658607.085	24150031.7	-18762167	-1692961690
57	PC44	-31.46134	-60.26922	1896.15037	989.81581	3632.37905	-31140.9301	-218920.657	979735.337	13194177.6	-30823784.9	-795202810
58	PC45	-45.94805	-47.24515	2170.8225	2111.22315	2232.10433	-97006.5831	-105456.107	4457263.17	4982289.76	-204802544	-235389034
59	PC46	-2.91324	-15.38882	44.8313045	8.48695759	236.815825	-24.7245302	-3644.31644	72.0284491	56081.735	-209.836039	-863031.805
60	PC47	-16.01250	-22.45867	359.619439	256.400103	504.391922	-4105.60622	-11327.9725	65741.0128	254411.211	-1052677.86	-5713737.8
61	PC48	-38.90822	-23.11131	899.219951	1513.84945	534.132716	-58901.1851	-12344.5075	2291740.17	285297.758	-89167526.9	-6593605.34
62	PC49	-61.71265	-38.80662	2394.85938	3808.45096	1505.95387	-235029.595	-58440.9816	14504298.7	2267897.05	-895098688	-88009422.2
63	PC51	-49.97019	-7.81979	390.75645	2497.01972	61.149138	-124776.546	-478.173505	6235107.49	3739.21708	-311569496	-29239.8976
64	PC52	-81.37789	-13.38199	1088.9982	6622.36071	179.077695	-538913.73	-2396.41617	43855661.4	32068.8207	-356888111	-429144.684
65	PC54	5.61136	15.29537	85.8278449	31.4873798	233.9483	176.687076	3578.32547	991.455084	54731.8069	5563.41305	837143.16
66	PC55	-24.93281	21.30765	-531.259518	621.644931	454.015888	-15499.3539	9674.01098	386442.421	206130.426	-9635094.81	4392154.68
67	PC61	-42.85558	27.31705	-1170.68791	1836.60059	746.221143	-78708.5806	20384.5592	3373101.74	556845.994	-144556226	15211389.1
68	PC63	-69.52854	37.96906	-2639.93314	4834.21764	1441.64941	-336116.087	54738.0708	23369660.2	2078353.02	-162485831	78913107.5
69	PC64	-64.10789	50.72449	-3251.83985	4109.82135	2572.97374	-263471.968	130512.777	16890631.5	6620193.87	-108282271	335805948
70	PC65	17.87473	42.87605	766.397863	319.506032	1838.35554	5711.08459	78821.4215	102084.105	3379551.1	1824725.98	144901797
71	PC66	33.56468	41.81055	1403.35775	1126.58786	1748.12197	37813.5627	73089.9386	1269200.2	3055930.43	42600300.5	127770128
72	PC67	52.33816	30.95174	1619.9571	2739.28317	958.010121	143369.045	29652.0788	7503672.27	917783.391	392728412	28406991.6
73	PC68	39.94629	50.94835	2035.19759	1595.70622	2595.73422	63742.546	132248.372	2546278.33	6737836.15	101714377	343281625
74	PC69	7.88088	58.80544	463.438703	62.1082958	3458.07961	489.46813	203353.888	3857.44041	11958314.6	30400.0314	703213932
75	PC70	11.74011	75.07210	881.35482	137.830222	5635.81998	1618.1422	423092.833	18997.1701	31762466.9	223028.898	2384475045

76	PC71	-7.53771	88.82495	-669.536555	56.8170469	7889.87149	-428.270328	700817.429	3228.17682	62250072.1	-24333.0553	5529359454
77	PC73	-22.83029	82.42985	-1881.89721	521.222065	6794.67994	-11899.65	560084.438	271672.441	46167675.4	-6202360.17	3805594494
78	PC75	-38.60821	45.85086	-1770.2195	1490.59375	2102.30123	-57549.1541	96392.3165	2221869.73	4419670.47	-85782409.4	202645686
79	PC76	-39.87817	70.49572	-2811.24013	1590.26831	4969.64634	-63416.9873	350338.79	2528953.3	24697384.7	-100850025	1741059882
80	PC77	-56.98405	85.42264	-4867.72777	3247.18176	7297.02718	-185037.563	623331.315	10544189.4	53246605.7	-600850599	4548465552
81	PC78	-54.96791	97.51161	-5360.00916	3021.47095	9508.51381	-166083.938	927190.476	9129286.68	90411834.8	-501817793	8816203445
82	PC79	-40.81920	101.7553	-4153.57053	1666.20695	10354.1449	-68013.2321	1053589.31	2776245.61	107208316	-113324120	1.0909E+10
83	VIRG	41.37004	22.12566	915.339417	1711.48035	489.544767	70804.0133	10831.5004	2929164.98	239654.079	121179677	5302504.33
84	ZOOL	-25.69314	69.08739	-1775.07183	660.137357	4773.06726	-16961.0004	329758.752	435781.331	22782171.1	-11196590	1573960705

Cuadro 17: Valores de las variables independientes utilizadas en la regresión polinomial múltiple.
Fuente: Elaboración propia.

- Aplicación del Análisis de Datos

La herramienta utilizada fue también Análisis de datos de Excel, opción “Regresión”, con la cual se procedió a solucionar la ecuación polinómica de quinto grado.

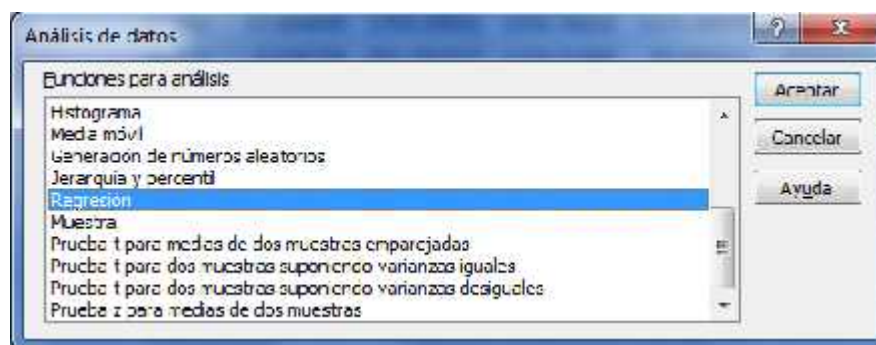


Figura 13: Herramienta de Análisis de datos, Regresión de Excel.
Fuente: Elaboración propia.

Se consideraron como valores de entrada:

- Rango Y de entrada : Los valores (Y: N) – Ondulación pseudo geoidal.
- Rango X de entrada : Todos los valores (X_1) y (X_2) elevados hasta la quinta potencia (Cuadro 17)
- Nivel de confianza : 95%
- Residuales : Todas las residuales del análisis de regresión.



Figura 14: Valores de entrada y opciones de salida para la regresión polinomial múltiple.
Fuente: Elaboración propia.

Como resultados se generaron una serie de valores estadísticos que denotan las características de la regresión, al igual que la calidad de ajuste, detallados a continuación.

- Coeficientes determinados

	Coeficientes
0 : Intercepción	48.07898240280980
1 : (X_1)	0.004444671587068
2 : (X_2)	-0.00182346566382
3 : $(X_1)(X_2)$	0.00002428053885
4 : $(X_1)^2$	-0.00005111401556
5 : $(X_2)^2$	0.00005406609294
6 : $(X_1)^3$	0.00000002161786
7 : $(X_2)^3$	-0.00000043385498
8 : $(X_1)^4$	0.00000001958295
9 : $(X_2)^4$	-0.00000000357293
10 : $(X_1)^5$	-0.00000000005039
11 : $(X_2)^5$	0.00000000004014

Cuadro 18: Valores de los coeficientes determinados en la regresión polinomial múltiple.
Fuente: Elaboración propia.

Con los resultados de los coeficientes , podemos definir la ecuación de predicción del valor de la ondulación pseudo geoidal (N):

$$Y = a_0 + a_1X_1 + a_2X_2 + a_3X_1X_2 + a_4X_1^2 + a_5X_2^2 + a_6X_1^3 + a_7X_2^3 + a_8X_1^4 + a_9X_2^4 + a_{10}X_1^5 + a_{11}X_2^5$$

Dónde:

Y : N – Ondulación pseudo geoidal

X_1 : Coordenada Este

X_2 : Coordenada Norte

Para tener certeza de la calidad del ajuste de la regresión polinomial, es importante revisar e interpretar los valores estadísticos determinados.

- Ondulaciones pseudo geoidales estimadas y residuales

Observación	Estación	N (m)	Pronóstico Y: N (m)	Residuos (m)	Observación	Estación	N (m)	Pronóstico Y: N (m)	Residuos (m)
1	ALAC	48.312	48.266	0.046	43	PC28	47.950	48.039	-0.089
2	ARUN	47.734	47.852	-0.118	44	PC29	48.547	48.604	-0.057
3	BEVI	48.103	47.983	0.120	45	PC30	48.499	48.577	-0.078
4	BPUL	48.309	48.221	0.088	46	PC32	48.473	48.613	-0.140
5	CALV	48.345	48.243	0.102	47	PC33	48.514	48.481	0.033
6	CHUQ	48.374	48.372	0.002	48	PC34	48.575	48.454	0.121
7	CMAT	47.753	47.918	-0.165	49	PC35	48.686	48.410	0.276
8	COL6	47.843	47.895	-0.052	50	PC36	48.450	48.344	0.106
9	FELC	48.306	48.149	0.157	51	PC37	48.206	48.316	-0.110
10	HB83	47.732	47.945	-0.213	52	PC38	48.123	48.211	-0.088
11	HUMB	47.991	47.957	0.034	53	PC39	47.852	48.201	-0.349
12	IGM1	48.173	48.121	0.052	54	PC40	48.347	48.286	0.061
13	INGA	48.390	48.463	-0.073	55	PC41	48.497	48.471	0.026
14	JARD	48.235	48.127	0.108	56	PC43	48.172	48.361	-0.189
15	LACA	48.320	48.333	-0.013	57	PC44	48.435	48.277	0.158
16	LAGU	48.598	48.631	-0.033	58	PC45	47.872	48.140	-0.268
17	LAIK	48.124	48.150	-0.026	59	PC46	48.069	48.109	-0.040
18	MAME	48.190	48.133	0.057	60	PC47	48.010	48.077	-0.067
19	MUNA	48.608	48.409	0.199	61	PC48	47.863	47.974	-0.111
20	P028	48.155	48.079	0.076	62	PC49	48.123	48.158	-0.035

21	P043	47.928	47.989	-0.061	63	PC51	48.183	47.891	0.292
22	P-27	48.142	48.116	0.026	64	PC52	48.555	48.467	0.088
23	PC01	48.464	48.643	-0.179	65	PC54	48.153	48.087	0.066
24	PC02	48.531	48.561	-0.030	66	PC55	47.916	47.912	0.004
25	PC03	48.606	48.511	0.095	67	PC61	48.048	47.818	0.230
26	PC04	48.200	48.259	-0.059	68	PC63	47.836	47.972	-0.136
27	PC05	48.308	48.258	0.050	69	PC64	47.701	47.864	-0.163
28	PC06	48.569	48.516	0.053	70	PC65	48.189	48.144	0.045
29	PC07	48.582	48.558	0.024	71	PC66	48.285	48.209	0.076
30	PC08	48.804	48.452	0.352	72	PC67	48.336	48.322	0.014
31	PC09	48.150	48.209	-0.059	73	PC68	48.223	48.250	-0.027
32	PC10	48.105	48.086	0.019	74	PC69	47.987	48.099	-0.112
33	PC11	48.088	47.957	0.131	75	PC70	47.982	48.112	-0.130
34	PC13	48.094	48.088	0.006	76	PC71	48.106	47.986	0.120
35	PC14	47.784	48.200	-0.416	77	PC73	47.781	47.872	-0.091
36	PC19	47.935	47.764	0.171	78	PC75	47.878	47.815	0.063
37	PC20	47.739	47.835	-0.096	79	PC76	48.089	47.775	0.314
38	PC21	47.776	47.822	-0.046	80	PC77	47.693	47.735	-0.042
39	PC23	47.765	47.787	-0.022	81	PC78	47.802	47.715	0.087
40	PC25	48.069	48.215	-0.146	82	PC79	47.684	47.742	-0.058
41	PC26	48.483	48.389	0.094	83	VIRG	48.334	48.231	0.103
42	PC27	47.835	48.053	-0.218	84	ZOOL	47.929	47.867	0.062

Cuadro 19: Valores de N estimados mediante regresión polinomial y residuales por estación.
Fuente: Elaboración propia.

Los valores de los residuos, corresponden a la diferencia entre el valor determinado para N (ondulación pseudo geoidal), y el valor de N pronosticado mediante la regresión polinomial múltiple. Vemos que el promedio de residuos es igual a 0 m, el valor absoluto máximo es de 0.416 m y como mínimo de 0.002 m, aspecto que demuestra que la calidad de la regresión fue muy alta y que las diferencias de valores en ninguna caso pasan los 42 centímetros.

Resumen	(m)
Promedio	0.000
Valor máximo	0.416
Valor mínimo	0.002

Cuadro 20: Valores de promedio, valor máximo y mínimo de los residuos de la regresión polinomial.
Fuente: Elaboración propia.

- Análisis de Varianza

El análisis de varianza nos permitió realizar la prueba de hipótesis, para determinar si existe una relación entre las variables independientes, con la variable dependiente, recordando que para la investigación es importante rechazar la hipótesis nula.

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F (Alfa 5%)
Regresión	11	5.058092	0.459827	21.547265	1.924308
Residuos	72	1.536506	0.021340		
Total	83	6.594599			

Cuadro 21: Análisis de varianza para la regresión polinomial.
Fuente: Elaboración propia.

Comparando los valores obtenidos tenemos:

-) Si $F < F_{Crítica}$, entonces se acepta la hipótesis nula (H_0)
-) Si $F > F_{Crítica}$, entonces se rechaza la hipótesis nula (H_0)

$$F (21.547265) > F_{Crítica} (1.924308)$$



Figura 15: Valores de F y $F_{Crítica}$ del análisis de varianza – Regresión polinomial.
Fuente: Elaboración propia.

Por tanto, se rechazó la hipótesis nula (H_0), determinando que si existe relación entre las variables independientes (Coordenada Este y Coordenada Norte), con la variable dependiente (N – Ondulación pseudo geoidal).

- **Correlación**

<i>Estadísticas de la regresión</i>		<i>Porcentaje (%)</i>
Coeficiente de determinación R^2	0.767005	77
R^2 ajustado	0.731409	73
Coeficiente de correlación múltiple	0.875788	88
Error típico	0.146083	
Observaciones	84	

Cuadro 22: Valores de correlación de la regresión polinomial
Fuente: Elaboración propia.

La interpretación de estos valores corresponden a:

) Coeficiente de determinación R^2

El valor obtenido oscila entre $0 < R^2 < 1$, interpretándose como la proporción de las variables en los datos de la predicción de (N). Si el valor es 1 significaría que el modelo esta explicado al 100%, en el caso de los resultados analizados observamos que corresponde a un 77%, lo cual implica que la ecuación definida permite generar datos confiables para predicciones de N.

) Coeficiente de determinación R^2 ajustado

Para fines de predicción o pronósticos se recomienda que el coeficiente de determinación ajustado deba estar al menos en un 0.7 (un 70%), en el caso de la regresión polinomial este valor se encuentra en 0.73, es decir en un 73%, lo cual significa que el modelo se encuentra ajustado para predecir la variable dependiente.

) Coeficiente de correlación múltiple

Para este coeficiente los resultados muestran un 0.875788 (un 88%), de intensidad de relación entre las variables independientes y dependientes del modelo generado.

Recordando que mientras más cerca se encuentre al valor de 1, el modelo resulta mejor ajustado.

J) Error típico – Desviación estándar

Llamado también error estándar, representa la desviación estándar de la distancia entre los valores de datos y los valores ajustados, es decir, que para la regresión realizada la desviación estándar para el valor de N corresponde a 0.146 m.

Utilice S (Error típico), para evaluar qué tan bien el modelo describe la respuesta. S se mide en las unidades de la variable de respuesta y representa la distancia que separa a los valores de los datos de los valores ajustados. Mientras más bajo sea el valor de S, mejor describirá el modelo la respuesta.¹¹

Verificada la calidad de la regresión polinomial, establecemos que el modelo determinado mediante el polinomio de quinto orden, permite predecir de manera confiable los valores de N – Ondulación pseudo geoidal, con una desviación estándar de 0.146 m.

Por tanto, el modelo está definido mediante:

$$\begin{aligned} N = & 48.07898240280980 + 0.00444671587068X_1 - 0.00182346566382X_2 + \\ & 0.00002428053885X_1X_2 - 0.00005111401556X_1^2 + 0.00005406609294X_2^2 + \\ & 0.00000002161786X_1^3 - 0.00000043385498X_2^3 + 0.00000001958295X_1^4 - 0.00000000357293X_2^4 - \\ & 0.00000000005039X_1^5 + 0.00000000004014X_2^5 \end{aligned}$$

Siendo:

X_1 : *Coordenada Este*
 X_2 : *Coordenada Norte*

¹¹ Soporte de Minitab 19, Interpretación de los estadísticos de la regresión de mejores subconjuntos, 2020.

3.5. DETERMINACIÓN DE VALORES DE ONDULACIÓN GEOIDAL A PARTIR DE MODELOS GLOBALES

Para realizar una comparación de los resultados de ondulación pseudo geoidal que puede predecir el modelo determinado con regresión polinomial múltiple, con valores de ondulación geoidal determinadas a través de modelos geoidales globales, se determinaron mediante el servicio interactivo de cálculo definido por el usuario a partir de coordenadas geodésicas del Centro Internacional de Modelos Globales de la Tierra (ICGEM), cuyo acceso es abierto y gratuito¹².

El ICGEM, cuenta con una trayectoria de 15 años de recopilación y distribución exitosa de modelos gravitacionales globales, además de prestar servicios asociados. Su objetivo se centra en proporcionar a la comunidad científica un archivo de última generación de modelos de campos gravitatorios globales estáticos y temporales de la Tierra, así también, desarrollar y operar servicios interactivos de cálculo y visualización de funciones del campo gravitatorio en el usuario.

Debido al interés de científicos y estudiantes de todo el mundo, el equipo de ICGEM ha desarrollado una interfaz web para calcular las funciones del campo de gravedad (por ejemplo, ondulación geoide, anomalía de altura, anomalía de gravedad), a partir de las representaciones armónicas esféricas del campo de gravedad global de la Tierra en cuadrículas de libre elección con respecto a un sistema de referencia de preferencia del usuario. Este servicio es el único servicio en línea disponible en todo el mundo que calcula una variedad de funciones de campo de gravedad con los gráficos GMT (Wessel et al., 2013), proporcionados para los valores de la cuadrícula y la opción de descargar los valores calculados. Los investigadores y estudiantes interesados han utilizado ampliamente el servicio ICGEM para calcular las funciones del campo gravitatorio en cuadrícula. Los resultados calculados no solo se proporcionan en formato ASCII, sino que también se visualizan utilizando el software Generic Mapping Tools (GMT) (Wessel y Smith, 1998; Wessel et al., 2013) con las estadísticas básicas proporcionadas (Ince et al., 2019).

¹² <http://icgem.gfz-potsdam.de/home>

Otra alternativa de cálculo corresponde al servicio de cálculos mediante puntos definidos por el usuario, el cual fue utilizado para la investigación.

Gravity field functional	Definition See also Barthelmes (2013).	Static model	Temporal model	Top. model	Celestial model
height_anomaly	It is an approximation to geoid height according to Molodensky's theory, defined on the Earth's surface (Eqs. 81 and 119), where the height (elevation) used in the calculation is taken from etopo1m.	x	x	x	-
height_anomaly_ell	It is the generalized pseudo height anomaly which is defined on the ellipsoid, therefore, the h value used in the calculation is set to zero.	x	x	x	x
geoid	It is a particular equipotential surface of the gravity potential of the Earth that is equal to the undisturbed sea surface and its continuation below the continents. Here it is approximated by the height anomaly plus a topography-dependent correction term (Eqs. 71, 117).	x	x	x	-
gravity_disturbance	It is defined as the magnitude of the gradient of the potential at a given point on the Earth's surface minus the magnitude of the gradient of the normal potential at the same point (Eqs. 87, 121–124).	x	x	x	-
gravity_disturbance_sa	It is calculated by spherical approximation on $h = 0$ or above at an arbitrary height over the ellipsoid, $h > 0$.	x	x	x	x
gravity_anomaly	It is defined (according to Molodensky's theory) as the magnitude of the gradient of the potential on the Earth's surface minus the magnitude of the gradient of the normal potential on the telluroid (Eqs. 101 and 121–124).	x	x	x	-
gravity_anomaly_cf	The classical gravity anomaly is defined as the magnitude of the gradient of the downward continued potential on the geoid minus the magnitude of the gradient of the normal potential on the ellipsoid (Eqs. 93, 121–124). This type of gravity anomaly is also known as the free-air gravity anomaly.	x	x	x	-
gravity_anomaly_sa	The gravity anomaly calculated by spherical approximation (Eqs. 100, 104, or 126). Unlike the classical gravity anomaly, the Molodensky gravity anomaly and the spherical approximation can be generalized to 3-D space; hence here it can be calculated on $h = 0$ or above the ellipsoid, $h > 0$.	x	x	-	-
gravity_anomaly_bg	The (simple) Bouguer gravity anomaly is defined by the classical gravity anomaly minus the attraction of the Bouguer plate. Here it is calculated by the spherical approximation of the classical gravity anomaly minus $2\pi G\rho H$ (Eqs. 107, 126). The topographic heights $H(\lambda, \varphi)$ are calculated from the spherical harmonic model of etopo1m up to the same maximum degree as the gravity field model. For $H \geq 0$ (rock) $\rightarrow \rho = 2670 \text{ kg m}^{-3}$, and for $H < 0$ (water) $\rightarrow \rho = (2670-1025) \text{ kg m}^{-3}$ is used. The density contrast between ice and rock is not been taken into account.	x	x	-	-
gravity_earth	The gravity is defined as the magnitude of the gradient of the potential (including the centrifugal potential) at a given point. Here it will be calculated on the Earth's surface (Eqs. 7, 121–124).	x	x	x	-
gravity_ell	The magnitude of the gradient of the potential (including the centrifugal potential) calculated on or above the ellipsoid (Eq. 7, 121–124).	x	x	x	-
potential_ell	The potential of the gravity field of the Earth without the centrifugal potential (gravitational field). Here it can be calculated on or above the ellipsoid (Eq. 108).	x	x	x	x
gravitation_ell	The magnitude of the gradient of the potential calculated on or above the ellipsoid without the centrifugal potential (Eqs. 7, 122).	x	x	x	x
second_r_derivative (vertical gravity gradient)	The second derivative of the disturbance potential in radial direction calculated on or above the ellipsoid.	x	x	x	x
water_column	The variable thickness of a fictitious water layer which is distributed over the reference ellipsoid and produces the disturbance potential or the geoid undulations. To calculate water_column from a gravity field model, Earth's elastic deformation due to the load of the water layer is considered.	x	x	-	-
vertical_deflection_abs	Magnitude of the deflection of the vertical. It is the angle between the vector of gravity and the vector of normal gravity both at the same point (h, λ, φ) .	x	x	-	-
vertical_deflection_ew	It is the east-west component of the angle between the vector of gravity and the vector of normal gravity both at the same point (h, λ, φ) .	x	x	-	-
vertical_deflection_ns	It is the north-south component of the angle between the vector of gravity and the vector of normal gravity both at the same point (h, λ, φ) .	x	x	-	-

Figura 16: Lista de funciones de campo de gravedad disponibles en el servicio de cálculo ICGEM.
Fuente: ICGEM, Sinem Ince, 2019.

3.5.1. Cálculos de ondulación geoidal mediante puntos definidos por el usuario

Con los datos definidos para el uso en la generación del modelo pseudo geoidal (84 estaciones del modelo, 8 estaciones de validación, además de las 21 estaciones depuradas), se generó el archivo de texto (*.txt separado por espacios), con valores de posición en latitud y longitud geodésica, donde los valores deben estar representados en grados.

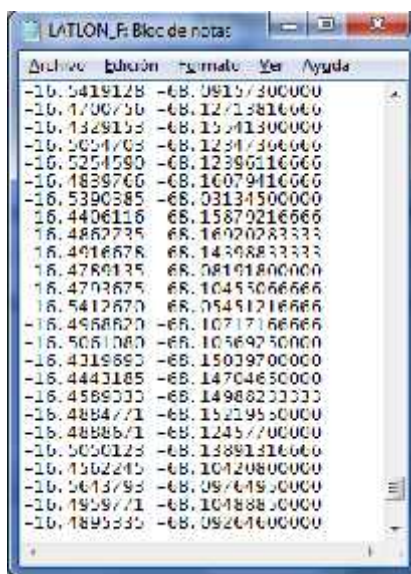


Figura 17: Valores de latitud y longitud en grados de las estaciones utilizadas en el modelo.
Fuente: Elaboración propia.

Este archivo de texto, fue utilizado en la interfaz ofrecida por el ICGEM para realizar cálculos en función de puntos definidos por el usuario, teniendo el cuidado de que el formato se del archivo coincida con lo exigido en el servicio.

Las configuraciones básicas del servicio se resumen en:

- Cargar el archivo txt que contiene los valores de latitud y longitud de los puntos.
- Selección del modelo (larga data, serie o relacionado con la topografía), definiéndose posteriormente el modelo global en específico (por ejemplo EIGEN.6C4).
- Selección funcional del producto ofrecido por el ICGEM, que en el caso de la investigación corresponde a “geoides”.
- Sistema de referencia, siendo para Bolivia el WGS84.

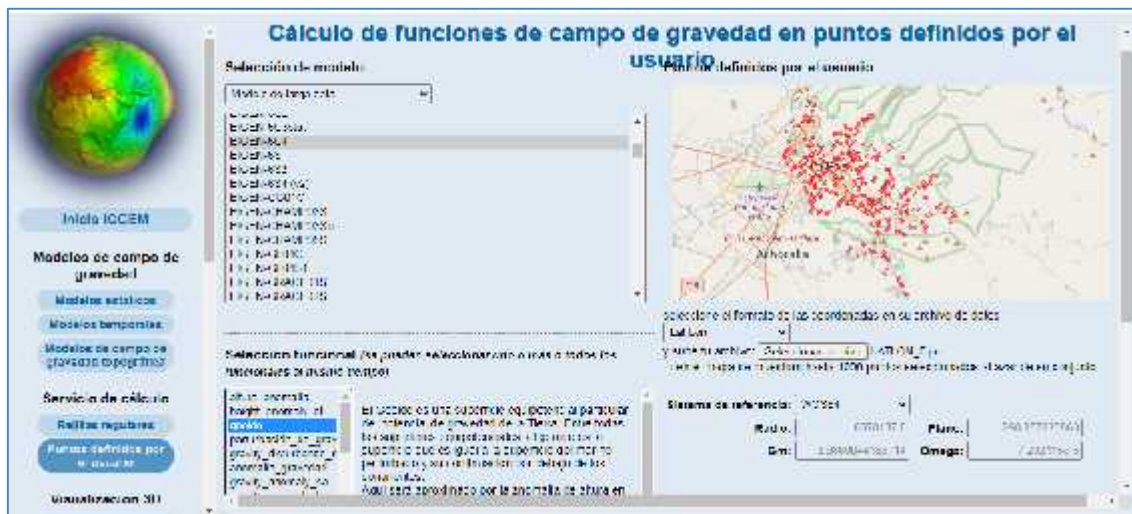


Figura 18: Instantánea del servicio de cálculo del valor de puntos en ICGEM.
Fuente: Elaboración propia.

Como resultado del cálculo, el servicio devuelve un archivo en formato ASCII con valores de la ondulación geoidal, que puede ser abierto con un programa de visualización de texto como WordPad o Bloc de notas (Anexo 2).

```

generated_by      ICGEM (hosted at GFZ Potsdam)
generating date   2020/10/21
-----
modelname         EIGEN 6C4
max used degree   2190
-----
tide_system       tide_free
zero degree term included
-----
relsysname        WGS84
max degree refpot 0
gnarcoefpot       3.98600441600E+14 [m**3/s**2]
radiusrefpot      6378137.000 m
llalrefpot        3.352810664747480E-05
(1/299.25722356300)
omegarcoefpot     7.29211500000E-05 1/s
-----
topo_shm         etopot ==> for Bouguer anomaly and geoid
crust density     2670.0 [kg/m**3] ==> for Bouguer anomaly
and geoid
water_density     1020.0 [kg/m**3] ==> for Bouguer anomaly

```

Figura 19: Instantánea del archivo devuelto por el servicio de cálculo del ICGEM.
Fuente: Elaboración propia.

La ventaja de realizar el cálculo mediante la opción de puntos definidos por el usuario y no así por rejillas regulares que deben cubrir el área de estudio, es que el usuario puede preparar la lista de puntos particulares en uno de los formatos permitidos y los cálculos se realizan directamente en esos puntos particulares. Se pueden introducir diferentes alturas para diferentes puntos en el cálculo de puntos, que es diferente al cálculo de la cuadrícula donde se supone que la altura es la misma para todos los puntos de la cuadrícula y, en consecuencia, ofrece resultados más rápidamente. Para los cálculos de puntos, después de que el usuario carga el archivo de texto del conjunto de puntos de datos en un formato predefinido, los puntos se muestran en el mapa, siendo otra manera de verificar que efectivamente se realiza los cálculos para los puntos deseados.



Figura 20: Instantánea del mapeo de la ubicación de los puntos definidos por el usuario.
Fuente: Servicio de cálculo ICGEM.

El cálculo de ondulación geoidal se realizó para los modelos geoidales globales: Earth Gravitational Model 1996 (EGM96), que es modelo geopotencial de la superficie terrestre constituido de la suma de armónico esféricos de hasta 360 en orden y grado desarrollado mediante la colaboración de distintos centros de investigación, fundamentalmente la NASA, (National Aeronautics and Space Administration), a través del Centro de vuelo espacial Goddard (GSFC), sumando esfuerzos con la Agencia Nacional de Mapeo e Imágenes (NIMA por sus siglas en inglés hoy llamado NGA (National Geospatial Intelligence Agency), la Universidad del Estado de Ohio en Estados Unidos (OSU), y la Universidad de Texas at Austin. Para su elaboración se obtuvieron datos de gravedad de diferentes partes del mundo continental y oceánica; y de datos de diferentes misiones satelitales. El Earth Gravitational

Model 2008 (EGM08), modelo geoidal que agrupa a las ondulaciones geoidales para tener un solo compendio de datos de distintos puntos, dentro de una misma zona, es así que el EGM08, contiene esféricos armónicos al grado y orden 2159, además de contener también coeficientes adicionales extendidos al grado 2190 y orden 2159. Incorpora también anomalías de la gravedad, obtenidas cada 5'x5' (9.26 x 9.26 km aproximadamente), empleando algoritmos que calculan anomalías de gravedad por área media, esto minimiza efectos aleatorios durante el proceso de análisis armónico, haciendo también un énfasis especial en la calibración del error en cuanto a la estimación del modelo. Incluye también como modelo de referencia al Dynamic Ocean Topography (DOT), que toma en cuenta elevaciones y depresiones oceánicas (Tierra, 2015).

Y finalmente, para el European Improved Gravity model of the Earth by New techniques (EIGEN-6C4), que es un modelo de campo de gravedad combinado global estático hasta grado y orden 2190. Fue elaborado conjuntamente por Geo Forschungs Zentrum (GFZ), Potsdam y GRGS (Tarrío Mosquera, 2017).

El desarrollo del Earth Gravitational Model 2008 (EGM2008) (Pavlis et al., 2012), fue un hito muy importante en términos de la entrega de modelos de campo gravitatorio global estático de alta resolución. El grado armónico esférico y el orden de expansión llegaron hasta 2190 utilizando una combinación de datos de gravedad y topografía disponibles en todo el mundo. La mejora se debió a la introducción de la cobertura mundial de datos terrestres de la Agencia Nacional de Inteligencia Geoespacial (NGA). Después del lanzamiento de EGM2008, diferentes centros de procesamiento también pudieron aprovechar el uso de las rejillas EGM2008 para los componentes de alta frecuencia del campo gravitatorio y desarrollar diferentes modelos de campo gravitatorio global de alta resolución "combinados". Como ejemplo, Uno de los modelos de campo gravitatorio global estático de alta resolución desarrollado por GFZ es EIGEN-6C4 (Förste et al., 2014), que también es una combinación de datos satelitales y terrestres y expandido hasta un grado armónico esférico y un orden de 2190. EIGEN-6C4 contiene datos de la misión GOCE, que aún no estaban disponibles para EGM2008. Con la última versión de la serie EIGEN, la resolución espacial de Potsdam Gravity Potato se ha incrementado hasta 9 km de media longitud de onda (Ince et al., 2019).

CAPÍTULO IV RESULTADOS

4.1. COMPARACIÓN DE VALORES DE ONDULACIÓN PSEUDO GEOIDAL CON LAS ESTACIONES DE VALIDACIÓN

Para validar el modelo pseudo geoidal determinado mediante la regresión polinomial múltiple, se utilizó las ocho estaciones de la Red Geodesia Municipal de La Paz, definidas anteriormente como estaciones de validación.

Para tal efecto se procedió a utilizar las coordenadas Este y Norte de las ocho estaciones a fin de predecir en el modelo el valor de la ondulación pseudo geoidal (N), cuyos resultados se muestran a continuación.

No.	ESTACIÓN	ESTE (m)	NORTE (m)	N (m)	N Estimada (m)	DIFERENCIA (m)
1	085X	600085.966	8169107.842	47.820	47.779	0.041
2	ACHA	591213.178	8178692.407	48.394	48.372	0.022
3	BAIR	591909.120	8173101.400	47.895	48.188	-0.293
4	JUPA	599074.723	8165614.003	47.828	47.757	0.071
5	PC12	596977.025	8174169.317	47.973	47.982	-0.009
6	PC24	595282.368	8167209.210	47.971	48.045	-0.074
7	PC31	591057.266	8181723.486	48.363	48.561	-0.198
8	PC42	595624.980	8180386.566	48.561	48.382	0.179
					PROMEDIO	0.111
					VALOR MÁXIMO	0.293
					VALOR MÍNIMO	0.009

Cuadro 23: Comparación de valores de ondulación pseudo geoidal de las estaciones de validación.
Fuente: Elaboración propia.

Los resultados son analizados en valores absolutos y nos muestran que las diferencias determinadas alcanzan un promedio de 0.111 metros, un valor máximo de 0.293 metros y 0.009 metros como valor mínimo, aspecto que en función al promedio guarda relación con la desviación estándar determinada para el modelo que se estableció de 0.146 metros.

Estos resultados permiten establecer la confianza del modelo generado a un 95%, y en relación a lo descrito durante la interpretación de los resultados estadísticos con un coeficiente de correlación múltiple al 88%

4.2. COMPARACIÓN DE VALORES DE ONDULACIÓN PSEUDO GEOIDAL CON MODELOS GLOBALES

A fin de generar mayor sustento en los resultados obtenidos con el modelo pseudo geoidal de la Ciudad de La Paz, se realizó comparaciones para las estaciones de validación, para los valores de ondulaciones geoidales determinadas a partir de modelos geoidales globales, cuyos valores fueron calculados anteriormente. Los valores de ondulación geoidal calculados corresponden a los modelos EGM966, EGM2008 y EIGEN-6C4.

4.2.1. Comparación con el modelo global EGM96

Las comparaciones corresponden a ondulaciones pseudo geoidales (N), determinadas a partir de los valores de las estaciones de validación (altura elipsoidal y altura nivelada), valores de N estimados a partir del modelo polinomial múltiple definido y los valores determinados para cada estación con el EGM96.

No.	ESTACIÓN	ESTE (m)	NORTE (m)	N (m)	N Estimada (m)	EGM96 (m)	DIFERENCIA N - N EGM96	DIFERENCIA N Est- N EGM96
1	085X	600085.966	8169107.842	47.820	47.779	44.419	3.401	3.360
2	ACHA	591213.178	8178692.407	48.394	48.372	45.066	3.328	3.306
3	BAIR	591909.120	8173101.400	47.895	48.188	45.056	2.839	3.132
4	JUPA	599074.723	8165614.003	47.828	47.757	44.555	3.273	3.202
5	PC12	596977.025	8174169.317	47.973	47.982	44.620	3.353	3.362
6	PC24	595282.368	8167209.210	47.971	48.045	44.841	3.130	3.204
7	PC31	591057.266	8181723.486	48.363	48.561	45.050	3.313	3.511
8	PC42	595624.980	8180386.566	48.561	48.382	44.660	3.901	3.722
PROMEDIO							3.317	3.350
VALOR MÁXIMO							3.901	3.722
VALOR MÍNIMO							2.839	3.132

Cuadro 24: Comparación de valores de ondulación pseudo geoidal con el modelo EGM96.
Fuente: Elaboración propia.

Los resultados nos muestran que las comparaciones con el valor de N y N estimada, en comparación con el modelo EGM96, son similares, en promedios y valores máximos superan los 3 metros de diferencia, solo para el valor mínimo de comparación entre N y la ondulación geoidal de EGM96, alcanza un valor inferior a los 3 metros (2.839), que caracterizan a los valores de la comparativa realizada.

4.2.2. Comparación con el modelo global EGM08

Las comparativas realizadas son similares a la anterior, solo que en este caso utilizamos el modelo geoidal EGM08, considerado uno de los más “precisos” en el ámbito de su aplicación y comparación con otros modelos globales.

No.	ESTACIÓN	ESTE (m)	NORTE (m)	N (m)	N Estimada (m)	EGM08 (m)	DIFERENCIA N - N EGM08	DIFERENCIA N Est - N EGM08
1	085X	600085.966	8169107.842	47.820	47.779	45.782	2.038	1.997
2	ACHA	591213.178	8178692.407	48.394	48.372	46.162	2.232	2.210
3	BAIR	591909.120	8173101.400	47.895	48.188	46.064	1.831	2.124
4	JUPA	599074.723	8165614.003	47.828	47.757	45.796	2.032	1.961
5	PC12	596977.025	8174169.317	47.973	47.982	45.830	2.143	2.152
6	PC24	595282.368	8167209.210	47.971	48.045	45.914	2.057	2.131
7	PC31	591057.266	8181723.486	48.363	48.561	46.236	2.127	2.325
8	PC42	595624.980	8180386.566	48.561	48.382	45.906	2.655	2.476
PROMEDIO							2.139	2.172
VALOR MÁXIMO							2.655	2.476
VALOR MÍNIMO							1.831	1.961

Cuadro 25: Comparación de valores de ondulación pseudo geoidal con el modelo EGM08.

Fuente: Elaboración propia.

En este caso los resultados varían en comparación al anterior, los valores de promedios y valores máximos se encuentran por encima de los 2 metros, y en los valores mínimos inferiores a este.

4.2.3. Comparación con el modelo global EIGEN-6C4

Esta comparativa también es similar a las anteriores, solo que en este caso utilizamos el modelo geoidal EIGEN-6C4, considerado uno de los más recientes con mejoras en comparación al EGM08 respecto al uso de datos de la misión GOCE.

No.	ESTACIÓN	ESTE (m)	NORTE (m)	N (m)	N Estimada (m)	EIGEN-6C4 (m)	DIFERENCIA N - N EIGEN-6C4	DIFERENCIA N _{Est} - N _{EIGEN-6C4}
1	085X	600085.966	8169107.842	47.820	47.779	45.751	2.069	2.028
2	ACHA	591213.178	8178692.407	48.394	48.372	46.090	2.304	2.282
3	BAIR	591909.120	8173101.400	47.895	48.188	46.006	1.889	2.182
4	JUPA	599074.723	8165614.003	47.828	47.757	45.760	2.068	1.997
5	PC12	596977.025	8174169.317	47.973	47.982	45.785	2.188	2.197
6	PC24	595282.368	8167209.210	47.971	48.045	45.868	2.103	2.177
7	PC31	591057.266	8181723.486	48.363	48.561	46.154	2.209	2.407
8	PC42	595624.980	8180386.566	48.561	48.382	45.843	2.718	2.539
PROMEDIO							2.194	2.226
VALOR MÁXIMO							2.718	2.539
VALOR MÍNIMO							1.889	1.997

Cuadro 26: Comparación de valores de ondulación pseudo geoidal con el modelo EIGEN-6C4.
Fuente: Elaboración propia.

Los resultados son muy similares a la comparativa con el modelo EGM08, los valores de promedios y valores máximos se encuentran por encima de los 2 metros, y en los valores mínimos inferiores a este. Resaltando además que en comparación al EGM08 las diferencias son un poco mayores (en 5 a 6 centímetros), con una tendencia constante en casi todos los valores analizados.

Los resultados son claros, para todos los modelos globales comparados, las diferencias entre la ondulación pseudo geoidal y geoidal van alrededor de 2 a 3 metros, aspecto que confirma para nuestro territorio y sobre todo para la ciudad de La Paz, que asumir que el uso de un modelo geoidal global, permitirá obtener una altura nivelada para ser utilizada en proyectos y trabajos de topografía y geodesia (recordando que los modelos geoidales nos proporcionan alturas ortométricas), no es apropiado.

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

Se aplicó de manera adecuada la regresión polinomial múltiple a los datos correspondientes a la Red Geodésica Municipal de La Paz 2005 – 2006 en cuanto a altura elipsoidal y altura nivelada para generar el Modelo Pseudo Geoidal Local para el Municipio de la Ciudad de La Paz, con una desviación estándar igual a 0.146 metros.

Se delimitó como área de aplicación y estudio para el modelo pseudo geoidal local, la parte urbana de la ciudad de La Paz, en la cual se entran distribuidas las estaciones de la Red Geodésica del Municipio.

Se realizó la recopilación de información correspondiente a la Red Geodésica Municipal de las campañas 2005 – 2006 y 2012 – 2013, definiendo como marco de datos la información de coordenadas Este y Norte (UTM Z19 Sur), y alturas niveladas de la red establecida en la campaña 2005 – 2006, por contar con información confiable en cuanto a las alturas niveladas en sus estaciones.

Se realizó la depuración de la información a través del análisis de valores atípicos de N (ondulación pseudo geoidal), mediante el programa Minitab 18, estableciéndose 21 estaciones depuradas, 84 estaciones utilizadas en la generación del modelo y 8 estaciones de validación.

Se aplicó la regresión lineal múltiple para generar el modelo pseudo geoidal local cuyos resultados y valores estadísticos (análisis de correlación - R^2 ajustado < 0.7), determinan que el modelo no es acorde para realizar predicciones o estimaciones de valores de N, considerada en la investigación como la variable dependiente.

Se aplicó la regresión polinomial múltiple a través del polinomio de quinto grado, estableciéndose los coeficientes de la regresión, que permitieron pronosticar los valores de N, cuyos residuos en comparación a los valores iniciales de las estaciones, tienen como

promedio 0 metros, valor máximo igual a 0.416 metros y mínimo de 0.002 metros. Asimismo, el análisis de varianza que rechaza la hipótesis nula, establece que si existe relación entre las variables independientes (coordenada Este y coordenada Norte), con la variable dependiente N (ondulación pseudo geoidal), así también, los valores de correlación se encuentran dentro de los rangos recomendados, con un coeficiente de determinación R^2 con 77%, R^2 ajustado igual a 73% y el coeficiente de correlación múltiple con un 88%, que establecen que el modelo se encuentra ajustado para las predicción de valores de N, siendo estos confiables con una alta intensidad de relación entre variables independientes y dependientes.

Se realizó la comparación de valores de ondulación pseudo geoidal determinados mediante el modelo determinado por regresión polinomial múltiple de quinto grado, de las estaciones de validación, cuyos resultados en promedio alcanzan 0.111 metros, valor máximo de 0.293 metros y mínimo de 0.009 metros, que en función al valor del promedio de diferencias guarda relación con la desviación estándar definida igual a 0.146 metros. Lo cual representa que el Modelo Pseudo Geoidal Local de la Ciudad de La Paz es confiable para determinar o estimar valores de N (ondulación pseudo geoidal), de puntos que se encuentren dentro del área de estudio, simplemente introduciendo valores de las coordenadas Este y Norte en proyección UTM.

Esta metodología permite reducir costos y tiempos para levantamientos y trabajos que requieran datos de alturas niveladas, que son aplicadas a proyectos de ingeniería, cartografía, gestión territorial, medio ambiente, y todo aquel que requiera una altura geométrica referida al nivel medio del mar, con punto datum vertical "Arica".

En cuanto a las precisiones alcanzadas con el modelo, permite ser aplicados para proyectos y trabajos topográficos, teniendo que hacer consideraciones en caso de aplicación geodésica.

Respecto a la comparación de los resultados del modelo con valores obtenidos de modelos geoidales globales como el EGM96, EGM08 y EIGEN-6C4, y siendo que estos proporcionan alturas ortométricas, es evidente que al ser comparadas con alturas del tipo geométrico, como las niveladas, discrepan en valor alcanzando diferencias entre 2 y 3 metros.

Rescatando únicamente que el modelo EGM08 es quien presenta diferencias menores incluso respecto al EIGEN-6C4.

5.2. RECOMENDACIONES

Es importante considerar que el establecimiento de la red de nivelación de la Ciudad de La Paz, corresponde a una nivelación geométrica determinada mediante alturas niveladas, mismas que fueron definidas para una determinada época, que con el transcurso del tiempo van sufriendo modificaciones, tanto respecto al nivel medio de los mares, como también, respecto a la posición horizontal de la estación, debido al desplazamiento de las placas tectónicas sobre las cuales se encuentran emplazadas, es decir, si obtenemos una coordenada con un punto para estimar su ondulación pseudo geoidal, se deberá verificar que este corresponda al ITRF y época igual que de las estaciones de la Red Geodésica Municipal 2005 – 2006, siendo que en base a esta se generó el modelo pseudo geoidal local de la Ciudad de La Paz, esta consideración de no ser tomada en cuenta podría generar un grado mayor de error.

Los valores de ondulación pseudo geoidal a ser estimados por el modelo, deben ser determinados para puntos que se encuentren dentro del área de estudio.

Se recomienda realizar la verificación de la precisión del modelo, mediante resultados de levantamientos altimétricos aplicando nivelación geométrica, misma que dará fe de las diferencias que puedan existir.

Para generar el modelo pseudo geoidal local, se debe tener mucho cuidado con los valores atípicos de ondulación pseudo geoidal determinada mediante valores de alturas elipsoidales y alturas niveladas, siendo que afectan directamente al ajuste de la regresión en relación a los datos considerados.

CAPÍTULO VI

BIBLIOGRAFÍA

Acero Velo, L., & Serruto Flores, J. C. (2018). *Modelo Geoidal con receptores GNSS en método cinemático y nivelación geométrica para la estimación de alturas ortométricas en la ciudad Universitaria UNA - Puno*. Universidad Nacional del Altiplano. <http://repositorio.unap.edu.pe/handle/UNAP/7690>

Canizares, E. C., Tierra, A. R., & Leiva, C. A. (2015). *Sistema de Alturas para el Ecuador*. 7.

Carrión Sánchez, J. L. (2013). *Generación de una malla de ondulaciones geoidales por el método GPS/Nivelación y redes neuronales artificiales a partir de datos dispersos* [Magister en Geomática, Universidad Nacional de La Plata]. <https://doi.org/10.35537/10915/29002>

Cornero Cecilia, C., Ayelen, P., Cristina, P. M., & Laura, B. (2015). *Comparación de modelos geopotenciales recientes en Argentina*. 11.

Ince, E. S., Barthelmes, F., Reißland, S., Elger, K., Förste, C., Flechtner, F., & Schuh, H. (2019). ICGEM – 15 years of successful collection and distribution of global gravitational models, associated services, and future plans. *Earth System Science Data*, 11(2), 647-674. <https://doi.org/10.5194/essd-11-647-2019>

Instituto Geográfico Militar. (2006). *Informe Técnico de nivelación geométrica de puntos de control vertical—Gobierno Municipal de la Ciudad de La Paz*.

Instituto Geográfico Militar. (2013). *Informe Técnico Fortalecimiento y actualización de la base catastral cartográfica y geodésica del Municipio de La Paz.*

Joaquín Amat Rodrigo. (2016). *Introducción a la Regresión Lineal Múltiple.* 81.

Laura Sánchez. (2006). *Modernización de los sistemas de alturas existentes en América Latina y el Caribe.* 19.

Leal Vélez Anderson. (2017). *Modelo Geoidal para la Zona Norte de Colombia: Propuesta Metodológica* [Universidad Distrital Francisco José de Caldas].
<http://repositorios.rumbo.edu.co/handle/123456789/121030>

Tarrío Mosquera, J. (2017). *Evaluación del Modelo Global de Geopotencial EIGEN-6C4 mediante GNSS/Nivelación en la Región Metropolitana de Chile.* Simposio SIRGAS 2017, 1, 1.

Tierra, A. (2015). *Evaluación del EGM08 y EGM96 en el Ecuador a partir de datos GPS y Nivelación geométrica.* Centro de Investigaciones Científicas. Escuela Politécnica del Ejército., 9.

Tierra, A., & Acurio, V. (2016, junio). *Modelo neuronal para la predicción de la altura geoidal local en el Ecuador.* Junio 2016, 13, 72.

Vinuesa, P. (2016, octubre 22). *Regresión lineal simple y polinomial: Teoría y práctica.* Centro de Ciencias Genómicas, UNAM, v 0.7.
https://www.ccg.unam.mx/~vinuesa/R4biosciences/docs/Tema9_regresion.html

Páginas electrónicas consultadas

https://es.wikipedia.org/wiki/Coeficiente_de_determinaci%C3%B3n

https://es.wikipedia.org/wiki/Coeficiente_de_determinaci%C3%B3n_corregido

<https://economipedia.com/definiciones/r-cuadrado-ajustado-coeficiente-de-determinacion-ajustado.html>

<https://aprendeia.com/algorithmo-regresion-polinomial-machine-learning/>