



**ESPE**

**UNIVERSIDAD DE LAS FUERZAS ARMADAS**  
**INNOVACIÓN PARA LA EXCELENCIA**

**DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA  
CONSTRUCCIÓN**

**CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE**

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO  
DE INGENIERO GEÓGRAFO Y DEL MEDIO AMBIENTE**

**TEMA: “GENERACIÓN DE UN MODELO DE PREDICCIÓN DE LA  
VARIABLE ONDULACIÓN GEOIDAL, PARA LA ZONA RURAL DEL  
CANTÓN GUAYAQUIL, MEDIANTE EL USO DEL MÉTODO  
COKRIGING”**

**AUTOR: PALACIOS OREJUELA, IVÁN FERNANDO**

**DIRECTOR: MSC. LEIVA GONZÁLEZ, CÉSAR ALBERTO**

**SANGOLQUÍ – ECUADOR**

**2019**



DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA  
CONSTRUCCIÓN

CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE

CERTIFICACIÓN

Certifico que el trabajo de titulación: “*GENERACIÓN DE UN MODELO DE PREDICCIÓN DE LA VARIABLE ONDULACIÓN GEOIDAL, PARA LA ZONA RURAL DEL CANTÓN GUAYAQUIL, MEDIANTE EL USO DEL MÉTODO COKRIGING*” fue realizado por el señor *Palacios Orejuela, Iván Fernando* el mismo que ha sido revisado en su totalidad, analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas – ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Sangolquí, 15 de enero de 2019

  
Ing. Leiva González, César Alberto. **MSc.**





DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA  
CONSTRUCCIÓN

CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE

AUTORÍA DE RESPONSABILIDAD

Yo, *Palacios Orejuela, Iván Fernando*, declaro que el contenido, ideas y criterios del trabajo de titulación: “*GENERACIÓN DE UN MODELO DE PREDICCIÓN DE LA VARIABLE ONDULACIÓN GEOIDAL, PARA LA ZONA RURAL DEL CANTÓN GUAYAQUIL, MEDIANTE EL USO DEL MÉTODO COKRIGING*” es de mi autoría y responsabilidad, cumpliendo con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas – ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Consecuentemente, el contenido de la investigación mencionada es veraz.

Sangolquí, 15 de enero de 2019

A handwritten signature in blue ink, appearing to read 'Iván Fernando Palacios Orejuela', is written over a horizontal dotted line.

Palacios Orejuela, Iván Fernando





DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA  
CONSTRUCCIÓN

CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE

AUTORIZACIÓN

Yo, *Palacios Orejuela, Iván Fernando*, autorizo a la Universidad de las Fuerzas Armadas – ESPE publicar el trabajo de titulación: “*GENERACIÓN DE UN MODELO DE PREDICCIÓN DE LA VARIABLE ONDULACIÓN GEOIDAL, PARA LA ZONA RURAL DEL CANTÓN GUAYAQUIL, MEDIANTE EL USO DEL MÉTODO COKRIGING*” en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

Sangolquí, 15 de enero de 2019



Palacios Orejuela, Iván Fernando



## DEDICATORIA

En primer lugar, quiero agradecer a Dios por permitirme llegar hasta aquí, y a la madre que tengo en el cielo, la Purísima de Macas, a ella, le agradezco por cuidarme, protegerme y escucharme en mis oraciones que siempre le he tenido presente desde el día que dejé atrás mi tierra querida, para emprender esta etapa de mi vida.

A mis padres, quienes fueron, son y serán siempre mi mayor orgullo y motivación, gracias por su amor incondicional y por estar ahí cuando más lo necesite, todo lo bueno de mi persona ha sido por su educación y ejemplo, y cada logro alcanzado siempre será para ustedes dos.

A mi hermano, que desde el cielo me ha acompañado a lo largo de mi vida y siempre vivirá en mi corazón, a él le dedico este trabajo.

A toda mi familia, por estar pendiente y confiar en mí, gracias a todos ustedes, en especial a mis tíos Teresita y Huguito, por abrirme las puertas de su hogar, y a mi tía Lolita por su apoyo y amor brindado, sin ustedes no hubiese logrado este objetivo.

A mi personita favorita, por creer en mí, brindarme su amor, apoyo, comprensión y estar presente en los momentos difíciles.

## **AGRADECIMIENTOS**

Al Ing. César Leiva, por su confianza, paciencia y consejos para el desarrollo de este trabajo, mi admiración y respeto como profesional, persona y amigo.

Al Dr. Marco Luna, por las enseñanzas y pautas esenciales en la elaboración de este proyecto de investigación, así como su paciencia y consejos brindados a mi persona.

A mis profesores de la universidad, gracias por depositar en mí su conocimiento, mi gratitud a todos ellos.

Gracias a mis amigos/as que me acompañaron en esta linda etapa de mi vida, por sus palabras de aliento en los buenos y malos momentos.

Mi agradecimiento al Instituto Geográfico Militar, por proporcionar la información requerida para este trabajo, y su atención afable ante las consultas realizadas.

## ÍNDICE

DEDICATORIA.....	iv
AGRADECIMIENTOS .....	v
ÍNDICE .....	vi
ÍNDICE DE TABLAS .....	xii
ÍNDICE DE FIGURAS.....	xvi
RESUMEN.....	xxiv
ABSTRACT .....	xxv
CAPÍTULO I INTRODUCCIÓN .....	1
1.1 Antecedentes .....	1
1.2 Planteamiento del problema.....	3
1.3 Justificación e importancia.....	4
1.4 Objetivos .....	5
1.4.1 Objetivo general .....	5
1.4.2 Objetivos específicos.....	5
1.5 Metas .....	6
CAPÍTULO II MARCO TEÓRICO .....	7
2.1 Fundamentos teóricos.....	7
2.1.1 Geodesia.....	7

2.1.2	Sistemas geodésicos de referencia .....	8
2.1.3	Superficies de referencia de la Tierra.....	10
2.1.4	Relación entre las superficies de referencia .....	12
2.1.5	Sistemas de coordenadas .....	13
2.1.6	Sistemas de alturas .....	21
2.1.7	Técnicas de nivelación .....	28
2.1.8	Ondulación geoidal .....	31
2.1.9	Modelos geopotenciales globales.....	33
2.1.10	Anomalías de gravedad .....	35
2.1.11	Técnica Remove – Restore.....	41
2.2	Fundamento legal .....	42
2.3	Hipótesis.....	42
<b>CAPÍTULO III GEOESTADÍSTICA .....</b>		<b>43</b>
3.1	Estadística espacial.....	43
3.2	Geoestadística.....	43
3.3	Criterios geoestadísticos fundamentales .....	44
3.3.1	Variable regionalizada.....	44
3.3.2	Teoría de momentos de una función aleatoria .....	45
3.3.3	Hipótesis de estacionariedad .....	47

3.3.4	Funciones de correlación espacial.....	49
3.3.5	Isotropía y anisotropía.....	50
3.4	Análisis exploratorio de datos.....	51
3.5	Análisis estructural de datos.....	52
3.5.1	Estimación del semivariograma.....	52
3.5.2	Características del semivariograma.....	53
3.5.3	Modelos teóricos de semivarianza.....	55
3.5.4	Modelamiento del semivariograma experimental.....	58
3.6	Predicción espacial.....	59
3.6.1	Criterios de una predicción espacial.....	59
3.6.2	Ponderaciones Kriging.....	60
3.6.3	Validación cruzada.....	61
3.7	Geoestadística univariada.....	63
3.7.1	Kriging ordinario.....	64
3.7.2	Kriging universal.....	65
3.7.3	Kriging residual.....	66
3.8	Geoestadística multivariada.....	68
3.8.1	Momentos cruzados de segundo orden.....	68
3.8.2	Propiedades del semivariograma cruzado.....	69

3.8.3 Estimador de semivariograma cruzado .....	70
3.8.4 Modelo de correogionalización lineal (MCL).....	70
3.8.5 Pesos Cokriging.....	73
3.8.6 Validación del modelo de cross – semivariograma.....	73
3.8.7 Criterios adicionales del Cokriging.....	74
3.8.8 Cokriging ordinario .....	75
3.8.9 Cokriging universal .....	78
3.8.10 Cokriging residual .....	79
CAPÍTULO IV METODOLOGÍA .....	80
4.1 Recopilación y validación de información.....	80
4.4.1 Proyecto Guayaquil .....	80
4.4.2 Área de estudio.....	80
4.4.3 Datos de posicionamiento GPS .....	80
4.4.4 Datos de altura nivelada .....	81
4.4.5 Obtención de la base de datos .....	82
4.2 Selección de posibles variables auxiliares .....	83
4.3 Determinación de la variable auxiliar .....	85
4.3.1 Gráfico cuartil – cuartil .....	86
4.3.2 Gráfico de dispersión .....	90

4.3.3	Gráfico de correlación espacial o correlograma.....	94
4.3.4	Coefficiente de correlación de Pearson .....	98
4.4	Análisis exploratorio de datos .....	99
4.4.1	Análisis exploratorio de Cokriging con el 5% de datos .....	100
4.4.2	Análisis exploratorio de Cokriging con el 10% de datos .....	106
4.4.3	Análisis exploratorio de Cokriging con el 15% de datos .....	111
4.4.4	Análisis exploratorio de Cokriging con el 20% de datos .....	116
4.4.5	Análisis exploratorio del total de datos .....	121
4.4.6	Análisis exploratorio de Kriging con 299 datos.....	124
4.5	Definición de la resolución del modelo de ondulación geoidal .....	127
4.6	Análisis estructural de datos.....	129
4.6.1	Obtención, predicción y validación del modelo con 5% de datos.....	129
4.6.2	Obtención, predicción y validación del modelo con 10% de datos.....	143
4.6.3	Obtención, predicción y validación del modelo con 15% de datos.....	151
4.6.4	Obtención, predicción y validación del modelo con 20% de datos.....	160
4.6.5	Obtención, predicción y validación del modelo con 299 de datos .....	170
4.7	Validación del modelo de ondulación geoidal .....	178
CAPÍTULO V ANÁLISIS Y DISCUSIÓN DE RESULTADOS.....		182
CAPÍTULO VI CONCLUSIONES Y RECOMENDACIONES .....		203

6.1 Conclusiones .....	203
6.2 Recomendaciones.....	205
REFERENCIAS BIBLIOGRÁFICAS .....	206

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> <i>Modelos teóricos acotados de semivariogramas</i> .....	56
<b>Tabla 2.</b> <i>Modelos teóricos no acotados de semivariogramas</i> .....	57
<b>Tabla 3.</b> <i>Coefficientes de correlación de Pearson</i> .....	99
<b>Tabla 4.</b> <i>Resumen de los datos de las variables con Cokriging al 5% de datos</i> .....	101
<b>Tabla 5.</b> <i>Estadísticos descriptivos de las variables con Cokriging al 5% de datos</i> .....	101
<b>Tabla 6.</b> <i>Pruebas de normalidad de las variables con Cokriging al 5% de datos</i> .....	104
<b>Tabla 7.</b> <i>Resumen de los datos de las variables con Cokriging al 10% de datos</i> .....	106
<b>Tabla 8.</b> <i>Estadísticos descriptivos de las variables con Cokriging al 10% de datos</i> .....	107
<b>Tabla 9.</b> <i>Pruebas de normalidad de las variables con Cokriging al 10% de datos</i> .....	109
<b>Tabla 10.</b> <i>Resumen de los datos de las variables con Cokriging al 15% de datos</i> .....	111
<b>Tabla 11.</b> <i>Estadísticos descriptivos de las variables con Cokriging al 15% de datos</i> .....	112
<b>Tabla 12.</b> <i>Pruebas de normalidad de las variables con Cokriging al 15% de datos</i> .....	114
<b>Tabla 13.</b> <i>Resumen de los datos de las variables con Cokriging al 20% de datos</i> .....	116
<b>Tabla 14.</b> <i>Estadísticos descriptivos de las variables con Cokriging al 20% de datos</i> .....	116
<b>Tabla 15.</b> <i>Pruebas de normalidad de las variables con Cokriging al 20% de datos</i> .....	119
<b>Tabla 16.</b> <i>Resumen del total de datos con transformación logarítmica</i> .....	121
<b>Tabla 17.</b> <i>Estadísticos descriptivos del total de datos con transformación logarítmica</i> .....	121
<b>Tabla 18.</b> <i>Resumen del total de datos con transformación box-cox</i> .....	121
<b>Tabla 19.</b> <i>Estadísticos descriptivos del total de datos con transformación box-cox</i> .....	122
<b>Tabla 20.</b> <i>Pruebas de normalidad para el total de puntos con transformación logarítmica</i> ....	123
<b>Tabla 21.</b> <i>Pruebas de normalidad para el total de puntos con transformación box-cox</i> .....	124

<b>Tabla 22.</b> <i>Resumen de los datos de la variable con Kriging (299 datos)</i> .....	124
<b>Tabla 23.</b> <i>Estadísticos descriptivos de la variable con Kriging (299 datos)</i> .....	124
<b>Tabla 24.</b> <i>Pruebas de normalidad de la variable con Kriging (299 datos)</i> .....	126
<b>Tabla 25.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging ordinario al 5% de datos</i> ...	130
<b>Tabla 26.</b> <i>Estadísticos de la validación cruzada con Cokriging ordinario al 5% de datos</i> .....	132
<b>Tabla 27.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging universal al 5% de datos</i> ...	135
<b>Tabla 28.</b> <i>Estadísticos de la validación cruzada con Cokriging universal al 5% de datos</i> .....	136
<b>Tabla 29.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging residual al 5% de datos</i> .....	139
<b>Tabla 30.</b> <i>Parámetros del modelo Exponencial ajustado – Cokriging residual al 5% de datos</i>	139
<b>Tabla 31.</b> <i>Parámetros del modelo Gaussiano ajustado – Cokriging residual al 5% de datos</i> ..	139
<b>Tabla 32.</b> <i>Estadísticos de la validación cruzada con Cokriging residual al 5% de datos</i> .....	142
<b>Tabla 33.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging ordinario al 10% de datos</i> .	144
<b>Tabla 34.</b> <i>Estadísticos de la validación cruzada con Cokriging ordinario al 10% de datos</i> .....	145
<b>Tabla 35.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging universal al 10% de datos</i> .	146
<b>Tabla 36.</b> <i>Estadísticos de la validación cruzada con Cokriging universal al 10% de datos</i> .....	147
<b>Tabla 37.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging residual al 10% de datos</i> ...	149
<b>Tabla 38.</b> <i>Parámetros del modelo Gaussiano ajustado – Cokriging residual al 10% de datos</i>	149
<b>Tabla 39.</b> <i>Estadísticos de la validación cruzada con Cokriging residual al 10% de datos</i> .....	151
<b>Tabla 40.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging ordinario al 15% de datos</i> .	152
<b>Tabla 41.</b> <i>Estadísticos de la validación cruzada con Cokriging ordinario al 15% de datos</i> .....	154
<b>Tabla 42.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging universal al 15% de datos</i> .	155
<b>Tabla 43.</b> <i>Estadísticos de la validación cruzada con Cokriging universal al 15% de datos</i> .....	155
<b>Tabla 44.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging residual al 15% de datos</i> ...	157

<b>Tabla 45.</b> <i>Parámetros del modelo Gaussiano ajustado – Cokriging residual al 15% de datos</i>	157
<b>Tabla 46.</b> <i>Estadísticos de la validación cruzada con Cokriging residual al 15% de datos</i> .....	159
<b>Tabla 47.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging ordinario al 20% de datos.</i>	160
<b>Tabla 48.</b> <i>Estadísticos de la validación cruzada con Cokriging ordinario al 20% de datos</i> .....	162
<b>Tabla 49.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging universal al 20% de datos</i> .	163
<b>Tabla 50.</b> <i>Estadísticos de la validación cruzada con Cokriging universal al 20% de datos</i> .....	163
<b>Tabla 51.</b> <i>Parámetros del modelo Potencia ajustado – Cokriging residual al 20% de datos</i> ...	166
<b>Tabla 52.</b> <i>Parámetros del modelo Esférico ajustado – Cokriging residual al 20% de datos</i> ....	166
<b>Tabla 53.</b> <i>Parámetros del modelo Gaussiano ajustado – Cokriging residual al 20% de datos</i>	166
<b>Tabla 54.</b> <i>Estadísticos de la validación cruzada con Cokriging residual al 20% de datos</i> .....	169
<b>Tabla 55.</b> <i>Parámetros del modelo Potencia ajustado – Kriging ordinario (299 datos)</i> .....	171
<b>Tabla 56.</b> <i>Estadísticos de la validación cruzada con Kriging ordinario (299 datos)</i> .....	173
<b>Tabla 57.</b> <i>Parámetros del modelo Circular ajustado – Kriging universal (299 datos)</i> .....	174
<b>Tabla 58.</b> <i>Estadísticos de la validación cruzada con Kriging universal (299 datos)</i> .....	175
<b>Tabla 59.</b> <i>Parámetros del modelo Estable ajustado – Kriging residual (299 datos)</i> .....	176
<b>Tabla 60.</b> <i>Estadísticos de la validación cruzada con Kriging residual (299 datos)</i> .....	178
<b>Tabla 61.</b> <i>Resumen de métodos y modelos teóricos usados</i> .....	179
<b>Tabla 62.</b> <i>Resultados de la validación del modelo generado con 16 puntos muestrales</i> .....	180
<b>Tabla 63.</b> <i>Resultados de la validación del modelo generado con 33 puntos muestrales</i> .....	181
<b>Tabla 64.</b> <i>Resultados de la validación del modelo generado con 49 puntos muestrales</i> .....	181
<b>Tabla 65.</b> <i>Resultados de la validación del modelo generado con 66 puntos muestrales</i> .....	181
<b>Tabla 66.</b> <i>Resultados de la validación del modelo generado con 299 puntos muestrales</i> .....	181
<b>Tabla 67.</b> <i>Resultados de la comparación entre el modelo de 49 puntos y otros modelos</i> .....	196

<b>Tabla 68.</b> <i>Resultados de la comparación entre el modelo de 66 puntos y otros modelos.....</i>	196
--	-----

## ÍNDICE DE FIGURAS

<i><b>Figura 1.</b></i> Relación entre el Geoide, Elipsoide y la superficie topográfica del planeta. ....	13
<i><b>Figura 2.</b></i> Sistema de coordenadas cartesianas. ....	15
<i><b>Figura 3.</b></i> Sistema de coordenadas geográficas.....	16
<i><b>Figura 4.</b></i> Ejemplo de sistema de coordenadas planas proyectadas según el plano.....	18
<i><b>Figura 5.</b></i> Ejemplo de sistema de coordenadas planas proyectadas según el punto origen .....	18
<i><b>Figura 6.</b></i> Sistema de coordenadas planas UTM.....	20
<i><b>Figura 7.</b></i> Ejemplo del sistema de coordenadas planas UTM. ....	21
<i><b>Figura 8.</b></i> Sistema de altura ortométrica. ....	25
<i><b>Figura 9.</b></i> Sistema de altura elipsoidal. ....	26
<i><b>Figura 10.</b></i> Sistema de altura nivelada. ....	27
<i><b>Figura 11.</b></i> Nivelación geométrica simple. ....	29
<i><b>Figura 12.</b></i> Nivelación geométrica compuesta. ....	29
<i><b>Figura 13.</b></i> Nivelación trigonométrica. ....	30
<i><b>Figura 14.</b></i> Ondulación Geoidal. ....	31
<i><b>Figura 15.</b></i> Reducción de gravedad por Aire libre. ....	36
<i><b>Figura 16.</b></i> Reducción de gravedad por Bouguer.....	38
<i><b>Figura 17.</b></i> Modelo de isostasia de Airy – Heiskanen.....	39
<i><b>Figura 18.</b></i> Semivariograma para un fenómeno estacionario de segundo orden.....	53
<i><b>Figura 19.</b></i> Vecindario de búsqueda por octantes. ....	62
<i><b>Figura 20.</b></i> Mapa de ubicación de la zona de estudio. ....	81
<i><b>Figura 21.</b></i> Diagrama de caja de los valores de ondulación geoidal. ....	83

<b>Figura 22.</b> qq – plot entre Anomalía de Aire libre (2.5') y ondulación geoidal.....	87
<b>Figura 23.</b> qq – plot entre Anomalía de Bouguer (2.5') y ondulación geoidal.....	87
<b>Figura 24.</b> qq – plot entre Anomalía Isostática (2') y ondulación geoidal.....	88
<b>Figura 25.</b> qq – plot entre DEM ALOS (12.5m) y ondulación geoidal.....	88
<b>Figura 26.</b> qq – plot entre DEM SIGTIERRAS (4m) y ondulación geoidal.....	89
<b>Figura 27.</b> qq – plot entre ondulación geoidal EGM08 (2.5') y ondulación geoidal.....	89
<b>Figura 28.</b> qq – plot entre ondulación geoidal EGM08 (1') y ondulación geoidal.....	90
<b>Figura 29.</b> Scatter – plot entre Anomalía de Aire libre (2.5') y ondulación geoidal.....	91
<b>Figura 30.</b> Scatter – plot entre Anomalía de Bouguer (2.5') y ondulación geoidal.....	91
<b>Figura 31.</b> Scatter – plot entre Anomalía Isostática (2') y ondulación geoidal.....	92
<b>Figura 32.</b> Scatter – plot entre DEM ALOS (12.5m) y ondulación geoidal.....	92
<b>Figura 33.</b> Scatter – plot entre DEM SIGTIERRAS (4m) y ondulación geoidal.....	93
<b>Figura 34.</b> Scatter – plot entre ondulación geoidal EGM08 (2.5') y ondulación geoidal.....	93
<b>Figura 35.</b> Scatter – plot entre ondulación geoidal EGM08 (1') y ondulación geoidal.....	94
<b>Figura 36.</b> Cross – correlación entre Anomalía de Aire libre (2.5') y ondulación geoidal.....	95
<b>Figura 37.</b> Cross – correlación entre Anomalía de Bouguer (2.5') y ondulación geoidal.....	95
<b>Figura 38.</b> Cross – correlación entre Anomalía Isostática (2') y ondulación geoidal.....	96
<b>Figura 39.</b> Cross – correlación entre DEM ALOS (12.5m) y ondulación geoidal.....	96
<b>Figura 40.</b> Cross – correlación entre DEM SIGTIERRAS (4m) y ondulación geoidal.....	97
<b>Figura 41.</b> Cross – correlación entre ondulación geoidal EGM08 (2.5') y ondulación geoidal.....	97
<b>Figura 42.</b> Cross – correlación entre ondulación geoidal EGM08 (1') y ondulación geoidal.....	98
<b>Figura 43.</b> Histograma de Ondulación geoidal real.....	102
<b>Figura 44.</b> Histograma de Ondulación geoidal EGM08 (1').....	102

<b>Figura 45.</b> Diagrama de caja de Ondulación geoidal real. ....	103
<b>Figura 46.</b> Diagrama de caja de Ondulación geoidal EGM08 (1'). ....	103
<b>Figura 47.</b> Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real. ....	105
<b>Figura 48.</b> Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real. ....	105
<b>Figura 49.</b> Gráfico de dispersión de la coordenada Este vs Ondulación geoidal EGM08 (1'). ..	105
<b>Figura 50.</b> Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal EGM08 (1').	106
<b>Figura 51.</b> Histograma de Ondulación geoidal real. ....	107
<b>Figura 52.</b> Histograma de Ondulación geoidal EGM08 (1'). ....	108
<b>Figura 53.</b> Diagrama de caja de Ondulación geoidal real. ....	108
<b>Figura 54.</b> Diagrama de caja de Ondulación geoidal EGM08 (1'). ....	109
<b>Figura 55.</b> Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real. ....	110
<b>Figura 56.</b> Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real. ....	110
<b>Figura 57.</b> Gráfico de dispersión de la coordenada Este vs Ondulación geoidal EGM08 (1'). ..	110
<b>Figura 58.</b> Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal EGM08 (1').	111
<b>Figura 59.</b> Histograma de Ondulación geoidal real. ....	112
<b>Figura 60.</b> Histograma de Ondulación geoidal EGM08 (1'). ....	113
<b>Figura 61.</b> Diagrama de caja de Ondulación geoidal real. ....	113
<b>Figura 62.</b> Diagrama de caja de Ondulación geoidal EGM08 (1'). ....	114
<b>Figura 63.</b> Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real. ....	115
<b>Figura 64.</b> Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real. ....	115
<b>Figura 65.</b> Gráfico de dispersión de la coordenada Este vs Ondulación geoidal EGM08 (1'). ..	115
<b>Figura 66.</b> Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal EGM08 (1').	116
<b>Figura 67.</b> Histograma de Ondulación geoidal real. ....	117

<b>Figura 68.</b> Histograma de Ondulación geoidal EGM08 (1').....	117
<b>Figura 69.</b> Diagrama de caja de Ondulación geoidal real. ....	118
<b>Figura 70.</b> Diagrama de caja de Ondulación geoidal EGM08 (1'). ....	118
<b>Figura 71.</b> Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real. ....	119
<b>Figura 72.</b> Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real.....	120
<b>Figura 73.</b> Gráfico de dispersión de la coordenada Este vs Ondulación geoidal EGM08 (1'). .	120
<b>Figura 74.</b> Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal EGM08 (1').	120
<b>Figura 75.</b> Histograma de Ondulación geoidal real con transformación logarítmica. ....	122
<b>Figura 76.</b> Histograma de Ondulación geoidal EGM08 (1') con transformación logarítmica. ...	122
<b>Figura 77.</b> Histograma de Ondulación geoidal real con transformación box-cox. ....	123
<b>Figura 78.</b> Histograma de Ondulación geoidal EGM08 (1') con transformación box-cox. ....	123
<b>Figura 79.</b> Histograma de la Ondulación real con 299 datos. ....	125
<b>Figura 80.</b> Diagrama de caja de Ondulación geoidal real con 299 datos. ....	125
<b>Figura 81.</b> Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real. ....	126
<b>Figura 82.</b> Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real.....	127
<b>Figura 83.</b> Cross – semivariograma ajustado al modelo Potencia. ....	130
<b>Figura 84.</b> Visualización del modelo de predicción y varianza de la ondulación geoidal. ....	131
<b>Figura 85.</b> Histograma de validación cruzada con 16 datos.....	133
<b>Figura 86.</b> Cross – semivariograma ajustado al modelo Potencia. ....	135
<b>Figura 87.</b> Histograma de validación cruzada con 16 datos.....	136
<b>Figura 88.</b> Cross – semivariograma ajustado al modelo Potencia. ....	137
<b>Figura 89.</b> Cross – semivariograma ajustado al modelo Exponencial .....	138
<b>Figura 90.</b> Cross – semivariograma ajustado al modelo Gaussiano.....	138

<b>Figura 91.</b> Visualización del modelo Potencia de predicción y varianza de residuales. ....	140
<b>Figura 92.</b> Visualización del modelo Exponencial de predicción y varianza de residuales.....	140
<b>Figura 93.</b> Visualización del modelo Gaussiano de predicción y varianza de residuales. ....	141
<b>Figura 94.</b> Histograma de validación cruzada del modelo Potencia con 16 datos. ....	141
<b>Figura 95.</b> Histograma de validación cruzada del modelo Exponencial con 16 datos.....	142
<b>Figura 96.</b> Histograma de validación cruzada del modelo Gaussiano con 16 datos. ....	142
<b>Figura 97.</b> Cross – semivariograma ajustado al modelo Potencia. ....	143
<b>Figura 98.</b> Visualización del modelo de predicción y varianza de la ondulación geoidal. ....	144
<b>Figura 99.</b> Histograma de validación cruzada con 33 datos.....	145
<b>Figura 100.</b> Cross – semivariograma ajustado al modelo Potencia. ....	146
<b>Figura 101.</b> Histograma de validación cruzada con 33 datos.....	147
<b>Figura 102.</b> Cross – semivariograma ajustado al modelo Potencia. ....	148
<b>Figura 103.</b> Cross – semivariograma ajustado al modelo Gaussiano.....	148
<b>Figura 104.</b> Visualización del modelo Potencia de predicción y varianza de residuales. ....	149
<b>Figura 105.</b> Visualización del modelo Gaussiano de predicción y varianza de residuales. ....	150
<b>Figura 106.</b> Histograma de validación cruzada del modelo Potencia con 33 datos. ....	150
<b>Figura 107.</b> Histograma de validación cruzada del modelo Gaussiano con 33 datos. ....	151
<b>Figura 108.</b> Cross – semivariograma ajustado al modelo Potencia .....	152
<b>Figura 109.</b> Visualización del modelo de predicción y varianza de la ondulación geoidal. ....	153
<b>Figura 110.</b> Histograma de validación cruzada con 49 datos.....	153
<b>Figura 111.</b> Cross – semivariograma ajustado al modelo Potencia. ....	154
<b>Figura 112.</b> Histograma de validación cruzada con 49 datos.....	155
<b>Figura 113.</b> Cross – semivariograma ajustado al modelo Potencia .....	156

<b>Figura 114.</b> Cross – semivariograma ajustado al modelo Gaussiano.....	157
<b>Figura 115.</b> Visualización del modelo Potencia de predicción y varianza de residuales. ....	158
<b>Figura 116.</b> Visualización del modelo Gaussiano de predicción y varianza de residuales. ....	158
<b>Figura 117.</b> Histograma de validación cruzada del modelo Potencia con 49 datos. ....	159
<b>Figura 118.</b> Histograma de validación cruzada del modelo Gaussiano con 49 datos. ....	159
<b>Figura 119.</b> Cross – semivariograma ajustado al modelo Potencia .....	160
<b>Figura 120.</b> Visualización del modelo de predicción y varianza de la ondulación geoidal. ....	161
<b>Figura 121.</b> Histograma de validación cruzada con 66 datos.....	161
<b>Figura 122.</b> Cross – semivariograma ajustado al modelo Potencia .....	162
<b>Figura 123.</b> Histograma de validación cruzada con 66 datos.....	163
<b>Figura 124.</b> Cross – semivariograma ajustado al modelo Potencia .....	164
<b>Figura 125.</b> Cross – semivariograma ajustado al modelo Esférico.....	165
<b>Figura 126.</b> Cross – semivariograma ajustado al modelo Gaussiano.....	165
<b>Figura 127.</b> Visualización del modelo Potencia de predicción y varianza de residuales. ....	167
<b>Figura 128.</b> Visualización del modelo Esférico de predicción y varianza de residuales. ....	167
<b>Figura 129.</b> Visualización del modelo Gaussiano de predicción y varianza de residuales. ....	168
<b>Figura 130.</b> Histograma de validación cruzada del modelo Potencia con 66 datos. ....	168
<b>Figura 131.</b> Histograma de validación cruzada del modelo Esférico con 66 datos. ....	169
<b>Figura 132.</b> Histograma de validación cruzada del modelo Gaussiano con 66 datos. ....	169
<b>Figura 133.</b> Semivariograma ajustado con el modelo Potencia. ....	171
<b>Figura 134.</b> Visualización del modelo de predicción de la ondulación geoidal.....	172
<b>Figura 135.</b> Histograma de validación cruzada del modelo Potencia con 299 datos. ....	172
<b>Figura 136.</b> Semivariograma ajustado con el modelo Circular.....	174

<b>Figura 137.</b> Histograma de validación cruzada del modelo Circular con 299 datos.....	174
<b>Figura 138.</b> Semivariograma ajustado con el modelo Estable. ....	176
<b>Figura 139.</b> Visualización del modelo de predicción de residuales de la ondulación geoidal. ..	177
<b>Figura 140.</b> Histograma de validación cruzada del modelo Estable con 299 datos. ....	177
<b>Figura 141.</b> Comparación entre valores reales y predichos – Cokriging ordinario 16 puntos. ..	184
<b>Figura 142.</b> Comparación entre valores reales y predichos – Cokriging residual 16 puntos. ....	184
<b>Figura 143.</b> Comparación entre valores reales y predichos – Cokriging universal 16 puntos. ..	185
<b>Figura 144.</b> Comparación entre valores reales y predichos – Cokriging ordinario 33 puntos. ..	185
<b>Figura 145.</b> Comparación entre valores reales y predichos – Cokriging residual 33 puntos. ....	186
<b>Figura 146.</b> Comparación entre valores reales y predichos – Cokriging universal 33 puntos. ..	186
<b>Figura 147.</b> Comparación entre valores reales y predichos – Cokriging ordinario 49 puntos. ..	187
<b>Figura 148.</b> Comparación entre valores reales y predichos – Cokriging residual 49 puntos. ....	187
<b>Figura 149.</b> Comparación entre valores reales y predichos – Cokriging universal 49 puntos. ..	188
<b>Figura 150.</b> Comparación entre valores reales y predichos – Cokriging ordinario 66 puntos. ..	188
<b>Figura 151.</b> Comparación entre valores reales y predichos – Cokriging residual 66 puntos. ....	189
<b>Figura 152.</b> Comparación entre valores reales y predichos – Cokriging universal 66 puntos. ..	189
<b>Figura 153.</b> Comparación entre valores reales y predichos – Kriging ordinario 299 puntos.....	190
<b>Figura 154.</b> Comparación entre valores reales y predichos – Kriging residual 299 puntos.....	190
<b>Figura 155.</b> Comparación entre valores reales y predichos – Kriging universal 299 puntos.....	191
<b>Figura 156.</b> Mapa variográfico con Kriging ordinario – 299 puntos. ....	192
<b>Figura 157.</b> Mapa variográfico con Kriging residual – 299 puntos. ....	192
<b>Figura 158.</b> Mapa variográfico con Kriging universal – 299 puntos. ....	193
<b>Figura 159.</b> Mapa variográfico con Cokriging ordinario – 66 puntos. ....	193

<b>Figura 160.</b> Mapa variográfico con Cokriging residual – 66 puntos. ....	194
<b>Figura 161.</b> Mapa variográfico con Cokriging universal – 66 puntos. ....	194
<b>Figura 162.</b> Anomalía de Aire libre en la zona de estudio (puntos azules).....	197
<b>Figura 163.</b> Anomalía de Bouguer en la zona de estudio (puntos azules) .....	198
<b>Figura 164.</b> Anomalía Isostática en la zona de estudio (puntos azules).....	198
<b>Figura 165.</b> Modelo de ondulación geoidal generado con Cokriging universal con 66 puntos.	201
<b>Figura 166.</b> Modelo de errores de ondulación geoidal (Cokriging universal con 66 puntos)....	202

## RESUMEN

La ondulación geoidal es la separación entre el Elipsoide y Geoide, sin embargo, este último es complicado de establecer debido al desconocimiento de la gravedad media. Para fines prácticos y del estudio, se considera al nivel medio del mar como una aproximación cercana al Geoide, es decir que la ondulación geoidal está definida por la diferencia entre alturas elipsoidales y niveladas. En Ecuador, se trabaja con alturas niveladas geoméricamente, pero su obtención es muy costosa, lo que lleva a buscar métodos para optimizar su realización. Cokriging es una técnica de interpolación no tradicional que mejora la eficiencia de las estimaciones, gracias a que utiliza tanto la correlación espacial de las muestras y la relación entre fenómenos para predecir el valor de la variable de interés. El objetivo fue generar un modelo predictivo de ondulación geoidal mediante Cokriging para la zona rural del cantón Guayaquil. Se partió de un total de 332 puntos que fueron divididos en 16, 33, 49, 66 muestras y se aplicó Cokriging (y Kriging) ordinario, residual y universal con la ondulación geoidal derivada del EGM08 como variable auxiliar. Los modelos generados fueron validados con puntos excluidos del modelamiento y se calcularon indicadores estadísticos para facilitar la toma de decisión. El modelo de Cokriging universal con 66 puntos fue el mejor, con un RMSE final de 8cm. y RSR de 2cm. A través de Cokriging se alcanzó mayores precisiones que con Kriging, lo que reduce presupuestos y potencia el uso de datos GNSS/nivelación para obtener la ondulación geoidal.

### **PALABRAS CLAVE:**

- **ALTURAS NIVELADAS**
- **COKRIGING**
- **EGM08**
- **GNSS/NIVELACIÓN**
- **NIVEL MEDIO DEL MAR**

## ABSTRACT

The geoid undulation is the separation between the Ellipsoid and Geoid, however the latter is complicated to establish due to the ignorance of the average gravity. For practical and study purposes, it's considered to mean sea level as a close approximation to the Geoid, it is to say that the geoidal undulation is defined as the difference between the ellipsoidal and levelled heights. In Ecuador, work is done with geometrically levelled heights, but obtaining them is very expensive, which leads looking for methods to optimize their performance. Cokriging is a non-traditional interpolation technique that improves the efficiency of estimates, thanks to the fact that it uses both the spatial correlation of the samples and the relationship between phenomena to predict the value of the variable of interest. The objective was to generate a predictive model of the geoid undulation through Cokriging for the rural area of the canton of Guayaquil. We started from a total of 332 points that were divided into 16, 33, 49, 66 samples and applied ordinary, residual and universal Cokriging (and Kriging) with the geoid undulation derived from the EGM08 as auxiliary variable. The generated models were validated with points excluded from modeling and statistical indicators were calculated to facilitate decision making. The universal Cokriging model with 66 points was the best, with a final RMSE of 8cm. and RSR of 2cm. Through Cokriging, greater precision was achieved than with Kriging, which reduces budgets and enhances the use of GNSS/levelling data to obtain the geoid undulation.

### KEYWORDS:

- **LEVELLED HEIGHTS**
- **COKRIGING**
- **EGM08**
- **GNSS/LEVELLING**
- **MEAN SEA LEVEL**

## CAPÍTULO I

### INTRODUCCIÓN

#### 1.1 Antecedentes

En la actualidad, con el desarrollo de las tecnologías geoespaciales, es posible determinar la ubicación de un objeto en el espacio. Gracias a técnicas GNSS (Global Navigation Satellite System), se ha resuelto la obtención de coordenadas en la componente horizontal, sean estas: latitud ( $\varphi$ ), longitud ( $\lambda$ ) geográfica, o coordenadas planas proyectadas  $x, y$ ; pero a pesar de esto, aún no es posible conseguir la coordenada en la componente vertical en forma directa, debido a que, por definición las alturas que se obtienen con GNSS están referidas a un Elipsoide, es decir, son alturas de carácter geométrico (Rapp, 1984), las mismas, que no pueden ser usadas para los trabajos sobre la superficie física del planeta.

Como una alternativa para la obtención de la coordenada vertical referida al nivel medio del mar, se utilizan los modelos geopotenciales globales, entre los más conocidos se encuentran el EGM96 (Earth Gravitational Model 1996), y el EGM08 (Earth Gravitational Model 2008), con los cuales se puede transformar las alturas elipsoidales en alturas sobre el nivel medio del mar, mediante la ondulación geoidal, que se define como la diferencia entre las alturas elipsoidales y ortométricas, es decir, Elipsoide y Geoide respectivamente. Para este trabajo de investigación se considera al Geoide muy próximo con el nivel medio del mar, por lo que se entiende a la ondulación geoidal como la diferencia de alturas elipsoidales y alturas referidas al nivel medio del mar (alturas niveladas).

Lamentablemente, estos modelos geopotenciales, no brindan las precisiones requeridas para trabajos de ingeniería, ya que, en Ecuador existen variaciones entre los valores obtenidos con GPS/nivelación y los modelos en el rango de metros, que alcanzan diferencias desde 4 metros (Tierra, 2009), hasta 32 metros (Drewes & Sánchez, 2002) según el área de trabajo.

Ante la necesidad de contar con coordenadas verticales referidas al nivel medio del mar, el Instituto Geográfico Militar (IGM), el cual es el ente rector de las actividades relacionadas a la cartografía en el país, lleva a cabo campañas de nivelación, materializado en placas que conforman la Red de Control Básico Vertical del IGM, a lo largo y ancho del territorio nacional. Estas coordenadas verticales están referidas al datum vertical del país, ubicado en La Libertad, provincia de Santa Elena.

La técnica de nivelación geométrica es usada por el IGM para la obtención de la coordenada vertical (Cañizares, 2015), la cual se ejecuta a través de métodos ópticos de medición y varían de acuerdo al campo de gravedad del sitio de trabajo. A pesar de ser una técnica muy precisa, su realización es demandante en costos, personal y tiempos de ejecución, por lo que se buscan métodos que permitan estimar el valor de la coordenada vertical de forma más eficiente. La geoestadística es una posible alternativa para dar solución a la obtención del valor de la componente vertical mediante el modelamiento de la ondulación geoidal, como variable objetivo.

Según Matheron (1970), quien es considerado el padre de la geoestadística, se refiere a esta, como la aplicación de la teoría de las variables regionalizadas a la estimación de valores. Se la puede definir como una rama de la estadística que estudia los fenómenos espaciales (Journel & Huijbregts, 1978), para estimar, predecir y simular los mismos (Myers, 1987). Esta predicción espacial se lleva a cabo a través de un predictor espacial, cuya función permite mediante sus

parámetros, interpolar con valores muestreados a un punto desconocido. El Kriging, es conocido como el mejor predictor lineal insesgado (BLUE, por sus siglas en inglés Best Linear Unbiased Estimator), debido a su característica de insesgamiento y varianza mínima para datos aleatoriamente dispersos (Oliver, 2010).

Existen varios métodos dentro del Kriging (simple, ordinario, residual, bloque, etc.), que realizan un análisis univariante (una variable), sin embargo, se puede realizar un estudio más detallado a través de métodos multivariante, como el Cokriging, que es semejante al Kriging, pero se diferencian en que este último utiliza únicamente la correlación espacial para determinar los coeficientes en el estimador lineal, en cambio, el Cokriging usa tanto la correlación espacial y la correlación entre las variables aleatorias al mismo tiempo para predecir el valor de la variable de interés (Giraldo, 2007). Esta técnica es recomendable usarla cuando la variable objetivo tiene altos costos de muestreo (Bogaert, Mahau, & Beckers, 1995), como es el caso de la ondulación geoidal. Definir una metodología clara para el uso correcto de este método geoestadístico, podría optimizar recursos, y lograr altas precisiones para los trabajos que requieran coordenadas sobre el nivel medio del mar.

## **1.2 Planteamiento del problema**

La obtención de alturas referidas al nivel medio del mar requiere un trabajo demandante en costo y tiempo, ya que se realiza mediante nivelación diferencial, por lo que, en labores donde se necesitan una gran cantidad de datos precisos de la coordenada vertical, utilizar este método para su cálculo, sería poco práctico en términos de eficiencia. Actualmente, con la tecnología GNSS, en el Ecuador no es posible conseguir la coordenada vertical de forma directa sobre el nivel de referencia oficial del país (datum vertical), ya que, por principio esta es de carácter geométrico, sin

embargo, se puede conseguir alturas referidas al nivel medio del mar a partir de la altura elipsoidal determinada mediante posicionamiento satelital, y la ondulación geoidal a través de la expresión:

$$H \cong h - n \quad (1)$$

Donde:  $n$  es la ondulación geoidal;  $h$  es la altura elipsoidal;  $H$  es la altura ortométrica (altura referida al nivel medio del mar).

Es por ello que se propone el uso de técnicas geoestadísticas que permiten generar un modelo de precisión de la variable ondulación geoidal a través del método Cokriging, partiendo de pocos datos dispersos de ondulación geoidal y una variable auxiliar densamente muestreada y correlacionada espacialmente con la primera, que permita conseguir alturas referidas al nivel medio del mar.

### **1.3 Justificación e importancia**

El proyecto de investigación es de importancia, debido a que permite optimizar recursos para obtener alturas referidas al nivel medio del mar, ya que se conseguirán coordenadas verticales precisas con el uso de pocos datos dispersos de ondulación geoidal combinados con datos de posicionamiento satelital, logrando los mismos o hasta mejores resultados que con el estimador Kriging.

Existen varios estudios acerca de la aplicación de métodos geoestadísticos en el país, de los cuales, pocos han estudiado la variable objeto del presente proyecto de investigación, sin embargo, el valor agregado de este trabajo es que, se propone un método no tradicional (Cokriging) como una solución eficiente, que permita determinar valores de altura precisos con pocos datos de una variable costosa de obtener y difícil de muestrear como lo es la ondulación geoidal.

De esta forma, resultan beneficiarios los profesionales que hacen uso indispensable de valores precisos de alturas sobre el nivel medio del mar, como topógrafos e ingenieros y demás profesionales que estén relacionados a las geociencias, y permitirá facilitar los trabajos que utilizan coordenadas verticales referidas al nivel medio del mar como, por ejemplo: construcción de carreteras, drenajes, topografía, sistemas de alcantarillado, entre otros.

Además, el presente proyecto se relaciona con el noveno punto de los Objetivos de Desarrollo Sostenible: Industria, innovación e infraestructura, y también con el quinto objetivo del Plan Nacional del Buen Vivir 2017 – 2021: Impulsar la productividad y competitividad para el crecimiento económico sostenible de manera redistributiva y solidaria.

También, abre nuevos campos de investigación con distintas variables de la naturaleza que pueden ser analizadas tanto en las geociencias, u otras ramas afines. Es novedoso debido a que se propone una nueva metodología para la modelación de esta variable física de la Tierra. Respecto al financiamiento, el presente proyecto no requiere de valor alguno ya que la información necesaria para su desarrollo, fue levantada y proporcionada por el Instituto Geográfico Militar.

## **1.4 Objetivos**

### **1.4.1 Objetivo general**

Determinar un modelo de predicción espacial de la variable ondulación geoidal, para la zona rural del cantón Guayaquil, mediante el uso del método Cokriging.

### **1.4.2 Objetivos específicos**

- Definir la variable auxiliar que mejor correlación espacial presente con la ondulación geoidal para ser usada en la predicción del modelo.

- Realizar el análisis exploratorio de los datos de entrada para la generación del modelo.
- Efectuar el análisis estructural de los datos una vez caracterizada las variables objetivo y auxiliar respectivamente.
- Ajustar el cross – semivariograma experimental a un modelo teórico que mejor represente el comportamiento espacial de la ondulación geoidal.
- Ejecutar la validación cruzada del modelo de predicción obtenido.
- Predecir los valores de ondulación geoidal en la zona de estudio a través del modelo teórico que mejor se ajuste a la variable bajo estudio.
- Validar el modelo de ondulación geoidal elaborado, mediante una muestra de puntos que no fueron incluidos para la generación del modelo.

### **1.5 Metas**

- Un modelo de predicción de la variable ondulación geoidal de la zona rural del cantón Guayaquil.
- Una tabla comparativa de las variables auxiliares consideradas, con su respectivo valor de correlación que presenta con la variable objeto del estudio.
- Una tabla resumen con los resultados obtenidos del análisis exploratorio de los datos que serán usados para el modelo de predicción de la ondulación geoidal.
- Un cuadro de los resultados alcanzados en el análisis estructural del modelo de predicción.
- Un cuadro comparativo con las diferencias encontradas entre los modelos teóricos aplicados en el ajuste del modelo experimental resultante.
- Un archivo .ascii con los valores de ondulación geoidal predichos para la zona de estudio.
- Una tabla resumen con los resultados de la validación del modelo de predicción generado.

## CAPÍTULO II

### MARCO TEÓRICO

#### 2.1 Fundamentos teóricos

##### 2.1.1 Geodesia

A lo largo de la historia, varios autores han contribuido para definir a esta ciencia; desde la primera concepción dada por Helmert en 1887, en la que se habla sobre el tamaño, forma y campo de gravedad del planeta, ha venido evolucionando esta idea gracias a los avances tecnológicos y teorías que proporcionan nuevos enfoques más amplios que competen a la Geodesia.

Según la Asociación Internacional de Geodesia (IAG), es una geociencia que trata a la Tierra como un sistema dinámico complejo constituido en muchas capas, rodeado de atmósfera y océanos, que se ocupa: del monitoreo de la Tierra sólida, variaciones temporales en la gravedad del campo terrestre, diferenciaciones en la rotación de la Tierra, variaciones en la Tierra líquida, define las órbitas para los satélites científicos, se encarga de observar la atmósfera con técnicas geodésicas satelitales, y, determina posiciones y sus cambios con el tiempo de puntos en o sobre la superficie del planeta con máxima precisión (IAG, 2012).

Para motivos del estudio, se adoptará esta definición ya que engloba a las anomalías de gravedad, modelos geopotenciales, técnicas GNSS y la ondulación geoidal, que son elementos claves en el desarrollo de la propuesta de investigación.

## 2.1.2 Sistemas geodésicos de referencia

### 2.1.2.1 Sistema de referencia

Un sistema de referencia es la definición conceptual de parámetros, constantes, estándares, modelos, origen, orientación, escala, plano principal, entre otros, que sirven de base para representar un sistema de coordenadas (Seeber, 2003), y describir cuantitativamente la superficie del planeta y sus cambios a través del tiempo. El sistema de referencia no se determina mediante mediciones, sino que se establece convencionalmente para su uso, a esto se conoce como un sistema de referencia convencional (Drewes, 2014).

En la actualidad, existen varios sistemas de referencia, entre los que se destacan:

- International Terrestrial Reference System (ITRS): es el sistema de referencia establecido por el Servicio Internacional de Rotación de la Tierra (del inglés International Earth Rotation Service, IERS), elaborado para definir al Sistema de Referencia Celeste Internacional (International Celestial Reference System: ICRS), y del ITRS, incluyendo la relación entre estos (rotación y orientación del planeta en el espacio) (Drewes, 2014).
- World Geodetic System 1984 (WGS84): fue creado en un inicio con el propósito de establecer las coordenadas de las órbitas para los satélites Doppler (WGS72), y posteriormente utilizado en el cálculo de las órbitas para los satélites de la constelación NAVSTAR GPS. Actualmente, WGS84 adoptó el ITRS desde el año 2002 (Drewes, 2014).

A nivel mundial, el ITRS es el sistema de referencia convencional más usado para las tareas geodésicas, el mismo que está definido de la siguiente manera (McCarthy, 2000):

- Es un sistema geocéntrico, con centro de masas definido para toda la Tierra, abarcando los océanos y la atmósfera;
- Unidad de medida es el metro (SI);
- La evolución temporal de su orientación se da sin una rotación global residual, respecto a la corteza de la Tierra;
- Sus ejes están dispuestos según la orientación del Bureau International de l'Heure (BIH), en época 1984.0, basado en la orientación media del eje polar, llamado Polo Terrestre convencional (Conventional Terrestrial Pole, CTP), u Origen Internacional Convencional (Conventional International Origin, CIO), y una longitud “cero” fijada en el ecuador dada por el Observatorio Medio de Greenwich (Greenwich Mean Observatory, GMO).

### **2.1.2.2 Marco de referencia**

Un marco de referencia es la materialización (física y matemática) del sistema de referencia en base a las definiciones establecidas por este. Su materialización se realiza mediante un conjunto de puntos con alta precisión sobre la superficie terrestre en varios sitios del planeta. Estos puntos fiduciales poseen coordenadas en tres dimensiones y velocidades lineales fijados en una época específica (Drewes, 2014).

Para que un sistema de referencia pueda ser usado, debe estar representado en la realidad física del planeta a través de un marco de referencia, el mismo que integra los puntos que serán usados para el control, mantenimiento y actualización del sistema de referencia materializado, y juntos, el marco y sistema de referencia, se fusionan para dar origen a una plataforma global de georreferenciación (Drewes & Sánchez, 2002)

Al igual que los sistemas de referencia, existen varios marcos de referencia, de estos, los más representativos son:

- **International Terrestrial Reference Frame (ITRF):** está representado mediante coordenadas cartesianas y sus variaciones en el tiempo con gran precisión (Seeber, 2003). Es el marco que materializa al ITRS y está formado con más de 900 puntos observados con técnicas como VLBI (Very Long Baseline Interferometry), LLR (Lunar Laser Ranging), GPS (Global Positioning System), SLR (Satellite Laser Ranging) y DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite), en 500 lugares distintos alrededor del planeta Tierra , además, el ITRF es empleado por el Servicio Internacional GNSS (International GNSS Service, IGS) en el cálculo de órbitas precisas para los satélites GPS (Drewes, 2014).
- **Sistema de Referencia Geocéntrico para las Américas (SIRGAS):** es marco de referencia para Latinoamérica y el Caribe, que corresponde a una densificación del ITRF. Las coordenadas dentro del marco SIRGAS se encuentran definidas en una época de referencia, y sus velocidades son obtenidas de forma individual en cada estación, o con un modelo de velocidades continental. Las realizaciones de las estaciones SIRGAS que se encuentren en diferentes épocas o soluciones del ITRF, representan el mismo sistema de referencia, cuyas diferencias entre coordenadas están en el orden del milímetro (SIRGAS, 2014).

### **2.1.3 Superficies de referencia de la Tierra**

Existen varias superficies de referencia que se han estudiado y definido, como el Cuasigeoide, Teluroide, Esferoide, Geoide, Elipsoide, superficie real (física) del planeta, entre otros, de las cuales, estas tres últimas superficies son de interés para el presente estudio.

### 2.1.3.1 Geoide

Al igual que la Geodesia, el concepto de Geoide ha venido cambiando con el tiempo desde que fue propuesta por Gauss y después por Listing (Kopeikin, Mazurova, & Karpik, 2015), como la forma aproximada del planeta. A esta proposición se incorporaron nuevos enfoques y principios para su definición, como las que se indican en los trabajos de Heiskanen y Moritz (1985), Zakatov (1997), o Torge (2001).

Una definición moderna es propuesta por Achola, Fukuda y Kuroishi (2012), quienes definieron al Geoide como una superficie de nivel cerrada y continua, que se extiende por dentro del cuerpo sólido de la Tierra y sirve como superficie de referencia para un riguroso sistema de altura ortométrica.

Para fines prácticos, en la mayoría de trabajos de ingeniería, topografía y geociencias en general, se considera la superficie potencial o Geoide coincidente o muy próximo con el nivel medio del mar local (Achola & Fukuda, 2015; Banerjee, Foulger, Satyaprakash & Dabral, 1999), ya que no existe un único nivel medio del mar en todo el mundo, sino que varía de acuerdo al país en el cual se define el datum vertical, sea esto, por el uso de diferentes mareógrafos, o registro del nivel medio del mar en periodos de tiempo cambiantes. En el caso de Ecuador el datum vertical oficial se encuentra en La Libertad, provincia de Santa Elena, obtenido entre 1960 y 1961 con datos de un mareógrafo de las campañas 1950 a 1959 y 1949 a 1967, con diferencia de 0.8 mm entre estos (Paredes, 1986).

Ante esta premisa, en el presente estudio se considerará al nivel medio del mar local como una superficie muy aproximada con el Geoide.

### **2.1.3.2 Elipsoide**

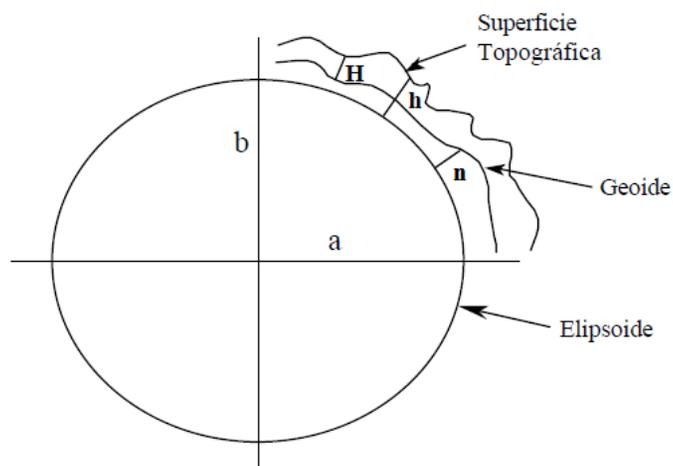
Es una superficie matemática, generada por un Elipsoide de revolución biaxial sobre su semieje menor ( $b$ ) (Rapp, 1984). Todo Elipsoide está definido por su semieje mayor ( $a$ ) y su achatamiento ( $f$ ), sobre el cual se calculan coordenadas geográficas, además de ser la forma más conveniente para realizar cálculos. Algunos ejemplos de Elipsoides representativos son: GRS80 (Geodetic Reference System 1980), usado por el ITRF; el WGS84, utilizado por el marco de referencia WGS84 (Drewes & Sánchez, 2002).

### **2.1.3.3 Superficie física de la Tierra**

Se entiende por superficie física de la Tierra, a la que comprende tanto el fondo oceánico y la parte de los continentes, es decir, la superficie topográfica o terrestre. Es en esta superficie donde se plasman las observaciones geodésicas por medio de los sistemas de observación geodésicos alrededor de la Tierra, para el estudio de la geodinámica del planeta y las leyes que lo rigen (Seeber, 2003).

### **2.1.4 Relación entre las superficies de referencia**

Tanto el Geoide, Elipsoide y la superficie física del planeta, se encuentran relacionadas mediante la ondulación geoidal ( $n$ ), que es la diferencia existente entre el Elipsoide y Geoide; sin embargo, como se señaló en el planteamiento del problema, para motivos prácticos y de este estudio de investigación, se considera coincidente o muy aproximado al Geoide con el nivel medio del mar, por lo que, ahora la ondulación geoidal representa la distancia entre el Elipsoide y el nivel medio del mar. Esto se aprecia de mejor forma en la Figura 1, a continuación:



**Figura 1.** Relación entre el Geoide, Elipsoide y la superficie topográfica del planeta  
Fuente: adaptado de (Rapp, 1984)

### 2.1.5 Sistemas de coordenadas

Una vez definido el sistema de referencia geodésico, se puede establecer un sistema de coordenadas para parametrizar el espacio e identificar la posición de cualquier punto sobre este. En Geodesia, un sistema de coordenadas es utilizado para designar unívocamente el lugar de un objeto en una, dos o tres dimensiones (punto, línea, polígono) (Drewes, 2014).

Se pueden expresar los sistemas de coordenadas según diversos criterios, pero todos poseen las siguientes características básicas:

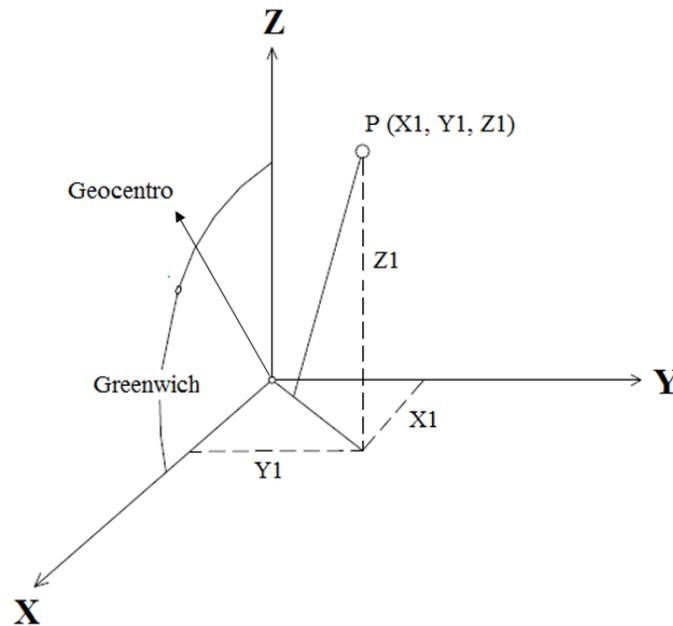
- Tipo de coordenada (espacial, plano, rectilíneo o curvilíneo);
- Origen definido;
- Orientación de los ejes establecido;
- Unidad de medida a usar.

Una premisa que se tiene que tener clara, es que, las coordenadas no se miden, sino se calculan en un sistema de coordenadas. En Geodesia, los sistemas de coordenadas principales son:

- Coordenadas cartesianas tridimensionales globales  $(X, Y, Z)$ , o locales denotadas como  $(e, n, u)$  del inglés east, nort, up;
- Coordenadas geográficas, geodésicas o elipsoidales  $(\varphi, \lambda, h)$ ;
- Coordenadas esféricas  $(\vartheta, \lambda, r)$ ;
- Coordenadas astronómicas  $(\Phi, \Lambda, \omega)$ ;
- Coordenadas planas  $(x, y, z)$ .

### 2.1.5.1 Cartesianas

Para propósitos prácticos, el sistema de coordenadas cartesianas es el más usado debido a que se aproxima a la forma de la Tierra (Seeber, 2003). Es un sistema geocéntrico, su eje X está en dirección del meridiano de Greenwich, el eje Z se orienta con el eje de rotación de la Tierra, y el eje Y es ortogonal a los dos anteriores para completar un sistema de “mano derecha” (ver Figura 2). Su unidad de medida es el metro (SI), y además, este sistema es usado para definir las orbitas de los satélites (GPS) y cambios en el sistema geodésico de referencia (Drewes, 2014).



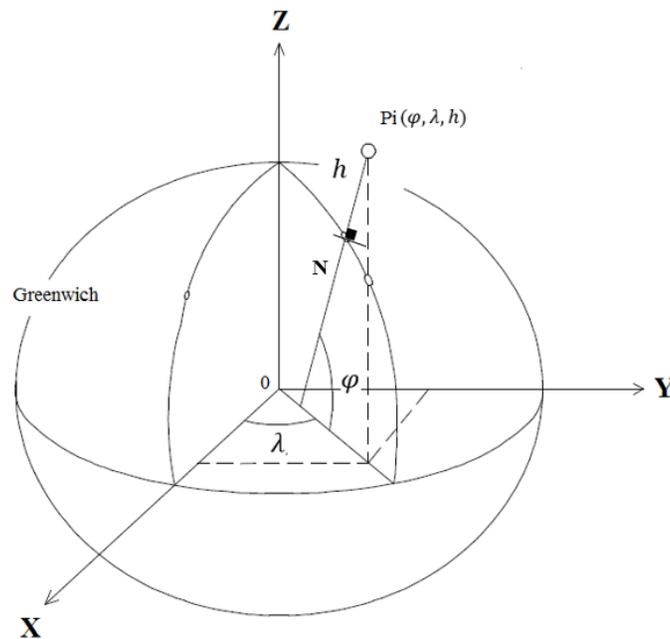
**Figura 2.** Sistema de coordenadas cartesianas

### 2.1.5.2 Geográficas

El sistema de coordenadas geográficas son las más usadas en los sistemas clásicos geodésicos, y también son muy utilizadas en la aviación, navegación, ingeniería, cartografía, entre otros. Su origen se encuentra en el geocentro, el eje vertical (Z) se direcciona al polo norte, el cual coincide con el semieje menor ( $b$ ) del Elipsoide de referencia, el eje X se orienta sobre el ecuador con el meridiano de Greenwich, el mismo que se sobrepone con el semieje mayor ( $a$ ) del Elipsoide de referencia (Drewes, 2014). En este sistema, las coordenadas se representan por latitud ( $\varphi$ ), longitud ( $\lambda$ ) y altura elipsoidal ( $h$ ) (ver Figura 3), cuyas definiciones se presentan a continuación:

- Latitud geodésica: es el ángulo medido en el plano meridiano que se forma entre la normal (N) al Elipsoide de referencia en el punto  $P_i$  y el plano del ecuador.
- Longitud geodésica: corresponde el ángulo medido en el plano ecuatorial formado por el meridiano que pasa el punto  $P_i$  y el meridiano de Greenwich.

- Altura elipsoidal: se representa por la distancia (altura) que separa al Elipsoide y al punto  $P_i$ , que se mide a lo largo de la normal al Elipsoide ( $N$ ), y permite al sistema de coordenadas geográficas extenderse hacia el espacio.



**Figura 3.** Sistema de coordenadas geográficas

Fuente: adaptado de (Drewes, 2014)

### 2.1.5.3 Planas

El sistema de coordenadas planas es uno de los más comunes para la representación de la realidad sobre mapas, cartas, o planos, basado en principios matemáticos y geométricos para trasladar la superficie del Elipsoide sobre una superficie de desarrollo (cono, cilindro, plano) a través de proyecciones cartográficas (Drewes, 2014).

Como se sabe, la superficie del planeta es curva (suponiendo una representación matemática más próxima con un Elipsoide de revolución), al proyectar sobre una superficie plana siempre va a existir distorsiones (anaformismos), sean de tipo angular, superficial (áreas) y distancias, por lo

que se trata de encontrar la superficie desarrollable que menos distorsión presente, según los fines que se necesiten.

En este sistema de coordenadas, generalmente se utilizan tres tipos de proyecciones:

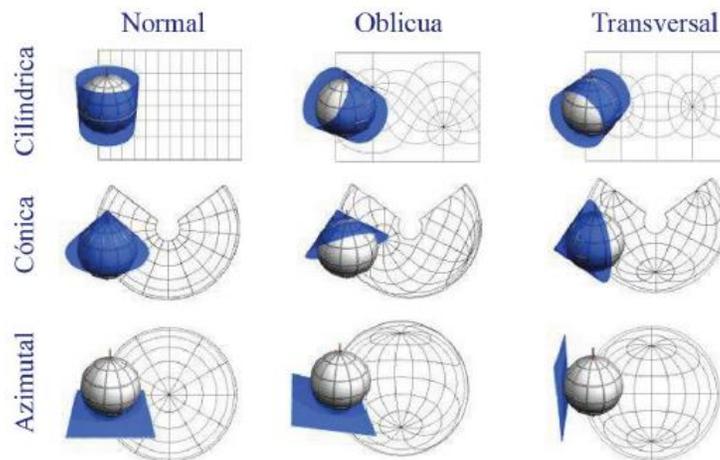
- Sobre un cono;
- Sobre un cilindro;
- Sobre un plano (acimutal).

Una clasificación usual en este tipo de sistema de coordenadas se puede representar de dos formas (otras clasificaciones no tan usadas son en base a la forma de meridianos y paralelos, y, según el uso):

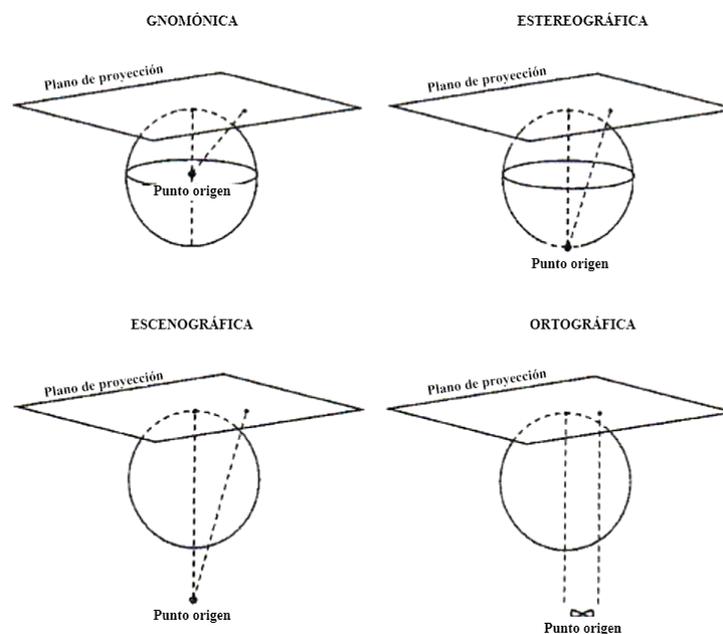
- Clasificación geométrica: proyecciones naturales, acimutales o perspectivas, y desarrollos.
  - Proyecciones naturales: a) sobre un plano: polar o normal (tangente al polo norte), transversal o ecuatorial (tangente al ecuador), oblicua (tangente a cualquier punto); b) por el punto de origen: gnomónica (origen en el centro de la Tierra), ortográfica (rayos desde el infinito), estereográfica (origen diametralmente opuesto al punto de tangencia), escenográfica (origen fuera del planeta a una distancia  $D$  conocida);
  - Desarrollos: a) según la superficie de revolución auxiliar: cilíndrica, cónica, b) tangentes y secantes, c) según la posición de la superficie auxiliar: directas o polares, transversas, oblicuas.
- Clasificación en función de la deformación:
  - Equivalentes o equiáreas (conserva el área proyectada);
  - Conformes u ortomorfa (no hay distorsión angular);

- Equidistantes o automecoicas (no hay distorsión en distancias);
- Afilácticas o de mínima deformación (mantiene un poco de todas las anteriores).

Lo expuesto anteriormente, se puede comprender de mejor forma, en las Figuras 4 y 5, a continuación:



**Figura 4.** Ejemplo de sistema de coordenadas planas proyectadas según el plano  
Fuente: (Drewes, 2014)



**Figura 5.** Ejemplo de sistema de coordenadas planas proyectadas según el punto origen  
Fuente: adaptado de (López, 2015)

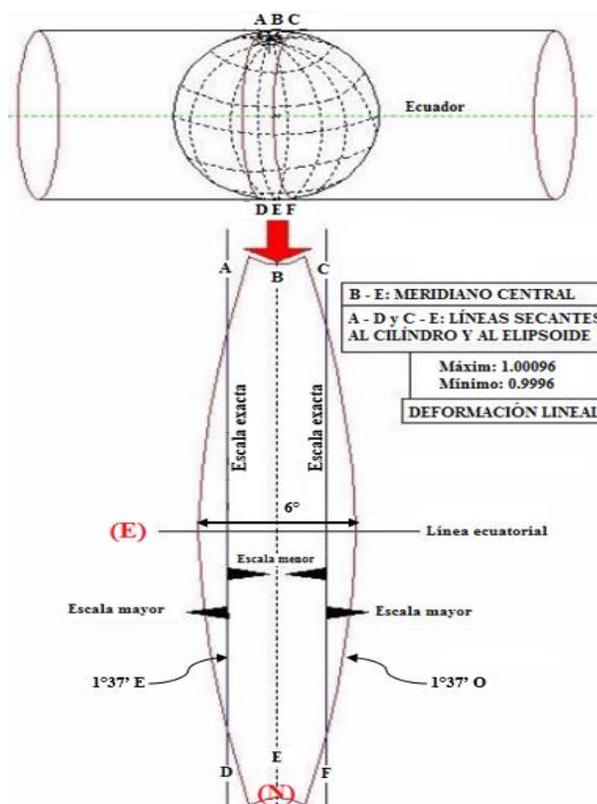
### 2.1.5.3.1 Proyección Universal Transversa de Mercator (UTM)

Dentro del sistema de coordenadas planas, las más usadas son las denominadas coordenadas planas proyectadas, las cuales se basan en una proyección cilíndrica transversal cuya representante principal es la denominada Universal Transversa de Mercator o UTM. Estas coordenadas UTM se derivan del sistema Gauss – Kruger, que posee como característica geométrica principal (es una proyección matemática, no es geométrica), la no distorsión angular, es decir, es una proyección conforme. El prefijo “Universal” se utiliza en primera instancia para diferenciarla de su proyección origen (Transversa de Mercator o Gauss – Kruger), y además, porque su uso es global. Presenta las siguientes características:

- Su unidad de medida es el metro (SI);
- Divide al planeta entre  $84^{\circ}$  N (norte) y  $80^{\circ}$  S (sur), en 60 zonas o husos de  $6^{\circ}$  cada uno, las mismas que van numeradas del 1 – 60, partiendo del meridiano  $180^{\circ}$  O (oeste);
- Cumple con las condiciones de contorno, es decir, tanto el meridiano central de cada zona y el ecuador, son líneas rectas; y es conforme;
- Los meridianos centrales se ubican cada  $3^{\circ}$  de longitud;
- A medida que se aleja del meridiano central de cada huso, aumenta la deformación lineal, por lo que se utiliza el factor de escala 0.9996 en el meridiano central, de modo que el cilindro de proyección es secante al Elipsoide de referencia, con lo que se crean dos líneas a  $1^{\circ}37'$  al este y oeste del meridiano central, en las que no presentan distorsión, es decir, escala verdadera con factor de escala 1 (Figura 6);
- Las coordenadas son siempre positivas, lo que le diferencia del sistema Gauss – Kruger, en que las coordenadas al oeste del meridiano central y al sur del ecuador son negativas. Para

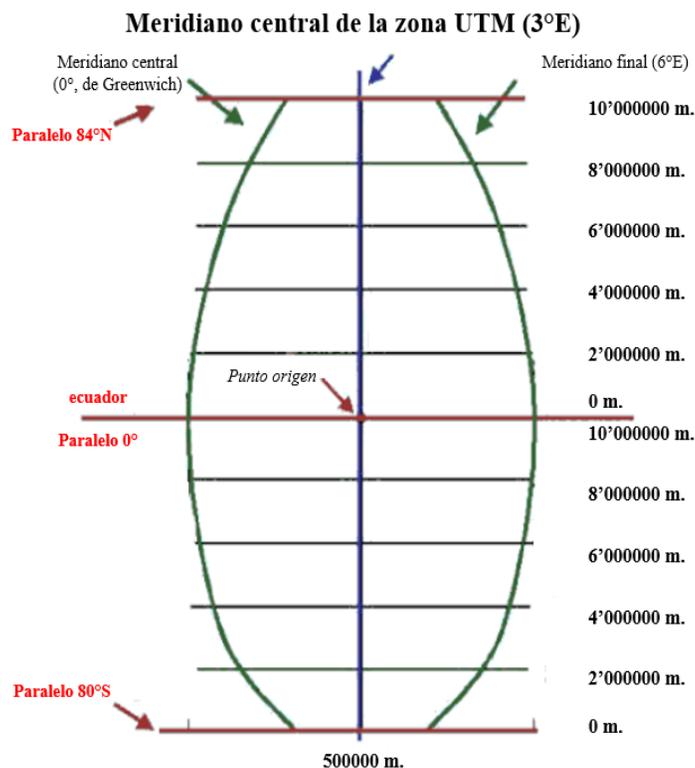
evitar esto, cuando se interseca el meridiano central del huso (eje Y, denominado Norte) con el ecuador (eje X, llamado Este) en el hemisferio norte, se desplaza el origen 500000 metros al oeste (falso este), con coordenadas origen (500000 m. E; 0 m. N); mientras que en el hemisferio sur, se considera en Este igual al hemisferio norte, y en Norte se desplaza 10'000000 metros hacia el polo sur (falso norte), por lo que las coordenadas origen serían (500000 m. E; 10'000000 m. N), de esta manera se garantiza que todas las coordenadas sean positivas (Figura 7).

En el caso de Ecuador, todo el territorio nacional se encuentra cubierto por cuatro zonas UTM, la región insular comprende las zonas 15 y 16, mientras que la parte continental del país abarca las zonas 17 y 18.



**Figura 6.** Sistema de coordenadas planas UTM

Fuente: adaptado de (Fernández, 2001)



**Figura 7.** Ejemplo del sistema de coordenadas planas UTM  
Fuente: adaptado de (Fernández, 2001)

### 2.1.6 Sistemas de alturas

Para conocer la altura a la que se encuentra un punto sobre la superficie terrestre (superficie topográfica o física de la Tierra), es necesario definir un nivel de referencia (datum vertical) sobre el cual se puede obtener la distancia vertical (altura) del punto en cuestión. En la práctica, este procedimiento se realiza a través de técnicas de nivelación, las mismas que, según el objetivo y precisión del trabajo requerido, puede ser de tipo barométrico, trigonométrico, y geométrico, pero debido a que estas medidas se realizan sobre un campo de gravedad terrestre, se debe considerar ciertas correcciones del tipo físico (gravimétricas) a las mediciones realizadas.

En las rutinas tradicionales de nivelación para determinar la altura de un punto, se utilizan instrumentos cuya proyección horizontal (horizonte astronómico) se relaciona en forma tangencial

con la superficie local equipotencial del área de trabajo, y la dirección que sigue la componente vertical (línea de la plomada) es coincidente con el vector de fuerza gravitatoria, el mismo que se proyecta perpendicularmente a ambas superficies, de esta forma, en las diferencias entre alturas se consideran tanto la variación de la superficie terrestre y las distorsiones de gravedad del planeta (Hofmann & Moritz, 2005). Es así que, en Geodesia las alturas usadas se pueden clasificar en base al modelo matemático o físico, su método de determinación y fin de aplicación. En grandes rasgos, se dividen en alturas de tipo físico y geométrico.

Las aturas geométricas son las proporcionadas directamente por los sistemas GNSS, o a través de la nivelación tradicional. Por su parte, con las alturas físicas se puede cuantificar la diferencia de potencial y así conocer las separaciones verdaderas entre las superficies de referencia, las cuales, si se considera una poligonal cerrada las diferencias serán cero, independientemente de la trayectoria seguida, debido a que el valor de potencial es único y depende solo de su posición (Hofmann & Moritz, 2005).

Las diferencias de potencial entre un punto cualquiera y el Geoide puede ser expresado de la siguiente manera:

$$\int_0^A g dn = W_o - W_A = C \quad (2)$$

donde:

$g$ : valor de gravedad observado en el punto A;

$C$ : número geopotencial;

$dn$ : diferencia de altura (desnivel);

$W_0$ : valor de potencial en el Geoide;

$W_A$ : valor de potencial en la superficie que atraviesa al punto A.

El número geopotencial se encuentra en unidades de geopotencial (g.p.u.), es independiente de la línea de nivelación seguida, fue adoptado en el encuentro de una subcomisión de la IAG (International Association of Geodesy) en Florencia en 1955. Generalmente se encuentra expresado en  $(m^2/s^2)$ , que no son unidades de longitud, por lo que su interpretación práctica es poco aplicable, sin embargo, el número geopotencial puede expresarse en unidades de longitud al ser dividido por un valor de gravedad establecido. Precisamente, este valor de gravedad que se utilice para la división, será el que sirva para distinguir o clasificar los diversos tipos de sistemas de alturas, según la ecuación (3):

$$H = \frac{C}{g} \quad (3)$$

donde:

$H$ : tipo de altura según el valor de gravedad establecido;

$g$ : valor de gravedad establecido.

Como se mencionó, el valor de gravedad usado en (3), distingue los diversos tipos de alturas, por ejemplo, existen alturas dinámicas, normales, ortométricas. En este estudio, se presta atención únicamente a las alturas ortométricas (altura física), elipsoidales y niveladas (alturas geométricas).

### 2.1.6.1 Ortométricas

Las alturas ortométricas son calculadas utilizando la expresión en (3), considerando  $g$  como el valor medio de gravedad ( $g'$ ) entre el punto sobre la superficie terrestre (P) y el Geoide ( $P_0$ ),

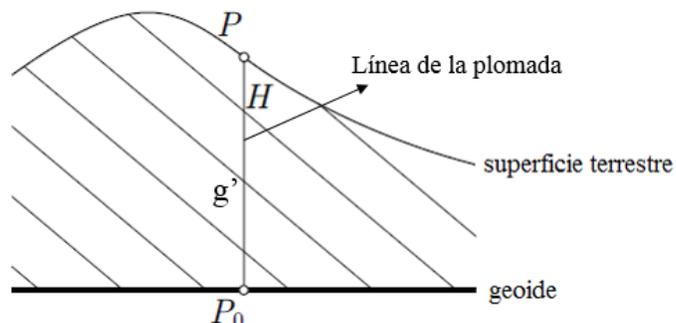
medido sobre la línea de la plomada que es perpendicular al Geoide y que pasa por el punto en cuestión (Figura 8). Dado esto, (3) puede expresarse de la siguiente forma:

$$H = \frac{C}{g'} \quad (4)$$

Este tipo de altura es el que representaría el valor real de cota (altura), sin embargo, el problema en el sistema, se debe a que resulta imposible conocer el valor exacto de gravedad media, ya que si bien se efectúa la medición de gravedad sobre la superficie terrestre, se necesita continuar el valor de gravedad hasta el Geoide siguiendo la línea de la plomada para obtener el valor medio de gravedad, y no se conoce a ciencia cierta la estructura y composición de los materiales de la Tierra (que tampoco son homogéneos), ni es posible llegar hasta el Geoide, por lo que, para su uso se manejan modelos de distribución de densidades de masa.

El valor de altura ortométrica dependerá del modelo e hipótesis usados para la distribución de densidades de masas terrestres. Entre las hipótesis más utilizadas para el cálculo de alturas ortométricas se encuentran la de Vignal, Helmert, Aire Libre, Baranov, entre otras.

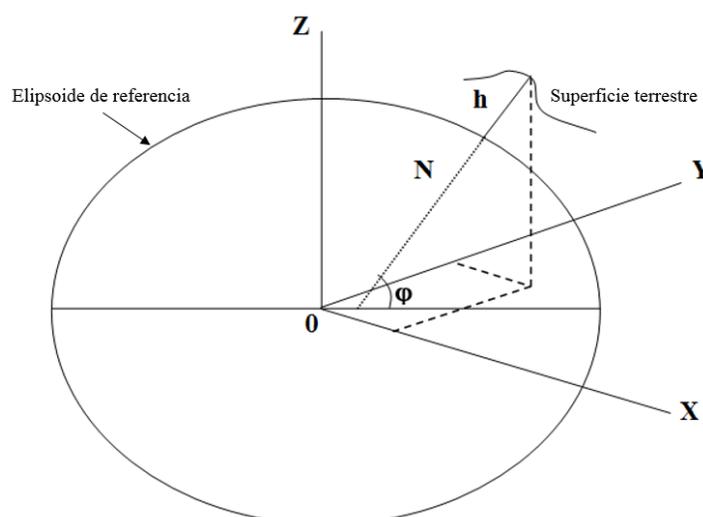
Un punto importante a señalar es que, las alturas ortométricas para fines prácticos, son asumidas como alturas niveladas sobre el nivel medio del mar, ya que, se considera coincidente o muy poca separación entre el Geoide con el nivel medio del mar y se simplifica la dificultad de obtener un modelo de densidades preciso. También, en el contexto de Ecuador, no existen todavía muchas mediciones de gravedad como para hacer uso de estas alturas en un sentido estricto, por lo que el IGM apoya con datos sobre alturas niveladas para suplir esta necesidad de contar con alturas sobre la superficie física del planeta.



**Figura 8.** Sistema de altura ortométrica  
Fuente: adaptado de (Hofmann & Moritz, 2005)

### 2.1.6.2 Elipsoidales

Las alturas elipsoidales, como su nombre lo indica, representan la separación entre un punto sobre la superficie terrestre y un Elipsoide de referencia, calculada sobre la línea normal (N) al Elipsoide (ver Figura 9) (Drewes, Sánchez, Bliztkow & Freitas, 1998). Las coordenadas son de tipo geocéntricas cartesianas, definidas sobre un Elipsoide en específico, por lo que la dirección del vector de altura y su magnitud dependerán del Elipsoide de referencia que se use (Sánchez & Martínez, 1997). Este tipo de alturas son fácilmente obtenidas a través de posicionamiento satelital sobre el punto que se requiera conocer, por lo que es importante llevar un registro oficial de la cantidad de puntos (marcas, hitos, mojones) medidos en un proyecto.



**Figura 9.** Sistema de altura elipsoidal  
Fuente: adaptado de (Drewes & Sánchez, 2002)

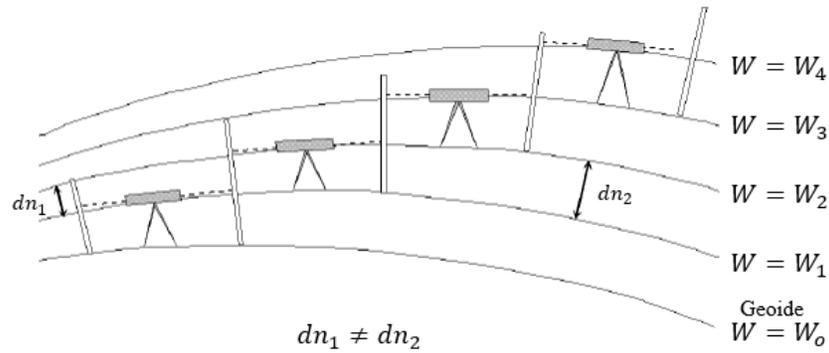
A pesar de su amplio uso y relativa facilidad de obtención, las alturas elipsoidales no consideran el campo de gravedad terrestre (altura geométrica), por lo que, podría darse el caso de presentar alturas iguales en puntos ubicados a distintos niveles (Drewes, Sánchez, Bliztkow & Freitas, 1998), además, al tratarse de un valor referido a una superficie matemática “no real”, su uso para trabajos sobre la superficie topográfica del planeta se ve limitada.

### 2.1.6.3 Niveladas

Las alturas niveladas son calculadas mediante trabajos en campo con técnicas de nivelación geométrica o trigonométrica (depende del orden de precisión requerido), a través de instrumentos ópticos, como el nivel (instrumento de medición), o estación total. Generalmente es más usada la nivelación geométrica ya que es más precisa que la trigonométrica.

Mediante la nivelación se obtienen desniveles entre las superficies equipotenciales del campo de gravedad de la Tierra, estos desniveles (diferencias de altura) no son fijos, sino que cambian en función al campo de gravedad local y distribución de densidad de las masas. Además, se sabe que

las superficies equipotenciales no son equidistantes, por lo que el valor de altura para un punto, dependerá de la línea de nivelación que se siga (Figura 10).



**Figura 10.** Sistema de altura nivelada  
Fuente: (Drewes & Sánchez, 2002)

Para obtener el valor de diferencia de altura entre dos puntos (superficies equipotenciales), se aplica una resta simple:

$$\Delta n_A^B = l_A - l_B \quad (5)$$

donde:

$l_A$ : valor de la lectura medido en A;

$l_B$ : valor de la lectura medido en B.

En una línea de nivelación, por lo general existen varios desniveles (deltas), por lo que la altura nivelada es el resultado de una suma de deltas de altura entre los puntos en cuestión, partiendo desde un valor conocido de altura referido al datum vertical oficial. Esta sucesión de desniveles se puede representar matemáticamente como:

$$H_B = \frac{\Delta n_A^B - \Delta n_B^A}{2} + H_A \quad (6)$$

donde:

$H_A$ : es la altura nivelada para el punto A;

$H_B$ : es la altura nivelada para el punto B;

$\Delta n_A^B$ : es el desnivel de altura calculado desde A hacia B;

$\Delta n_B^A$ : es el desnivel de altura calculado desde B hacia A.

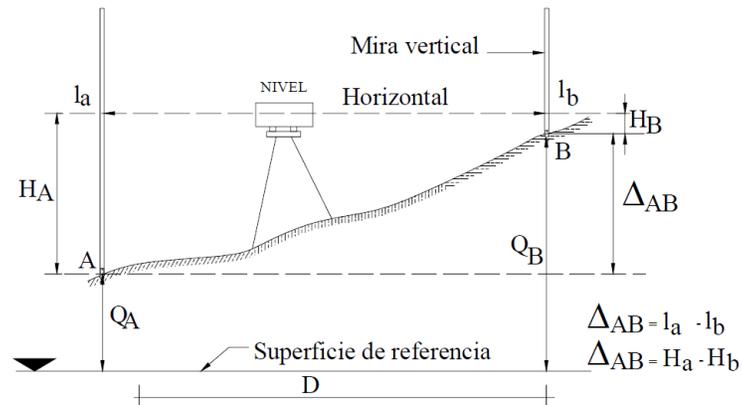
Debido a que los desniveles de altura están en función del campo de gravedad local y la densidad de las masas internas, se aplica líneas de nivelación locales, de máximo 10 km., distancia en la cual, las variaciones del campo de gravedad, en teoría, no son muy significativas (SIRGAS, 2002).

### **2.1.7 Técnicas de nivelación**

Se entiende por nivelación a los procedimientos para obtener el valor de altura de uno o más puntos, respecto a una superficie de referencia horizontal denominado plano de comparación. Como se mencionó anteriormente, existen varias técnicas de nivelación como: trigonométrica, barométrica, geométrica, de las cuales, las más usadas para el cálculo de alturas niveladas son la trigonométrica y geométrica, que serán estudiadas a continuación.

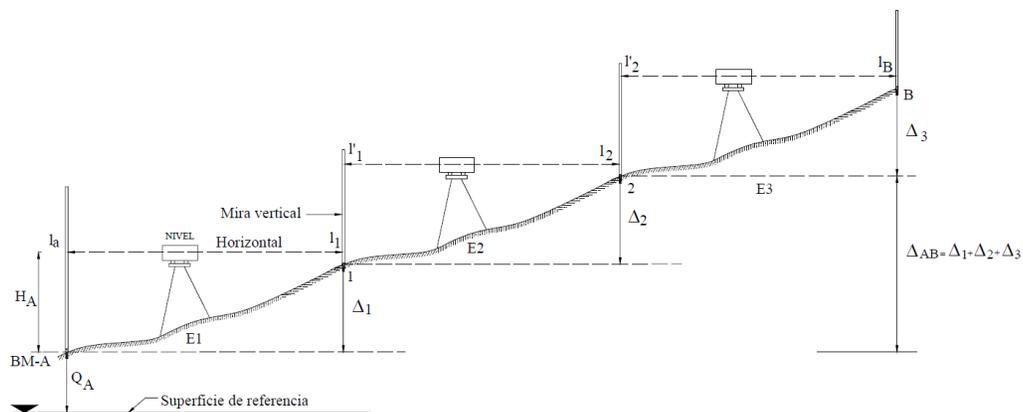
#### **2.1.7.1 Nivelación geométrica**

La nivelación geométrica o también llamada nivelación diferencial permite obtener el desnivel entre dos puntos mediante la visual horizontal (horizonte astronómico), usando el nivel o equialtímetro y miras verticales. Cuando el delta de altura se puede determinar con una sola estación del nivel, se conoce como nivelación geométrica simple (ver Figura 11), y cuando la distancia entre puntos es mayor al alcance visual, se colocan varias estaciones intermedias, tomando esto por nombre nivelación geométrica compuesta (ver Figura 12) (Farjas, 2009).



**Figura 11.** Nivelación geométrica simple

Fuente: (Navarro, 2010)



**Figura 12.** Nivelación geométrica compuesta

Fuente: (Navarro, 2010)

La nivelación diferencial es más precisa que la trigonométrica, ya que la primera elimina los errores sistemáticos del nivel, la esfericidad de la superficie terrestre y la refracción. En el país, el IGM adopta esta técnica para dar puntos de precisión, basado en tres rangos u órdenes de precisión, que son:

- Alturas de primer orden:  $4mm\sqrt{k}$ ;
- Alturas de segundo orden:  $8mm\sqrt{k}$ ;
- Alturas de tercer orden:  $12mm\sqrt{k}$ .

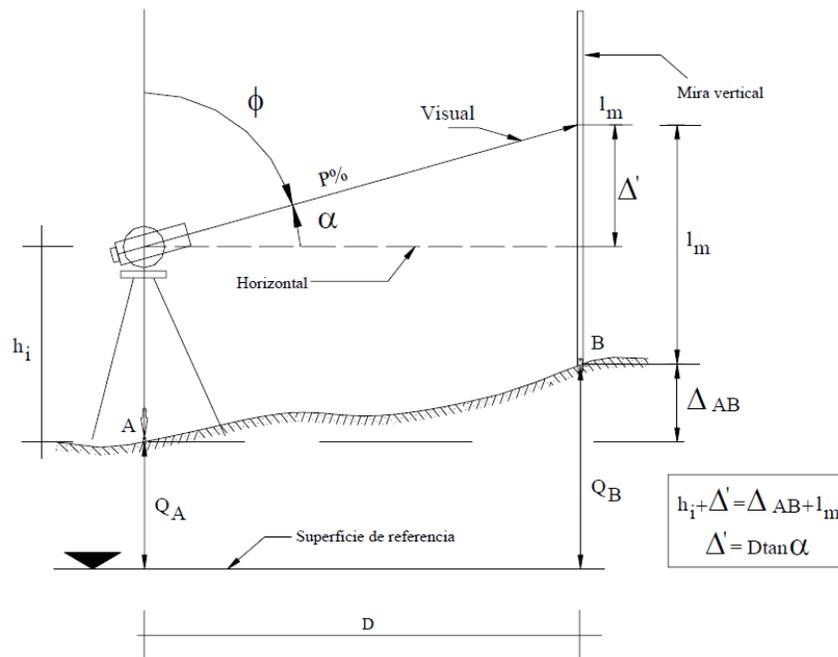
Donde  $k$  es la distancia de la línea de nivelación en kilómetros. Estos órdenes de precisión hacen referencia a la tolerancia de cierre de cada trayectoria nivelada. La mayoría de puntos de la Red de Control Básico Vertical son de segundo orden.

Los datos que son utilizados en este proyecto, corresponden a alturas niveladas de segundo orden.

### 2.1.7.2 Nivelación trigonométrica

La nivelación trigonométrica o por pendientes, es una técnica de nivelación que consiste en medir los ángulos verticales y distancias para obtener el valor del desnivel entre los puntos (Figura 13). Puede darse al igual que la nivelación geométrica, de forma simple o compuesta (Farjas, 2009).

Esta técnica es más adecuada usarla cuando el terreno a nivelar es muy escarpado, montañoso o boscoso, ya que proporciona de manera rápida valores de alturas.



**Figura 13.** Nivelación trigonométrica

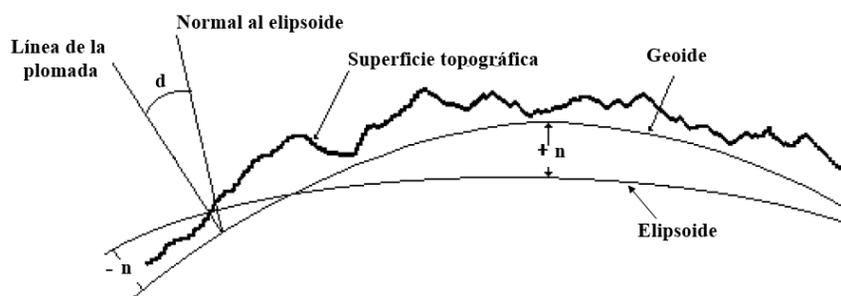
Fuente: (Navarro, 2010)

De igual forma, el IGM trabaja con precisiones de cierre en las alturas obtenidas por este método, como son:

- Alturas de primer orden:  $12mm\sqrt{k}$ ;
- Alturas de segundo orden:  $16mm\sqrt{k}$ ;
- Alturas de tercer orden:  $24mm\sqrt{k}$ .

### 2.1.8 Ondulación geoidal

Se entiende por ondulación geoidal a la distancia entre un Elipsoide de referencia (global) y el Geoide, es decir, la diferencia de la altura elipsoidal y ortométrica, medida a lo largo de la normal al elipsoide (ver Figura 14) (Mena, 2008). Como ya se mencionó anteriormente, para efectos prácticos, se considera al nivel medio del mar coincidente con el Geoide (datum natural de alturas), o que su separación es ínfima, por lo que se utiliza a las alturas referidas al nivel medio del mar como una aproximación de alturas ortométricas, resultando la ondulación geoidal en la diferencia entre la altura elipsoidal y nivelada, cuya representación matemática se observa en (1).



**Figura 14.** Ondulación Geoidal

Fuente: adaptado de (Hansen, 2008)

Para tener un valor de nivel medio del mar confiable, es recomendado tener observaciones del mareógrafo  $\geq 18$  años (Achola & Fukuda, 2015). Al garantizar lo mejor posible un modelo de

ondulación geoidal, se puede potenciar el uso de técnicas de posicionamiento satelital (GNSS) combinándolas con datos de nivelación para generar modelos que ofrezcan valores precisos de alturas sobre la superficie terrestre.

Además, se debe tomar en cuenta que el modelo matemático mostrado en (1) corresponde a un comportamiento ideal de ondulación geoidal (modelo determinístico), algo que en la realidad no es posible alcanzar, ya que tanto la obtención de altura elipsoidal como altura nivelada poseen errores subyacentes en su medición, por lo que se puede inferir que el valor de ondulación geoidal tiene una componente aleatoria de error, cambiando su ecuación determinística a estocástica como se muestra a continuación (Banerjee *et al.*, 1999; Leiva, 2014):

$$n = h - H + \varepsilon \quad (7)$$

El término  $\varepsilon$  es una cantidad pequeña que representa la variabilidad aleatoria (ruido estocástico) en la determinación de la ondulación geoidal, que según Torge (2001) se debe a la deflexión de la vertical ( $d$ ) y a la curvatura presente en la línea de la plomada, lo cual está relacionado a los errores cometidos en la nivelación geométrica y posicionamiento satelital.

Cabe señalar que la nivelación diferencial (geométrica) es la técnica más usada para la obtención de alturas niveladas, y a través de la combinación de datos GPS (comúnmente utilizado) con datos de nivelación es posible obtener la ondulación geoidal (y por ende la separación del geoide) sobre la superficie terrestre del planeta (Milbert, 1991; Zilkoski, 1990). En el caso de Ecuador, el IGM utiliza la nivelación geométrica para la obtención de alturas niveladas, y como se verá más adelante, los datos utilizados en este proyecto para modelar la ondulación geoidal, son de este tipo de nivelación.

Debido a la complejidad del cálculo del Geoide, y por consiguiente, de la ondulación geoidal, se han desarrollado algunos métodos para su obtención, los más conocidos son: gravimétricos, astrogeodésicos, geométricos y combinados (gravimétricos y geométricos). De estos, el método geométrico es objeto de interés para el estudio, por ser el que encaja con la realidad de los datos usados para el modelamiento de la ondulación geoidal en el Ecuador.

#### **2.1.8.1 Método geométrico**

Es el más antiguo de los métodos empleados en la determinación de la ondulación geoidal. Su nombre hace alusión a que utiliza la geometría en sus cálculos, y justamente, al no considerar ningún componente físico, su aplicación se limita a áreas no muy grandes (geoide local o puntual), de por lo general inferiores a  $2^\circ \times 2^\circ$  de extensión (Núñez, 2006).

Su procedimiento se fundamenta en la obtención de las diferencias de cota en varios sitios, entre la altura elipsoidal (obtenida mediante GPS) y la altura nivelada (considerada igual que la ortométrica para fines prácticos y en este estudio), para posteriormente interpolar los valores de ondulación geoidal en la zona de estudio.

#### **2.1.9 Modelos geopotenciales globales**

En la actualidad, gracias al desarrollo tecnológico de plataformas satelitales, es posible obtener modelos globales de ondulación geoidal, siendo el más usado el obtenido a través de modelos geopotenciales, los cuales se basan en determinar coeficientes (en función de seis elementos orbitales) para un desarrollo en serie de funciones esféricas armónicas del potencial gravitacional de la Tierra (Tierra, 2009).

En un caso ideal utópico, el campo gravitacional no tiene perturbaciones, por lo que la órbita de los satélites correspondería a una elipse kepleriana, sin embargo, en la realidad, esta órbita es deformada debido al campo gravitatorio del planeta, pero gracias a su misma perturbación es posible obtener datos sobre el comportamiento del campo en mención (Pacino, 1999). En el modelo geopotencial, las incógnitas corresponden a los coeficientes armónicos, mientras que los cambios de los elementos orbitales son las observaciones del modelo. Entre los más reconocidos se encuentran el EGM96 (Earth Gravitational Model 1996) (Lemoine *et al.*, 1998), EGM08 (Earth Gravitational Model 2008) (Pavlis, Holmes, Kenyon, & Factor, 2008), OSU91 (Ohio State University) (Rapp & Pavlis, 1990), entre otros.

#### **2.1.9.1 Modelo EGM96**

El EGM96 es un modelo geopotencial que fue desarrollado por la NASA (National Aeronautics and Space Administration), la NGA (National Geospatial –Intelligence Agency) en conjunto con la Universidad Estatal de Ohio, y consiste en coeficientes armónicos esféricos completos hasta el orden y grado 360. Presenta una solución compuesta de tres partes: a) solución combinada de grado y orden 70; b) solución diagonal de bloques con grados desde 71 al 359; c) solución de cuadratura de grado 360. Los cálculos de ondulación geoidal están referidos al Elipsoide WGS84, por lo que puede ser determinada la ondulación geoidal en un punto cualquiera con coordenadas latitud y longitud dentro de este sistema. La resolución del modelo es una grilla de 15' x 15' (NASA, 2005).

#### **2.1.9.2 Modelo EGM08**

El EGM08 es el modelo a nivel global más moderno de los EGM, fue realizado por la NGA a través de una serie de armónicos esféricos completos de orden y grado 2159, con coeficientes adicionales de grado 2190 y orden 2159. Al igual que el EGM96, los valores de ondulación geoidal

están referidos al WGS84. Posee una mejor resolución con datos en una grilla de 2.5' x 2.5' y derivados cada 1'. Además de proporcionar un modelo de ondulación geoidal, el EGM08 posee valores de anomalías de gravedad de tipo Aire libre y Bouguer cada 2.5' (NGA, 2008), que serán variables de interés para este estudio.

### **2.1.9.3 Modelo WGM12**

El WGM12 (World Gravity Map 2012), es el primer mapa y grilla de alta resolución respecto a anomalías de gravedad (Aire libre, Bouguer, Isostática) para todo el planeta, fue desarrollado por la BGI (Bureau Gravimétrique International), la CGMW de la UNESCO (Commission for the Geological Map of the World), IAG, IUGG (International Union of Geodesy and Geophysics), IUGS (International Union of Geological Sciences) y otras organizaciones científicas (Bonvalot *et al.*, 2012).

Los datos de anomalía de gravedad del modelo son derivados del EGM08, DTU10 (Dometric Turbo Unit 10000), e incluye correcciones de terreno del ETOPO1; los productos del modelo fueron calculados en series de armónicos esféricos según la metodología presentada por Balmino *et al.*, (2011) llegando a grado 32400. La resolución de este modelo es cada 2' (BGI, 2012).

### **2.1.10 Anomalías de gravedad**

La anomalía de gravedad es una variación de las cantidades medias de gravedad referidas a la gravedad normal, luego de aplicarse ciertas correcciones. Para hablar de anomalías de gravedad, se debe en primera instancia aplicar ciertas hipótesis de reducción de gravedad, las mismas que son igualmente utilizadas para el cálculo de anomalías  $\Delta g$ , o perturbaciones  $\delta g$  de gravedad, y sirven para los propósitos siguientes (Hofmann & Moritz, 2005):

- Permite la interpolación o extrapolación de valores de gravedad;
- Ayuda al estudio de la corteza terrestre;
- Se utiliza para determinar el geoide.

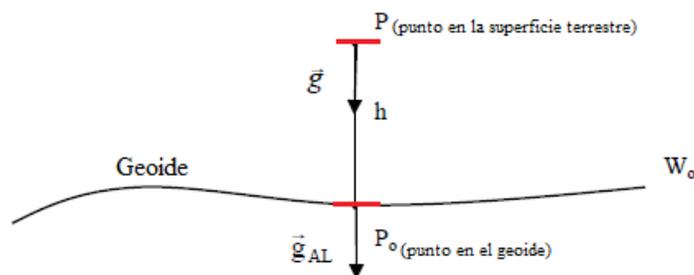
Entre las más destacadas se encuentran las anomalías de Aire libre, Bouguer e Isostática, las mismas que son de interés para el desarrollo del trabajo.

### 2.1.10.1 Aire libre

La anomalía de Aire libre se obtiene a través de una reducción de Aire libre, la cual se basa a reducir el valor de gravedad sin considerar las masas existentes entre el punto sobre la superficie terrestre y el Geoide, esto se logra sumando la corrección de Aire libre ( $F$ ) y el valor medido de gravedad ( $g$ ) en el punto. Esta corrección se obtiene de  $F = 0.30856 * h$  en mgals (miligales) siendo  $h$  la altura del punto (positiva o negativa) en metros (Hofmann & Moritz, 2005).

La reducción de Aire libre será entonces  $g_{AL} = g + F$ , y la anomalía de Aire libre se obtiene de la diferencia entre  $g_{AL}$  y la gravedad normal medida en el Elipsoide de referencia  $\gamma$  (ver Figura 15), expresado de la siguiente forma:

$$\Delta g_{AL} = g_{AL} - \gamma = g + F - \gamma \quad (8)$$



**Figura 15.** Reducción de gravedad por Aire libre  
Fuente: adaptado de (Heiskanen & Moritz, 1985)

### 2.1.10.2 Bouguer

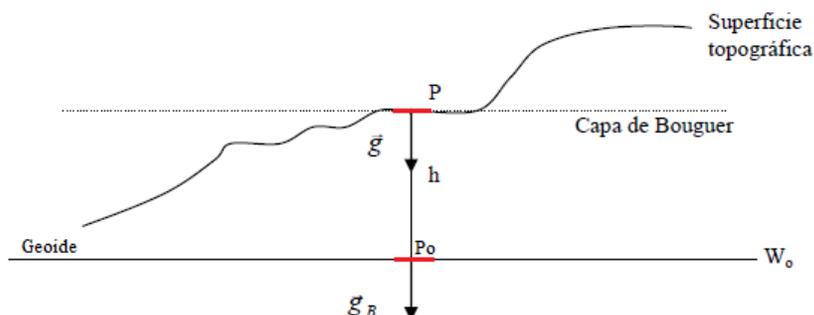
La reducción de gravedad por Bouguer remueve todas las masas topográficas que se encuentren fuera del Geoide, para lo cual en primer lugar se aplica una reducción por Aire libre. Su hipótesis asume que la superficie alrededor de la estación de gravedad en el punto P, es completamente horizontal y plana con una densidad constante ( $\rho$ ) de las masas existentes entre el Geoide y la superficie topográfica (Figura 16) (Hofmann & Moritz, 2005).

Para eliminar las masas, se construye una “capa de Bouguer” con grosor igual a la altura del punto P, en la que se aplica una corrección topográfica ( $A_t$ ), que representa la componente de atracción gravitatoria perpendicular al punto en donde se estudia la masa que hay que sumar o restar para formar la capa de Bouguer; una vez definida esta capa se procede a restar su efecto gravitatorio en el punto de estudio, eliminando así las masas entre el Geoide y la superficie topográfica (Hofmann & Moritz, 2005). Lo mencionado en estos párrafos, se expresa matemáticamente de la siguiente forma:

- Reducción de gravedad por Aire libre:  $g_{AL} = g + F$ ;
- Anomalía de Aire libre:  $\Delta g_{AL} = g + F - \gamma$ ;
- Corrección topográfica siempre positiva  $A_t$ ;
- Atracción de gravedad de la capa de Bouguer para densidad constante  $\rho = 2.67 \frac{g}{cm^3}$ :  $A_B = 2\pi G\rho = 0.1119 * h$  mgales, con  $h$  en metros,  $G$  constante gravitacional;
- Reducción de gravedad por Bouguer:  $g_B = g + A_t - A_B + F = g_{AL} + A_t - A_B$ ;

Finalmente, la anomalía de gravedad por Bouguer se obtiene como:

$$\Delta g_B = g_B - \gamma = g + A_t - A_B + F - \gamma = \Delta g_{AL} + A_t - A_B \quad (9)$$



**Figura 16.** Reducción de gravedad por Bouguer

Fuente: adaptado de (Heiskanen & Moritz, 1985)

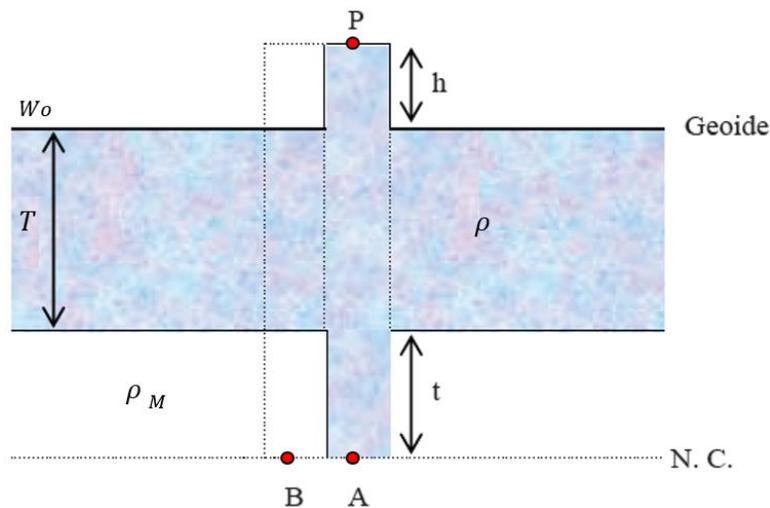
### 2.1.10.3 Isostática

Empíricamente, se asumiría que las masas topográficas están superpuestas sobre una capa o corteza terrestre homogénea, sin embargo, se ha demostrado que es todo lo contrario, ya que si la corteza fuese homogénea, la reducción de gravedad por Bouguer sería mínima, debido que se eliminarían las distorsiones del campo de gravedad terrestre y por ende sus anomalías estarían cercanas al cero.

La anomalía de Bouguer calculada en zonas de montaña presenta un valor negativo que asciende sistemáticamente 100 mgals por cada 1000 metros de altura aproximadamente, por lo que se deduce la existencia de alguna clase de deficiencia de masa y que es compensada de cierta forma. Esta compensación de masas se denomina “compensación isostática” la cual se puede estudiar por dos teorías: a) teoría de Pratt – Hayford y b) teoría de Airy – Heiskanen, de las cuales se estudiará la segunda por ser la teoría que utiliza el modelo WGM12 para el cómputo de sus anomalías de gravedad.

### 2.1.10.3.1 Teoría de Airy – Heiskanen

Como premisa, se tiene un valor de densidad homogénea para la corteza  $\rho = 2.67 \frac{g}{cm^3}$ , además, en zonas montañosas la compensación isostática es lograda a través de la formación de las raíces de la corteza correspondientes al manto superior con densidad  $\rho_M = 3.27 \frac{g}{cm^3}$ , por lo que habrá una diferencia de densidades que será constante  $\Delta\rho = \rho_M - \rho = 0.6 \frac{g}{cm^3}$ . Tomando un punto P para el estudio, ubicado a cierta altura  $h$  separado del Geoide sobre una elevación en forma de cilindro con sección definida  $S$ , y una corteza de grosor  $T$  ubicado sobre el manto superior considerado homogéneo, el valor de  $T$  cambia según la zona de montaña u oceánica en la que se encuentre (ver Figura 17) (Hofmann & Moritz, 2005).



**Figura 17.** Modelo de isostasia de Airy – Heiskanen

Fuente: adaptado de (Heiskanen & Moritz, 1985)

Según esta teoría, los puntos A y B poseen la misma presión  $P_A = P_B$  (y todos los puntos que estén al nivel de compensación N.C.), además,  $h$  produce un incremento en el grosor de la corteza que es representado por  $t$  (cuyo valor depende del área de montaña en que se encuentre), y como resultado el grosor final viene dado por la suma de  $h + T + t$ . Si se aplica isostasia en A y B, las

masas de cada columna del cilindro  $M_A$  y  $M_B$ , con sección  $S$  establecida y volumen igual, serán las mismas. Además, como  $\Delta\rho$  es un valor constante, la raíz  $t$  es la incógnita del modelo propuesto, el mismo que resulta de la deducción del planteamiento de isostasia como:

$$t = \frac{h}{\Delta\rho} * \rho = 4.45 * h \quad (10)$$

Es decir que  $t$  es 4.45 veces más que el valor de  $h$ . Precisamente, en la zona raíz ( $t$ ) es donde se produce los defectos de masas, por lo que para un punto P sobre la superficie terrestre, su valor de anomalía de Bouguer será negativa (Hofmann & Moritz, 2005).

Por otro lado, en la parte del cilindro donde exista  $\Delta\rho$ , habrá también  $t$ , y la expresión de atracción de gravedad de compensación de un punto P para obtener la restauración de la masa se expresa como:

$$A_c = 2\pi G\Delta\rho \left[ b + \sqrt{a^2 + (c - b)^2} - \sqrt{a^2 + c^2} \right] \quad (11)$$

Donde  $a = b$  (radio y altura del cilindro respectivamente) para este modelo y  $c = h + T + t$ . Esta reducción de gravedad aplicando  $A_c$  y después de una reducción por Bouguer se conoce como “gravedad reducida isostáticamente” ( $g_I$ ), que en áreas de montaña, la reducción está dada por  $g_I = g_B + A_c$ . Finalmente, la anomalía isostática se obtiene de la siguiente expresión:

$$\Delta g_I = \Delta g_B + A_c \quad (12)$$

### 2.1.11 Técnica Remove – Restore

Este apartado, tiene por objeto dar una base teórica no tan detallada, ya que no está dentro del desarrollo del estudio, sin embargo, se utilizará a manera de analogía el principio en el que se basa la técnica de Remove – Restore.

El principio fundamental de esta técnica es el uso de campos aproximados (modelos), que hacen las veces de variables auxiliares, para remover tendencias de las variables bajo estudio, especialmente aquellos modelos que están formulados con expresiones de armónicos esféricos de potencial con grados 180 o mayor (Heiskanen & Moritz, 1985) (como es el caso de los modelos geopotenciales usados en este estudio, EGM08 y WGM12).

Para aplicar esta técnica se debe considerar dos cuestiones: a) lo que es removido antes de ingresar los datos de entrada, debe ser incorporado en la salida; b) es posible usar dos o más modelos auxiliares diferentes en forma simultánea. El objetivo de Remove – Restore es eliminar las principales tendencias como (Heiskanen & Moritz, 1985):

- Efectos topográficos locales que producen anomalías de Bouguer;
- Efectos a nivel regional produce anomalías de Bouguer e Isostáticas;
- Efectos globales, que son expresados a través de modelos terrestres que llevan a las llamadas anomalías residuales.

Para entender mejor el principio de Remove – Restore, se da el siguiente ejemplo: en el paso de “remove” una parte de la longitud de onda larga (dada por un modelo geopotencial global) y una parte de la longitud de onda corta (dada por datos combinados de GPS/nivelación) son eliminados,

se realiza el cálculo con los datos de entrada (en este caso usando técnicas geoestadísticas), y finalmente en “restore” las partes eliminadas se incorporan en el modelo resultante.

Este es el principio que se utilizará en el presente estudio, y justifica el uso de modelos globales como variables auxiliares que proveen una longitud de onda larga, combinados con datos de nivelación diferencial y posicionamiento GPS que proveen la longitud de onda corta para la predicción de un modelo de ondulación geoidal local.

## **2.2 Fundamento legal**

El presente proyecto de investigación se fundamenta en el “Reglamento a la Ley de la Cartografía Nacional”, en mención a lo señalado en el Título I de Generalidades, Capítulo II de las Cartas y Mapas, en su artículo 18, literales d) y e), en que, considera como plano de referencia de las alturas, al nivel medio de los mares con su origen en La Libertad, y el uso de la Proyección Universal Transversa de Mercator (UTM) para elaborar productos cartográficos respectivamente (IGM, 1978).

## **2.3 Hipótesis**

Es factible obtener un modelo de ondulación geoidal de alta precisión a través del método Cokriging utilizando muestras dispersas y en menor cantidad que el estimador Kriging.

## **CAPÍTULO III**

### **GEOESTADÍSTICA**

#### **3.1 Estadística espacial**

Se entiende por estadística espacial a la rama de la estadística que se encarga de analizar los datos de variables aleatorias medidos en distintos lugares, distribuidos aleatoriamente o por agregados espaciales, en un área geográfica. A la estadística espacial se la puede dividir en tres partes: lattices o enmallados, patrones espaciales y geoestadística (Giraldo, 2007), de las cuales, esta última se utilizará en el presente estudio.

#### **3.2 Geoestadística**

El padre de esta ciencia, Matheron (1970), la definió como la aplicación de la teoría de las variables regionalizadas, para ser usada (en un principio), en la estimación de recursos mineros. Actualmente, la geoestadística se la considera una rama de la estadística que se encarga del estudio de los fenómenos espaciales (Journel & Huijbregts, 1978), para estimar, predecir y simular su comportamiento que ocurren en la naturaleza (Myers, 1987), mediante la descripción de la continuidad espacial de la variable (Giraldo, 2007).

Es una herramienta que permite identificar la continuidad espacial de los fenómenos naturales, incorporando variaciones a las técnicas comunes de regresión, para aprovechar al máximo su continuidad en el espacio. Su propósito fundamental es interpolar los valores de un fenómeno en un área geográfica, esto se logra siempre y cuando exista continuidad espacial, caso contrario las predicciones pueden arrojar resultados poco lógicos. El proceso de predicción consta de dos fases que son, el análisis estructural de los datos y la interpolación como tal.

La geoestadística moderna, estudia al fenómeno de la naturaleza de interés (variable) considerándolo como una variable aleatoria, es decir, que en cada punto “ $x$ ” donde se realizó una medición, existe un conjunto de valores correspondientes a una propiedad  $Z(x)$ , con un valor observado  $z(x)$ , el cual se obtiene de una función de probabilidad. En el punto  $x$ , la propiedad estudiada  $Z(x)$ , representa una variable aleatoria con media  $\mu$  y varianza  $\sigma^2$  (Oliver, 2010).

A menudo existe confusión entre el uso adecuado de las palabras estimación y predicción; la primera hace referencia exclusivamente a la inferencia de los parámetros fijos y desconocidos, por otro lado, la segunda hace relación a cantidades aleatorias. A pesar de estas distinciones, en geoestadística se utiliza más frecuentemente el término estimación, pudiéndose utilizar ambas palabras como sinónimos, gracias a las razones que menciona Matheron (1989), por lo que en el estudio, se intercambiarán estos términos.

### **3.3 Criterios geoestadísticos fundamentales**

#### **3.3.1 Variable regionalizada**

Se dice que una variable  $Z(x)$  es regionalizada cuando se encuentra distribuida continuamente en el espacio presentando una estructura de correlación (Díaz, 2002; Matheron, 1969; Matheron, 1970). Generalmente, la relación de valores medidos en dos lugares cercanos, son semejantes, de forma inversa sucede con los lugares más alejados, cuyos valores serán menos similares.

Matemáticamente, una variable regionalizada es un proceso estocástico (variable aleatoria)  $Z$  establecida en un espacio euclidiano  $d$ -dimensional  $R^d$ ,  $\{Z(x): x \in D \subset R^d\}$ , donde generalmente es un espacio tridimensional, es decir en un punto  $x = (x_1, x_2, x_3)$  (Giraldo, 2007).

Si la variable es regionalizada, debe existir un proceso estocástico subyacente en el que sus valores medidos en puntos cercanos entre sí, están autocorrelacionados, por lo que, es posible establecer su covarianza espacial, la cual para un par de puntos viene dada por:

$$C(x_1, x_2) = E[\{Z(x_1) - \mu(x_1)\}\{Z(x_2) - \mu(x_2)\}] \quad (13)$$

Donde  $\mu(x_1)$  y  $\mu(x_2)$  es el valor medio de la variable  $Z$  medido en el par de puntos  $x_1, x_2$ , y  $E$  representa la esperanza matemática del proceso. Sin embargo, existe una única realización (muestra de la función aleatoria) de  $Z$  para cada punto  $x$ , en la que se desconoce el valor medio, por lo que, para ser válida esta autocorrelación, se debe adoptar hipótesis de que el proceso es estacionario (Oliver, 2010; Oliver & Webster, 2015).

### 3.3.2 Teoría de momentos de una función aleatoria

Dada una función aleatoria  $Z(x)$  definida en un espacio tridimensional, presenta un vector aleatorio  $\{Z(x_1), Z(x_2), \dots, Z(x_n)\}$  caracterizado por su función de distribución n-variada:

$$F_{Z(x_1), Z(x_2), \dots, Z(x_n)}(z_1, z_2, \dots, z_n) = \Pr[Z(x_1) \leq z_1, Z(x_2) \leq z_2, \dots, Z(x_n) \leq z_n] \quad (14)$$

Para todo valor de  $n$  en cualquier punto, establece la ley de probabilidad espacial para la función estocástica  $Z(x)$ , la misma que es imposible determinarla completamente, por lo que en la práctica se estima los primeros momentos de la distribución de  $Z(x)$ . En geoestadística lineal, se emplean hasta momentos de orden dos, ya que muchas de las veces, debido a la información disponible de la variable bajo estudio, no se puede estimar momentos superiores al grado dos (Díaz, 2002).

Se presenta a continuación los momentos de distribución para  $Z(x)$  (Díaz, 2002; Giraldo, 2007):

- Momento de primer orden, conocido como esperanza matemática, representa la media en la que se distribuyen los valores de las realizaciones de la función aleatoria:

$$m(x) = E[Z(x)] \quad (15)$$

- Momentos de segundo orden:
  - Varianza o varianza a priori, mide la dispersión de  $Z(x)$  alrededor de su media, representando su carácter estocástico:

$$\sigma^2(x) = Var[Z(x)] = E[\{Z(x) - m(x)\}^2] \quad (16)$$

- Covarianza o función autocovarianza, indica la relación existente entre  $Z(x_1)$ ,  $Z(x_2)$ :

$$C(x_1, x_2) = E[\{Z(x_1) - m(x_1)\}\{Z(x_2) - m(x_2)\}] \quad (17)$$

- Semivariograma o función de semivarianza:

$$2\gamma(x_1, x_2) = Var[Z(x_1) - Z(x_2)] \quad (18)$$

$$\gamma(x_1, x_2) = \frac{1}{2}E[\{Z(x_1) - Z(x_2)\}^2] \quad (19)$$

- Correlograma o coeficiente de correlación lineal, es adimensional con valores entre  $[-1,1]$ , mientras más cercano a la unidad señala correlación y que ambos son proporcionales e independientes (Emery, 2007):

$$\rho(x_1, x_2) = \frac{C(x_1, x_2)}{\sqrt{Var[Z(x_1)], Var[Z(x_2)]}} \quad (20)$$

Cabe señalar que el semivariograma o variograma  $2\gamma(x_1, x_2)$ , y la varianza son siempre positivos, en cambio, la covarianza y correlograma se pueden presentar valores negativos. Se diferencian en que la covarianza y correlograma representa cuan semejantes son las realizaciones

en dos sitios diferentes, mientras el semivariograma expresa cuan diferentes son estos (Emery, 2007).

### 3.3.3 Hipótesis de estacionariedad

Se conoce como función aleatoria estrictamente estacionaria si (14) es constante para cualquier valor de un vector de traslación  $h$  (lag o paso), o en otras palabras, la función de distribución del vector aleatorio  $\{Z(x_1), Z(x_2), \dots, Z(x_n)\}$  es similar al vector  $\{Z(x_1 + h), Z(x_2 + h), \dots, Z(x_n + h)\}$  independientemente del valor de  $h$  (Díaz, 2002). En geoestadística se trabaja en los momentos descritos anteriormente, relacionando la estacionariedad en base a estos.

#### 3.3.3.1 Estacionariedad de segundo orden o fuerte

Una función aleatoria, para que sea definida como estacionaria de segundo orden debe cumplir lo siguiente (Díaz, 2002; Giraldo, 2007):

- La esperanza existe, es finita, constante e independiente de  $x$ :

$$E[Z(x)] = m; \forall x \quad (21)$$

- Cualquier par  $Z(x), Z(x + h)$ , existe su covarianza y depende únicamente del vector  $h$ :

$$C(h) = C(x, x + h) = E[Z(x), Z(x + h)] - m^2 \quad (22)$$

Si la covarianza existe, obviamente la varianza también, la cual es finita e independiente de  $x$ , por lo que:

$$\sigma^2 = C(0) = Var[Z(x)] \quad (23)$$

Lo mismo sucede para el semivariograma, que si es estacionario cumple con:

$$\gamma(h) = \gamma(x, x + h) = \frac{1}{2} E[\{Z(x), Z(x + h)\}^2] \quad (24)$$

Existen varias relaciones directas entre los momentos siempre y cuando sean estacionarias de segundo orden, tales como (Emery, 2007):

$$\gamma(h) = C(0) - C(h) = \sigma^2[1 - \rho(h)] \quad (25)$$

$$\sigma^2 = C(0), \quad \text{si } h = 0 \quad (26)$$

$$\rho(h) = \frac{C(h)}{C(0)} \quad (27)$$

$$\gamma(\infty) = C(0) = \sigma^2 \quad (28)$$

### 3.3.3.2 Estacionariedad débil o intrínseca

En la naturaleza, existen fenómenos físicos que poseen un incremento de varianza infinita en todo el espacio geográfico, y por ende no existe covarianza ni correlograma (Emery, 2007). Para dar solución a esta desventaja de la estacionariedad de segundo orden, Matheron formuló la hipótesis intrínseca, la misma que se basa en el supuesto que la media podría no ser constante en toda el área, pero es posible que lo sea para valores de  $h$  cortas, es decir sus incrementos  $Z(x + h) - Z(x)$  son estacionarios y muestran una varianza finita. Según esto, se debe cumplir lo siguiente (Giraldo, 2007):

$$E[Z(x + h) - Z(x)] = 0 \quad (29)$$

$$Var[Z(x + h) - Z(x)] = E[Z(x + h) - Z(x)]^2 = 2\gamma(h) \quad (30)$$

Se dice entonces que una función aleatoria que sea estacionaria de orden dos (o fuerte), es siempre intrínseca (o débil) también, pero el proceso inverso no siempre es factible.

### 3.3.3.3 No estacionariedad

Se tiene el caso de funciones aleatorias (variable regionalizada), conocidas como no estacionarias, que presentan esperanza dependiente de  $x$ , es decir, no constante (Díaz, 2002):

$$E[Z(x)] = m(x) \quad (31)$$

Donde  $m(x)$  representa una función de tendencia o deriva. Considerando una función aleatoria  $Z(x)$  resultado de la sumatoria de una parte determinística  $m(x)$  y un residuo estacionario  $R(x)$  de esperanza matemática nula, se tiene que:

$$Z(x) = m(x) + R(x) \quad (32)$$

Por lo que el semivariograma resultante va a depender de  $x$ :

$$\gamma(x + h, x) = \gamma_R(h) + \frac{1}{2} [m(x + h) - m(x)]^2 \quad (33)$$

Esto representa un fenómeno no estacionario, sin embargo, si se consideran a estas funciones como intrínsecas de un orden  $k$  específico, su comportamiento es estacionario (Díaz, 2002).

### 3.3.4 Funciones de correlación espacial

La dependencia espacial se puede representar con cualquiera de las tres funciones antes mostradas, que son: semivariograma, covarianza y correlograma, las mismas que poseen las siguientes expresiones matemáticas:

- Semivariograma: basado en la hipótesis intrínseca, se asume para una variable regionalizada, la varianza de sus incrementos es finita. De (30) se despeja  $\gamma(h)$ :

$$\gamma(h) = \frac{E[Z(x + h) - Z(x)]^2}{2} \quad (34)$$

- Covariograma: la función covarianza se calcula mediante la fórmula:

$$C(h) = \frac{\sum_{i=1}^n [Z(x+h) - m] * [Z(x) - m]}{n} \quad (35)$$

Donde  $m$  es el valor promedio en todos los puntos del área geográfica,  $n$  representa el número de pares de puntos.

- Correlograma: dado un fenómeno estacionario y asumiendo un valor de varianza para la variable en base a la varianza muestral, se tiene:

$$\rho(h) = \frac{C(Z(x+h), Z(x))}{S_{x+h} * S_x} = \frac{C(h)}{S_x^2} = \frac{C(h)}{C(0)} \quad (36)$$

Como se explicó anteriormente, si el fenómeno estudiado cumple con la hipótesis de estacionariedad se puede utilizar el semivariograma, correlograma y covariograma para representar la dependencia espacial de una variable regionalizada, siempre que esta sea de orden dos, pero si el proceso demuestra un comportamiento únicamente intrínseco, no existe la función covarianza ni correlograma, sin embargo, si es posible definir el semivariograma, gracias a que es la única función que no necesita estimar parámetros (Giraldo, 2007), siendo una herramienta poderosa y versátil, convirtiéndola en el núcleo de la geoestadística (Oliver, 2010).

### 3.3.5 Isotropía y anisotropía

Es muy importante tener claro estos dos términos, ya que en la naturaleza, los fenómenos pueden seguir alguno de estos dos comportamientos. En un caso, donde la esperanza matemática de la variable cambia según la dirección, también si la covarianza o semivariograma están dependiendo del sentido en que se determinaron, la estacionariedad no existirá, es decir es anisotrópico. El caso

contrario es cuando el proceso solo es dependiente del vector  $h$ , mas no de su orientación, por lo que es isotrópico (Emery, 2007; Giraldo, 2007).

La anisotropía se puede presentar de varias formas, de las cuales la anisotropía geométrica y zonal son las más importantes. La primera hace que el rango del variograma o el gradiente inicial varíe según la orientación, y a través de una transformación de coordenadas se elimina; la segunda no puede ser eliminada con tanta facilidad como la anterior, se presenta cuando la meseta cambia según la orientación, indicando zonas con orientación de preferencia. Generalmente para analizar anisotropía se utilizan las direcciones en  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$  (Emery, 2007).

### **3.4 Análisis exploratorio de datos**

Este paso constituye en primer peldaño en todo proceso geoestadístico, ya que permite, por medio de gráficos exploratorios, determinar: la forma de la distribución de los datos, valores anómalos o extremos, variabilidad, medidas de localización, correlación, entre otros supuestos que son indicadores y guías para la selección de los posibles métodos a seguir (Giraldo, 2007).

Además, es esencialmente útil cuando se trabaja con datos georreferenciados, para lo cual existen varias herramientas de análisis gráfico como diagrama de caja, tallos u hojas, histogramas, etc., que facilitan la identificación de valores externos o anómalos, su forma y variabilidad (Hoaglin, Mosteller, & Tukey, 1983). También se utilizan gráficos de dispersión para establecer interacciones entre variables o tendencias.

### 3.5 Análisis estructural de datos

El análisis estructural es uno de los más importantes dentro del proceso de predicción geoestadística, ya que es aquí en donde se construye el modelo probabilístico que explica el comportamiento de la variable aleatoria de interés.

Básicamente se estima y modela a través de una función que mejor represente la correlación espacial de una variable aleatoria regionalizada a través de supuestos de estacionariedad en base a funciones de semivarianza o covarianza, según el comportamiento del proceso (Díaz, 2002). Como se mencionó anteriormente, la semivarianza es la función más usada en geoestadística, la misma que será objeto de estudio en este trabajo.

#### 3.5.1 Estimación del semivariograma

El método de estimación por momentos propuesto por Matheron, es el más usado para la obtención del semivariograma empírico (también existe el método de máxima verosimilitud), llamado así porque representa el comportamiento real de los datos de una variable medida, cuya ecuación es igual a (34), expresado en sumatoria como:

$$\hat{\gamma}(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} [z(x_i) - z(x_i + h)]^2 \quad (37)$$

Donde  $z(x_i)$ ,  $z(x_i + h)$  representan los valores de la variable aleatoria  $Z$ , medida en los puntos  $x_i, x_i + h$ ;  $m(h)$  se refiere al número de pares comparados a una distancia  $h$ , cuya variación permite la construcción del variograma muestral (experimental) (Oliver, 2010).

Si los datos de la variable son medidos en una grilla o transecto regular, el valor de  $h$  se convierte en un escalar de múltiplos enteros, con un lag máximo de un tercio de la longitud del transecto. En

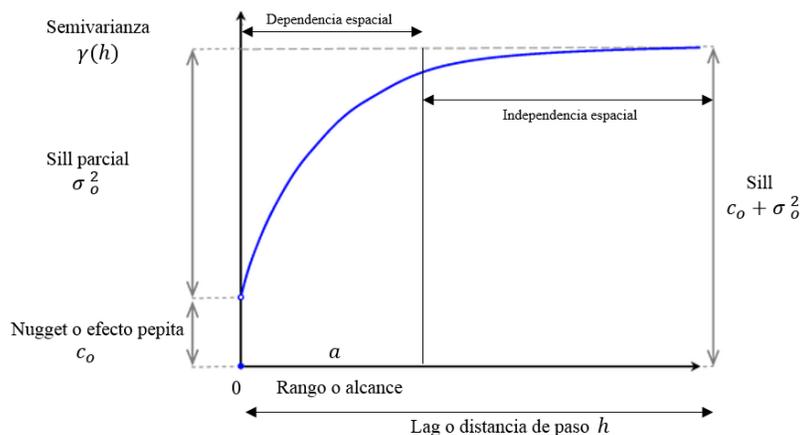
cambio, si los datos son medidos irregularmente o se desea obtener el variograma omnidireccional de datos regularmente muestreados, el valor del lag se considera dentro de sectores de búsqueda (vecindad), con tolerancias de distancia y orientación.

Según Oliver y Webster (2015), si se utiliza el método de momentos, para obtener un semivariograma confiable, se necesitan por lo menos 100 datos del valor medido de la variable bajo estudio.

El enfoque de este estudio no se basa en el cálculo del semivariograma empírico, sino que se utilizará el cross – semivariograma (semivariograma cruzado) para una optimización de datos obteniendo la misma confiabilidad que con el semivariograma simple, sin embargo, es necesario estudiar primero al semivariograma, sus características y modelos, para entender de mejor forma al cross – semivariograma, ya que es la extensión del primero en un caso multivariante.

### 3.5.2 Características del semivariograma

Todo semivariograma que represente una variable aleatoria regionalizada, y que además siga un comportamiento estacionario de segundo orden, va a presentar la siguiente forma:



**Figura 18.** Semivariograma para un fenómeno estacionario de segundo orden

Fuente: adaptado de (Oliver, 2010)

El semivariograma mostrado en la figura anterior, representa el comportamiento en la mayoría de los fenómenos naturales, ya que su variación es continua en el espacio geográfico (Oliver & Webster, 2015). Los elementos fundamentales que definen un semivariograma son: rango o alcance ( $a$ ), es la distancia en que las observaciones son dependientes espacialmente, más allá de este valor, se dice que la correlación espacial no existe (Wagner, 2003), además se puede dar el caso de que no exista un rango (observaciones independientes), por lo que se define un “rango efectivo”, entendiéndose al alcance en el cual el 95% del semivariograma llega al sill, también puede ocurrir que la meseta sea finita (estacionariedad fuerte) o infinita (estacionariedad débil) (Giraldo, 2007). Nugget o efecto pepita ( $c_o$ ) representa un salto (discontinuidad) en el origen del semivariograma, es la variación no correlacionada debido a la escala del muestreo, incluyendo los errores en la medición de la variable (Oliver & Webster, 2014), por tanto, este valor no debe ser mayor al 50% del valor del sill (Luna, 2017); en ciertos casos puede no existir efecto pepita. Sill o meseta ( $c_1 = c_o + \sigma_o^2$ ) denota la cota superior del semivariograma y el alcance a partir del cual las observaciones no están correlacionadas, está formado por la semivarianza a priori o sill parcial ( $\sigma_o^2$ ) y el Nugget; es la diferencia promedio cuadrática de dos puntos independientes (Wagner, 2003).

Con estos tres parámetros se define al semivariograma, pero hay que recordar que, el objetivo del análisis estructural es obtener un modelo empírico de semivariograma para posteriormente ajustarlo a un modelo teórico, ya que el empírico se construye a partir del conjunto de datos discretos medidos a intervalos de lag (diferentes en la mayoría de los casos, o regulares), de la variable de interés, por lo que no se puede estimar el valor de la variable más allá de estos, y es ahí donde entra en juego el modelo teórico, con el que se ofrece una expresión matemática que mejor

se ajusta al semivariograma muestral y permita determinar el valor de la variable a cualquier distancia de lag.

### 3.5.3 Modelos teóricos de semivarianza

Para seleccionar un modelo teórico al semivariograma experimental, no basta con observar un aparente acople visual de estos, sino que es necesario que se cumpla con ciertas condiciones.

En el caso de una variable aleatoria que sea estacionaria de segundo orden, con sus funciones  $\gamma(h)$  y  $C(h)$ , se tendrá que, dada una combinación lineal de la función  $Z(x_i)$ , cumple con:  $Y = \sum_{i=1}^n \lambda_i Z(x_i)$ , donde  $\lambda_i, i = 1, \dots, n$  representa las ponderaciones arbitrarias. Su covarianza es  $Var(Y) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(x_i, x_j)$ , que puede ser positiva o cero (Díaz, 2002).

Si no existe  $C(h)$ , se está en presencia de la hipótesis intrínseca, por lo que solo habrá el semivariograma, que cumple con:  $Var(Y) = C(0) \sum_{i=1}^n \lambda_i \sum_{j=1}^n \lambda_j - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(x_i, x_j)$ ; si la sumatoria de los pesos es cero, entonces:  $Var(Y) = - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(x_i, x_j)$ , lo que señala que  $\gamma(h)$  es condicionalmente semidefinida negativa (Díaz, 2002). Esto se puede resumir en las siguientes propiedades que deben ser cumplidas para que el semivariograma sea el adecuado (Emery, 2007):

- Debe ser nulo en el origen:  $\gamma(0) = 0$ ;
- Paridad:  $\gamma(h) = \gamma(-h)$ ;
- Es positivo:  $\gamma(h) \geq 0$ ;
- Su comportamiento hacia el infinito es:  $\lim_{|h| \rightarrow +\infty} \frac{\gamma(h)}{|h|^2} = 0$ ;

- Condicional negativo semidefinido:  $\forall k \in \mathbf{N}^*, \forall \lambda_1, \dots, \lambda_k \in \mathbf{R}$  tales que  $\sum_{i=1}^k \lambda_i = 0, \forall x_1, \dots, x_k \in \mathbf{D}, \sum_{i=1}^k \sum_{j=1}^n \lambda_i \lambda_j \gamma(x_i - x_j) \leq 0$ .

Esta última propiedad es suficiente para asegurar que  $\gamma$  es el semivariograma de una función estocástica. Es “condicional” ya que la desigualdad restringe los valores que pueden tomar los pesos, asegurando que su sumatoria sea cero.

Los modelos teóricos que cumplen estos requisitos, se les conoce con el nombre de modelos autorizados, los cuales pueden ser transitivos o acotados, y no acotados. Los primeros representan a funciones delimitadas, es decir que presentan un comportamiento estacionario de segundo orden (alcanzan una meseta), mientras que los últimos son funciones no delimitadas (sin meseta), es decir, son procesos que están regidos por la hipótesis intrínseca (Díaz, 2002). A continuación se presentan las expresiones matemáticas de los modelos teóricos acotados y no acotados más usuales.

**Tabla 1.**

*Modelos teóricos acotados de semivariogramas*

Modelo teórico	Fórmula matemática	
Gaussiano	$\gamma(h) = c_0 + c_1 \left[ 1 - \exp\left(\frac{-h}{a}\right)^2 \right]; h \geq 0$	(38)
Exponencial	$\gamma(h) = c_0 + c_1 \left[ 1 - \exp\left(\frac{-3h}{a}\right) \right]; h \geq 0$	(39)
Esférico	$\gamma(h) = \begin{cases} c_0 + c_1 \left[ \frac{3}{2}\left(\frac{h}{a}\right) - \frac{1}{2}\left(\frac{h}{a}\right)^3 \right]; h \leq a \\ c_0 + c_1; h > a \end{cases}$	(40)
Circular	$\gamma(h) = \begin{cases} \frac{2c_0 + c_1}{\pi} \left[ \frac{h}{a} \sqrt{1 - \left(\frac{h}{a}\right)^2} + \arcsin\left(\frac{h}{a}\right) \right]; h \leq a \\ c_0 + c_1; h > a \end{cases}$	(41)

Continúa 

$$\text{Pentaesférico} \quad \gamma(h) = \begin{cases} c_0 + c_1 \left[ \frac{15}{8} \left(\frac{h}{a}\right) - \frac{5}{4} \left(\frac{h}{a}\right)^3 + \frac{3}{8} \left(\frac{h}{a}\right)^5 \right]; & h \leq a \\ c_0 + c_1 & ; h > a \end{cases} \quad (42)$$

$$\text{Estable} \quad \gamma(h) = c_0 + c_1 \left[ 1 - \exp\left(\frac{-3h}{a}\right)^\theta \right]; h \geq 0, 0 < \theta \leq 2 \quad (43)$$

$$\text{Bessel*} \quad \gamma(h) = 1 - \left(\frac{h}{a}\right) K_1\left(\frac{h}{a}\right); h \geq 0, \quad (44)$$

$$\text{Cuadrático racional} \quad \gamma(h) = c_0 + c_1 \left[ \frac{19 \left(\frac{h}{a}\right)^2}{1 + 19 \left(\frac{h}{a}\right)^2} \right]; h \geq 0 \quad (45)$$

$$\text{Lineal con meseta} \quad \gamma(h) = \begin{cases} c_0 + c_1 \left(\frac{h}{a}\right) & ; 0 \leq h \leq a \\ c_0 + c_1 & ; h > a \end{cases} \quad (46)$$

\* $K_1$ : primera función modificada de Bessel de tipo dos

Fuente: (Díaz, 2002; Giraldo, 2007; Luna, 2017)

## Tabla 2.

### Modelos teóricos no acotados de semivariogramas

Modelo teórico	Fórmula matemática
Potencia	$\gamma(h) = \frac{1}{2}h^\theta; 0 < \theta < 2$ (47)
Pepita puro	$\gamma(h) = \begin{cases} 0 & ; h = 0 \\ c_0 & ; h > 0 \end{cases}$ donde $c_0 > 0$ (48)
Logarítmico	$\gamma(h) = \begin{cases} 0 & ; h = 0 \\ \log(h + a) & ; h > 0 \end{cases}$ (49)

Fuente: (Díaz, 2002)

Cabe señalar, que estos no solo los únicos modelos teóricos que existen, sino que de hecho hay varios más. En este trabajo serán usados algunos de los ya mostrados en las Tablas 1 y 2, centrandos su uso a través de librerías en el software libre R, con los paquetes RGeostats (Renard, Ors, & Desassis, 2017) en el que hay 28 modelos teóricos, y Gstat (Pebesma, 2014) que cuenta con 12 modelos teóricos, como se mostrará más adelante.

### 3.5.4 Modelamiento del semivariograma experimental

Para realizar la modelación del semivariograma experimental, se tiene que tener en cuenta que el modelo teórico debe ajustarse lo mejor posible al muestral, principalmente en los puntos iniciales del semivariograma, ya que son los valores más confiables (Alfaro, 2007). Según el comportamiento del semivariograma en el origen se puede tener un indicio de que modelo teórico usar, este comportamiento es de tres tipos (Ligas & Kulczycki, 2010):

- Lineal: variable regionalizada continua, no muy regular;
- Parabólico (cuadrático): variable regionalizada muy regular, o ligado también a la presencia de una tendencia (drift);
- Discontinuo: variable regionalizada muy errática, se debe a la distribución intermitente (discontinua) de sus datos, cuyos valores cambian con una escala pequeña y es poco apreciable su continuidad. A esto se conoce como efecto pepita o Nugget.

No existe un proceso definido para el modelamiento del semivariograma, pero se debe tener en cuenta ciertos aspectos que son útiles para esto, como por ejemplo (Díaz, 2002):

- La distancia de paso o lag ( $h$ ) debe contar con una tolerancia  $\pm\Delta h/2$  y con una dirección  $\beta$  de tolerancia  $\pm\Delta\beta/2$ , esto es útil principalmente para muestras que son distribuidas de forma aleatoria, por lo que no se encontrarán datos a un valor exacto de lag. El semivariograma resultante se considera regular o suavizado;
- Por principio, los datos más cercanos entre sí están más correlacionados que los más alejados, por lo que el semivariograma se estima con  $|h|$  inferiores a la mitad de la distancia máxima de separación entre muestras  $|h| < D_{max}/2$ ;

- Definir el número de intervalos (tamaño de lag), si bien es arbitrario, es aconsejable un máximo de 25 intervalos y mínimo 10 para establecer con precisión el rango y sill del semivariograma experimental;
- El número de pares en cada intervalo deben ser los suficientes para que el semivariograma estimado sea estable, siendo recomendable de 30 a 50 pares.

### **3.6 Predicción espacial**

#### **3.6.1 Criterios de una predicción espacial**

En la literatura de geoestadística, se considera a la técnica Kriging como el mejor predictor lineal insesgado, es decir que es un predictor espacial óptimo debido a que es insesgado y posee varianza mínima (Oliver, 2010). Para alcanzar la más alta predicción posible, la variable aleatoria que se estudie debe tener una distribución normal de sus datos, sin embargo, las variables no siempre cumplen esta condición sea por el comportamiento natural de la variable, o a la escasez de datos (Caro, 2012). Se puede optar por transformaciones de los datos para llegar a una distribución normal, como logarítmica, box – cox (cuadrática) (Oliver & Webster, 2014), pero se puede dar el caso que estas transformaciones no conduzcan a una normalización visible, por lo que es preferible trabajar con los datos crudos (sin transformar) (White, 1997).

Kriging aporta un componente principal a la predicción espacial, y es que ofrece varianzas Kriging o errores, que indican la incertidumbre del modelo de predicción realizado, esto es algo que le diferencia de los métodos de predicción determinísticos (Oliver & Webster, 2015), y además evita los errores sistemáticos en el cálculo de modelo final de predicción (Alfaro, 2007).

Sin embargo, como se aprecia en el título de este trabajo, se utilizará el método Cokriging para la realización de la predicción espacial, por las razones mencionadas en el problema planteado del estudio, y porque se ha demostrado que si se trabaja con variables auxiliares (covariables) que estén correlacionadas con la variable de interés, se logran resultados superiores a los obtenidos con Kriging (Hengl, 2009; Mulla & McBratney, 2002; Waller & Gotway, 2004).

### **3.6.2 Ponderaciones Kriging**

Los pesos o ponderaciones Kriging dependen de varios factores, como el alcance del semivariograma, la relación Nugget – sill, el modelo teórico usado, forma en que fue realizado el muestreo y la anisotropía presente (Oliver & Webster, 2015). Para definir los pesos que se asignan a cada punto, se necesita un radio de búsqueda óptimo, este radio (vecindario o vecindad) cumple el propósito de asignar mayor peso a los datos que se encuentren más cerca del punto a predecir, lo que indica que Kriging asigna las ponderaciones localmente (predictor local) (Oliver & Webster, 2014).

En un modelo de semivariograma constituido por un modelo pepita puro, el peso dado a los puntos será igual y corresponde al valor medio de la vecindad de búsqueda (Giraldo, 2007). La incidencia que tiene el rango sobre los pesos depende del modelo de semivariograma usado, mostrando un mayor peso a los puntos cercanos al lugar de predicción a medida que se incrementa el valor del alcance. Si se trabaja con datos distribuidos de forma discontinua (irregular), se tiene que el peso individual por punto será mayor en los aislados y menor en los agrupados (Oliver, 2010).

### 3.6.3 Validación cruzada

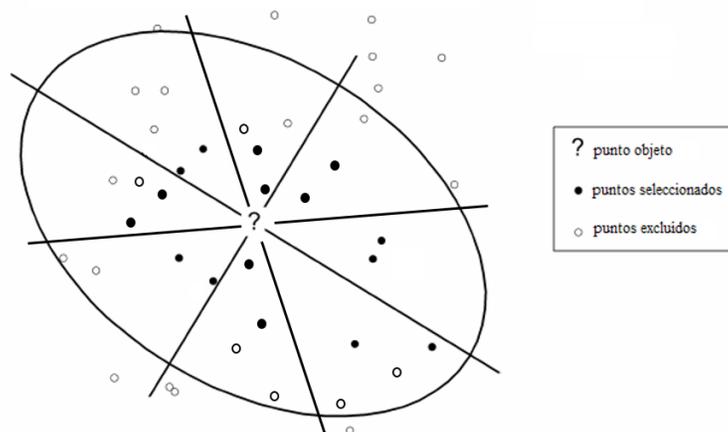
Este método consiste en eliminar un punto de los utilizados para generar el modelo de predicción, y con los datos restantes más el semivariograma ajustado se estima con Kriging el valor del punto excluido, cuya diferencia entre el valor predicho y observado (real) sería mínima, con lo que se infiere que el modelo teórico establecido representa adecuadamente la estructura espacial de autocorrelación de la variable (Giraldo, 2007). En este apartado, juega un papel importante la vecindad de búsqueda, que para la mayoría de predicciones hechas con Kriging se trabaja con un “vecindario móvil”, esto quiere decir que se desplaza en el área geográfica a través de los nodos de la grilla regular de predicción (Emery, 2007). Como se mencionó, Kriging es un estimador local, lo cual representa las siguientes ventajas (Oliver & Webster, 2014):

- Considera únicamente a los puntos más cercanos al punto o bloque objeto, reduciendo tiempos de cálculos con matrices más pequeñas;
- Se asume estacionariedad para la media dentro de la vecindad de búsqueda;
- Utiliza solamente las semivarianzas cercanas al eje de las ordenadas, asegurando un buen modelamiento del semivariograma con los primeros puntos aledaños al origen.

No existe un procedimiento estándar para definir el vecindario de búsqueda, sin embargo, se puede señalar algunas consideraciones importantes (Oliver & Webster, 2014):

- Si se cuenta con abundantes datos, un semivariograma acotado y efecto pepita pequeño, se recomienda que el vecindario de búsqueda esté cerca del rango (o en el rango efectivo), ya que los puntos más allá de este no poseerán pesos significativos;

- En caso de tener un semivariograma con un efecto pepita considerable, la vecindad debe superar al rango, ya que los puntos más allá del alcance podrían tener pesos importantes. Esto también es aplicable si se cuenta con datos escasos que presenten puntos fuera del rango;
- Se recomienda un número máximo y mínimo de puntos dentro del vecindario, entre 7 – 25. Según Alfaro (2007), propone un mínimo de 8, y Oliver (2010) señala un máximo de 16 – 20 puntos;
- Si las muestras son muy dispersas, es recomendable dividir la vecindad en octantes con un mínimo de dos puntos por cuadrante (ver Figura 19).



**Figura 19.** Vecindario de búsqueda por octantes

Fuente: adaptado de (Emery, 2007)

Una vez establecido el vecindario de búsqueda, se procede a examinar la bondad del ajuste del modelo de predicción a través de la denominada validación cruzada, que es el método más usado para la evaluación de valores predichos. El resultado de este proceso es un conjunto de residuales, a los que se aplican diversos indicadores estadísticos para facilitar su comprensión y tomar una decisión adecuada.

### 3.7 Geoestadística univariada

Antes de comenzar con la descripción de este apartado, hay que señalar que en la literatura científica, existen dos clasificaciones generales de métodos de interpolación, los llamados determinísticos (no geoestadísticos) y por otro lado, los estocásticos (estadísticos o geoestadísticos). Li y Heap (2008) mencionan 42 técnicas diferentes, lo que demuestra el avance en el estudio de interpolación de datos.

En este estudio, se tratan únicamente las técnicas geoestadísticas, siendo su mayor representante Kriging, dentro del cual existen subdivisiones con aplicaciones específicas. Una clasificación es mostrada en Díaz (2002), según:

- Forma del predictor:
  - Lineales: simple, ordinario, residual, universal;
  - No lineales: indicador, disyuntivo, probabilístico, lognormal, trans – gaussiano;
- Soporte para la medición de datos:
  - Puntual;
  - Bloque;
- Paramétricos y no paramétricos:
  - Paramétrico: disyuntivo, lognormal, trans – gaussiano;
  - No paramétrico: simple, ordinario, residual, universal, probabilístico, indicador.

El objetivo de este estudio no se centra en la utilización de métodos tradicionales, como los mencionados anteriormente, sino que propone el uso de nuevos enfoques para la predicción

espacial de variables en la naturaleza, en este caso, la ondulación geoidal. Este nuevo enfoque se abordará en la sección de Geoestadística multivariada.

A pesar de ello, es necesario contar con modelos para la comparación de resultados, y demostrar la mejora o no del enfoque propuesto, por lo que en esta sección se tratarán las técnicas de Kriging ordinario, universal y residual, las mismas que no se demostrará la deducción matemática de cada una, por las razones mencionadas, sino únicamente las fórmulas necesarias para entender el fundamento teórico del cual provienen.

### 3.7.1 Kriging ordinario

Esta técnica es una de las más usadas para la predicción de variables espaciales, el supuesto que se maneja es que la media es desconocida.

Se considera a la variable de interés  $Z$  como un conjunto de  $n$ -realizaciones representadas como combinaciones lineales de esta, resultando un predictor de la forma:

$$\hat{Z}(x_o) = \sum_{i=1}^n \lambda_i * Z(x_i) \quad (50)$$

Donde  $\hat{Z}(x_o)$  es el predictor de la variable  $Z$  en el punto  $x_o$  en el cual no existe el dato;  $\lambda_i$  corresponde los pesos asignados a los datos originales (conocidos), que son obtenidos en función de la distancia entre los puntos conocidos y el punto desconocido. Además, cumple con la condición de insesgamiento, lo cual se ve demostrado en la sumatoria de todos los pesos, siendo esta 1 (ecuación 52), con lo que se garantiza que la esperanza del predictor sea la misma que la esperanza de la variable estudiada (ecuación 51). Esto se expresa como:

$$E[\{\hat{Z}(x_o) - Z(x_o)\}] = 0 \quad (51)$$

$$\sum_{i=1}^n \lambda_i = 1 \quad (52)$$

El predictor Kriging  $\hat{Z}(x_o)$  es óptimo ya que minimiza la varianza del error de la predicción realizada:

$$Var[\hat{Z}(x_o)] = E[\{\hat{Z}(x_o) - Z(x_o)\}^2] \rightarrow \text{mínimo} \quad (53)$$

La varianza de Kriging ordinario expresada en función de semivarianza es:

$$\sigma_{KO}^2(x_o) = \sum_{i=1}^n \lambda_i \gamma(x_i, x_o) + \omega(x_o) \quad (54)$$

Donde  $\omega(x_o)$  es el multiplicador de Lagrange introducido para minimizar la función,  $\gamma(x_i, x_o)$  es la semivarianza entre el punto  $i$ -ésimo conocido y el punto donde se va a realizar la predicción. En Giraldo (2007) se puede observar la deducción de las fórmulas.

### 3.7.2 Kriging universal

Se diferencian dos partes de la variable, una determinística (no estacionaria) representada como una tendencia, y otra estocástica de media cero, basada en una función de semivariograma. Esta división de la variable se denota como:

$$Z(x) = m(x) + \varepsilon(x) \quad (55)$$

$$E[\varepsilon(x)] = 0 \quad (56)$$

$$Var[\varepsilon(x)] = \sigma^2 \quad (57)$$

$$E[Z(x)] = m(x) \text{ cte.} \quad (58)$$

La tendencia se expresa como:

$$m(x) = \sum_{l=1}^p a_l * f_l(x) \quad (59)$$

Siendo  $f_l(x)$  funciones determinísticas conocidas de la ubicación espacial seleccionadas para representar la tendencia,  $p$  son los términos usados en el ajuste de la tendencia,  $a_l$  son los coeficientes que hacen que estas funciones sean óptimas (parámetro desconocido). El predictor es igual a (50):

$$\hat{Z}(x_o) = \sum_{i=1}^n \lambda_i * Z(x_i) \quad (60)$$

La condición de insesgamiento viene dada por:

$$\sum_{i=1}^n \lambda_i * f_l(x_i) = \sum_{l=1}^p f_l(x_o) \quad (61)$$

Su varianza de predicción queda formulada como sigue:

$$\sigma_{KU}^2(x_o) = \sum_{i=1}^n \lambda_i \gamma(x_i, x_o) + \sum_{l=1}^p \omega_l(x_o) f_l(x_o) \quad (62)$$

Una deducción completa de las fórmulas se puede observar en (Giraldo, 2007) y (Luna, 2017).

### 3.7.3 Kriging residual

Esta técnica está bajo los mismos supuestos del Kriging universal (proceso no estacionario), por lo que a la variable aleatoria también se la divide en dos partes. Su hipótesis se basa en que la tendencia existente es conocida. Posteriormente, se obtienen los residuos generalmente a través de

una regresión lineal múltiple (Alsamamra, Ruiz, Pozo, & Tovar, 2009), con los que se estima mediante Kriging ordinario (Giraldo, 2007). La variable se expresa en forma similar a (55), como:

$$Z(x) = m^*(x) + \varepsilon^*(x) = \sum_{i=1}^n \alpha_i a_i(x) + \varepsilon^*(x) \quad (63)$$

$$m^*(x) = \hat{\alpha}(x) + \hat{\vartheta}_1(E_i) + \hat{\vartheta}_2(F_i) \quad (64)$$

$$\varepsilon^*(x) = \sum_{i=1}^n \lambda_i \varepsilon(x_i) \quad (65)$$

Donde  $m^*(x)$  es la tendencia;  $\alpha_i$  ( $\hat{\alpha}$ ,  $\hat{\vartheta}_1$ ,  $\hat{\vartheta}_2$ ) son coeficientes ajustados mediante mínimos cuadrados ordinarios (OLS);  $E_i$  y  $F_i$  son coordenadas de los puntos de muestreo (cualquier sistema de coordenada), representadas en (63) como  $a_i(x)$  que es la variable externa donde está la tendencia. Los pesos son estimados mediante Kriging ordinario. Una vez predicho  $m^*(x)$  se obtiene los residuales:

$$R(x) = Z(x) - m^*(x) \quad (66)$$

A través de Kriging ordinario se predice  $R(x)$  consiguiendo  $\hat{R}(x)$ . Finalmente es estimador viene dado separadamente por:

$$\hat{Z}(x_o) = m^*(x_i) + \hat{R}(x_i) \quad (67)$$

La varianza queda expresada de la siguiente forma:

$$\sigma_{KR}^2(x_o) = (c_o + c_1) - \mathbf{c}_o^T \mathbf{c}^{-1} \mathbf{c}_o + (\mathbf{q}_o - \mathbf{q}^T \mathbf{c}^{-1} \mathbf{c}_o)^T (\mathbf{q}^T \mathbf{c}^{-1} \mathbf{q})^{-1} (\mathbf{q}_o - \mathbf{q}^T \mathbf{c}^{-1} \mathbf{c}_o) \quad (68)$$

Donde  $c_0 + c_1$  es la variación de la meseta (para seguir la misma nomenclatura,  $c_1 = \sigma_0^2$ ),  $\mathbf{c}_0$  es el vector de covarianzas de residuales en los puntos no muestreados,  $\mathbf{q}$  es la matriz de predictores (muestras conocidas),  $\mathbf{c}$  es la matriz covarianza de residuales,  $\mathbf{q}_0$  es el vector  $(n + 1)$  donde  $n$  es el número de predictores (variables auxiliares).

La deducción de estas fórmulas puede verse en los trabajos de (Hengl, 2009), (Hengl, Heuvelink, & Rossiter, 2007), (Reyes, Giraldo, & Mateu, 2015) y (Alsamamra, Ruiz, Pozo, & Tovar, 2009).

### 3.8 Geoestadística multivariada

Una variante a la estimación de variables regionalizadas estocásticas, es considerar nuevas variables (auxiliares) para la predicción del proceso de interés. Esta estimación conjunta de variables se denomina Cokriging (o Kriging conjunto), el cual es una extensión de su similar en un caso univariable, Kriging, con la diferencia que el primero se basa únicamente en la correlación espacial de los datos medidos de la variable para estimar su valor en el espacio geográfico, en cambio Cokriging, además de esta relación espacial utiliza la correlación que existe entre las variables (objetivo y auxiliares) para predecir de mejor manera el fenómeno (variable o proceso) de interés.

En este acápite se muestra sus bases teóricas para el entendimiento y uso de la técnica Cokriging.

#### 3.8.1 Momentos cruzados de segundo orden

La definición de momentos de segundo orden (momentos cruzados), es de vital importancia para el análisis de correlación entre variables. Esta viene dada bajo el supuesto de que son funciones aleatorias relacionadas y estacionarias. Al igual que en el caso visto en la teoría de momentos para una sola variable, cuando se trabaja con dos o más variables, se pueden utilizar tanto las funciones

de covarianza y semivarianza, que en este caso, llevan el nombre de cross – covariograma o covarianza cruzado y cross – semivariograma o semivariograma cruzado.

Como se mencionó en los apartados anteriores, este estudio se enfocará más en el uso de semivariograma (semivariograma cruzado), sin embargo, algunas fórmulas se mostrarán expresadas en covarianzas, especialmente los errores de varianzas.

Considerando  $Z_i$  y  $Z_j$  funciones aleatorias (variable objetivo y auxiliar respectivamente) con un vector de separación  $h$ , la semivarianza cruzada se define como:

$$\gamma_{ij}(h) = \frac{1}{2} E\{[Z_i(x+h) - Z_i(x)][Z_j(x+h) - Z_j(x)]\} \quad (69)$$

Matricialmente (69) se representa como:

$$\Gamma(h) = [\gamma_{ij}(h)]_{i,j=1\dots N} \quad (70)$$

En términos de covarianza cruzada:

$$C_{ij}(h) = E\{[Z_i(x+h) - m_i][Z_j(x+h) - m_j]\} \quad (71)$$

Donde  $m_i = E[Z_i(x)]$  y  $m_j = E[Z_j(x)]$ . Teniendo en cuenta que si  $i = j$ , se convierte en un varianza y covarianza simple.

### 3.8.2 Propiedades del semivariograma cruzado

Entre las propiedades, se destacan las siguientes (Díaz, 2002; Emery, 2007):

- Simetría:  $\gamma_{ij}(h) = \gamma_{ji}(h)$  y  $\gamma_{ij}(h) = \gamma_{ij}(-h)$ ;
- Es nulo en el punto origen:  $\gamma_{ij}(0) = 0$ ;

- A diferencia del semivariograma simple, el cross – semivariograma puede tener valores positivos y negativos.
- Desigualdad de Cauchy – Schwarz:  $\forall i, j \in \{1, \dots, N\}, \gamma_{ii}(h)\gamma_{jj}(h) \geq |\gamma_{ij}(h)|^2$ ;
- Positividad: dado  $h$ , la matriz  $\Gamma(h)$  es positiva y simétrica, cumpliendo lo siguiente:  
 $\forall \lambda_1, \dots, \lambda_N \in \mathbf{R}, \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma_{ij}(h) \geq 0$ ,
- Relación con cross – covarianza:  $\gamma_{ij}(h) = C_{ij}(0) - \frac{1}{2}[C_{ij}(h) + C_{ij}(-h)]$ , o matricialmente:  
 $\Gamma(h) = C(0) - \frac{1}{2}[C(h) + C(-h)]$ .
- Coeficiente de correlación entre dos variables  $Z_i$  y  $Z_j$ :  $\rho_{ij} = \frac{C_{ij}(0)}{\sqrt{C_{ii}(0)C_{jj}(0)}}$ .

### 3.8.3 Estimador de semivariograma cruzado

De igual forma que en el caso univariante, el estimador de semivariograma cruzado presenta una forma análoga al primero, dado por (Oliver, 2010):

$$\hat{\gamma}_{ij}(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^N [Z_i(x_\alpha + h) - Z_i(x_\alpha)][Z_j(x_\alpha + h) - Z_j(x_\alpha)] \quad (72)$$

Donde  $N(h)$  es el número de pares  $Z_i(x_\alpha), Z_i(x_\alpha + h)$  y  $Z_j(x_\alpha), Z_j(x_\alpha + h)$  a un vector de distancia  $h = |h|$ .

### 3.8.4 Modelo de corregeionalización lineal (MCL)

Uno de los principales problemas que se encuentran en la modelación de semivarianzas cruzadas, es la falta de modelos teóricos para estas, por lo que, el análisis multivariado en Cokriging es más complejo que en Kriging, y esto se ve reflejado en el hecho que para modelar estos

semivariogramas hay que ajustar  $\frac{n(n+1)}{2}$  modelos de semivariogramas simples, donde  $n$  el número de variables (Díaz, 2002; Emery, 2007; Oliver, 2010).

Suponiendo que se trabajen con dos variables (una principal y una auxiliar), se deberán ajustar 3 modelos de semivariogramas a la vez, dos simples y un cruzado (Oliver, 2010). Esto en la práctica es el caso más usual, trabajar como máximo con una variable auxiliar, ya que, como se demuestra en las fórmulas anteriores, su cómputo es muy dispendioso a medida que se incrementa el número de variables (Giraldo, 2001).

Además, a pesar de que el semivariograma cruzado resulta ser una combinación lineal de semivariogramas simples, los modelos autorizados y sus combinaciones no garantizan que cumplan con las propiedades del cross – semivariograma, especialmente la desigualdad de Cauchy – Schwarz. La solución más utilizada y aceptada para su realización es mediante un modelo de correogionalización lineal, que es una generalización del concepto de un modelo anidado, pero en el caso multivariable (Goovaerts, 1997). El MCL trabaja bajo el supuesto que los semivariogramas simples y cruzados provienen de combinaciones lineales con un modelo base de la forma:

$$\forall i, j \in [1, N], \gamma_{ij}(h) = \sum_{u=0}^s b_{ij}^u \xi_u(h) \quad (73)$$

O en matrices:

$$\Gamma(h) = \sum_{u=0}^s B_u \xi_u(h) \quad (74)$$

Siendo  $\Gamma(h) = [\gamma_{ij}(h)]_{i,j=1\dots N}$  la matriz de semivariogramas cruzados y simples;  $\xi_u(h)$  el modelo de semivariograma básico;  $B_u = [b_{ij}^u]_{i,j=1\dots N}$  la matriz de corregeionalización. Una condición primordial para que el MCL sea coherente matemáticamente, es que cada matriz de corregeionalización  $B_u, u = 1, \dots, S$  se comporte simétrica y positivamente. En el caso de dos variables, esto se reduce a verificar que todos los valores de  $B_u$  sean  $\geq 0$ . Es decir, que cumple con la condición de positividad (Emery, 2007):

$$\forall u \in [1, S], |b_{ij}^u| \leq \sqrt{b_{ii}^u b_{jj}^u} \quad (75)$$

Lo que se indica en (75) se traduce en que, si un modelo básico está presente en el semivariograma cruzado con sill no nulo, también va a presentar un sill no nulo en los dos cross – semivariograma simples. El proceso anterior se complica para más de tres variables, ya que la desigualdad en (75) no es suficiente.

Un ejemplo de modelo de corregeionalización lineal para dos variables, se representa de la siguiente manera: Sean dos variables  $Z_i$  y  $Z_j$

$$\begin{cases} \gamma_i(h) = \alpha_0 \gamma_0(h) + \dots + \alpha_n \gamma_n(h) \\ \gamma_j(h) = \beta_0 \gamma_0(h) + \dots + \beta_n \gamma_n(h) \\ \gamma_{ij}(h) = \delta_0 \gamma_0(h) + \dots + \delta_n \gamma_n(h) \end{cases} \quad (76)$$

Donde  $\gamma_i(h)$ ,  $\gamma_j(h)$  son semivariogramas simples y  $\gamma_{ij}(h)$  semivariograma cruzado.  $\gamma_0(h), \dots, \gamma_n(h)$  son los modelos de semivariogramas básicos;  $\alpha, \beta, \delta$  son los coeficientes de la matriz de corregeionalización. Expresado en matrices, sería:

$$\Gamma(h) = \begin{pmatrix} \gamma_i(h) & \gamma_{ij}(h) \\ \gamma_{ij}(h) & \gamma_j(h) \end{pmatrix} = \sum_{u=0}^S B_u \xi_u(h) \quad (77)$$

$$B_u = \begin{pmatrix} \alpha_u & \delta_u \\ \delta_u & \beta_u \end{pmatrix} \quad (78)$$

$$\xi_u(h) = \begin{pmatrix} \xi_u(h) & 0 \\ 0 & \xi_u(h) \end{pmatrix} \quad (79)$$

Finalmente, para cerciorarse que el modelo es válido, basta con comprobar lo siguiente:

$$b_{ii}^u > 0 ; b_{jj}^u > 0 \quad \forall u = 0, \dots, S \quad (80)$$

$$|b_{ij}^u| \leq \sqrt{b_{ii}^u b_{jj}^u} \quad (81)$$

### 3.8.5 Pesos Cokriging

Sucede de forma análoga a los pesos Kriging, sin embargo, hay que tener presente que los pesos asignados a la variable principal son adimensionales, mientras que los pesos de la variable secundaria siguen una relación (unidad variable principal / unidad variable secundaria). Otro aspecto importante es que existen dos valores de ponderaciones, uno por cada variable (2 pesos), los cuales en Cokriging, condicionarán el insesgamiento del predictor (cambia la condición en Cokriging ordinario y Cokriging universal), como se verá más adelante.

### 3.8.6 Validación del modelo de cross – semivariograma

En el caso más usual, que se refiere a trabajar con una sola variable auxiliar, el semivariograma cruzado se valida por medio de la estimación por Cokriging de los valores  $Z_i$  y  $Z_j$  en los sitios donde se tiene un valor conocido, y se aplica el principio de “leave one out”, que es el mismo al

realizado en el caso univariable. Con los residuales de la predicción, se usan estadígrafos que permitan tomar decisiones respecto al modelo generado (Díaz, 2002).

### **3.8.7 Criterios adicionales del Cokriging**

El Cokriging posee en común con el Kriging algunas propiedades particulares, como por ejemplo:

- Cuando se trabaja por puntos (Kriging o Cokriging puntal), si se predice un valor en un sitio donde se conoce el valor de la variable, el resultado será el mismo, por lo que son interpoladores exactos (Alfaro, 2007; Díaz, 2002; Emery, 2007; Oliver, 2010; Oliver & Webster, 2015; Rivoirard, 2003).
- Presentan condiciones de insesgamiento, es decir, los valores de las predicciones resultantes varían menos que los reales desconocidos, por lo que presentan un suavizamiento (Emery, 2007; Oliver, 2010, Rivoirard, 2003).

También se destacan algunas situaciones por las que Cokriging es mejor que el predictor Kriging:

- Se recomienda usar este método cuando se posee pocos datos de la variable a predecir, es muy costosa o lleva mucho tiempo su obtención (Bogaert, Mahau, & Beckers, 1995; Isaaks & Srivastava, 1989).
- La varianza de estimación que se obtiene con Cokriging siempre es igual o menor que la de Kriging, gracias al análisis de correlación entre variables (Emery, 2007), por lo que sus resultados son más confiables debido al mayor aporte de las variables auxiliares (Hengl, 2009).

- Es una herramienta poderosa cuando los datos se presenten en condiciones de heterotopía (puntos muestrales ubicados en lugares diferentes), en casos donde las variables tienen buena correlación espacial, y la (o las) variable auxiliar es densamente muestreada y accesible (Rivoirard, 2003).

### 3.8.8 Cokriging ordinario

Consideremos las siguientes ecuaciones (Ligas & Szombara, 2018):

$$\begin{cases} Z_1(x) = u_{Z_1}(x) + \varepsilon_{Z_1}(x) \\ Z_2(x) = u_{Z_2}(x) + \varepsilon_{Z_2}(x) \\ Z_1(x_o) = u_{Z_1}(x_o) + \varepsilon_{Z_1}(x_o) \end{cases} \quad (82)$$

Donde  $Z_1$  y  $Z_2$  son las variables objeto y auxiliar respectivamente. Dado (82), se cumple que:

$$E[\varepsilon_{Z_1}(x_i)] = E[\varepsilon_{Z_2}(x_i)] = E[\varepsilon_{Z_1}(x_o)] = 0 \quad (83)$$

$$E[\varepsilon_{Z_1}(x_i) * \varepsilon_{Z_1}(x_i)^T] = \gamma_{11} \quad (84)$$

$$E[\varepsilon_{Z_2}(x_i) * \varepsilon_{Z_2}(x_i)^T] = \gamma_{22} \quad (85)$$

$$E[\varepsilon_{Z_1}(x_i) * \varepsilon_{Z_2}(x_i)^T] = \gamma_{12} \quad (86)$$

$$E[\varepsilon_{Z_2}(x_i) * \varepsilon_{Z_1}(x_i)^T] = \gamma_{21} \quad (87)$$

$$E[\varepsilon_{Z_1}(x_i) * \varepsilon_{Z_1}(x_o)] = \gamma_{10}; E[\varepsilon_{Z_2}(x_i) * \varepsilon_{Z_1}(x_o)] = \gamma_{20}; E[\varepsilon_{Z_1}(x_o)^2] = 0 \quad (88)$$

El estimador de cokriging ordinario, para obtener el valor de predicción en el punto  $x_o$  desconocido, se expresa de la siguiente forma (Giraldo, 2007; Ligas & Szombara, 2018; Oliver, 2010; Rivoirard, 2003):

$$\hat{Z}_1^*(x_o) = \sum_{i=1}^{n1} a_i Z_1(x_i) + \sum_{j=1}^{n2} b_j Z_2(x_j) \quad (89)$$

Donde  $a_i, i = 1, \dots, n1$  y  $b_j, j = 1, \dots, n2$  son los pesos de los observables primarios y secundarios respectivamente, los cuales se estiman de forma semejante en Kriging ordinario, y están inmersos en el MCL ajustado con los cross – semivariogramas cruzados y simples. Estos pesos serán los que determinen la condición de insesgamiento.

La ecuación (89) puede ser escrita matricialmente como:

$$\hat{Z}_1^*(x_o) = \lambda_1^T Z_1(x) + \lambda_2^T Z_2(x) \quad (90)$$

De donde  $\lambda_1, \lambda_2$  son las matrices de pesos. (89) conduce al sistema de ecuaciones siguiente (Rivoirard, 2003):

$$\left\{ \begin{array}{l} \sum_{\beta \in i} a_\beta C(Z_{1\alpha}, Z_{1\beta}) + \sum_{\beta \in j} b_\beta C(Z_{1\alpha}, Z_{2\beta}) + \omega_1 = C(Z_0, Z_{1\alpha}) \\ \sum_{\beta \in i} a_\beta C(Z_{1\beta}, Z_{2\alpha}) + \sum_{\beta \in j} b_\beta C(Z_{2\alpha}, Z_{2\beta}) + \omega_2 = C(Z_0, Z_{2\alpha}) \\ \sum_{\alpha \in i} a_\alpha = 1 \\ \sum_{\alpha \in j} b_\alpha = 0 \end{array} \right. \quad (91)$$

Las dos últimas sumatorias de (91) corresponden a la condición de insesgamiento, en el que los pesos de la variable principal suman uno, y los de la secundaria cero. En términos de matrices, se representa de la forma (Rivoirard, 2003):

$$\begin{pmatrix} C_{11} & C_{12} & 1 & 0 \\ C_{11} & C_{22} & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} * \begin{pmatrix} a_i \\ b_j \\ \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} C_{01} \\ C_{02} \\ 1 \\ 0 \end{pmatrix} \quad (92)$$

O en términos de cross – semivarianza (Ligas & Szombara, 2018):

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} & 1 & 0 \\ \gamma_{11} & \gamma_{22} & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} * \begin{pmatrix} a_i \\ b_j \\ \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} \gamma_{01} \\ \gamma_{02} \\ 1 \\ 0 \end{pmatrix} \quad (93)$$

El lado izquierdo de la matriz posee los valores calculados de semivarianzas simples y cruzadas para los pares considerados. Las últimas filas, hacen referencia a la condición de insesgamiento del predictor, los pesos son estimados en el modelo y  $\omega_1$ ,  $\omega_2$  son los multiplicadores de Lagrange. La parte derecha de la matriz posee las semivarianzas cruzadas para los puntos de predicción con las dos variables. Finalmente, la varianza de predicción Cokriging ordinario viene dada por (Oliver, 2010; Rivoirard, 2003):

$$\sigma_{CKO}^2(x_o) = C[Z_1(x_o), Z_1(x_o)] - \sum_{i=1}^{n1} a_i C[Z_1(x_i), Z_1(x_o)] - \sum_{j=1}^{n2} b_j C[Z_2(x_j), Z_2(x_o)] - \omega_1 \quad (94)$$

O en semivarianzas (Wendroth, 2018):

$$\sigma_{CKO}^2(x_o) = \gamma[Z_1(x_o), Z_1(x_o)] - \sum_{i=1}^{n1} a_i \gamma[Z_1(x_i), Z_1(x_o)] - \sum_{j=1}^{n2} b_j \gamma[Z_2(x_j), Z_2(x_o)] + \omega_1 \quad (95)$$

### 3.8.9 Cokriging universal

Dadas las siguientes ecuaciones (Ligas & Szombara, 2018):

$$\begin{cases} Z_1(x) = \mathbf{X}_{Z_1}(x)\zeta + \varepsilon_{Z_1}(x) \\ Z_2(x) = \mathbf{X}_{Z_2}(x)\zeta + \varepsilon_{Z_2}(x) \\ Z_1(x_o) = \mathbf{x}^T(x_o)\zeta + \varepsilon_{Z_1}(x_o) \end{cases} \quad (96)$$

Donde  $\mathbf{X}_{Z_1}$ ,  $\mathbf{X}_{Z_2}$  son matrices de diseño, que en su forma análoga con Kriging universal, representan las funciones determinísticas conocidas de la ubicación espacial seleccionadas para representar la tendencia;  $\zeta$  corresponde al vector de los coeficientes de la deriva presente;  $\mathbf{x}$  es el vector de diseño para el modelo de deriva.

El predictor de Cokriging universal es igual a (89), cuya representación matricial coincide con (90) (Knotters, Brus & Oude, 1995; Ligas & Szombara, 2018).

La condición de insesgamiento está dada por (Ligas & Szombara, 2018):

$$\lambda_1^T \mathbf{X}_{Z_1} = \mathbf{x}_{10}^T ; \lambda_2^T \mathbf{X}_{Z_2} = 0 \quad (97)$$

El sistema de ecuaciones que se debe resolver, es semejante al de Cokriging ordinario, de la siguiente forma (Knotters, Brus, & Oude, 1995):

$$\begin{pmatrix} C_{11} & C_{12} & \mathbf{X}_1 & 0 \\ C_{11} & C_{22} & 0 & \mathbf{X}_2 \\ \mathbf{X}_1^T & 0 & 0 & 0 \\ 0 & \mathbf{X}_2^T & 0 & 0 \end{pmatrix} * \begin{pmatrix} a_i \\ b_j \\ \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} C_{01} \\ C_{02} \\ \mathbf{x}_{10} \\ 0 \end{pmatrix} \quad (98)$$

O en términos de cross – semivarianza (Ligas & Szombara, 2018):

$$\begin{pmatrix} \gamma_{11} & \gamma_{12} & \mathbf{X}_1 & 0 \\ \gamma_{11} & \gamma_{22} & 0 & \mathbf{X}_2 \\ \mathbf{X}_1^T & 0 & 0 & 0 \\ 0 & \mathbf{X}_2^T & 0 & 0 \end{pmatrix} * \begin{pmatrix} a_i \\ b_j \\ \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} \gamma_{01} \\ \gamma_{02} \\ \mathbf{x}_{10} \\ 0 \end{pmatrix} \quad (99)$$

Por último, la varianza de predicción se expresa como (Knotters, Brus, & Oude, 1995):

$$\begin{aligned} \sigma_{CKU}^2(x_o) = & C[Z_1(x_o), Z_1(x_o)] - \sum_{i=1}^{n1} a_i C[Z_1(x_i), Z_1(x_o)] - \sum_{j=1}^{n2} b_j C[Z_2(x_j), Z_2(x_o)] \\ & - \omega_1 \mathbf{x}_{10} \end{aligned} \quad (100)$$

O en semivarianzas (Ligas & Szombara, 2018):

$$\begin{aligned} \sigma_{CKO}^2(x_o) = & \gamma[Z_1(x_o), Z_1(x_o)] - \sum_{i=1}^{n1} a_i \gamma[Z_1(x_i), Z_1(x_o)] - \sum_{j=1}^{n2} b_j \gamma[Z_2(x_j), Z_2(x_o)] \\ & + \omega_1 \mathbf{x}_{10} \end{aligned} \quad (101)$$

### 3.8.10 Cokriging residual

En este caso particular de Cokriging, se expresa de forma muy similar que Cokriging ordinario, tanto en su condición de insesgamiento, así como en la forma de estimar las ponderaciones para minimizar el error, al igual que su modelo de correogionalización lineal (Hengl, 2009), por lo que no se escribirá nuevamente las fórmulas, ya que su desarrollo es similar a las expuestas anteriormente.

## **CAPÍTULO IV**

### **METODOLOGÍA**

#### **4.1 Recopilación y validación de información**

##### **4.4.1 Proyecto Guayaquil**

Los datos para el presente estudio fueron recopilados por el IGM, para el denominado Proyecto Guayaquil, durante los años 2005 – 2006 en el que consta el control geodésico suplementario integrado por puntos GPS y altura nivelada.

##### **4.4.2 Área de estudio**

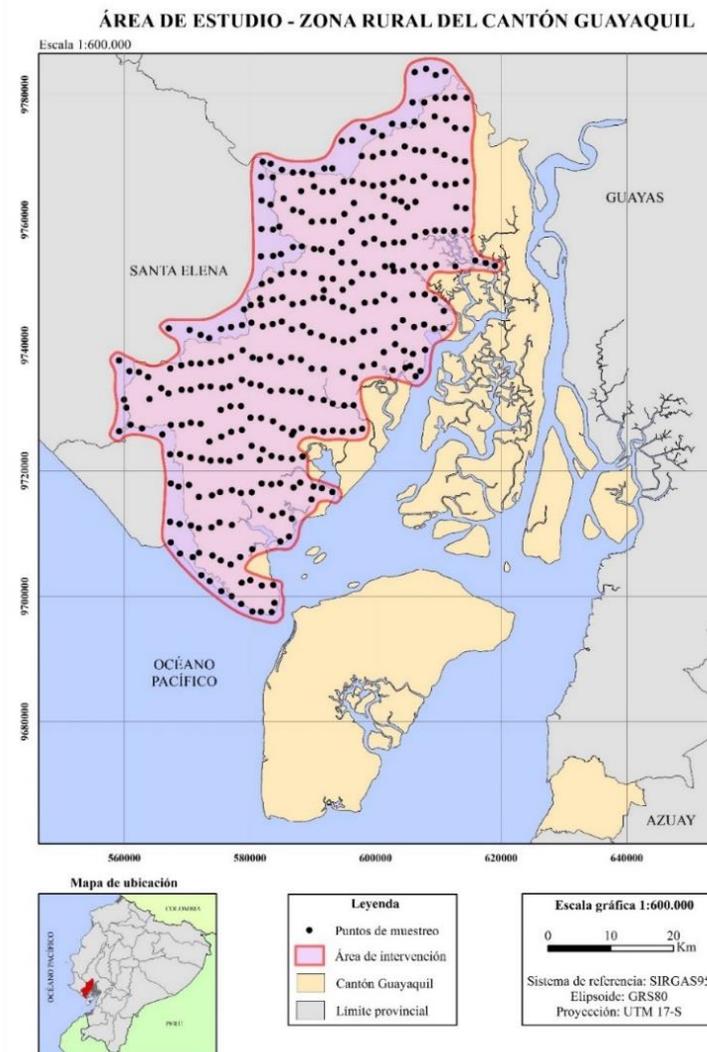
El presente estudio se ubica en la provincia de Guayas, cantón Guayaquil, más específicamente, en la zona rural del mencionado cantón (ver Figura 20). El área comprendida abarca aproximadamente 239825.09 ha., en las que se encuentran distribuidos los puntos pertenecientes al Proyecto Guayaquil.

##### **4.4.3 Datos de posicionamiento GPS**

La recolección de los datos fueron llevados a cabo por el IGM mediante posicionamiento satelital GPS con el método estático diferencial usando receptores doble frecuencia, para lograr precisiones relativas de 5mm + 1ppm en horizontal y 10 mm + 2ppm en vertical, correspondientes a latitud, longitud y altura elipsoidal respectivamente. Las coordenadas obtenidas están en el sistema de referencia SIRGAS95 con el Elipsoide GRS80, además, su representación en coordenadas planas se encuentra en UTM zona 17 sur (IGM, 2006).

#### 4.4.4 Datos de altura nivelada

A partir de puntos fijos, con altura conocida sobre el nivel medio del mar, el IGM realizó el arrastre de altura a través de nivelación diferencial de segundo orden ( $\pm 8.4\text{mm} \sqrt{k}$ ) y trigonométrica con recíprocas de tolerancia  $\pm 3\text{cm} \sqrt{k}$ , siendo  $k$  la distancia de la línea de nivelación seguida. Los puntos de control básico vertical, correspondientes a puntos de primer y segundo orden, están referidos al datum vertical oficial del país, ubicado en La Libertad (IGM, 2006).



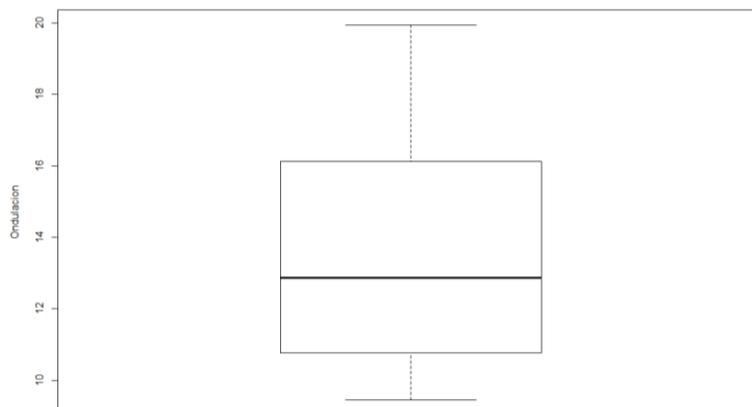
**Figura 20.** Mapa de ubicación de la zona de estudio

#### 4.4.5 Obtención de la base de datos

Como parte del convenio institucional que mantiene la Universidad de las Fuerzas Armadas – ESPE y el Instituto Geográfico Militar, se solicitó a éste los datos para la realización del proyecto, en el departamento de Geodesia del IGM, el mismo que posee almacenados los datos de todas las campañas geodésicas del país. En este caso, se solicitó la base de datos correspondiente al Proyecto Guayaquil, que consiste de un documento de Excel en el que constan los campos siguientes: nombre del punto, coordenadas geográficas latitud y longitud, coordenadas planas proyectadas este y norte, altura elipsoidal, altura nivelada, orden de precisión, técnica de nivelación.

Se procedió al filtrar la información para construir la base de datos del proyecto, que en este caso se escogieron los puntos realizados con nivelación diferencial de segundo orden; a continuación se procedió a generar una nueva base de datos con los siguientes campos heredados de la fuente original: ID del punto, coordenadas planas proyectadas este y norte, altura elipsoidal y un nuevo campo llamado ondulación geoidal, el cual fue obtenido en base a (1), por las explicaciones dadas anteriormente. Se procedió a verificar los valores de ondulación geoidal para detectar datos anómalos (mayores a 32 metros o negativos). De las novedades encontradas, todos fueron revisados con la ayuda del personal del IGM, siendo problemas en el tipo de los datos.

Una vez superado esta primera fase de verificación, se utilizó un diagrama de caja para constatar gráficamente que no existan valores atípicos, lo cual fue precisamente lo que sucedió. La base de datos final fue constituida por 332 puntos.



**Figura 21.** Diagrama de caja de los valores de ondulación geoidal

## 4.2 Selección de posibles variables auxiliares

En el Capítulo II, específicamente en la sección 2.1.9 donde se habla de los modelos geopotenciales globales, se hace especial énfasis en el EGM08 y WGM12, los cuales proveen de información respecto a valores de anomalías de gravedad como son Bouguer, Aire libre, Isostática, además del valor de ondulación geoidal. Adicionalmente, también se consideró un modelo digital de elevación (DEM), proveniente del satélite ALOS, y del programa nacional SIGTIERRAS.

En cuanto al EGM08, se escogió a este modelo ya que dispone de información libre a escala global, respecto a valores de ondulación geoidal, anomalías de Bouguer y Aire libre. Estas variables se descargaron de la plataforma oficial, para la zona de estudio, con una resolución de 2.5'x2.5' de las tres, y un derivado de este modelo solo para la ondulación geoidal cada 1'x1', es decir, se trabajó con 4 datos provenientes del EGM08.

Para el modelo WGM12 se seleccionó únicamente la anomalía Isostática, con una resolución de 2'x2' para la zona de estudio. A pesar de que esta plataforma dispone de otras anomalías (Bouguer, Aire libre), se decidió trabajar solo con una variable del WGM12, ya de que el modelo del EGM08

es más usado en el país y contiene las otras dos anomalías necesarias para el estudio (Bouguer y Aire libre).

El satélite ALOS (Advanced Land Observation Satellite) es parte de la JAXA (Japan Aerospace Exploration Agency), el cual posee múltiples recursos cartográficos, entre ellos, el denominado ALOS PALSAR DEM, que como su nombre lo indica utiliza la banda L del radar de apertura sintética (SAR) para elaborar el DEM (Alaska Satellite Facility, 2018). El modelo digital de elevaciones de ALOS tiene una resolución espacial de 12.5 m y 30 m (Alaska Satellite Facility, 2018), de los que se escogió el de mayor resolución para el proyecto.

El programa SIGTIERRAS (Sistema Nacional de Información y Gestión de Tierras Rurales e Infraestructura Tecnológica), fue creado para ayudar a la administración del sector rural del país. El modelo digital fue obtenido a partir de imágenes aéreas de escala 1:5000, cuenta con una resolución espacial de 3 m en la región Sierra, 4 m en la Costa y 5 m en el Oriente (MAGAP, 2016). Al igual que las anteriores variables, se escogió el DEM para el área del estudio.

En resumen, se seleccionó 7 posibles variables auxiliares, que son: anomalías de Bouguer, Aire libre, Isostática; ondulación geoidal, DEM del ALOS y SIGTIERRAS.

La razón del porqué se escogieron estas variables se explica a continuación. Respecto a las anomalías de gravedad, se ha demostrado que están relacionadas con la distribución de las densidades de masas internas de la Tierra (Del Cogliano, 2006), cuyos cambios se producen por el contraste de densidad lateral (Banerjee *et al.*, 1999); o asociados a corrientes convectivas (Introcaso, 2006), lo que provoca que el Geoide se deforme sobre y debajo de los excesos y deficiencias de masas respectivamente (García, 2014), y con esto, una distorsión de los campos

gravimétricos de la Tierra que se relaciona con los cambios de la ondulación geoidal (Heiskanen & Moritz, 1985). Además, existen varios estudios que han utilizado las anomalías de gravedad para el modelamiento del Geoide, especialmente, con las anomalías de Bouguer (Chicaiza, Leiva, Arranz, & Buenaño, 2017) y Aire libre (Yun, 1999).

La ondulación geoidal derivada de modelos geopotenciales globales, también ha sido utilizada para el modelamiento de ondulación geoidal local, basado en el principio de Remove – Restore, siendo esta la que proporcione información de la longitud larga de onda, a partir del análisis espectral de dichos modelos gravimétricos (Del Cogliano, 2006). Se mencionan algunos trabajos relacionados como por ejemplo (Achola & Fukuda, 2015), (Ligas & Szombara, 2018), (Banerjee *et al.*, 1999), (Achola, Fukuda, & Kuroishi, 2012), entre otros.

Finalmente, el modelo digital de elevaciones como una variable auxiliar para el modelamiento del Geoide, ha sido objeto de estudio por varios autores. Esto se debe a que un DEM proporciona información de las estructuras topográficas del terreno que se relacionan con valores elevados de anomalías gravimétricas (especialmente en áreas montañosas) (Sánchez, 2003), así como sus efectos sobre el Geoide, por ejemplo el “efecto topográfico directo” (Vaníček, Janák, & Véronneau, 2000). Algunos estudios relacionados a esta variable, se mencionan a continuación: (Al-Krargy, Hosny, & Dawod, 2015), (Kiamehr & Sjöberg, 2005), (Tsoulis, Novák, & Kadlec, 2009), (Singh & Srivastava, 2018), (Hoyer, Wildermann, Suárez, & Hernández, 2004), entre otros.

### **4.3 Determinación de la variable auxiliar**

Una vez seleccionadas las posibles variables auxiliares, se procedió a determinar cuál de éstas, es la que presenta las mejores condiciones para ser considerada en la predicción con Cokriging. Para ello, se utilizaron gráficos de dependencia espacial, como el gráfico cuartil – cuartil (qq –

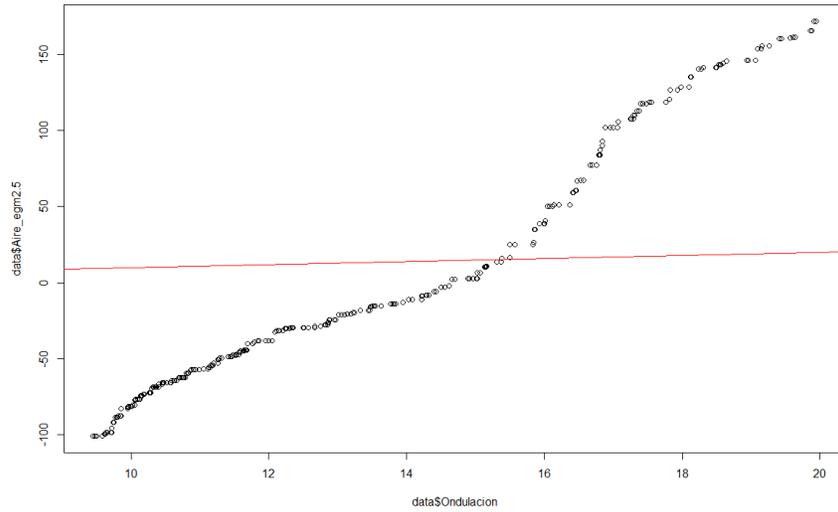
plot), dispersión (scatter – plot), además de usar coeficientes que indiquen el grado de correlación espacial con la variable objetivo, en este caso, el coeficiente de correlación de Pearson.

Para la realización de este apartado, así como en el resto del análisis geoestadístico del proyecto, se utilizó el software R, el cual es un lenguaje y entorno de programación disponible para todo público, bajo los términos y condiciones de la Free Software Foundation. Es una plataforma muy usada para realizar cálculos estadísticos y representación gráfica de estos, a través de paquetes (librerías) para realizar las operaciones. En este caso, se utilizarán diversos paquetes geoestadísticos como se verán en el transcurso del proyecto.

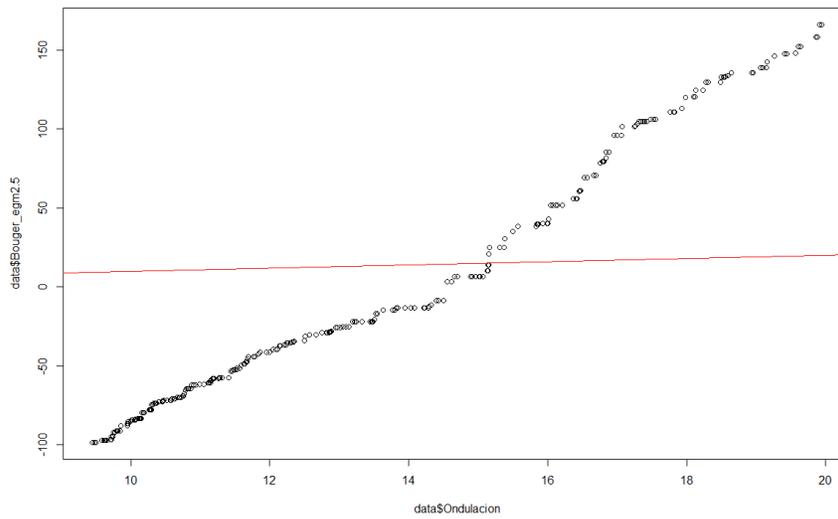
#### **4.3.1 Gráfico cuartil – cuartil**

El gráfico también llamado qq – plot sirve para comparar dos distribuciones de datos, en el cual, cada eje corresponde a los cuantiles de las variables. Si el gráfico resultante se asemeja a una línea de 45° de pendiente, es decir,  $x = y$ , siendo “x” los datos de la primera variable, “y” los datos de la segunda variable, las variables tendrán una correlación fuerte. En el caso de que no se asemeje a una línea recta, los datos todavía presentan una similar distribución de sus datos, pero con media y varianza distintas (Volfová & Šmejkal, 2012).

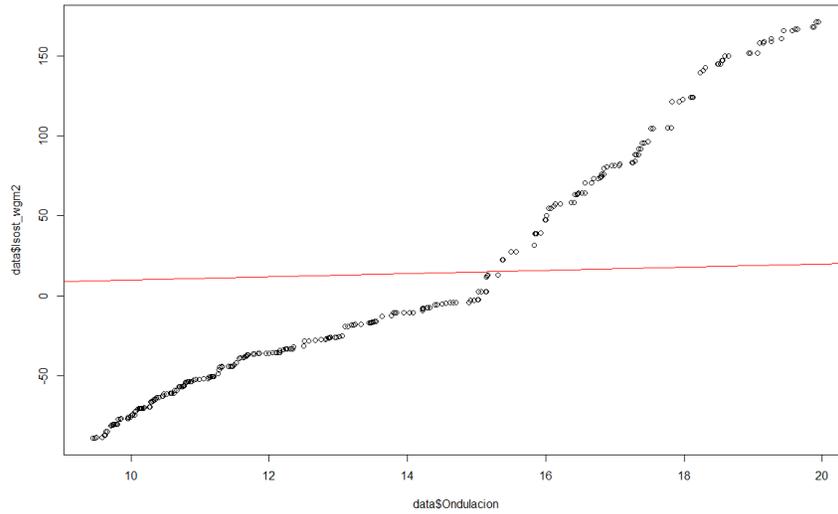
A continuación se presentan los resultados para las variables estudiadas:



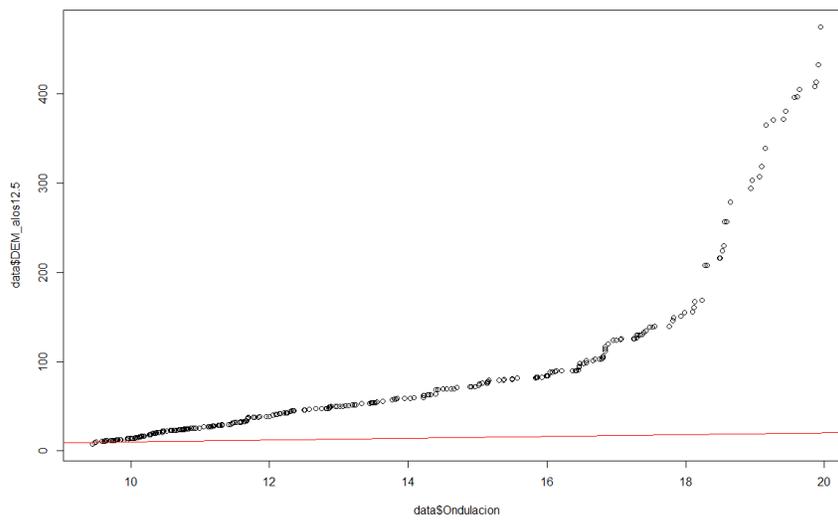
**Figura 22.** qq – plot entre Anomalía de Aire libre (2.5') y ondulación geoidal



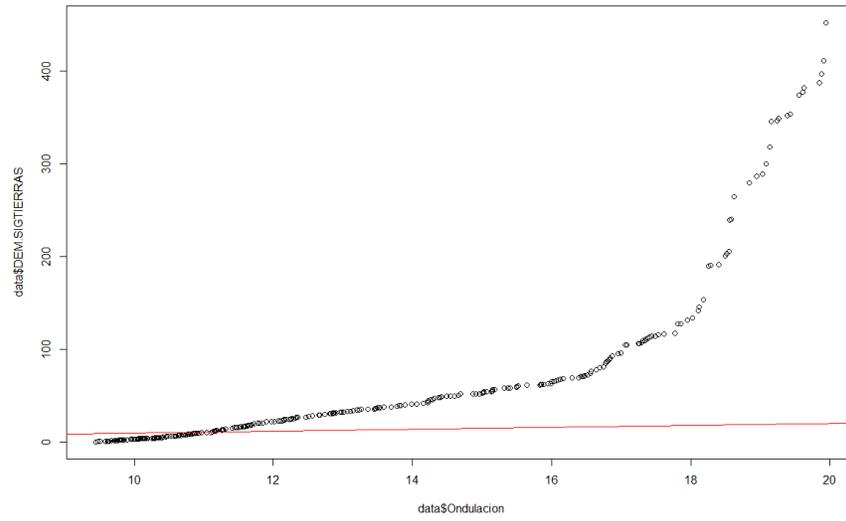
**Figura 23.** qq – plot entre Anomalía de Bouguer (2.5') y ondulación geoidal



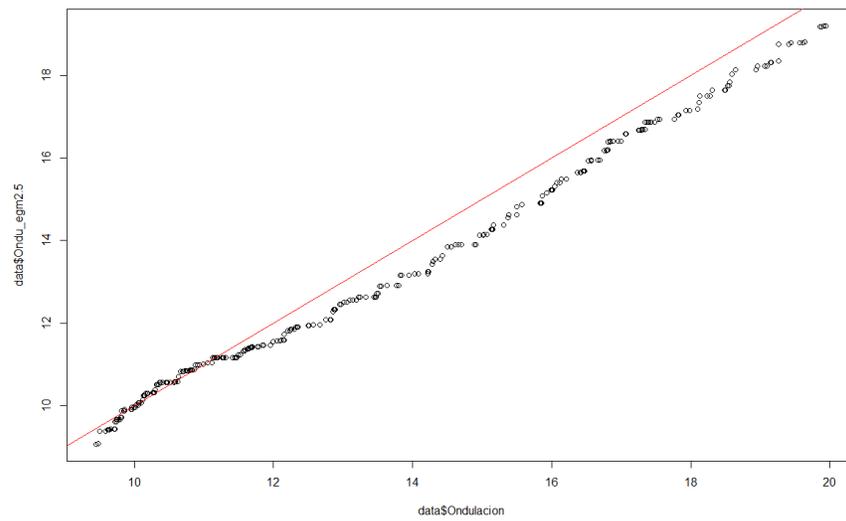
**Figura 24.** qq – plot entre Anomalía Isostática (2') y ondulación geoidal



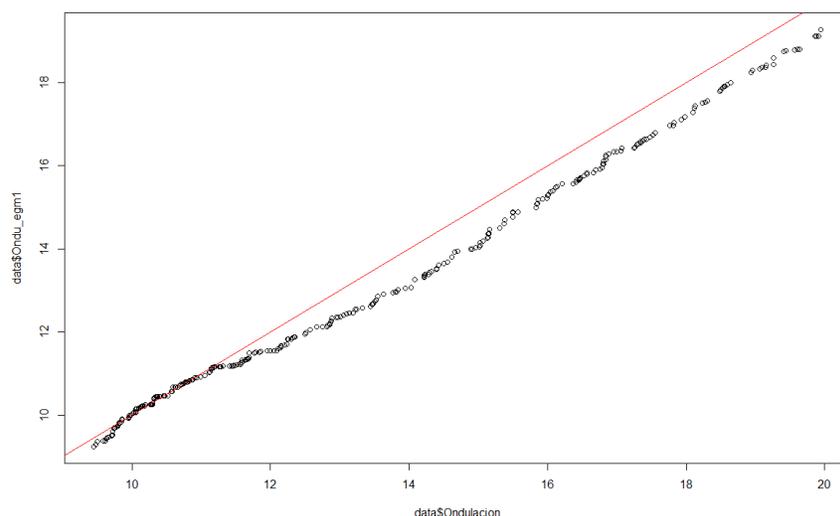
**Figura 25.** qq – plot entre DEM ALOS (12.5m) y ondulación geoidal



**Figura 26.** qq – plot entre DEM SIGTIERRAS (4m) y ondulación geoidal



**Figura 27.** qq – plot entre ondulación geoidal EGM08 (2.5') y ondulación geoidal



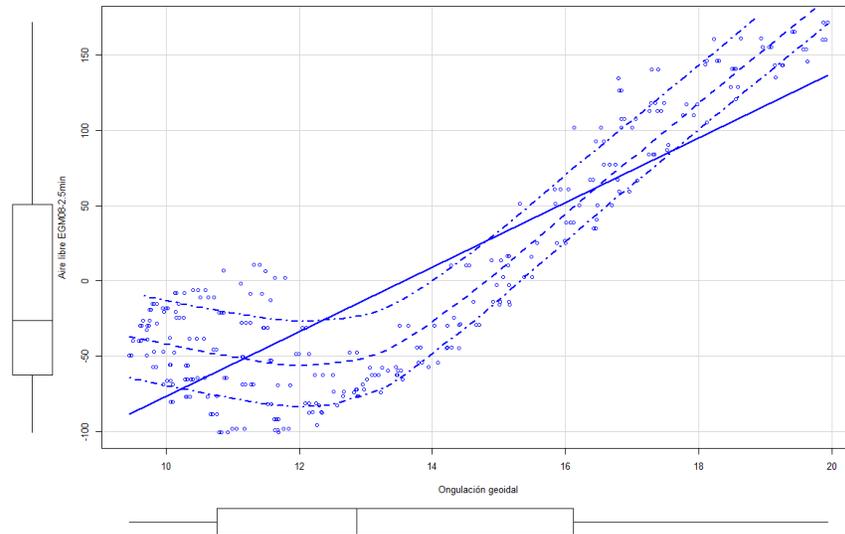
**Figura 28.** qq – plot entre ondulación geoidal EGM08 (1') y ondulación geoidal

De las figuras anteriores, se puede notar claramente, que la dependencia espacial es más fuerte con las variables ondulación geoidal del EGM08 en sus dos resoluciones (2.5'x2.5', 1'x1'), en comparación al resto de variables. En referencia a las anomalías de gravedad, las tres poseen una dependencia no muy clara con la variable objetivo, a pesar de ello, existe dependencia. Lo que llamó la atención es que tanto el DEM de ALOS como el de SIGTIERRAS, no describen ninguna relación con la ondulación geoidal.

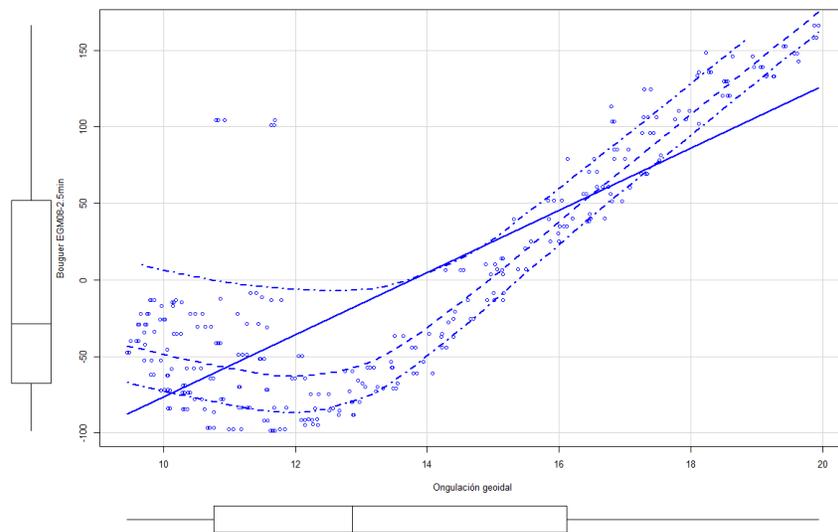
#### 4.3.2 Gráfico de dispersión

También conocido como scatter – plot, es una herramienta gráfica para evaluar la dependencia espacial entre dos variables a través de una regresión lineal. Se basa en comparar los valores de ambas variables medidos en los mismos lugares para generar una nube de puntos. Mientras la nube de puntos sea más estrecha, el grado de relación aumentará. El gráfico de dispersión resultante excluye de la nube de puntos a los errores de medición y valores atípicos, por lo que son fácilmente identificados (Volfová & Šmejkal, 2012).

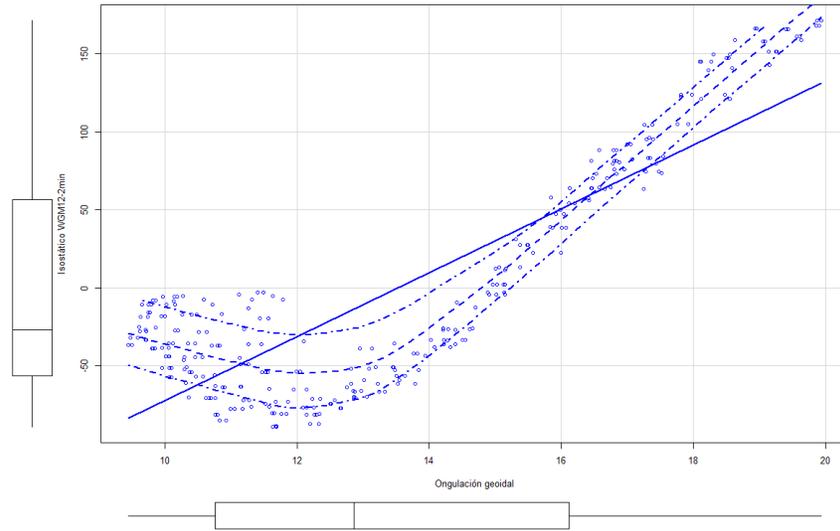
Los gráficos de dispersión resultantes se muestran en las Figuras 29, 30, 31, 32, 33, 34, 35:



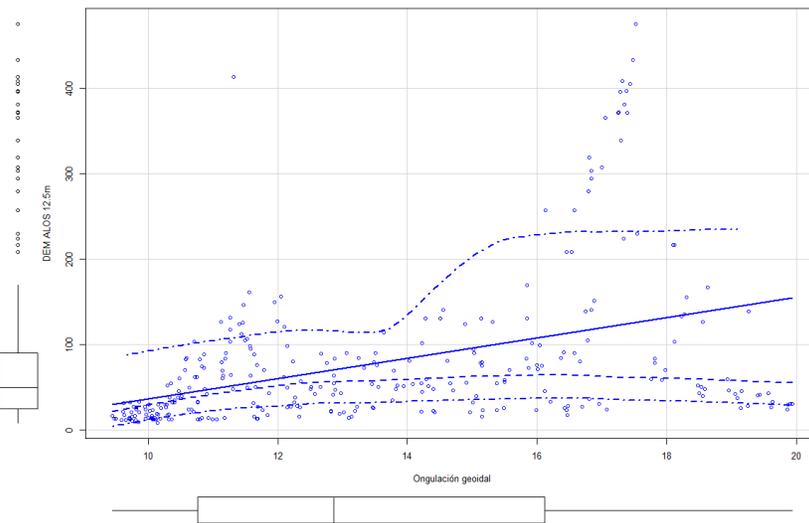
**Figura 29.** Scatter – plot entre Anomalía de Aire libre (2.5') y ondulación geoidal



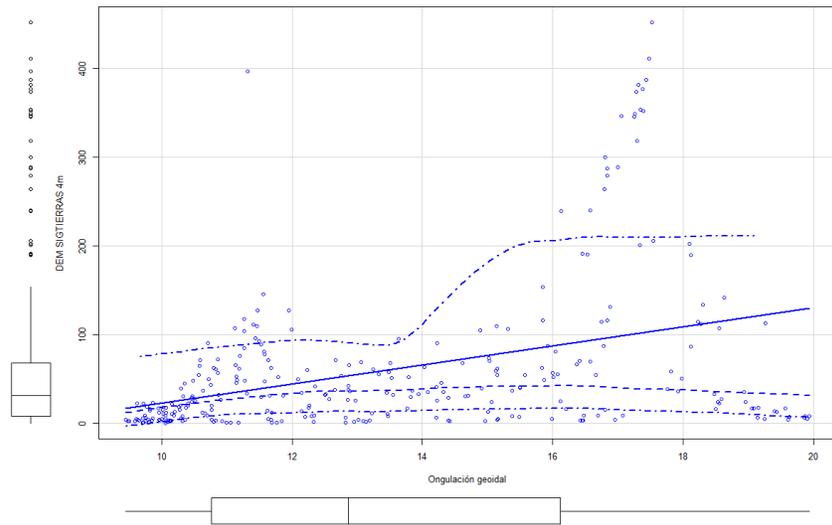
**Figura 30.** Scatter – plot entre Anomalía de Bouguer (2.5') y ondulación geoidal



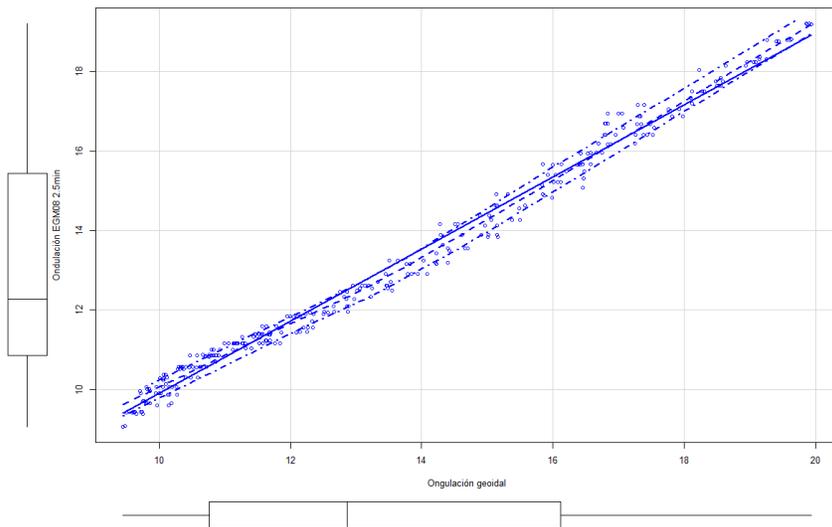
**Figura 31.** Scatter – plot entre Anomalía Isostática (2') y ondulación geoidal



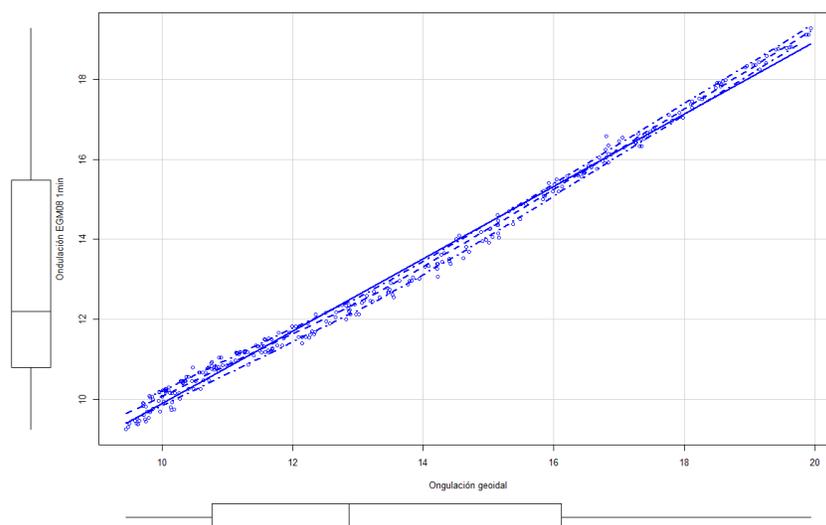
**Figura 32.** Scatter – plot entre DEM ALOS (12.5m) y ondulación geoidal



**Figura 33.** Scatter – plot entre DEM SIGTIERRAS (4m) y ondulación geoidal



**Figura 34.** Scatter – plot entre ondulación geoidal EGM08 (2.5') y ondulación geoidal



**Figura 35.** Scatter – plot entre ondulación geoidal EGM08 (1') y ondulación geoidal

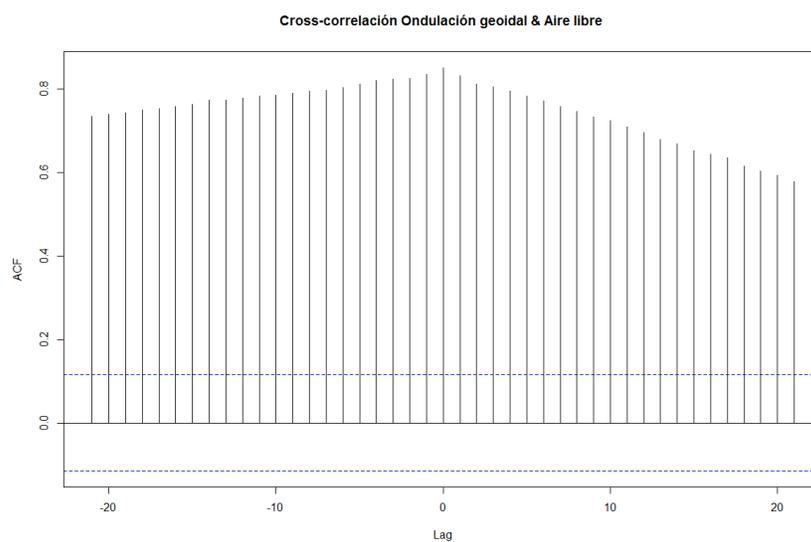
De los gráficos de dispersión ya se puede observar más claramente lo indicado con el gráfico qq – plot, es decir, que las variables que mejor dependencia espacial poseen con la variable objetivo (ondulación geoidal) son la ondulación derivada del modelo EGM08 en las dos resoluciones disponibles. Hay que enfatizar también, que con este último gráfico scatter – plot se observa una mejor afinidad con la ondulación del EGM08 a 1' de resolución espacial.

Por otro lado, se sigue corroborando que el DEM no representa dependencia espacial con la ondulación geoidal. Mientras que las anomalías de gravedad, si presentan dependencia, siendo la Isostática la que aparentemente posee mayor correlación que las otras dos anomalías.

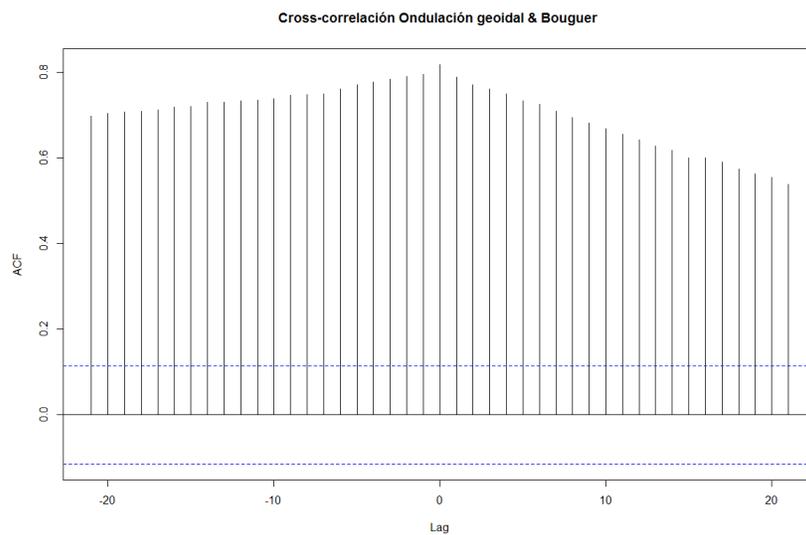
### 4.3.3 Gráfico de correlación espacial o correlograma

La autocorrelación o correlograma es una herramienta importante para el análisis de correlación e independencia de variables, supuestos que son necesarios para trabajar con más de un dato de entrada. Cuando se trabaja con dos variables se denomina cross – correlación o cross – correlograma (Venables & Ripley, 2002).

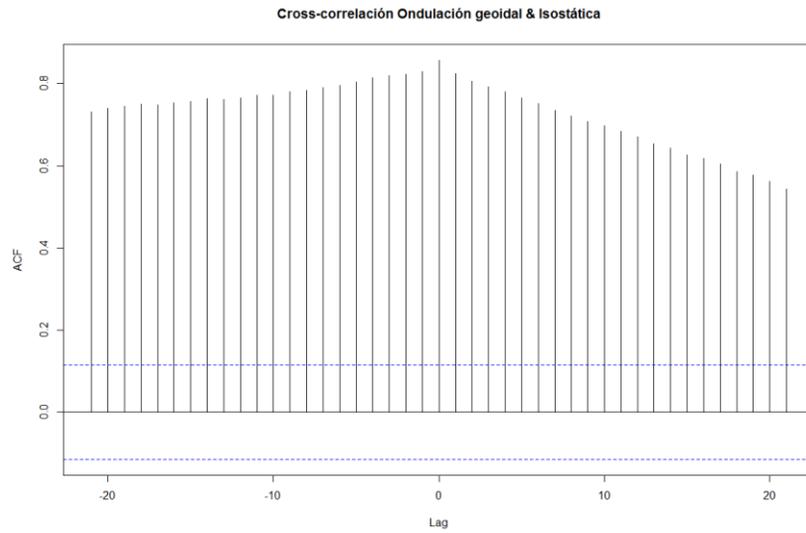
De igual forma que en los gráficos anteriores, se presenta a continuación la cross – correlación existente entre las variables auxiliares y la ondulación geoidal.



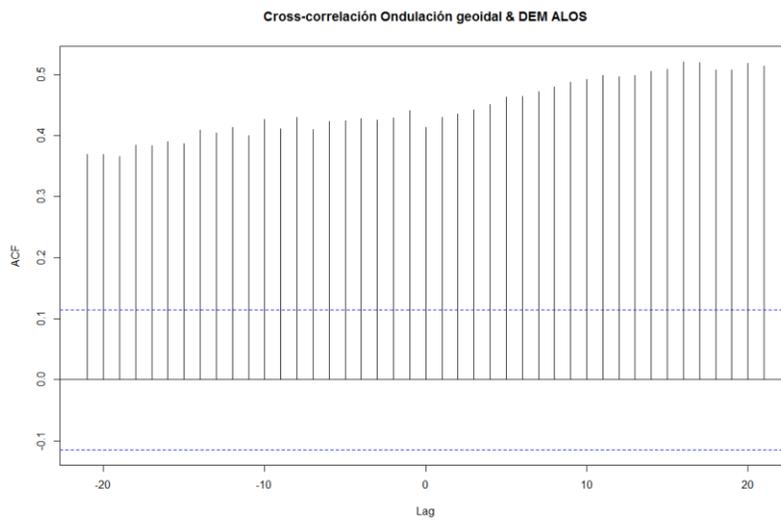
**Figura 36.** Cross – correlación entre Anomalía de Aire libre (2.5') y ondulación geoidal



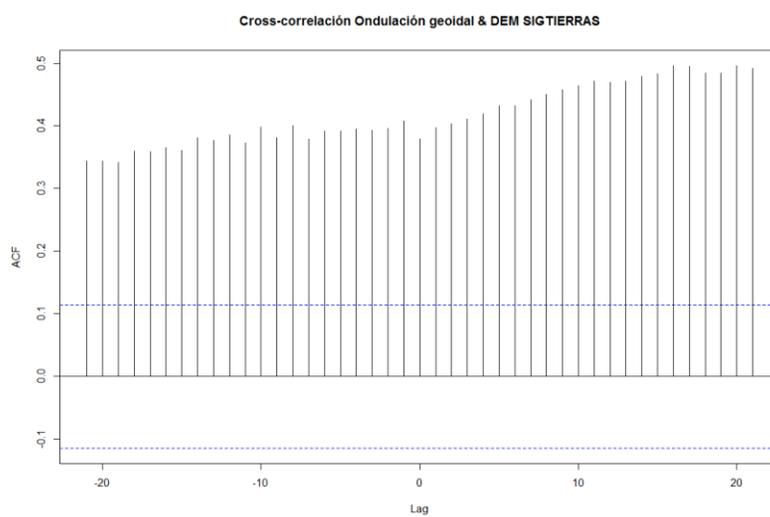
**Figura 37.** Cross – correlación entre Anomalía de Bouguer (2.5') y ondulación geoidal



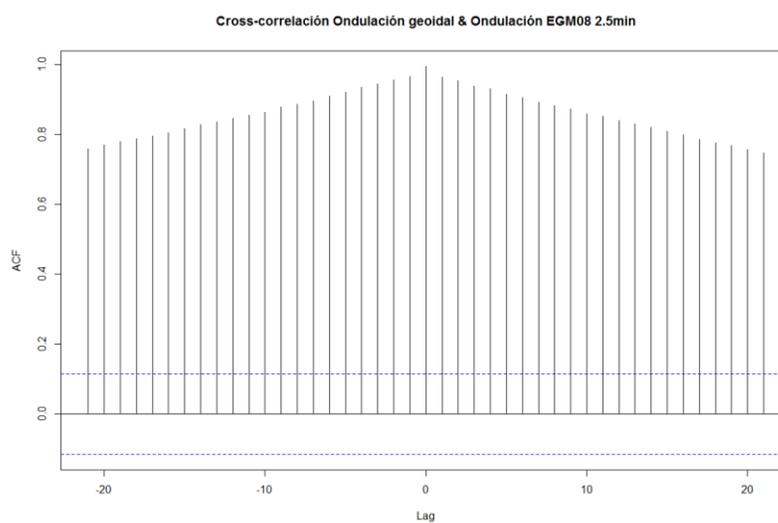
**Figura 38.** Cross – correlación entre Anomalía Isostática (2') y ondulación geoidal



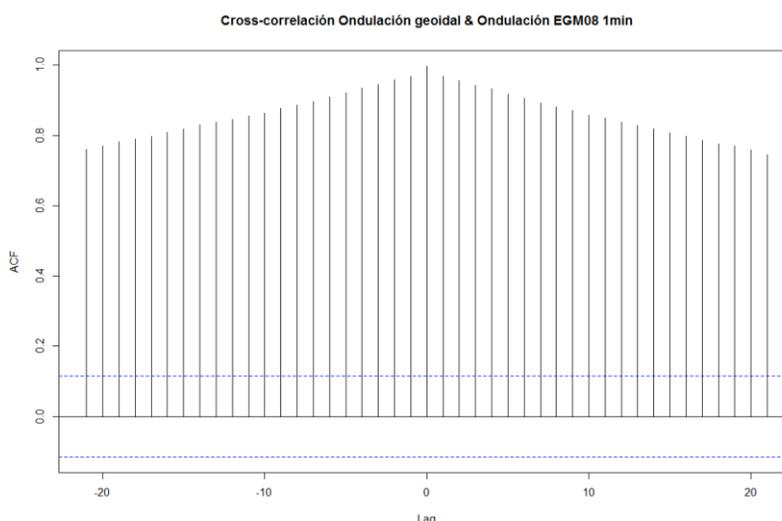
**Figura 39.** Cross – correlación entre DEM ALOS (12.5m) y ondulación geoidal



**Figura 40.** Cross – correlación entre DEM SIGTIERRAS (4m) y ondulación geoidal



**Figura 41.** Cross – correlación entre ondulación geoidal EGM08 (2.5') y ondulación geoidal



**Figura 42.** Cross – correlación entre ondulación geoidal EGM08 (1') y ondulación geoidal

Las gráficas de cross – correlación indican, en primera instancia, que en todos los casos, existe autocorrelación e independencia, ya los valores pasan el valor medio (franjaz azules) al 95% de confianza, pero lo más importante, es que observamos una tendencia de los datos, ya que no decrecen sino que su disminución es muy lenta, es decir es un proceso no estacionario (Venables & Ripley, 2002).

#### 4.3.4 Coeficiente de correlación de Pearson

El coeficiente de correlación generalmente representado como  $R^2$ , es una prueba estadística que utiliza el coeficiente de correlación del momento del producto de Pearson y sigue una distribución “t” con  $n - 2$  grados de libertad si son muestras independientes y distribuidas normalmente (Venables & Ripley, 2002).

En la Tabla 3 se muestran los resultados de  $R^2$  de la ondulación geoidal con las posibles variables auxiliares:

**Tabla 3.**  
*Coefficientes de correlación de Pearson*

Variable	$R^2$
Anomalía de gravedad de Aire libre (2.5'x2.5')	0.8501
Anomalía de gravedad de Bouguer (2.5'x2.5')	0.8187
Anomalía de gravedad Isostática (2'x2')	0.8536
DEM ALOS (12.5m)	0.4107
DEM SIGTIERRAS (4m)	0.3793
Ondulación geoidal derivada del EGM08 (2.5')	0.9940
Ondulación geoidal derivada del EGM08 (1')	0.9963

Como se puede apreciar, el valor de correlación indica que la ondulación geoidal derivada del modelo geopotencial EGM08, con una resolución espacial de 1 minuto es la variable que mayor dependencia espacial posee con la variable objetivo. Esto concuerda con las gráficas antes mostradas, señalando una superioridad con respecto a las demás posibles variables auxiliares.

En resumen, según los gráficos de dependencia espacial y al coeficiente de correlación de Pearson, la variable auxiliar seleccionada para formar parte del análisis multivariado mediante Cokriging es la ondulación geoidal derivada del EGM08 con resolución espacial de 1 min. A partir de ahora, las demás etapas del proceso geoestadístico se llevarán a cabo entre la variable objetivo y la variable auxiliar establecida.

#### **4.4 Análisis exploratorio de datos**

La variable objetivo de este estudio, como se ha venido mencionando en los anteriores capítulos, es la ondulación geoidal, esta variable física de la Tierra para ser obtenida requiere de un dispendioso trabajo de campo, lo que implica costos elevados, tiempo y personal necesario para su realización (generalmente mediante GPS de precisión y nivelación diferencial). Es por ello que Cokriging, al ser una herramienta geoestadística que permite disminuir los puntos muestrales para

este tipo de variables, como la ondulación geoidal, se procedió a separar la base de datos (332 puntos) en cuatro submuestras, que representan el 5%, 10%, 15% y 20% del total, con el fin de encontrar el número de datos óptimos necesarios para alcanzar mejores precisiones que su semejante univariable Kriging.

Adicionalmente, como se mencionó en el Capítulo III, sección 3.7 de Geoestadística univariada, se precedió a realizar la predicción con los tres tipos de Kriging, con el fin de comparar con los resultados del Cokriging según el tipo de predictor estudiado. El número de muestras tomadas para la realización de Kriging fue 299 puntos, es decir, que se separó el 10% del total de datos para validar el modelo final.

#### **4.4.1 Análisis exploratorio de Cokriging con el 5% de datos**

El número de datos que representan el 5% del total de los puntos muestrales, son 16 datos en los que constan las cantidades de ondulación geoidal real y ondulación geoidal derivada del EGM08 (variable auxiliar), los cuales fueron seleccionados aleatoriamente para reducir algún posible sesgo en la submuestra. Para abreviar los nombres de las variables, se designará como Ond-real a la ondulación geoidal medida por el IGM con nivelación geométrica de segundo orden, y Ond-egm, se referirá al valor de ondulación geoidal obtenido del modelo geopotencial global EGM08 a una resolución espacial de 1'.

Un resumen de los datos se presenta en la Tabla 4, mientras de los estadísticos descriptivos se muestran en la Tabla 5 a continuación:

**Tabla 4.***Resumen de los datos de las variables con Cokriging al 5% de datos*

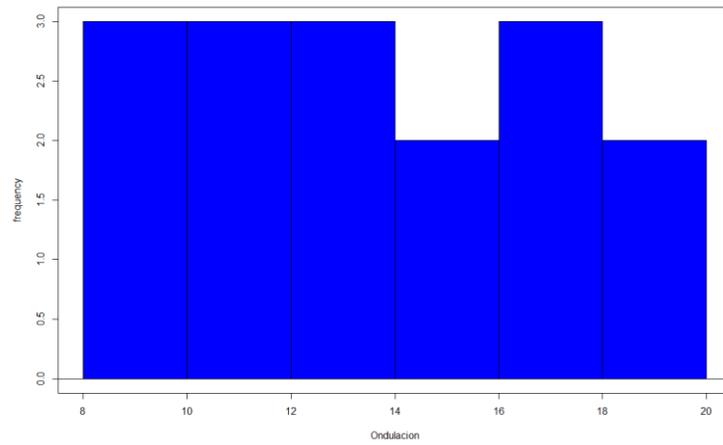
Variable	Min. (m)	1er cuartil (m)	Mediana (m)	3er cuartil (m)	Max. (m)
Ond-real	9.717	11.210	13.806	16.335	19.877
Ond-egm	9.61	10.98	12.96	15.72	19.11

**Tabla 5.***Estadísticos descriptivos de las variables con Cokriging al 5% de datos*

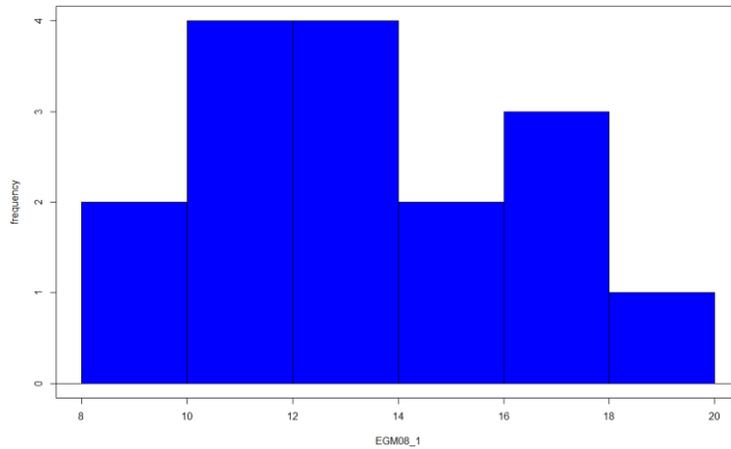
Variable	Media (m)	Desv. Est. (m)	Rango (m)	Coef. Varia.	Asimetría	Kurtosis	R <sup>2</sup>
Ond-real	13.837	3.233	5.125	0.234	0.301	-1.009	0.997
Ond-egm	13.364	2.975	4.736	0.223	0.461	-0.942	

Según los datos mostrados en la Tabla 5, el rango intercuartil no es muy grande en ambas variables; el coeficiente de variación es 23% y 22% lo que indica que no existe problema por variabilidad en los datos, e indica una leve carga (alargada), con valores mayor a la media; la asimetría (skewness) corresponde al tercer momento de probabilidad de la distribución y permite caracterizar a la misma, en este caso, los valores son positivos y la mediana es menor que la media de cada variable pero distintos de cero, por lo que la distribución de ambos datos no es muy simétrica; por otro lado el Kurtosis, que representa el cuarto momento de distribución de probabilidad de la distribución, proporciona información sobre la forma de esta, y según los resultados la forma de la distribución es más plana ya que presenta valores menores a 3 y negativos.

El histograma de distribución de los datos de ambas variables confirma lo mencionado en el anterior párrafo:

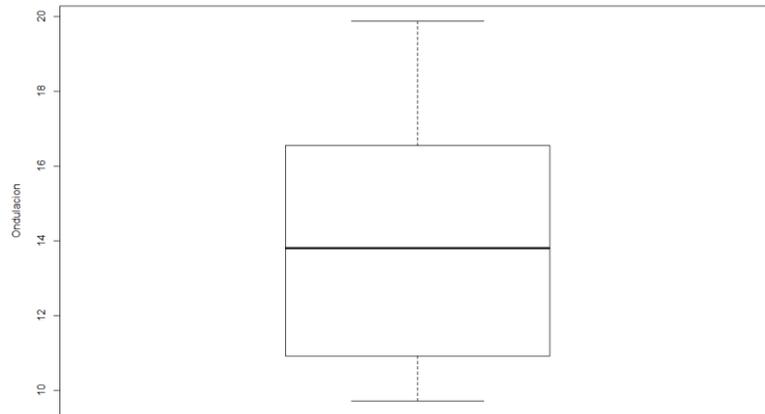


**Figura 43.** Histograma de Ondulación geoidal real

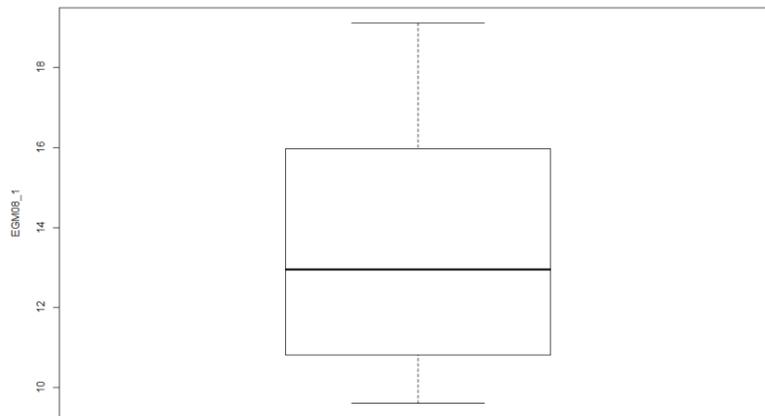


**Figura 44.** Histograma de Ondulación geoidal EGM08 (1')

También se realizó un diagrama de caja para corroborar la inexistencia de valores extremos, como se muestra a continuación:



**Figura 45.** Diagrama de caja de Ondulación geoidal real



**Figura 46.** Diagrama de caja de Ondulación geoidal EGM08 (1')

Un parámetro a tomar en cuenta es la normalidad de la distribución de los datos, por lo que se procedió a utilizar otras pruebas estadísticas para verificar lo expuesto, como el test Shapiro – Wilk y Kolmogorov – Smirnov.

Shapiro – Wilk se utiliza más para muestras pequeñas (<50), aunque su uso también ha sido probado para muestras de 3 – 5000 datos como máximo (Royston, 1995); para obtener este test se calcula la varianza muestral y su media, posteriormente se ordenan los datos y se obtienen las diferencias corregidas (primero menos último, segundo menos penúltimo, etc.).

Kolmogorov – Smirnov es otro test usado para verificar la normalidad y puede ser utilizado para cualquier tamaño de muestra, desde 4 datos en adelante. Esta prueba es independiente de la distribución de la población en cuyos casos se desconoce la media y la varianza, aplicando una corrección de Lilliefors. Se obtiene restando las funciones de distribución de la muestra y la función teórica de una población normal (Thode, 2002).

Los resultados de la prueba de normalidad se muestran en la Tabla 6:

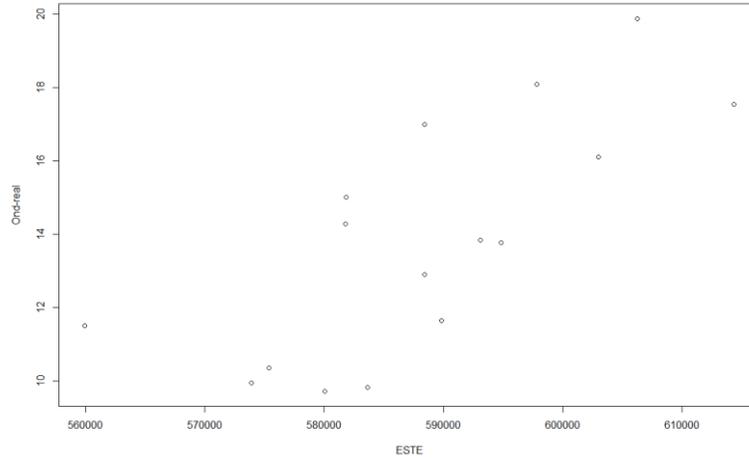
**Tabla 6.**

*Pruebas de normalidad de las variables con Cokriging al 5% de datos*

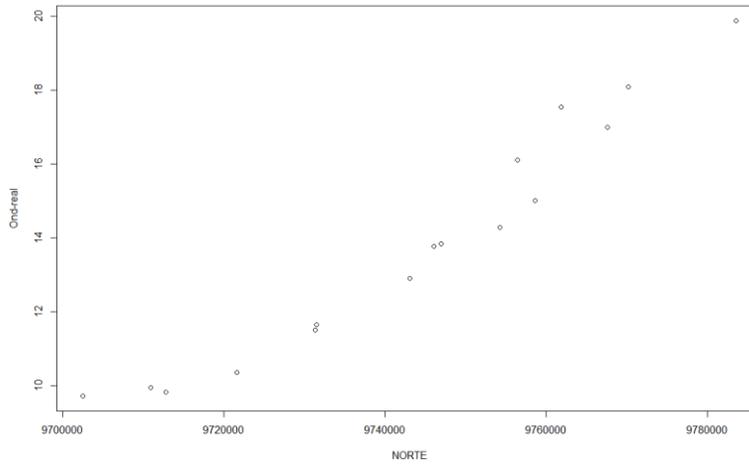
Variable	Estadístico W Shapiro-Wilk	Estadístico D Kolmogorov- Smirnov	p-value Shapiro-Wilk	p-value Kolmogorov- Smirnov
Ond-real	0.943	0.126	0.383	0.702
Ond-egm	0.937	0.139	0.317	0.558

Según los valores del p-value, el cual debe ser mayor a 0.05 para asegurar una distribución normal (Ahad *et al.*, 2011; Nornadiah *et al.*, 2011; Nornadiah *et al.*, 2012; Yap & Sim, 2011), se acepta que estas distribuciones son normales, a pesar de no ser muy notorio en los histogramas, esto último se puede asumir debido a la pequeña cantidad de puntos usados para las pruebas de normalidad.

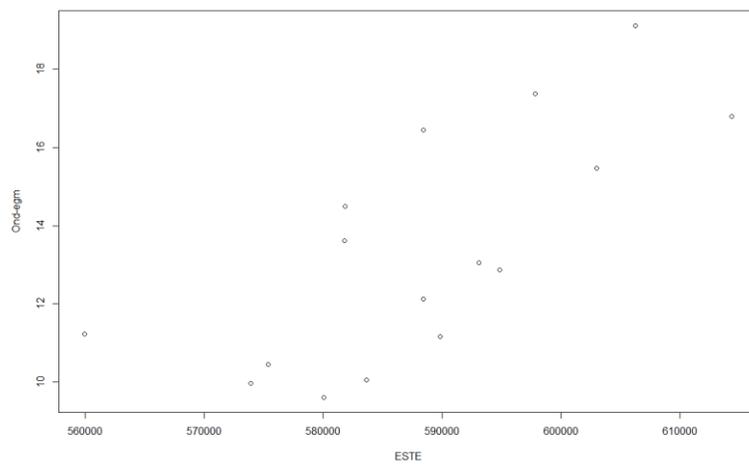
También es importante conocer si los datos son estacionarios o no, para lo cual se representa en un gráfico de dispersión en función de la locación geográfica de las muestras, como se observa en las Figuras 47, 48, 49, 50.



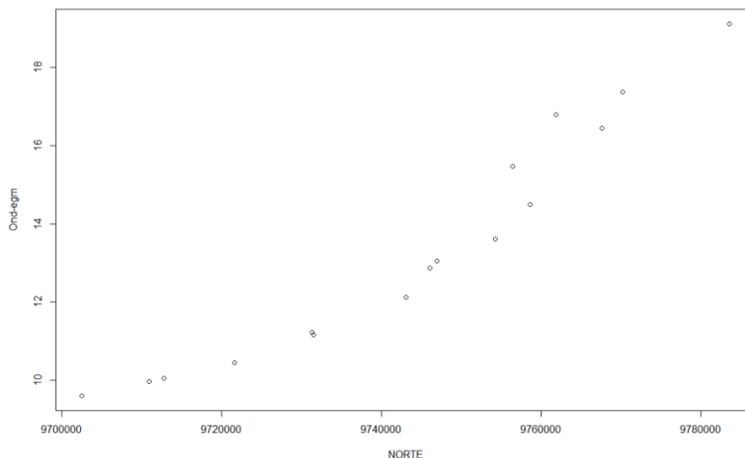
**Figura 47.** Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real



**Figura 48.** Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real



**Figura 49.** Gráfico de dispersión de la coordenada Este vs Ondulación geoidal EGM08 (1')



**Figura 50.** Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal EGM08 (1')

De las figuras anteriores, se puede notar claramente que la estacionariedad no está presente en el área geográfica en ambas variables, por lo que se puede inferir algún tipo de tendencia existente, principalmente en la dirección de la coordenada Norte, y una media no constante en el área de estudio, es decir, es un escenario de hipótesis intrínseca.

#### 4.4.2 Análisis exploratorio de Cokriging con el 10% de datos

Los datos correspondiente al 10% el total de muestras es de 33 puntos, con los cuales se realizó los mismos pasos del análisis estructural mostrado en 4.4.1. En las Tablas 7 y 8 se muestra el resumen de los datos y sus estadísticos principales.

**Tabla 7.**

*Resumen de los datos de las variables con Cokriging al 10% de datos*

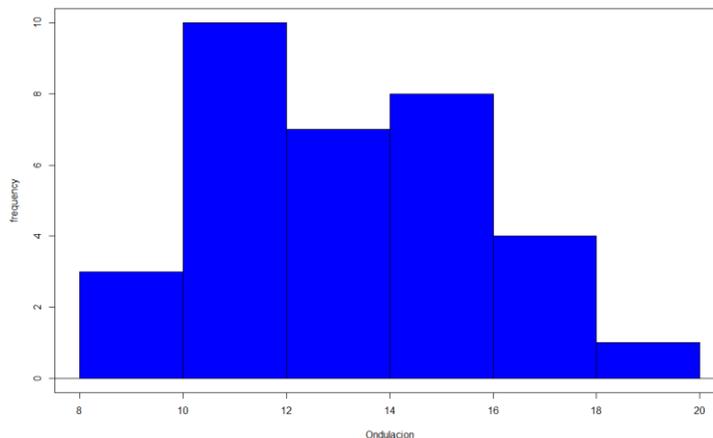
Variable	Min. (m)	1er cuartil (m)	Mediana (m)	3er cuartil (m)	Max. (m)
Ond-real	9.437	10.592	13.226	15.149	18.371
Ond-egm	9.298	10.464	12.649	14.152	17.600

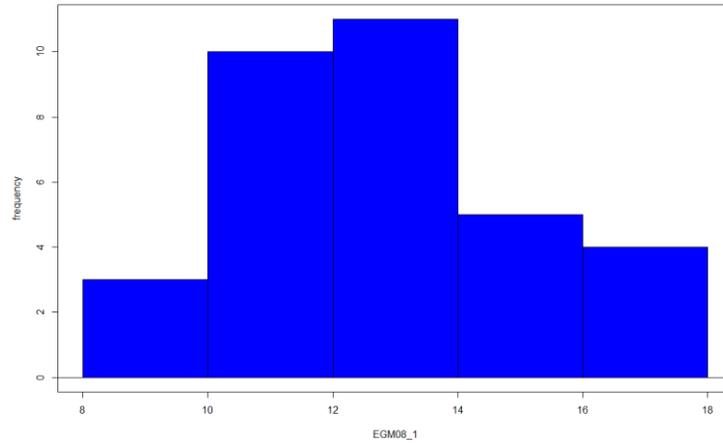
**Tabla 8.***Estadísticos descriptivos de las variables con Cokriging al 10% de datos*

Variable	Media (m)	Desv. Est. (m)	Rango (m)	Coef. Varia.	Asimetría	Kurtosis	R <sup>2</sup>
Ond-real	13.194	2.682	4.557	0.203	0.254	-1.097	0.995
Ond-egm	12.716	2.348	3.688	0.184	0.446	-0.752	

De lo observado en las tablas anteriores, se menciona que los rangos no son muy amplios al igual que en el caso de 16 puntos; los coeficientes de variación son menores a 100% (20.3% y 18.4%) por lo que no hay problema de variabilidad en los datos; la asimetría muestra que no son distribuciones simétricas en ambas variables; Kurtosis indica que la forma de la distribución es plana.

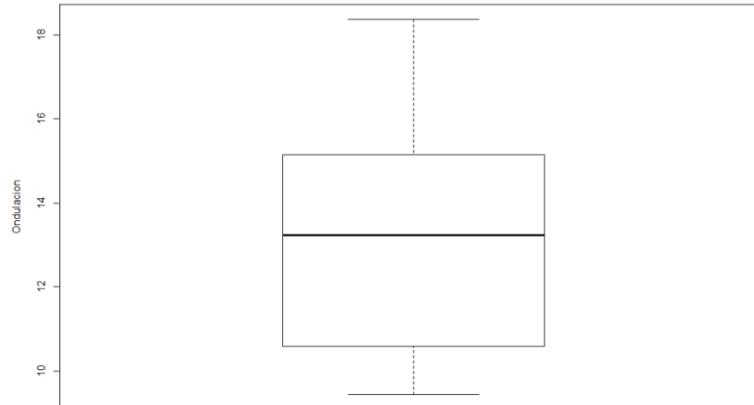
Los histogramas para la submuestra de 33 puntos son los siguientes:

**Figura 51.** Histograma de Ondulación geoidal real

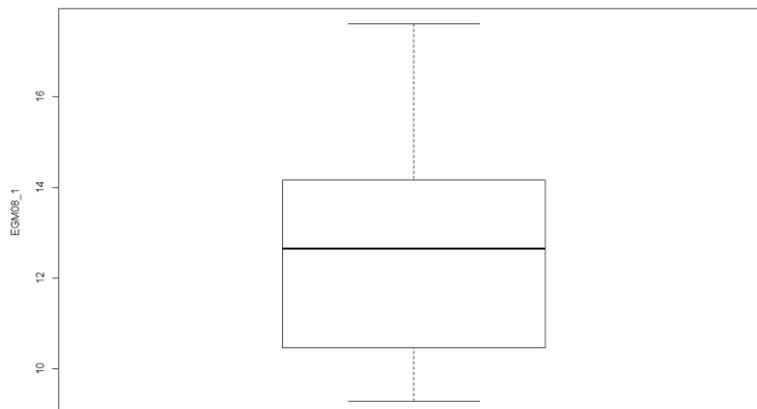


**Figura 52.** Histograma de Ondulación geoidal EGM08 (1')

Se puede observar en ambos histogramas, que la distribución está levemente cargada con valores mayores a sus respectivas medias. Los diagramas de caja confirman que ningún valor es extremo en los puntos utilizados.



**Figura 53.** Diagrama de caja de Ondulación geoidal real



**Figura 54.** Diagrama de caja de Ondulación geoidal EGM08 (1')

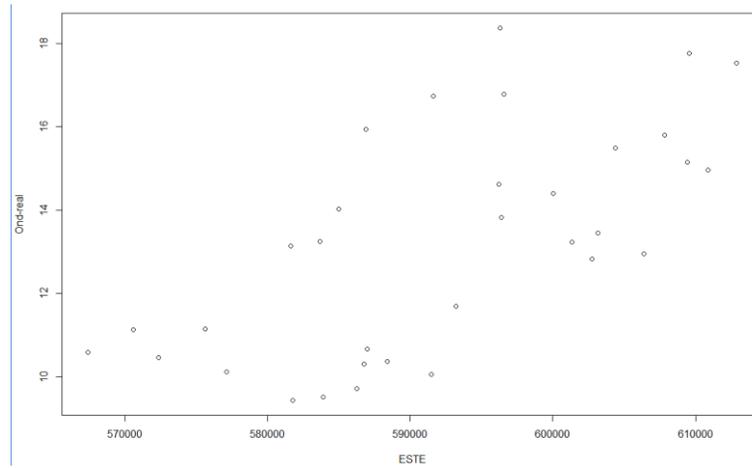
Del mismo modo, se realizaron las pruebas de normalidad para este conjunto de datos, resultando una distribución normal de los datos para las dos variables, cuyos valores se presenta en la Tabla 9:

**Tabla 9.**

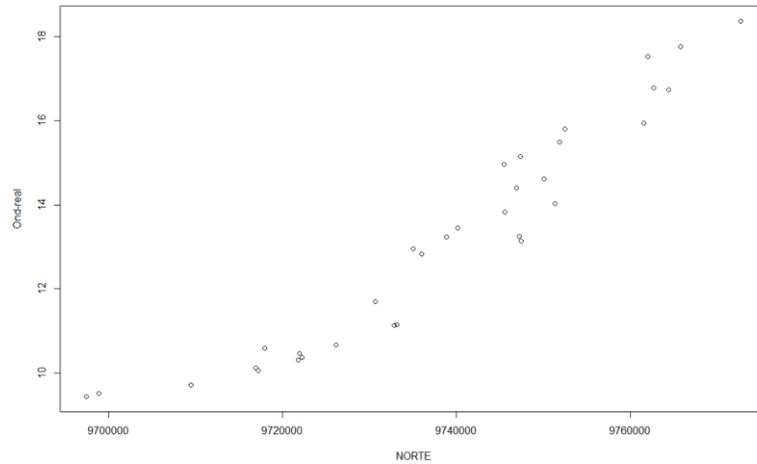
*Pruebas de normalidad de las variables con Cokriging al 10% de datos*

Variable	Estadístico W Shapiro-Wilk	Estadístico D Kolmogorov- Smirnov	p-value Shapiro-Wilk	p-value Kolmogorov- Smirnov
Ond-real	0.940	0.141	0.071	0.094
Ond-egm	0.947	0.123	0.113	0.225

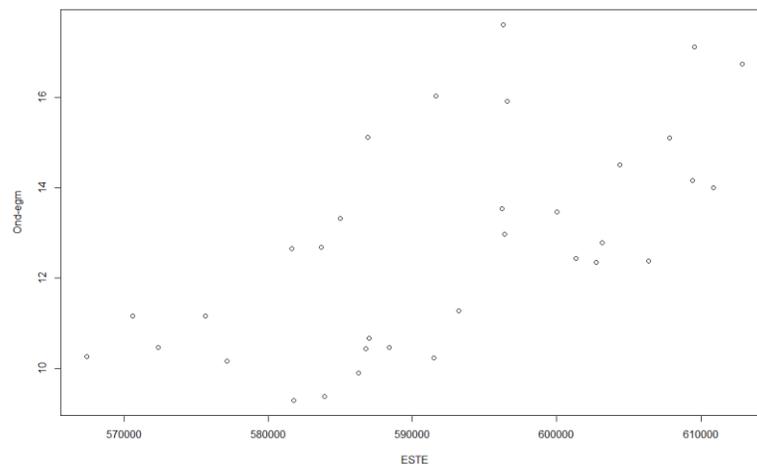
De igual forma, es necesario determinar si el proceso es estacionario o presenta no estacionariedad en la región, por lo que se realizaron los gráficos de dispersión en base a las coordenadas geográficas de los puntos muestreados, como se aprecia a continuación:



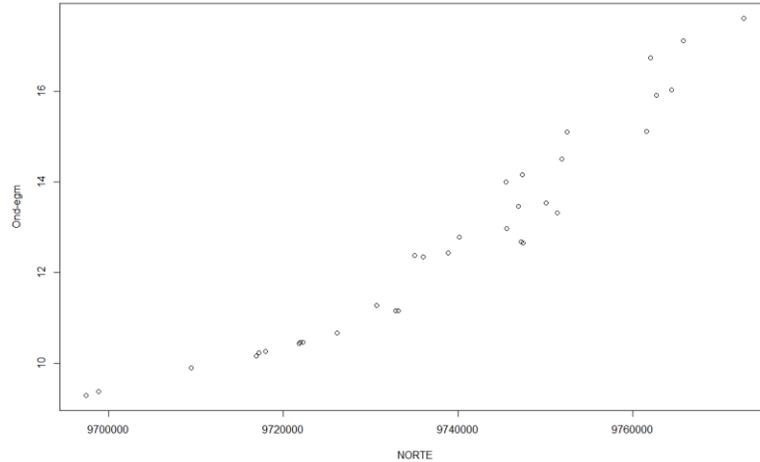
**Figura 55.** Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real



**Figura 56.** Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real



**Figura 57.** Gráfico de dispersión de la coordenada Este vs Ondulación geoidal EGM08 (1')



**Figura 58.** Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal EGM08 (1')

Al igual que en los gráficos de dispersión con 16 puntos, se puede observar que el proceso no es estacionario, más notorio en la dirección de la coordenada Norte, por lo que se dice que es un proceso no estacionario, relacionado con la hipótesis intrínseca.

#### 4.4.3 Análisis exploratorio de Cokriging con el 15% de datos

El desarrollo es similar a los dos anteriores, siendo en este caso evaluado para 49 puntos muestrales que corresponden al 15% del total de datos. En primero instancia se obtienen los cuadros resumen y estadísticos de las variables, los mismos que se observan a continuación en las Tablas 10 y 11:

**Tabla 10.**

*Resumen de los datos de las variables con Cokriging al 15% de datos*

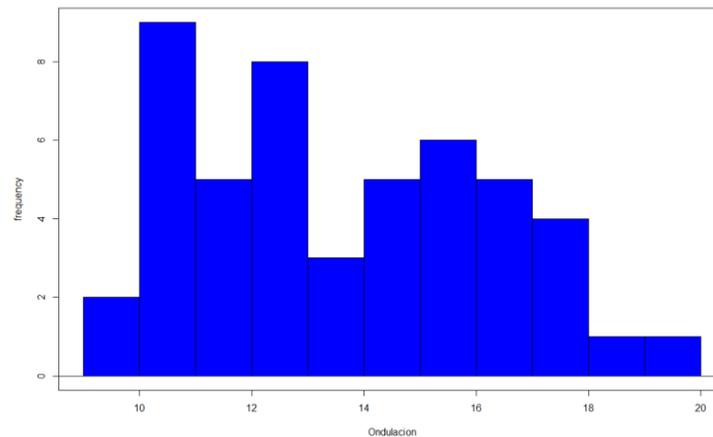
Variable	Min. (m)	1er cuartil (m)	Mediana (m)	3er cuartil (m)	Max. (m)
Ond-real	9.484	11.267	13.486	15.383	19.877
Ond-egm	9.298	10.976	12.745	14.386	19.112

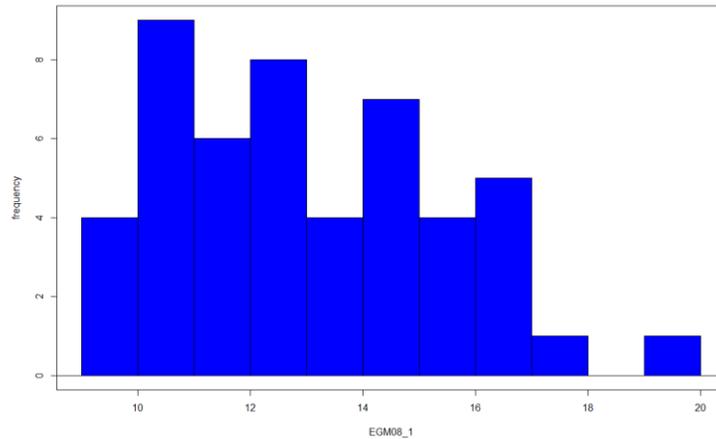
**Tabla 11.***Estadísticos descriptivos de las variables con Cokriging al 15% de datos*

Variable	Media (m)	Desv. Est. (m)	Rango (m)	Coef. Varia.	Asimetría	Kurtosis	R <sup>2</sup>
Ond-real	13.577	2.647	4.116	0.194	0.304	-0.807	0.996
Ond-egm	13.079	2.402	3.41	0.183	0.467	-0.564	

De estas tablas se puede decir lo siguiente, el rango intercuartil es pequeño al igual que los dos casos anteriores; su mediana es menor a la media en ambas variables; el coeficiente de variación es menor a 100% por lo que se dice que no existen gran variabilidad; la asimetría muestra que no son distribuciones simétricas y Kurtosis al igual que los anteriores casos, indica que la distribución es aplanada.

Los histogramas para el conjunto de datos son los siguientes:

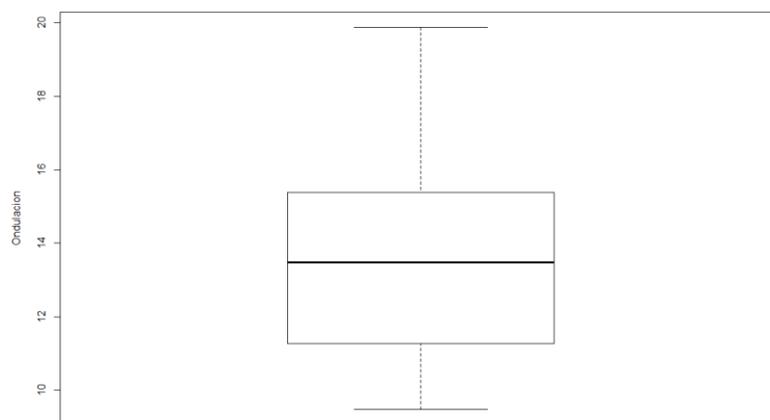
**Figura 59.** Histograma de Ondulación geoidal real



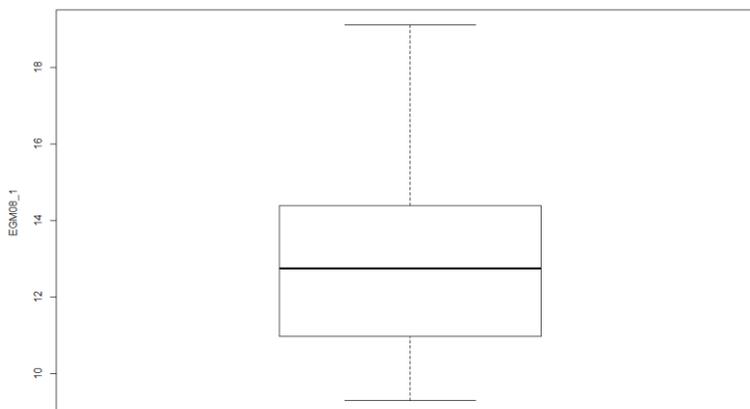
**Figura 60.** Histograma de Ondulación geoidal EGM08 (1')

De los histogramas mostrados, podemos constatar una distribución un tanto sesgada con valores superiores a la media concordante con la asimetría y Kurtosis, además de casi mostrar una aparente forma de normalidad.

A continuación se presentan las figuras de los diagramas de caja para las variables, con los que se verificó que no hay problemas de valores anómalos o extremos, reflejando los resultados del coeficiente de variación.



**Figura 61.** Diagrama de caja de Ondulación geoidal real



**Figura 62.** Diagrama de caja de Ondulación geoidal EGM08 (1')

En la Tabla 12 se ve reflejado los valores para las pruebas de normalidad efectuadas en ambas variables

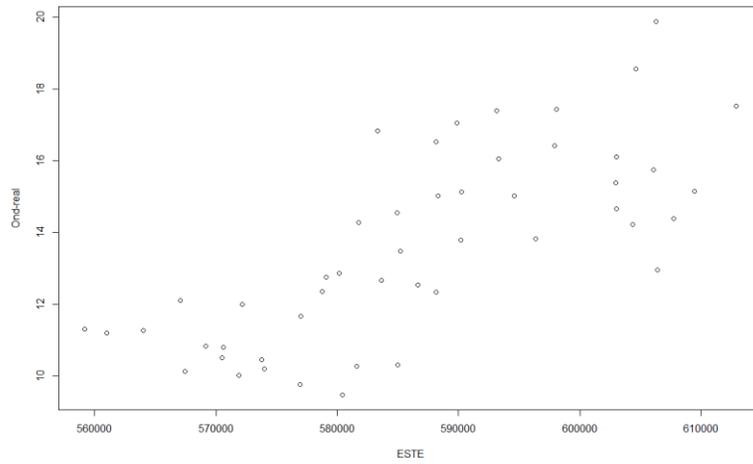
**Tabla 12.**

*Pruebas de normalidad de las variables con Cokriging al 15% de datos*

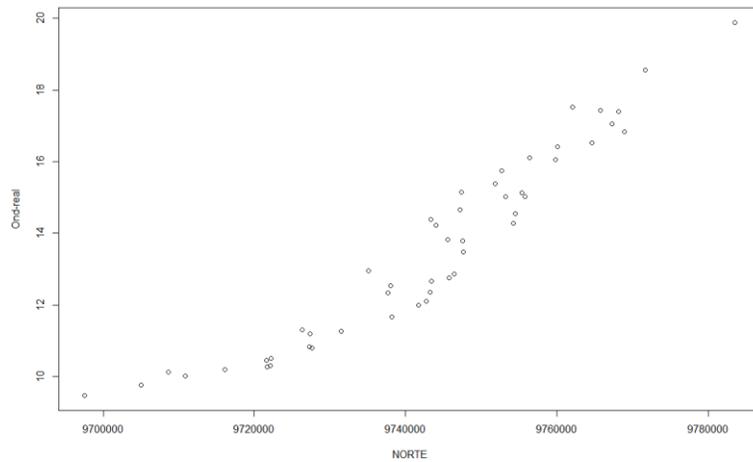
Variable	Estadístico W Shapiro-Wilk	Estadístico D Kolmogorov- Smirnov	p-value Shapiro-Wilk	p-value Kolmogorov- Smirnov
Ond-real	0.961	0.089	0.112	0.423
Ond-egm	0.961	0.105	0.104	0.191

En base a los resultados de normalidad efectuados para los 49 datos de la muestra, se dice que la distribución de estos es normal, lo que concuerda con la forma de los respectivos histogramas.

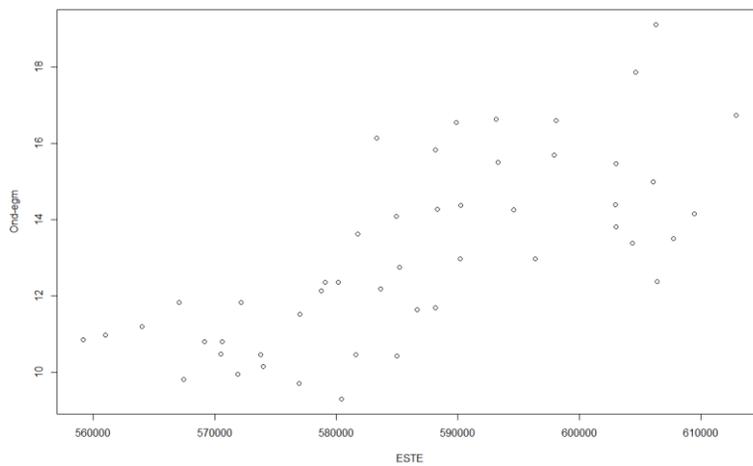
Además, es necesario constatar si el proceso se mantiene no estacionario, es decir si su media varía en la región es estudio, y a pesar de que en los dos casos anteriores se mantiene esta tendencia, se debe determinar si con el aumento del número de puntos existe algún cambio o persiste el fenómeno como intrínseco, para lo cual se realizó las gráficas de dispersión según su posición geográfica, como se muestra a continuación:



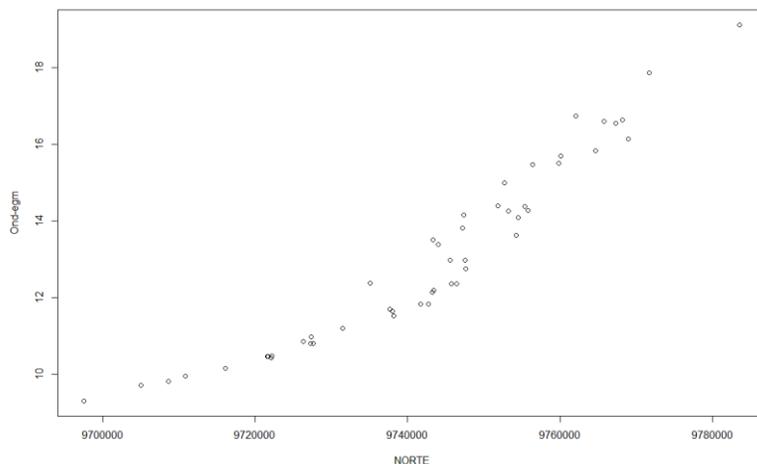
**Figura 63.** Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real



**Figura 64.** Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real



**Figura 65.** Gráfico de dispersión de la coordenada Este vs Ondulación geoidal EGM08 (1')



**Figura 66.** Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal EGM08 (1')

El observar las gráficas de dispersión se puede notar que el fenómeno continúa variando en el espacio, donde existe un mayor cambio en la componente Norte, manteniendo las características de media no constante y por ende, hipótesis intrínseca.

#### 4.4.4 Análisis exploratorio de Cokriging con el 20% de datos

La última submuestra seleccionada corresponde a 66 puntos (20% del total de datos), para lo cual se presentan las Tablas 13 y 14 de resumen y estadísticos respectivamente, para esta submuestra:

**Tabla 13.**

*Resumen de los datos de las variables con Cokriging al 20% de datos*

Variable	Min. (m)	1er cuartil (m)	Mediana (m)	3er cuartil (m)	Max. (m)
Ond-real	9.451	10.835	12.742	16.100	19.612
Ond-egm	9.252	10.803	12.202	15.310	18.803

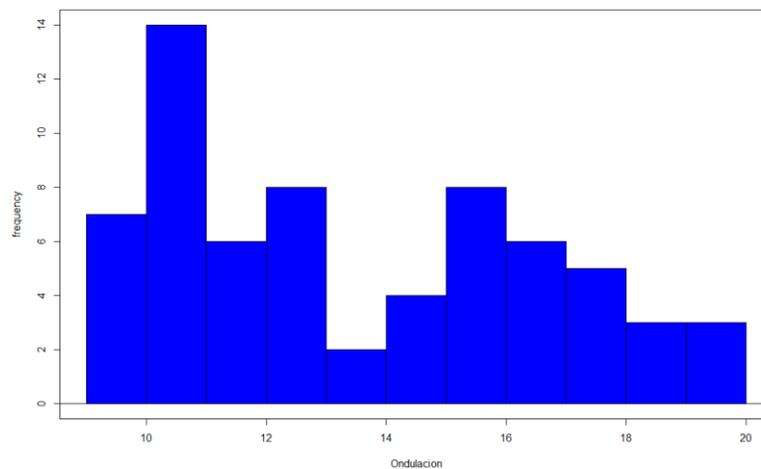
**Tabla 14.**

*Estadísticos descriptivos de las variables con Cokriging al 20% de datos*

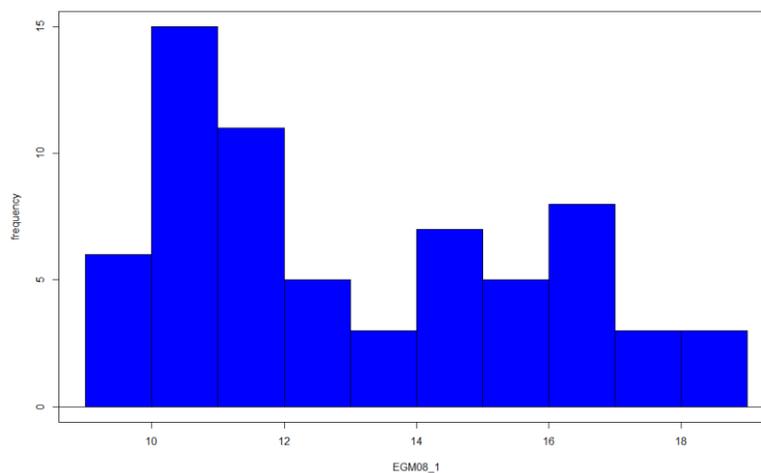
Variable	Media (m)	Desv. Est. (m)	Rango (m)	Coef. Varia.	Asimetría	Kurtosis	R <sup>2</sup>
Ond-real	13.506	3.062	5.264	0.226	0.353	-1.223	0.997
Ond-egm	13.072	2.772	4.506	0.212	0.463	-1.059	

Según los valores de los estadísticos se puede interpretar lo siguiente: el rango intercuartil al igual que los otros casos es pequeño; el coeficiente de variación es 22.6% y 21.2% por lo que se dice que no hay problema de variabilidad en los datos; su media es mayor a la mediana en ambos casos; la asimetría de la distribución al ser positiva pero distinta de cero se infiere que no es una distribución simétrica; los valores de Kurtosis sugieren que la forma de la distribución no presenta picos, sino es de forma plana.

Los histogramas para las dos muestras son los siguientes:



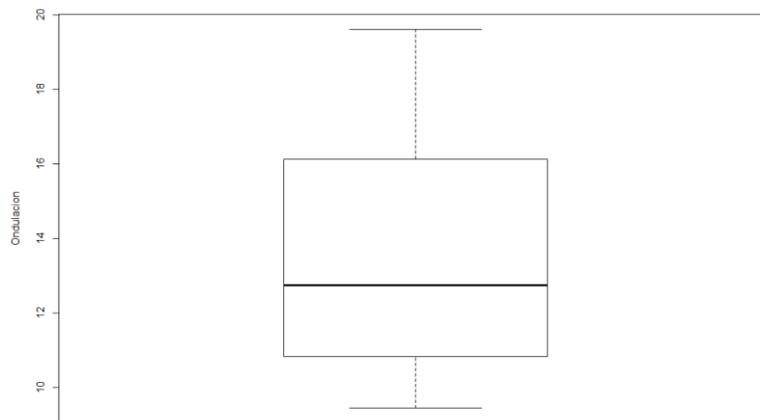
**Figura 67.** Histograma de Ondulación geoidal real



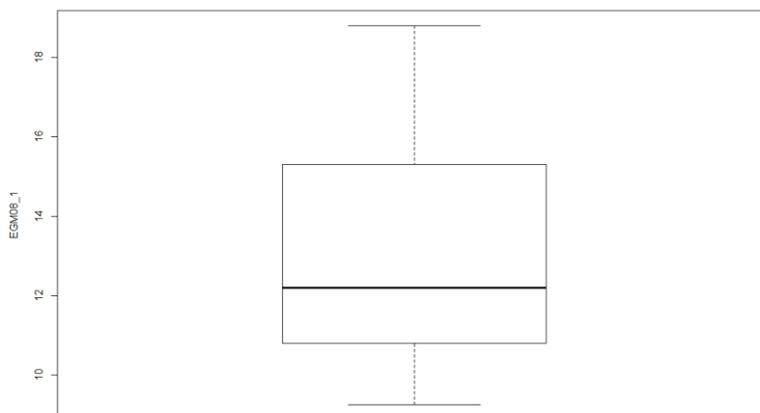
**Figura 68.** Histograma de Ondulación geoidal EGM08 (1')

Los histogramas comprueban que la distribución no es simétrica, y se nota que los datos están más concentrados por encima del valor medio en ambas variables.

También se comprobó la inexistencia de valores extremos en las muestras a través de los diagrama de caja de las Figuras 69, 70 respectivamente.



**Figura 69.** Diagrama de caja de Ondulación geoidal real



**Figura 70.** Diagrama de caja de Ondulación geoidal EGM08 (1')

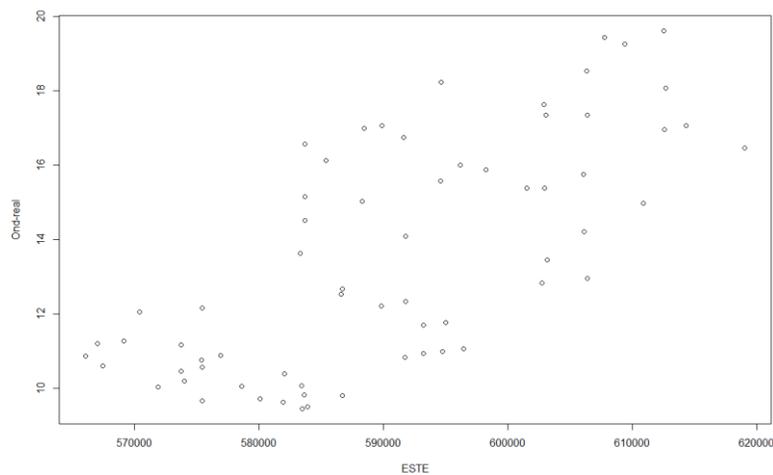
Para comprobar si la distribución de los datos sigue un comportamiento normal, se aplicó las pruebas de normalidad al igual que en los casos anteriores, como se observa en la Tabla 15 a continuación:

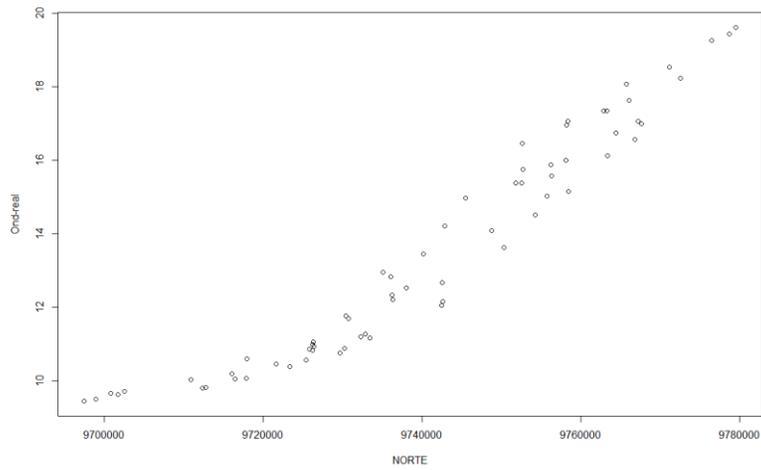
**Tabla 15.***Pruebas de normalidad de las variables con Cokriging al 20% de datos*

Variable	Estadístico W Shapiro-Wilk	Estadístico D Kolmogorov- Smirnov	p-value Shapiro-Wilk	p-value Kolmogorov- Smirnov
Ond-real	0.919	0.146	0.0003	0.001
Ond-egm	0.920	0.153	0.0004	0.0005

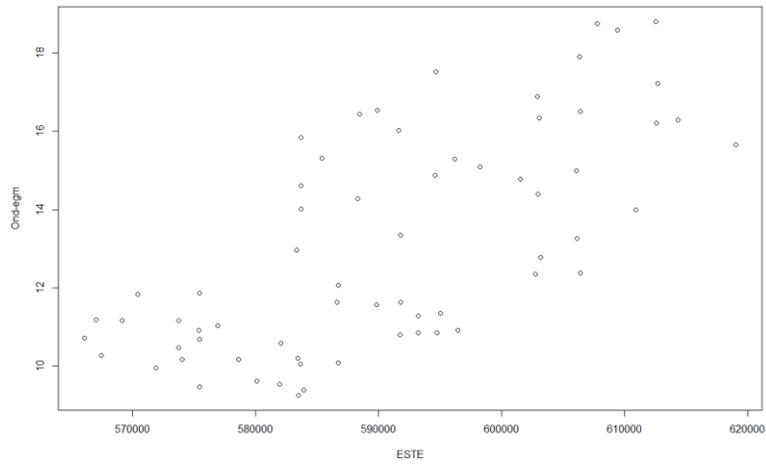
Al contrario de los casos anteriores, con 66 puntos muestrales el comportamiento de la distribución de los datos no es normal, esto hace suponer que a medida que se incrementa el número de puntos, existe más variabilidad en la distribución, esto también se ve reflejado en los valores de asimetría y Kurtosis, los cuales son los más altos en comparación con los demás conjuntos de datos al 5%, 10% y 15%.

Finalmente, se analizó el comportamiento estacionario de las variables en la región mediante los gráficos de dispersión en función de las coordenadas geográficas como se muestran a continuación:

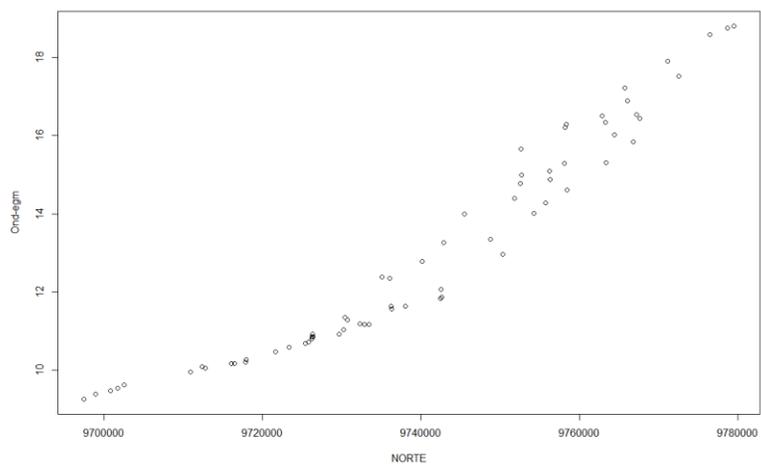
**Figura 71.** Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real



**Figura 72.** Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real



**Figura 73.** Gráfico de dispersión de la coordenada Este vs Ondulación geoidal EGM08 (1')



**Figura 74.** Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal EGM08 (1')

Con los gráficos de dispersión se confirma que el proceso estudiado es no estacionario, por lo que se trata de un caso de hipótesis intrínseca.

#### 4.4.5 Análisis exploratorio del total de datos

Para comprobar si el incremento de puntos hace que la distribución de los datos no siga un comportamiento normal, realizó los mismos pasos de las submuestras con una transformación logarítmica y box-cox (cuadrática) al total de datos. En las Tablas 16, 17, 18, 19 se muestra el resumen de los datos y los estadísticos descriptivos para cada variable y tipo de transformación aplicada:

**Tabla 16.**

*Resumen del total de datos con transformación logarítmica*

Variable	Min. (m)	1er cuartil (m)	Mediana (m)	3er cuartil (m)	Max. (m)
Ond-real	0.974	1.034	1.117	1.210	1.299
Ond-egm	0.966	1.034	1.094	1.192	1.284

**Tabla 17.**

*Estadísticos descriptivos del total de datos con transformación logarítmica*

Variable	Media (m)	Desv. Est. (m)	Rango (m)	Coef. Varia.	Asimetría	Kurtosis	R <sup>2</sup>
Ond-real	1.124	0.096	0.176	0.085	0.144	-1.30	0.995
Ond-egm	1.110	0.089	0.157	0.080	0.285	-1.184	

**Tabla 18.**

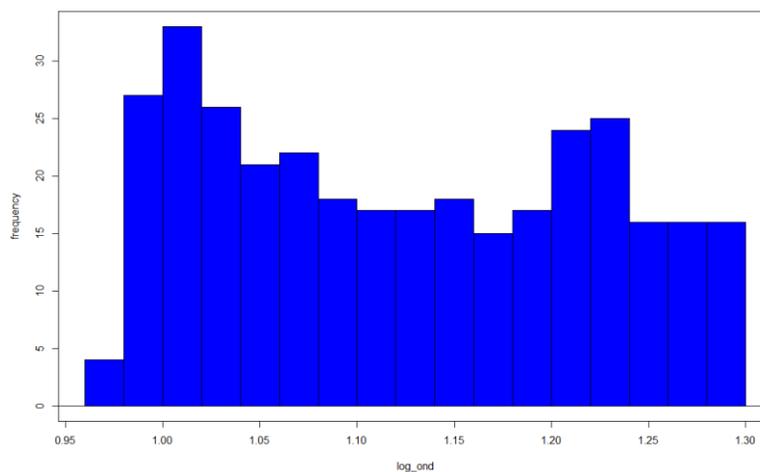
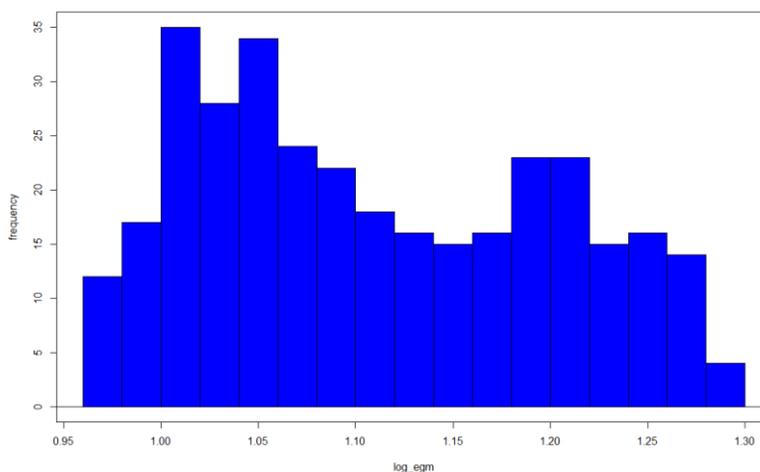
*Resumen del total de datos con transformación box-cox*

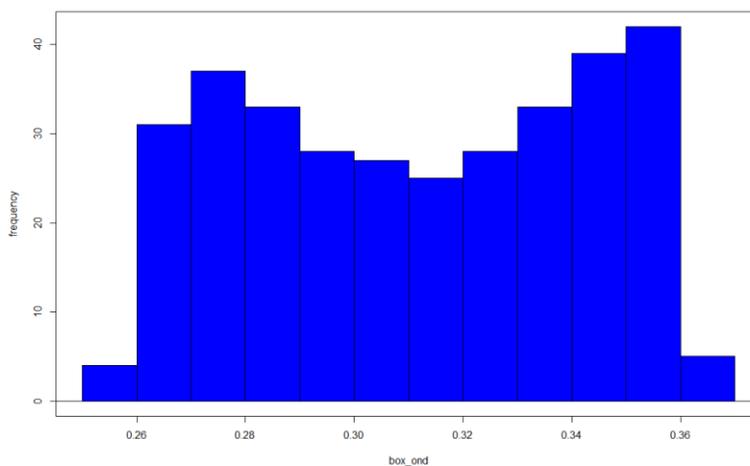
Variable	Min. (m)	1er cuartil (m)	Mediana (m)	3er cuartil (m)	Max. (m)
Ond-real	0.258	0.283	0.312	0.340	0.362
Ond-egm	0.067	0.082	0.101	0.114	0.131

**Tabla 19.***Estadísticos descriptivos del total de datos con transformación box-cox*

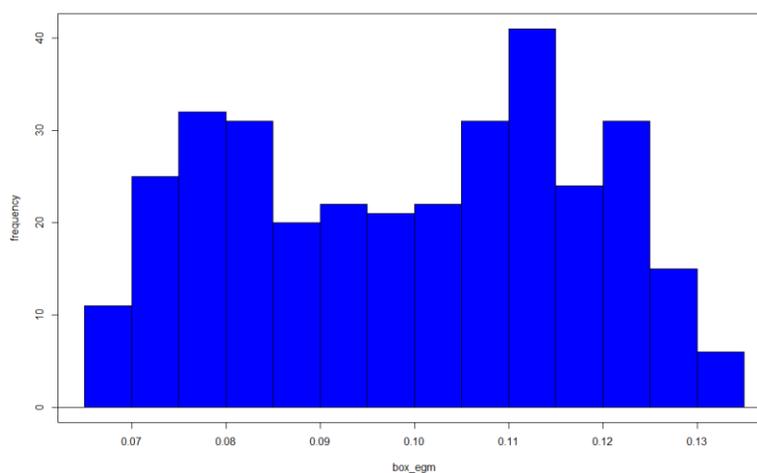
Variable	Media (m)	Desv. Est. (m)	Rango (m)	Coef. Varia.	Asimetría	Kurtosis	R <sup>2</sup>
Ond-real	0.311	0.031	0.057	0.099	-0.044	-1.341	0.995
Ond-egm	0.099	0.017	0.032	0.181	-0.079	-1.247	

Se procedió a graficar los histogramas para el total de muestras según las transformaciones efectuadas como se indican en las siguientes Figuras (75, 76, 77, 78):

**Figura 75.** Histograma de Ondulación geoidal real con transformación logarítmica**Figura 76.** Histograma de Ondulación geoidal EGM08 (1') con transformación logarítmica



**Figura 77.** Histograma de Ondulación geoidal real con transformación box-cox



**Figura 78.** Histograma de Ondulación geoidal EGM08 (1') con transformación box-cox

Las pruebas de normalidad se aplicaron para los dos tipos de transformaciones efectuadas, como se indica en las tablas siguientes:

**Tabla 20.**

*Pruebas de normalidad para el total de puntos con transformación logarítmica*

Variable	Estadístico W Shapiro-Wilk	Estadístico D Kolmogorov- Smirnov	p-value Shapiro-Wilk	p-value Kolmogorov- Smirnov
Ond-real (332)	0.937	0.103	1.451e-10	3.77e-09
Ond-egm (332)	0.942	0.109	4.229e-10	2.176e-10

**Tabla 21.***Pruebas de normalidad para el total de puntos con transformación box-cox*

Variable	Estadístico W Shapiro-Wilk	Estadístico D Kolmogorov- Smirnov	p-value Shapiro-Wilk	p-value Kolmogorov- Smirnov
Ond-real (332)	0.938	0.093	1.817e-10	2.382e-07
Ond-egm (332)	0.949	0.093	3.395e-09	1.948e-07

Los resultados confirman la suposición de que a medida que se incrementan los datos, la distribución no es normal, a pesar de las transformaciones aplicadas, por lo que se decidió trabajar con los datos sin transformarlos.

#### 4.4.6 Análisis exploratorio de Kriging con 299 datos

Como se mencionó anteriormente, se utilizará el método Kriging para comparar los resultados generados mediante Cokriging, por lo cual también se hizo en análisis exploratorio de datos con 299 puntos muestrales, dejando el 10% del total (33 puntos) para validar el modelo generado.

En las Tablas 22 y 23 se muestra el resumen y estadísticos de los datos analizados:

**Tabla 22.***Resumen de los datos de la variable con kriging (299 datos)*

Variable	Min. (m)	1er cuartil (m)	Mediana (m)	3er cuartil (m)	Max. (m)
Ond-real	9.451	10.766	12.861	16.123	19.945

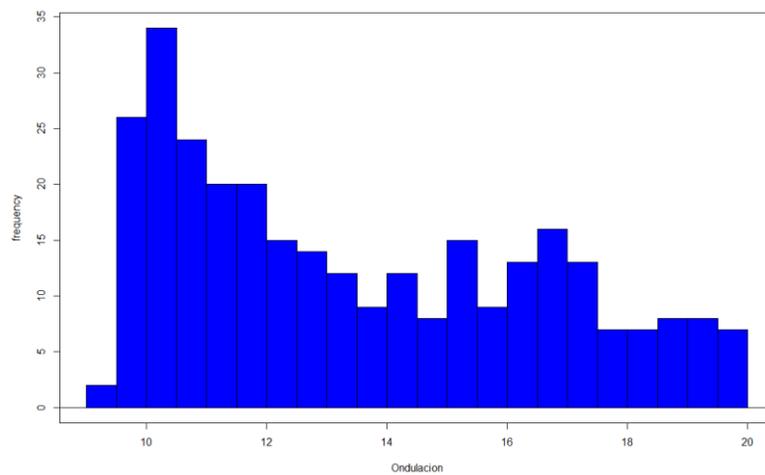
**Tabla 23.***Estadísticos descriptivos de la variable con kriging (299 datos)*

Variable	Media (m)	Desv. Est. (m)	Rango (m)	Coef. Varia.	Asimetría	Kurtosis
Ond-real	13.514	3.036	5.357	0.224	0.448	-1.083

En base a los estadísticos mostrados, se dice que el rango es un poco más grande comparado con las submuestras (lo cual es obvio ya que abarca más datos); el coeficiente de variación indica

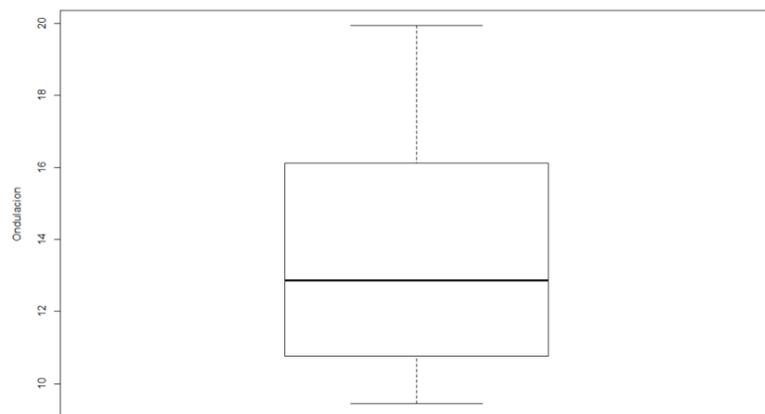
que no hay problema por variabilidad en los datos; su media es mayor a la mediana y sumado a que la asimetría muestra que la distribución no es simétrica, señalaría que la distribución está cargada ligeramente con valores superiores a la media; Kurtosis infiere que la forma de la distribución es en general plana.

El histograma comprueba lo expuesto anteriormente.



**Figura 79.** Histograma de la Ondulación real con 299 datos

El diagrama de caja comprueba que no existen valores extremos.



**Figura 80.** Diagrama de caja de Ondulación geoidal real con 299 datos

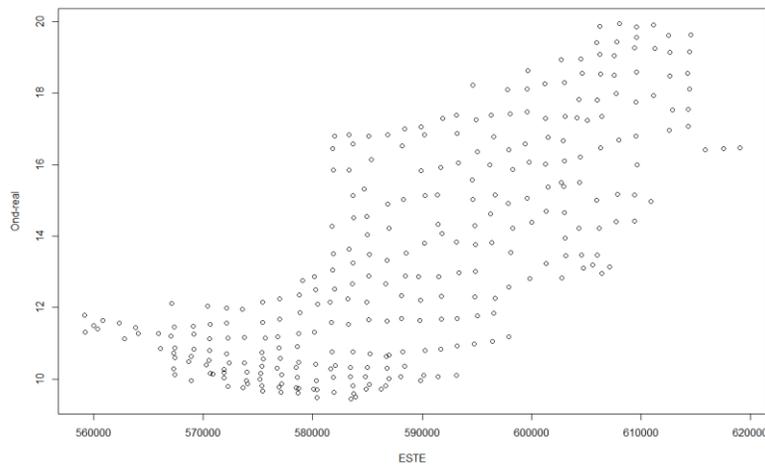
Las pruebas de normalidad indican que la distribución no es normal, sin embargo, como se demostró en el análisis con todos los datos de la muestra, no se aplicará ninguna transformación para alcanzar algún indicio de normalidad, sino que se trabajará con los datos crudos.

**Tabla 24.**

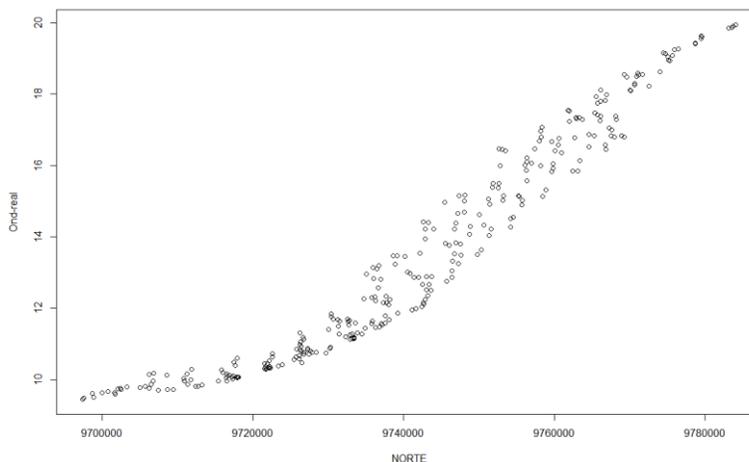
*Pruebas de normalidad de la variable con kriging (299 datos)*

Variable	Estadístico W Shapiro-Wilk	Estadístico D Kolmogorov- Smirnov	p-value Shapiro-Wilk	p-value Kolmogorov- Smirnov
Ond-real	0.921	0.127	1.835e-11	1.078e-12

Finalmente, se observa el comportamiento de la variable en el espacio a través del gráfico de dispersión mostrado a continuación:



**Figura 81.** Gráfico de dispersión de la coordenada Este vs Ondulación geoidal real



**Figura 82.** Gráfico de dispersión de la coordenada Norte vs Ondulación geoidal real

Como se puede notar, el comportamiento persiste independientemente de la cantidad de puntos analizados, y de manera más clara en la coordenada Norte donde se observa una tendencia en sentido creciente, por lo que se confirma la hipótesis intrínseca en el fenómeno.

#### **4.5 Definición de la resolución del modelo de ondulación geoidal**

Antes de pasar al análisis estructural, en el cual tiene como finalidad ajustar el cross – semivariograma empírico a un modelo teórico que mejor represente al fenómeno, se debe tomar muy en cuenta la resolución espacial con la que se generará el modelo de predicción con el modelo autorizado.

Cuando se trata de representar una variable, proceso o fenómeno natural, el formato más común, útil y usado es el ráster, el cual está formado por la unidad mínima conocida como pixel, tesal o celda (DeMers, 2001). A menudo se confunde el término píxel con el de grilla o cuadrícula, sin embargo, en términos prácticos significan o hacen referencia a lo mismo, pero quienes estudian la geoinformación los diferencian en que el píxel es una tecnología y la grilla es el modelo del ráster (De By, 2001).

Han sido numerosos los estudios que investigaron la importancia de la resolución espacial del modelo, en varias áreas del conocimiento como por ejemplo: variables del suelo (Florinsky & Kuryakova, 2000), datos meteorológicos (Noda & Niino, 2003), corrientes marinas (Molnár & Julien, 2000), por mencionar algunos. Esto sugiere dos puntos que hay que resaltar: primero, la resolución espacial es un factor determinante en la utilidad del mapa, ya que si la resolución es muy grande se puede obviar valores importantes de la variable y por ende no representar de manera correcta la dinámica en el espacio, y si la resolución es muy pequeña se sobrestima la funcionalidad del modelo debido a que solo mejoraría la presentación mas no su precisión; segundo, se puede optimizar la resolución del modelo bajo ciertos parámetros para mejorar su coste computacional y la representación misma de la variable en el área geográfica.

A pesar de ser un aspecto de gran trascendencia para la construcción de un correcto modelo espacial, a menudo la resolución es escogida según la variabilidad espacial de los datos disponibles, es decir sin ningún criterio técnico (Bishop, McBratney, & Whelan, 2001). Un ejemplo práctico de esto es cuando se utiliza ArcGIS y se emplea la herramienta Spatial Analyst o Geostatistical Analyst para generar un ráster a partir de una base de datos tipo punto, el programa considera la distancia más pequeña de la zona estudiada y divide para 250, estimando así la resolución del modelo (ESRI, 2002).

Se puede considerar varios criterios para definir la resolución del modelo ráster, entre los más importantes se tiene: escala del trabajo; precisión del posicionamiento GPS; tamaño del objeto mínimo; densidad de puntos; distancia de separación entre puntos; estructura espacial de dependencia; topografía del terreno. En el presente estudio, se siguió la metodología dada por Hengl (2006) en base a la densidad de puntos muestreados, según la fórmula siguiente:

$$Resolución = 0.05 * \sqrt{\frac{A}{N}} \quad (102)$$

Donde  $A$  es el área de la zona de estudio en  $m^2$ ,  $N$  es el número total de puntos. Aplicando este criterio resultó un tamaño de píxel de 135 m, con el cual se generará el modelo de predicción para la ondulación geoidal.

## 4.6 Análisis estructural de datos

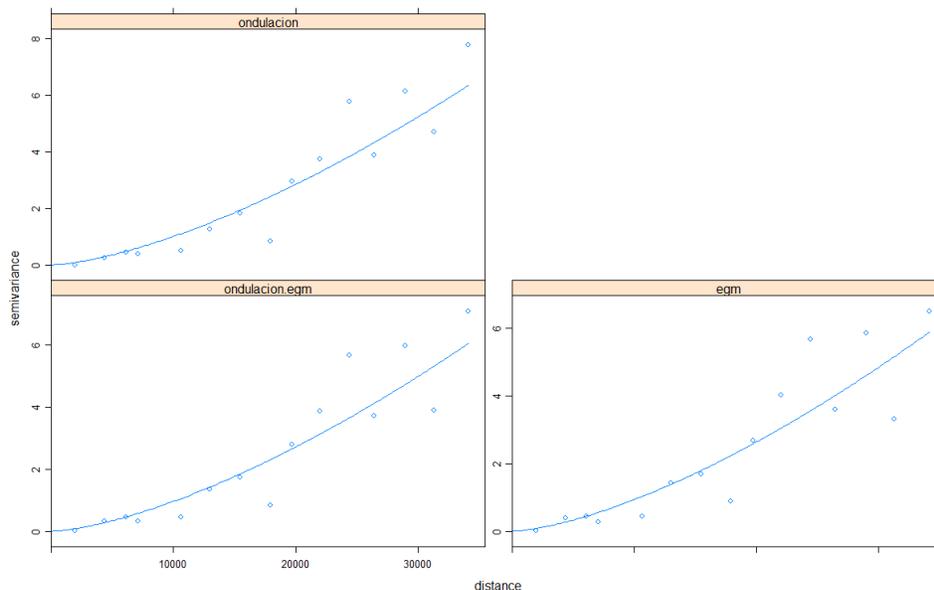
### 4.6.1 Obtención, predicción y validación del modelo con 5% de datos

#### 4.6.1.1 Cokriging ordinario

Cokriging ordinario, así como universal y residual se realizó en el software R, utilizando el paquete Gstat (Pebesma, 2014), además de otros paquetes complementarios como geoR (Ribeiro & Diggle, 2016), sp (Pebesma, *et al.*, 2018), epiR (Stevenson, *et al.*, 2018) y Lattice (Sarkar, 2018).

La variable auxiliar, es decir, la ondulación geoidal del EGM08 a 1 minuto, fue seleccionada para toda el área de trabajo (ver Figura 20) más un buffer de 5 km., por lo que el número de puntos que abarcó fue de 1487, con los cuales se procedió a realizar Cokriging.

El primer paso fue generar el semivariograma cruzado, que si se considera un modelo de correogionalización lineal de dos variables, se ajustaron 3 semivariogramas, 2 simples y uno cruzado. El modelo autorizado para este caso de Cokriging, fue el modelo potencia, ya que es el único que cumplió con los requisitos de hipótesis intrínseca, lo cual se puede observar en la Figura 83:



**Figura 83.** Cross – semivariograma ajustado al modelo Potencia

Como se puede observar, el modelo teórico presenta Nugget, y se ajusta bastante bien a la distribución de los datos. Cabe aquí señalar que el paquete Gstat ajusta todos los parámetros de los modelos teóricos menos el valor del rango (Reyes & Wendroth, 2017), por lo cual se optó por la herramienta “Solver” de Excel para ajustar este valor. El valor obtenido con esta herramienta fue de 26945.462 metros, sin embargo, al momento de contrastar resultados se obtuvieron los mismos valores entre el rango ajustado y sin ajustar (estimado 35000 metros), señalando la robustez que posee el estimador Cokriging, por lo que únicamente se trabajó con el rango estimado visualmente.

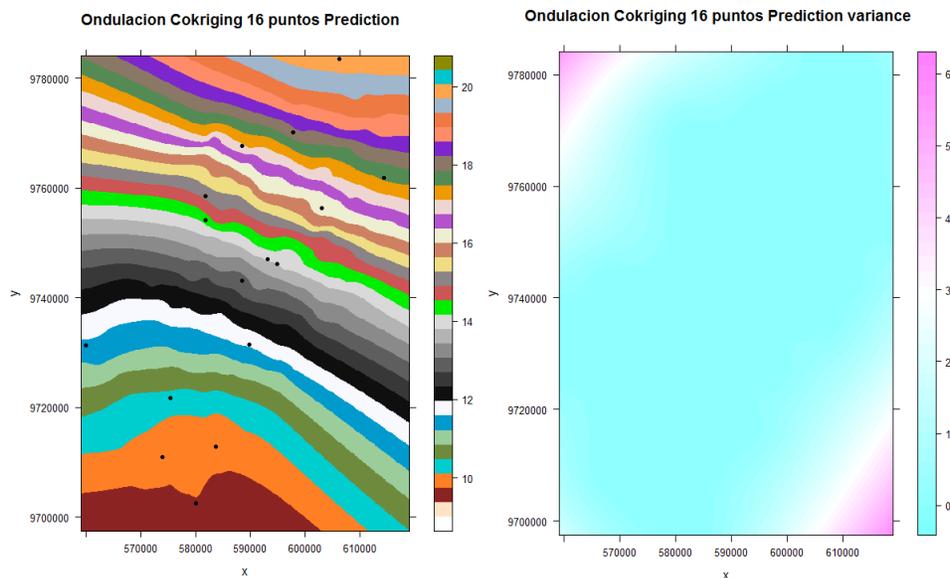
Los parámetros de los semivariogramas se muestran en la Tabla 25.

**Tabla 25.**

*Parámetros del modelo Potencia ajustado – Cokriging ordinario al 5% de datos*

Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	1.006178e-06		
Simple (Ond-egm)	9.346621e-07	35000	1.50
Cruzado	9.610894e-07		

Una vez ajustado el modelo teórico se procedió a realizar la predicción de la ondulación geoidal sobre una grilla con la resolución espacial calculada (135 metros), cuyo resultado se puede previsualizar dentro del software, mismo que se muestra en la Figura 84.



**Figura 84.** Visualización del modelo de predicción y varianza de la ondulación geoidal

Para comprobar que el modelo teórico representa correctamente la variabilidad espacial de la ondulación geoidal, se aplicó una cross – validación del modelo utilizando el principio de “leave one out”, estableciendo un vecindario de búsqueda móvil con los siguientes parámetros:

- Número mínimo de puntos dentro del vecindario: 8;
- Número máximo de puntos dentro del vecindario: 16;
- Se consideró trabajar por sectores;
- Número de sectores: 8;
- Número máximo de puntos dentro de cada sector: 2;
- Alcance de búsqueda: 35000 m.

Estos criterios se aplicaron a todas las predicciones hechas mediante Cokriging y Kriging.

Además se calculó los estadísticos: error medio (ME), raíz del error cuadrático medio (RMSE) y razón de la desviación estándar del RMSE (RSR) para la toma de decisiones. Las fórmulas de estos estadísticos son las siguientes:

$$ME = \frac{1}{n} \sum_{i=1}^n [z(x_i) - \hat{z}(x_i)] \quad (103)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [z(x_i) - \hat{z}(x_i)]^2} \quad (104)$$

$$RSR = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n [z(x_i) - \hat{z}(x_i)]^2}}{s(x_i)} \quad (105)$$

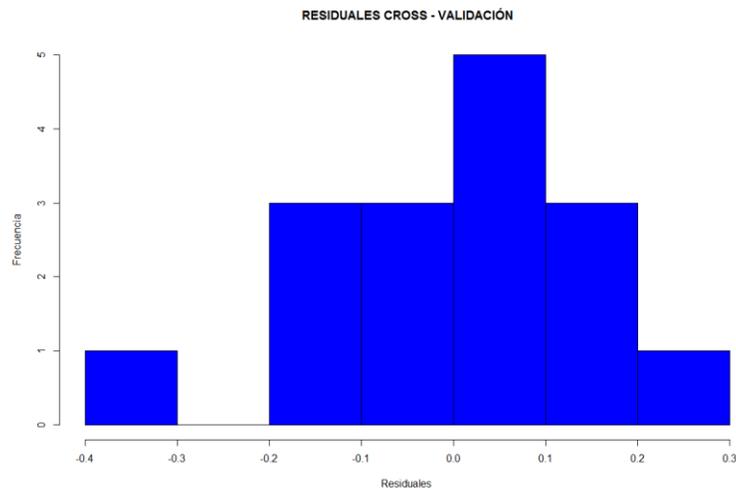
Donde  $n$  es el número de datos de la muestra,  $z(x_i)$  es el valor conocido de ondulación geoidal,  $\hat{z}(x_i)$  representa el valor de ondulación geoidal predicho,  $s(x_i)$  es la desviación estándar de los valores medidos de ondulación geoidal.

**Tabla 26.**

*Estadísticos de la validación cruzada con Cokriging ordinario al 5% de datos*

Parámetro	ME	RMSE	RSR
Valor	0.0064	0.1416	0.0438

El histograma de la cross – validación se muestra a continuación:



**Figura 85.** Histograma de validación cruzada con 16 datos

En el histograma anterior, se tiene que la mayoría de los errores de la validación cruzada están en un rango de  $[-0.28;0.28]$  metros al 95% de confianza. A pesar de esto, es un modelo bastante bueno considerando que solo se utilizó 16 puntos.

#### 4.6.1.2 Cokriging universal

Existen varias formas de trabajar con datos no estacionarios, es decir, con presencia de deriva o tendencia. En la variable estudiada, se observó a través de los gráficos de dispersión en base a la posición geográfica de los puntos muestrales, y al análisis del cross – correlograma o gráfico de cross – correlación la existencia de tendencia. Una forma de modelar esta tendencia es con Cokriging universal (Emery, 2007; Díaz, 2002; Giraldo, 2007; Oliver, 2010), para lo cual se siguió los mismos pasos que en Cokriging ordinario.

Para generar el semivariograma cruzado, en Cokriging universal es necesario definir una función de la variable que se estudia, en este caso dicha función corresponde a un residual ( $re$ ) más un valor ajustado ( $va$ ):

$$Ondulación = re + va \quad (106)$$

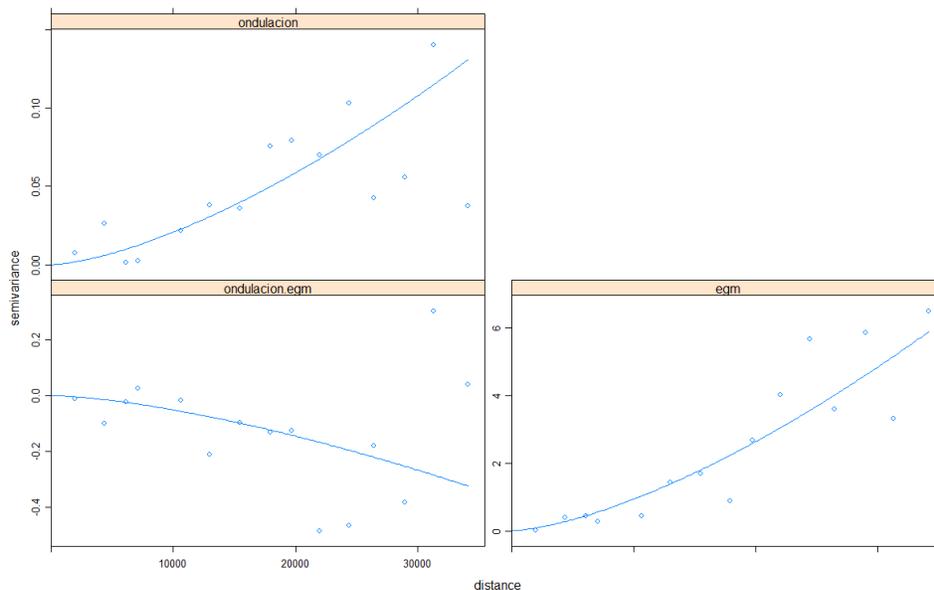
El valor ajustado viene dado por:

$$va = C_{egm} * EGM_1 + I \quad (107)$$

Siendo  $C_{egm}$  el coeficiente de la tendencia de ondulación geoidal EGM08,  $EGM_1$  es el valor de ondulación geoidal del modelo geopotencial para cada punto analizado,  $I$  es el coeficiente del intercepto de la función de Cokriging universal. Reescribiendo (106), la expresión resultante de este método es:

$$Ondulación = re + C_{egm} * EGM_1 + I \quad (108)$$

Con (108) se procedió a calcular el semivariograma cruzado empírico, para posteriormente ajustar con un modelo teórico autorizado, en que se notó dos importantes rasgos, el primero que el modelo autorizado sigue siendo no acotado, en este caso el modelo potencia, y segundo el semivariograma cruzado se comporta en sentido inverso, es decir que cuando el valor de una variable aumenta, el valor de otra variable disminuye, presentado un valor negativo de semivarianza.



**Figura 86.** Cross – semivariograma ajustado al modelo Potencia

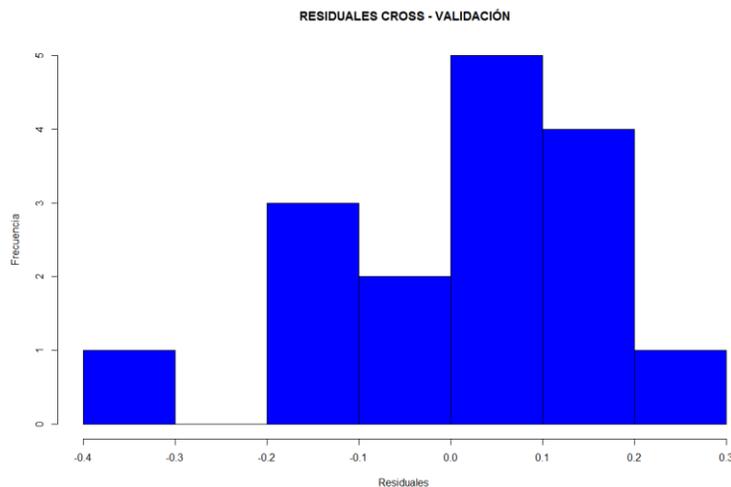
Los parámetros del modelo ajustado son los siguientes:

**Tabla 27.**

*Parámetros del modelo Potencia ajustado – Cokriging universal al 5% de datos*

Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	2.073783e-08		
Simple (Ond-egm)	9.346621e-07	35000	1.50
Cruzado	-5.137616e-08		

Con el modelo potencia ajustado se realizó la predicción sobre la grilla con resolución definida, para en forma consecutiva aplicar una validación cruzada cuyo histograma de residuales se puede apreciar en la siguiente figura:



**Figura 87.** Histograma de validación cruzada con 16 datos

El histograma de residuales de la validación cruzada se muestra similar al de Cokriging ordinario, con un intervalo de  $[-0.28;0.28]$  metros al 95% de confianza. Finalmente, se calcularon los estadísticos para verificar la precisión de la validación cruzada, como se muestra en la Tabla 28.

**Tabla 28.**

*Estadísticos de la validación cruzada con Cokriging universal al 5% de datos*

Parámetro	ME	RMSE	RSR
Valor	0.0114	0.1403	0.0434

#### 4.6.1.3 Cokriging residual

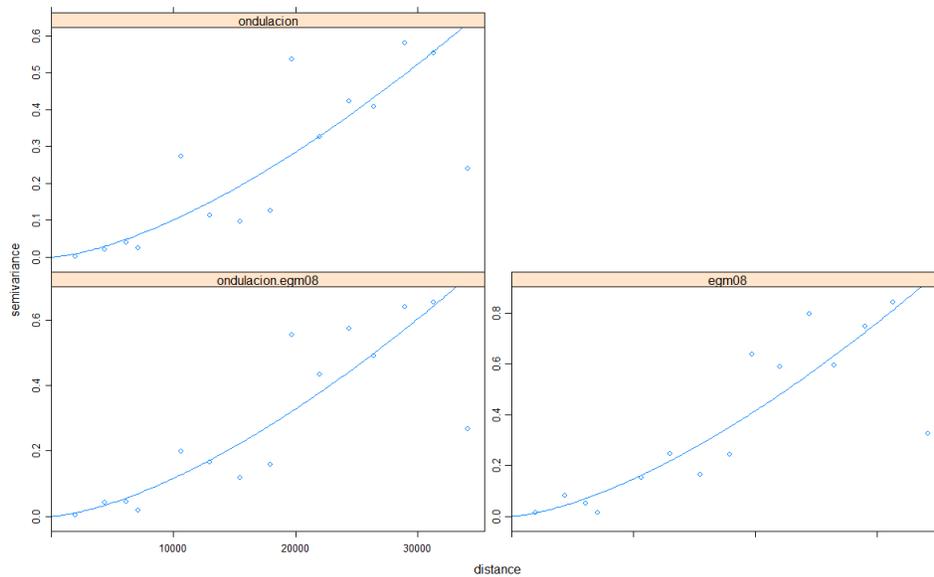
Otra forma de trabajar con tendencia es a través de regresiones lineales múltiples, que analizan la deriva bajo el supuesto de que este drift está subyacente en la locación espacial de las muestras tomadas de la variable (Díaz, 2002; Giraldo, 2007, Hengl, 2009). En este caso, la función de regresión utilizada para calcular los residuales es la siguiente:

$$\text{Regresión} = f(\text{Este, Norte}) \rightarrow re \quad (109)$$

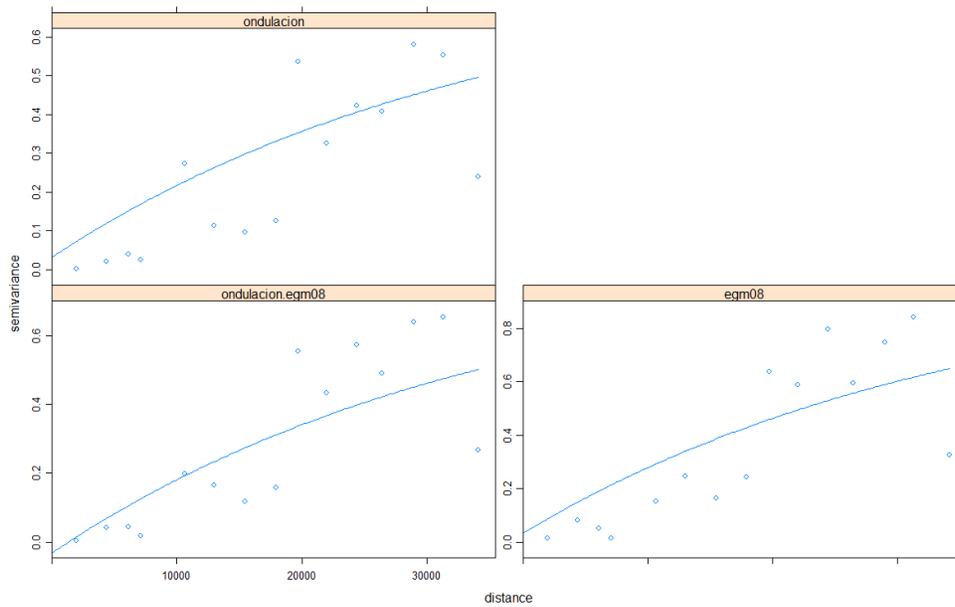
$$\text{Ondulación} = \text{Este} * C_{\text{Este}} + \text{Norte} * C_{\text{Norte}} + I + re \quad (110)$$

Donde  $C_{\text{Este}}$  es el coeficiente de la regresión para la coordenada Este,  $C_{\text{Norte}}$  el coeficiente de la regresión en la coordenada Norte,  $I$  y  $re$  son el intercepto y residual respectivamente.

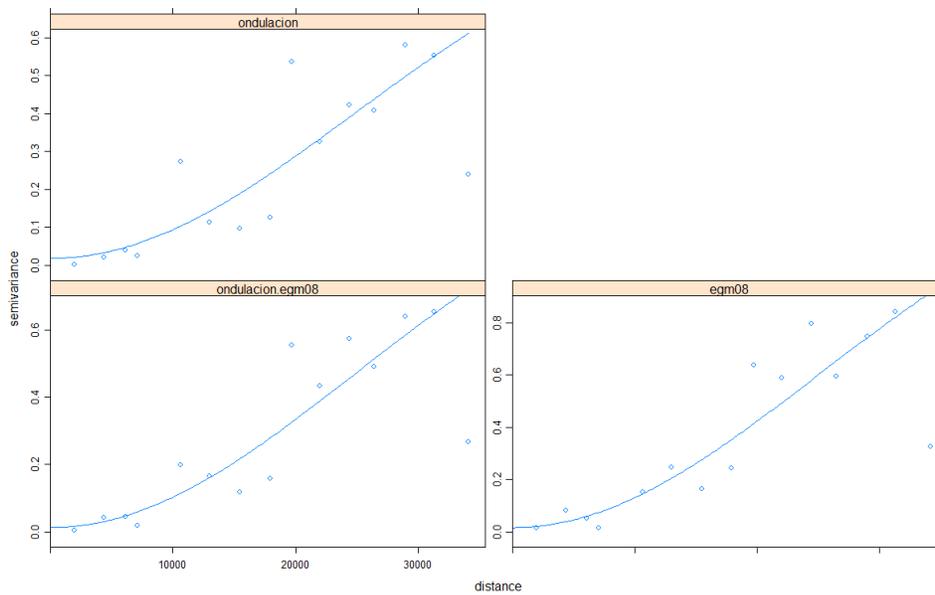
Mediante (110) se construyó el semivariograma cruzado de los residuales, que a diferencia de los casos con el estimador ordinario y universal, existió más de un modelo autorizado, siendo estos el potencia, exponencial y gaussiano. En el caso del modelo potencia, el semivariograma cruzado es el siguiente:



**Figura 88.** Cross – semivariograma ajustado al modelo Potencia



**Figura 89.** Cross – semivariograma ajustado al modelo Exponencial



**Figura 90.** Cross – semivariograma ajustado al modelo Gaussiano

Los parámetros de los modelos ajustados se presentan en las siguientes tablas:

**Tabla 29.***Parámetros del modelo Potencia ajustado – Cokriging residual al 5% de datos*

Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	1.006605e-07		
Simple (Ond-egm)	1.467032e-07	35000	1.50
Cruzado	1.163065e-07		

**Tabla 30.***Parámetros del modelo Exponencial ajustado – Cokriging residual al 5% de datos*

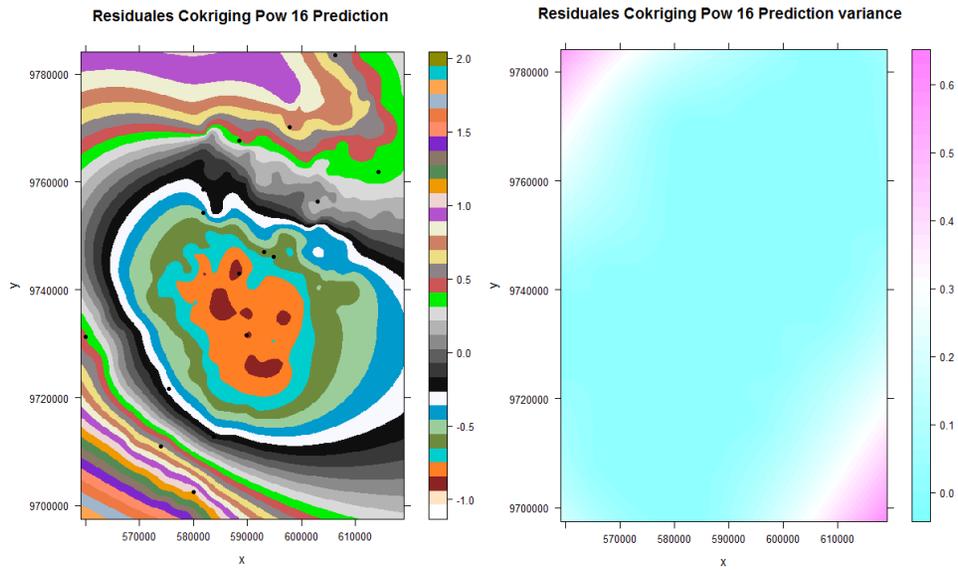
Semivariograma	Sill	Rango	Nugget
Simple (Ond-real)	0.74469895		0.03200762
Simple (Ond-egm)	0.98955769	35000	0.03200762
Cruzado	0.85844194		-0.03200762

**Tabla 31.***Parámetros del modelo Gaussiano ajustado – Cokriging residual al 5% de datos*

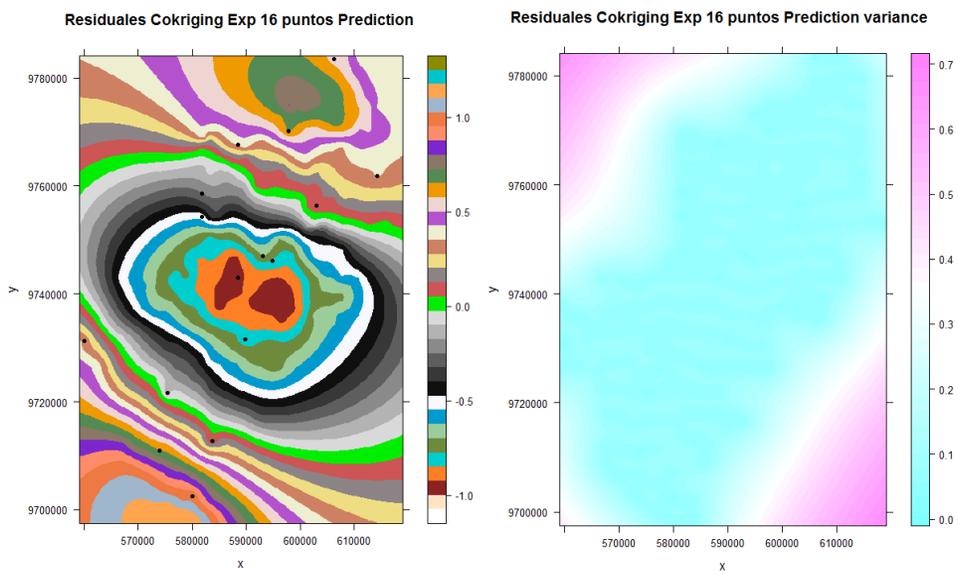
Semivariograma	Sill	Rango	Nugget
Simple (Ond-real)	0.96750586		0.01791954
Simple (Ond-egm)	1.46470542	35000	0.01520571
Cruzado	1.15649853		0.01225381

Una previsualización de la interpolación realizada se puede observar en las Figuras 91, 92 y

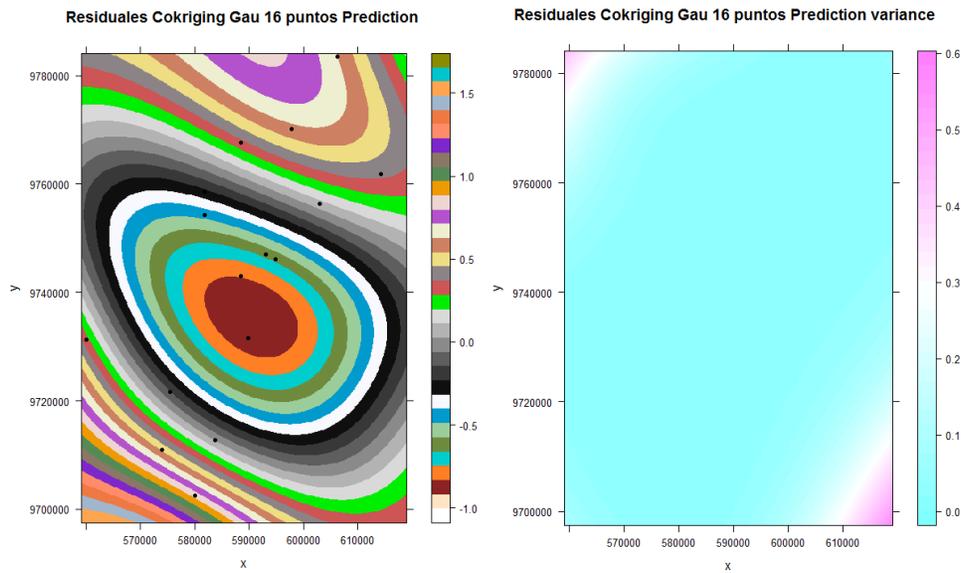
93.



**Figura 91.** Visualización del modelo Potencia de predicción y varianza de residuales

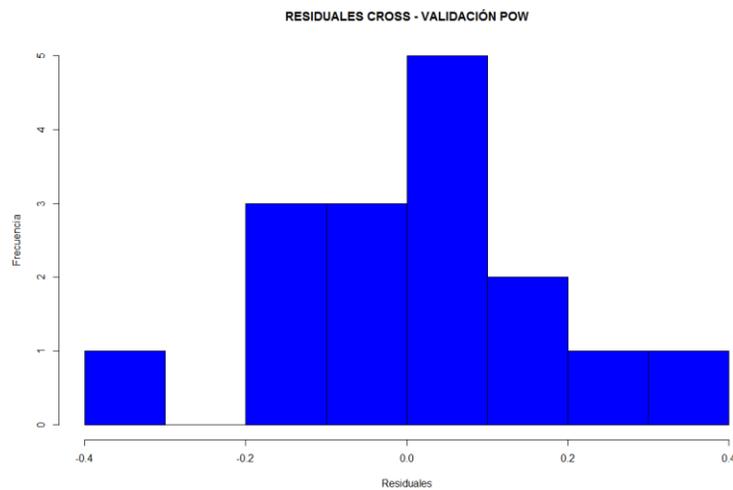


**Figura 92.** Visualización del modelo Exponencial de predicción y varianza de residuales

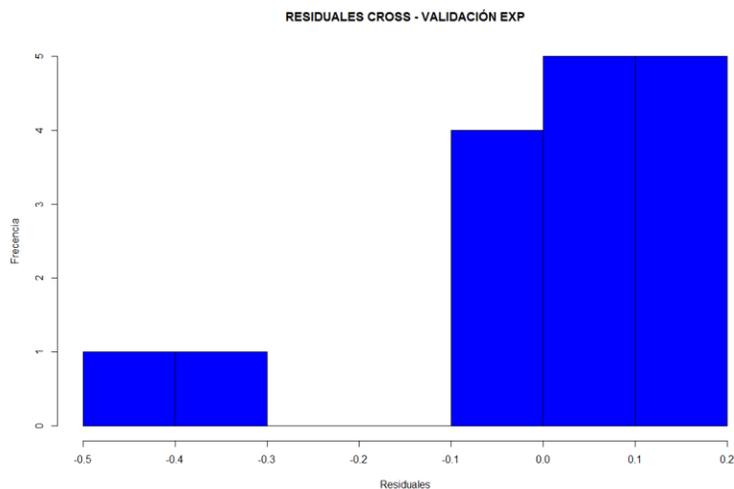


**Figura 93.** Visualización del modelo Gaussiano de predicción y varianza de residuales

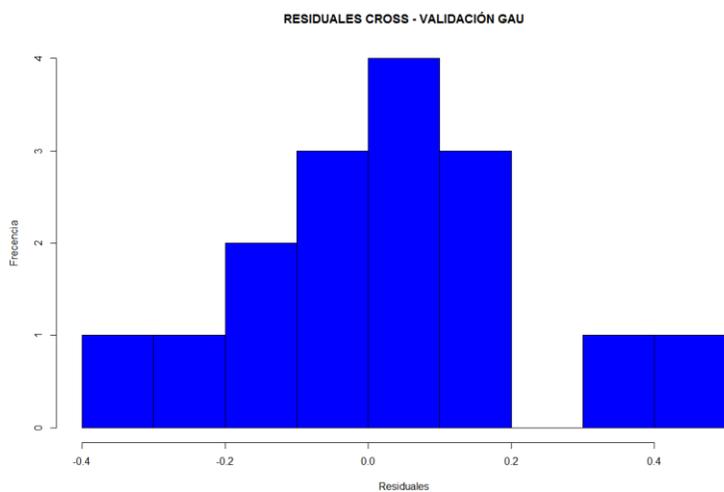
De igual forma se realizó la validación cruzada con cada modelo ajustado, cuyos histogramas se muestran a continuación:



**Figura 94.** Histograma de validación cruzada del modelo Potencia con 16 datos



**Figura 95.** Histograma de validación cruzada del modelo Exponencial con 16 datos



**Figura 96.** Histograma de validación cruzada del modelo Gaussiano con 16 datos

De los histogramas mostrados, se observa que el menor intervalo de los resultados de la validación cruzada es del modelo potencia  $[-0.32;0.32]$  al 95% de confianza, y además presenta menos datos extremos en comparación a los otros modelos usados. Los estadísticos para comprobar la validación cruzada se muestran a continuación:

**Tabla 32.**

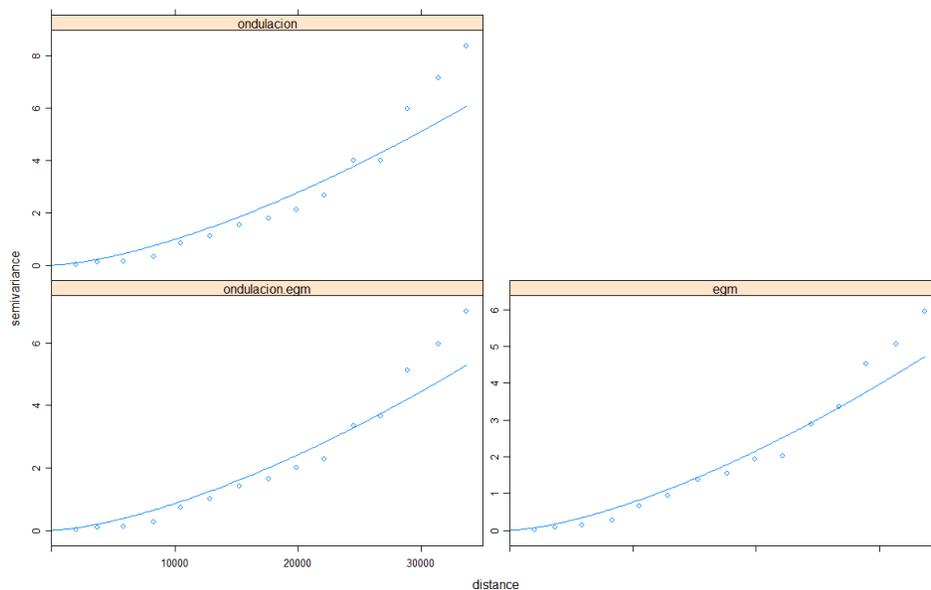
*Estadísticos de la validación cruzada con Cokriging residual al 5% de datos*

Modelo\Parámetro	ME	RMSE	RSR
Potencia	0.0116	0.1609	0.0498
Exponencial	0	0.1765	0.0546
Gaussiano	0.0161	0.1944	0.0601

## 4.6.2 Obtención, predicción y validación del modelo con 10% de datos

### 4.6.2.1 Cokriging ordinario

Con 33 puntos de muestra, se procedió a generar el semivariograma cruzado y ajustarlo a un modelo autorizado que represente la dinámica de la variable, en este caso al igual que con 16 puntos muestrales, el único modelo que se ajustó fue el potencia, cuyo semivariograma teórico ajustado es el siguiente:



**Figura 97.** Cross – semivariograma ajustado al modelo Potencia

Como se puede observar, el semivariograma cruzado es no acotado, presentando un pequeño valor de Nugget, y al igual que el caso con 16 datos, el rango estimado mediante Solver, no presentó

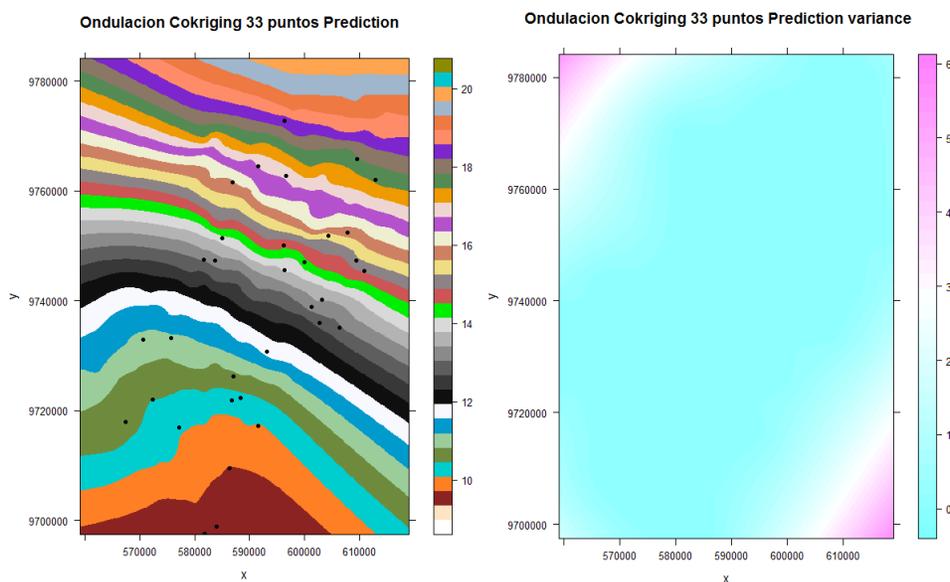
mejoras con el rango estimado a priori, por lo que se trabajó con este valor (35000 metros). Los parámetros del modelo ajustado se presentan en la Tabla 33.

**Tabla 33.**

*Parámetros del modelo Potencia ajustado – Cokriging ordinario al 10% de datos*

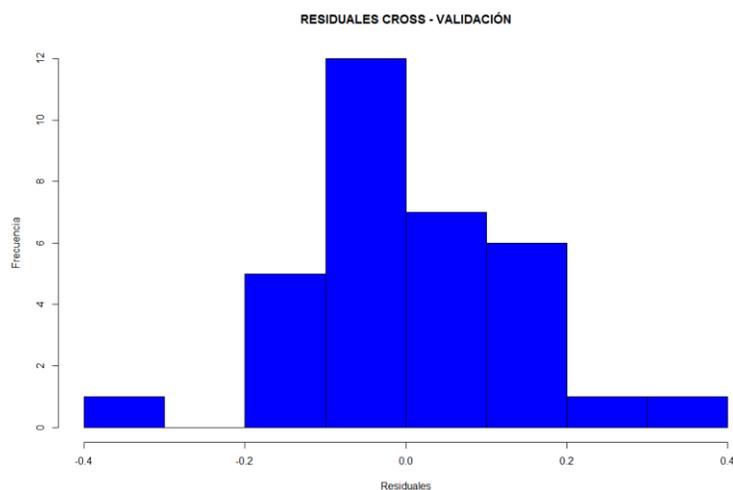
Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	9.828179e-07		
Simple (Ond-egm)	7.635417e-07	35000	1.52
Cruzado	8.558555e-07		

Una vez ajustado el modelo teórico se procedió a predecir los valores de ondulación geoidal sobre la grilla de puntos con resolución definida para el estudio, cuya previsualización es la siguiente:



**Figura 98.** Visualización del modelo de predicción y varianza de la ondulación geoidal

La validación cruzada que se efectuó al modelo de predicción se observa en su respectivo histograma de residuales en la Figura 99. El intervalo del error al 95% de confianza fue de  $[-0.26; 0.26]$  metros.



**Figura 99.** Histograma de validación cruzada con 33 datos

Los indicadores estadísticos de la cross – validación se muestran a continuación:

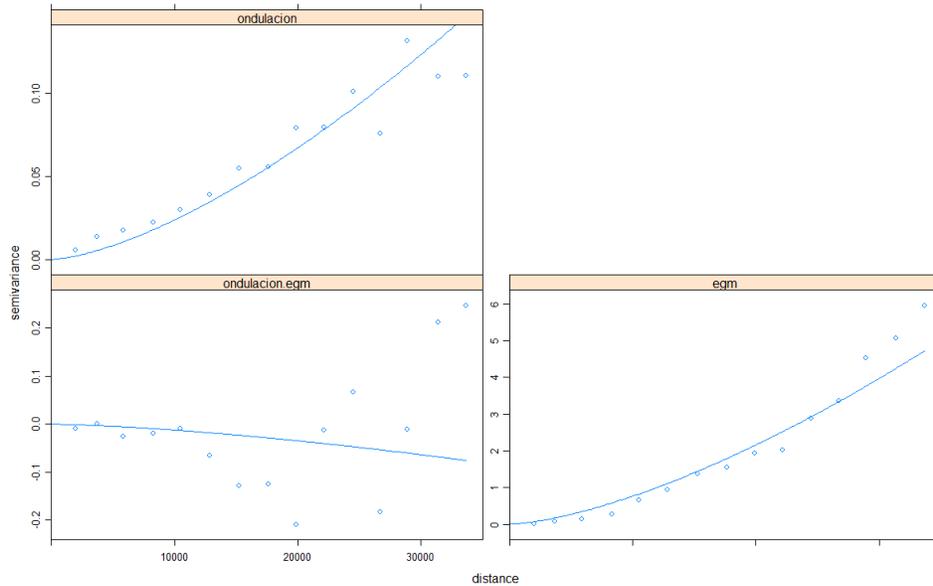
**Tabla 34.**

*Estadísticos de la validación cruzada con Cokriging ordinario al 10% de datos*

Parámetro	ME	RMSE	RSR
Valor	0.0095	0.1327	0.0495

#### 4.6.2.2 Cokriging universal

En el caso de Cokriging universal con 33 puntos, el modelo teórico que se ajustó fue el potencia, al igual que en el caso con 16 muestras. El semivariograma cruzado ajustado fue el siguiente:



**Figura 100.** Cross – semivariograma ajustado al modelo Potencia

Los parámetros del modelo ajustado se presentan en la Tabla 35.

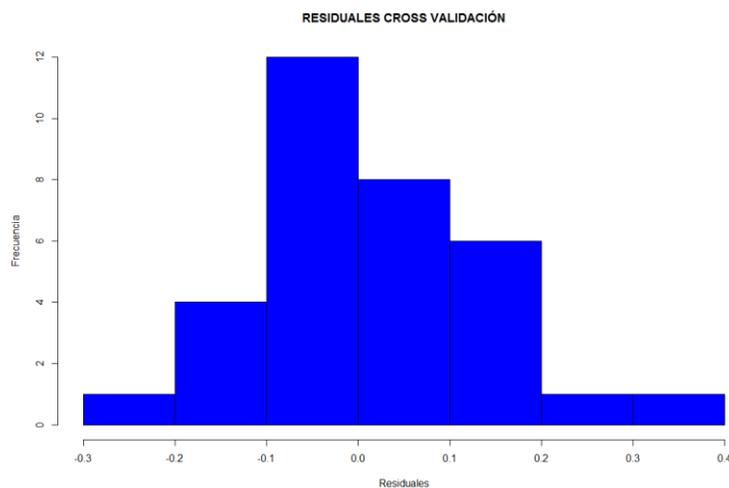
**Tabla 35.**

*Parámetros del modelo Potencia ajustado – Cokriging universal al 10% de datos*

Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	2.368298e-08		
Simple (Ond-egm)	7.635417e-07	35000	1.52
Cruzado	-1.220745e-08		

Este modelo surgió de la función expresada en (108), y al igual que en el anterior caso, el semivariograma cruzado presentó valores negativos. Con el modelo autorizado se procedió a generar el modelo de predicción de la variable sobre una grilla de puntos de la zona de estudio, a la resolución espacial definida anteriormente.

Posteriormente, se aplicó una validación cruzada para analizar la precisión del modelo, cuyo histograma se observa en la Figura 101 y sus respectivos indicadores estadísticos en la Tabla 36:



**Figura 101.** Histograma de validación cruzada con 33 datos

**Tabla 36.**

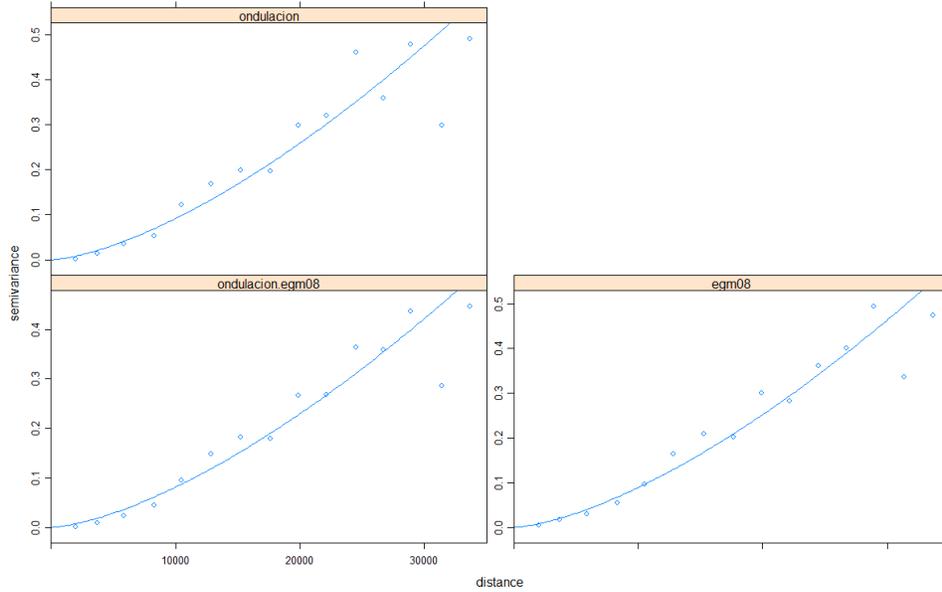
*Estadísticos de la validación cruzada con Cokriging universal al 10% de datos*

Parámetro	ME	RMSE	RSR
Valor	0.0119	0.1234	0.0460

El intervalo en que se encuentran agrupados la mayor cantidad de residuales es en un rango entre [-0.24;0.24] metros al 95% de confianza.

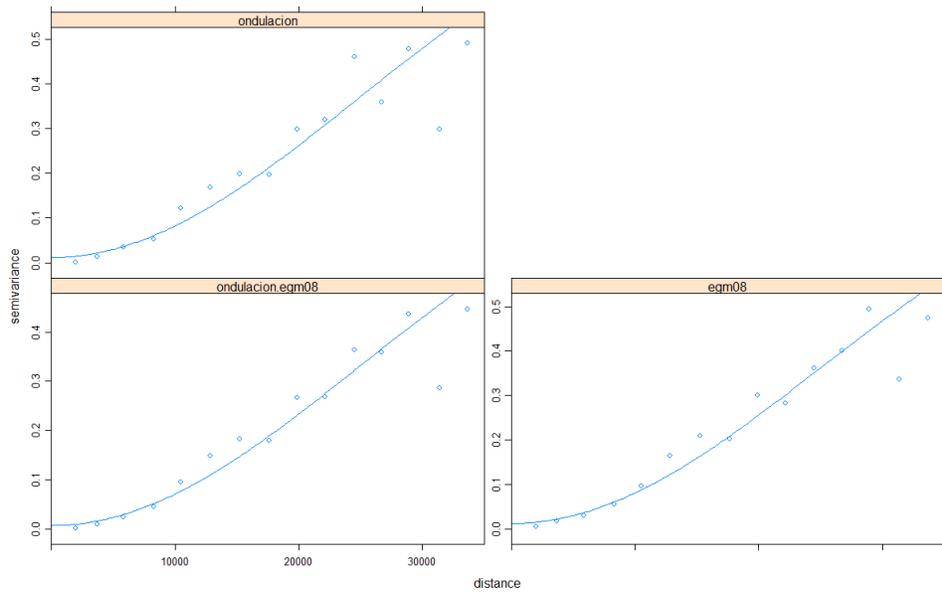
#### 4.6.2.3 Cokriging residual

Se siguió de forma similar que en el apartado mostrado de Cokriging residual con 16 puntos, empezando con la obtención de residuales según (109), con que se generó el semivariograma experimental para ajustarlo a un modelo teórico autorizado, que en este caso, presentó 2 modelos que cumplieron las condiciones del modelo de correogionalización lineal, el potencia y gaussiano. En el caso del modelo potencia, el semivariograma resultante fue el siguiente:



**Figura 102.** Cross – semivariograma ajustado al modelo Potencia

El modelo gaussiano presentó el siguiente semivariograma cruzado:



**Figura 103.** Cross – semivariograma ajustado al modelo Gaussiano

Los parámetros de los modelos ajustados, se presentan en las Tablas 37 y 38.

**Tabla 37.**

*Parámetros del modelo Potencia ajustado – Cokriging residual al 10% de datos*

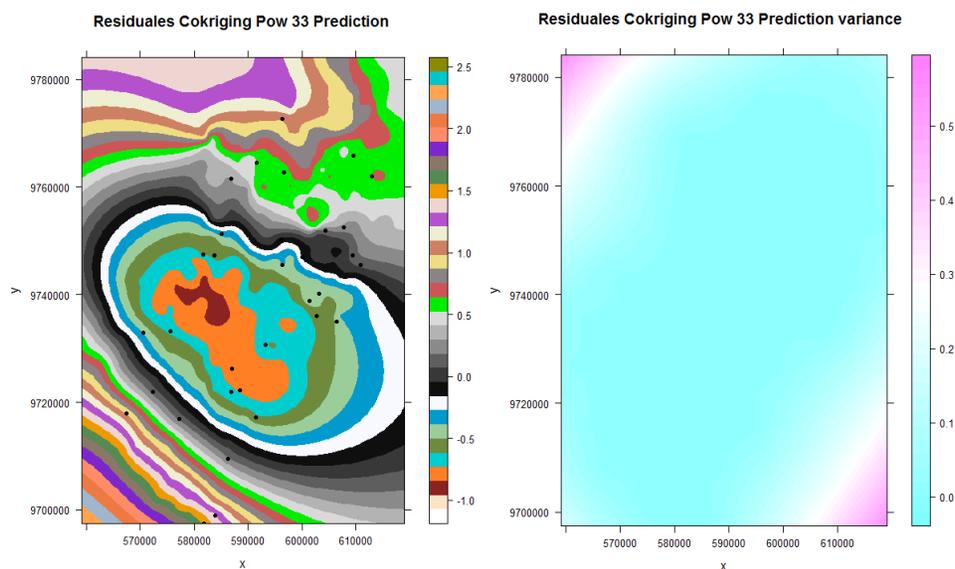
Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	9.163802e-08		
Simple (Ond-egm)	8.925811e-08	35000	1.52
Cruzado	8.128831e-08		

**Tabla 38.**

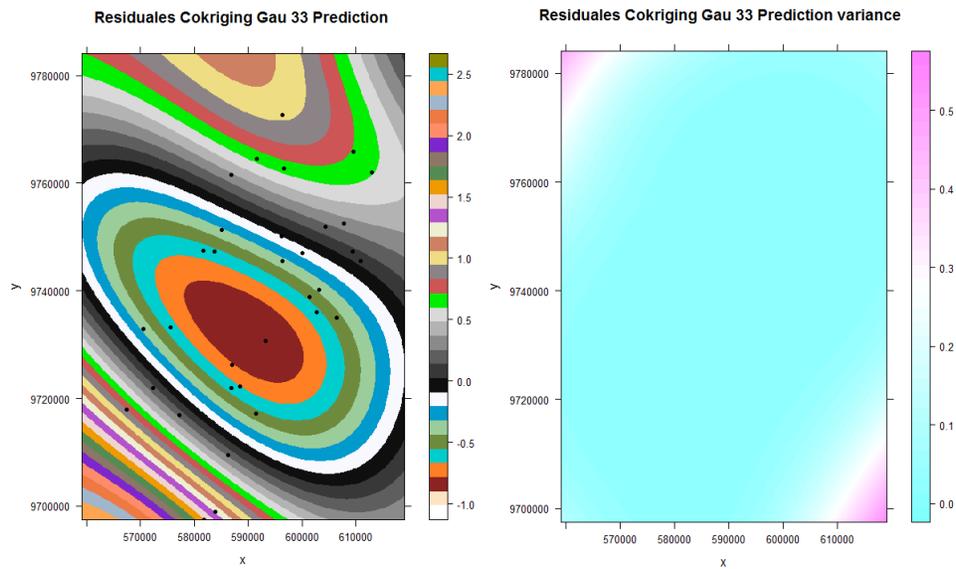
*Parámetros del modelo Gaussiano ajustado – Cokriging residual al 10% de datos*

Semivariograma	Sill	Rango	Nugget
Simple (Ond-real)	0.90007978		0.01182137
Simple (Ond-egm)	0.87675902	35000	0.01166512
Cruzado	0.81471058		0.00647737

La previsualización de los modelos predichos tanto para el modelo potencia y gaussiano se muestran en las Figuras 104 y 105:

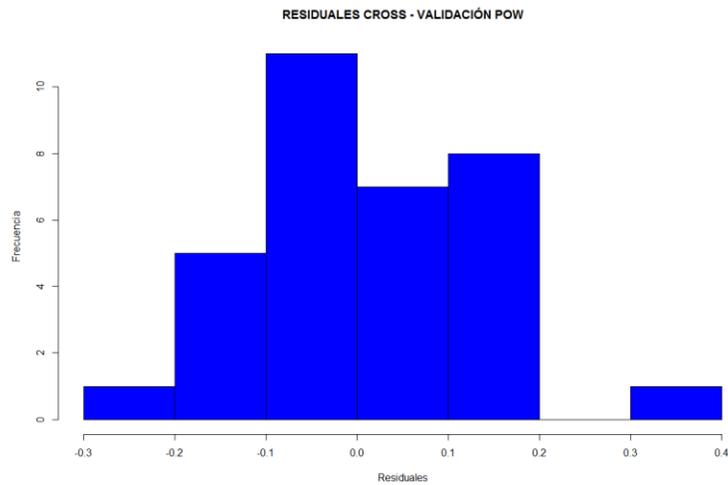


**Figura 104.** Visualización del modelo Potencia de predicción y varianza de residuales

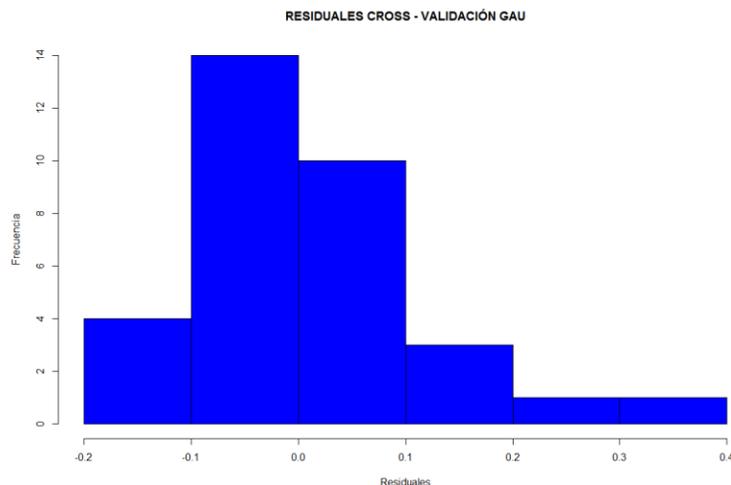


**Figura 105.** Visualización del modelo Gaussiano de predicción y varianza de residuales

Finalmente se procedió a realizar una validación cruzada de los modelos ajustados, con sus respectivos indicativos estadísticos, los mismos que se muestra en las siguientes figuras y tablas respectivamente.



**Figura 106.** Histograma de validación cruzada del modelo Potencia con 33 datos



**Figura 107.** Histograma de validación cruzada del modelo Gaussiano con 33 datos

**Tabla 39.**

*Estadísticos de la validación cruzada con Cokriging residual al 10% de datos*

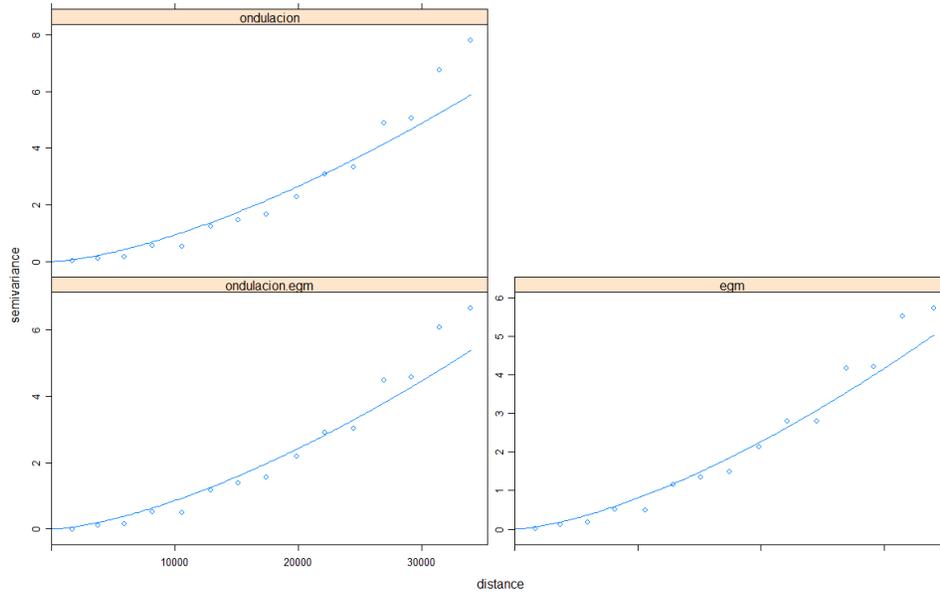
Modelo\Parámetro	ME	RMSE	RSR
Potencia	0.0107	0.1240	0.0462
Gaussiano	0.0046	0.1073	0.0400

Entre los modelos ajustados, se observa en los histogramas que el modelo gaussiano es ligeramente mejor que el potencia cuyo intervalo de confianza al 95% en que se encuentran los errores fue de  $[-0.21;0.21]$ , lo cual se confirmó también con los estadísticos usados.

#### 4.6.3 Obtención, predicción y validación del modelo con 15% de datos

##### 4.6.3.1 Cokriging ordinario

Con 49 puntos muestrales, el modelo autorizado para Cokriging ordinario fue el potencia, con el cual se procedió a ajustar el semivariograma cruzado empírico resultando de la siguiente forma:



**Figura 108.** Cross – semivariograma ajustado al modelo Potencia

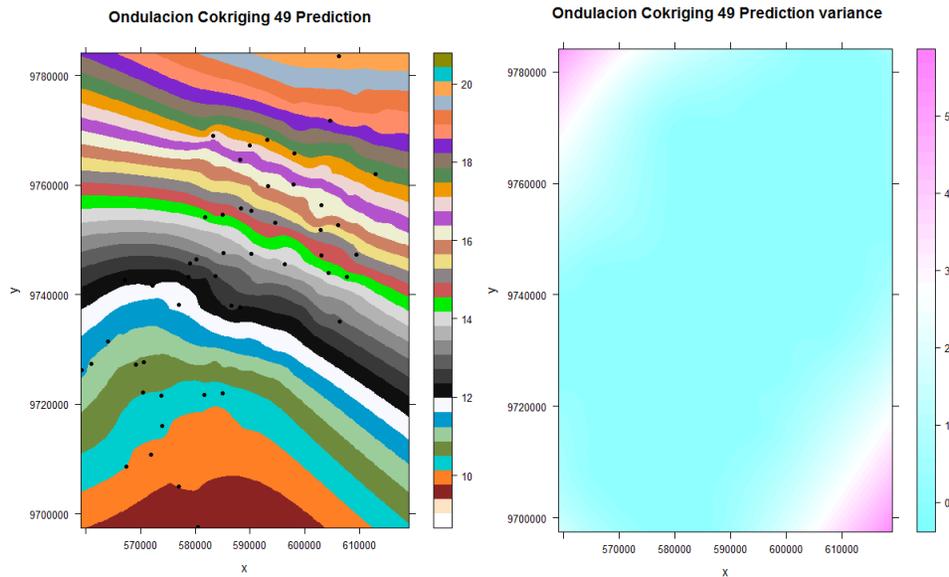
Los parámetros del modelo ajustado fueron los siguientes:

**Tabla 40.**

*Parámetros del modelo Potencia ajustado – Cokriging ordinario al 15% de datos*

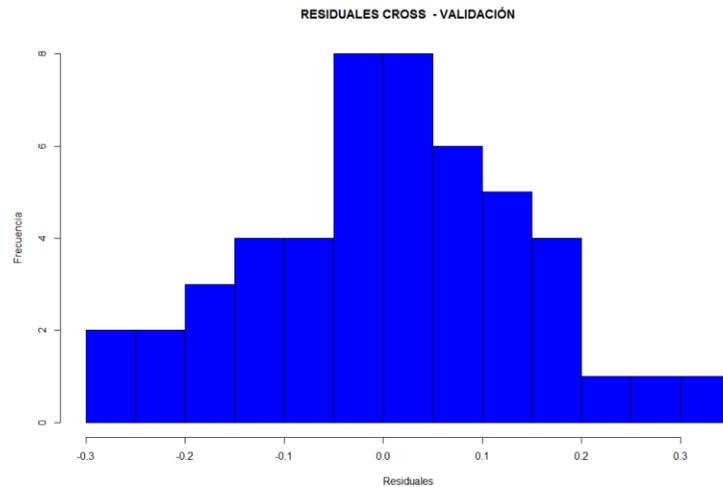
Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	9.396311e-07		
Simple (Ond-egm)	8.035252e-07	35000	1.56
Cruzado	8.598123e-07		

Con el modelo ajustado, se realizó la predicción sobre la grilla de puntos para el área de estudio, cuyos resultados se pueden previsualizar a continuación:



**Figura 109.** Visualización del modelo de predicción y varianza de la ondulación geoidal

La validación cruzada realizada mostró que los residuales se encuentran en un rango de  $[-0.26; 0.26]$  metros al 95% de confianza, los mismos que se pueden observar en el histograma siguiente:



**Figura 110.** Histograma de validación cruzada con 49 datos

Los estadísticos de la cross – validación fueron los siguientes:

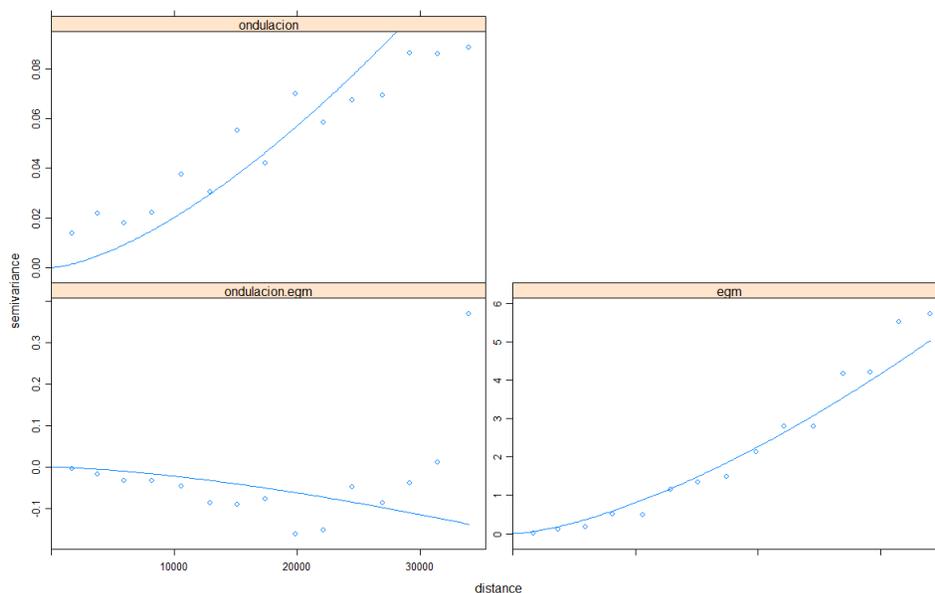
**Tabla 41.**

*Estadísticos de la validación cruzada con Cokriging ordinario al 15% de datos*

Parámetro	ME	RMSE	RSR
Valor	0.0044	0.1361	0.0514

#### 4.6.3.2 Cokriging universal

Al igual que los dos casos anteriores de Cokriging universal, se procedió a obtener la función de predicción de la variable para construir el semivariograma empírico en primera instancia, y consecutivamente ajustar con un modelo teórico que cumpla con el modelo de correogionalización lineal, que en este caso fue únicamente el modelo potencia, el mismo que se presenta a continuación:



**Figura 111.** Cross – semivariograma ajustado al modelo Potencia

Los parámetros del modelo ajustado se presentan en la Tabla 42.

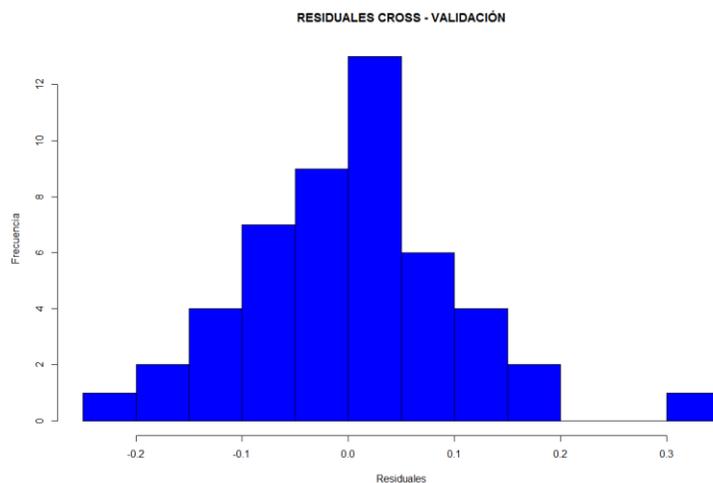
**Tabla 42.**

*Parámetros del modelo Potencia ajustado – Cokriging universal al 15% de datos*

Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	2.019291e-08		
Simple (Ond-egm)	8.035252e-07	35000	1.56
Cruzado	-2.203268e-08		

Como se puede observar en la Figura 111, el cross – semivariograma presenta valores negativos al igual que los casos con 16 y 33.

Finalmente se procedió a realizar la validación cruzada con el modelo autorizado, cuyo histograma presentó un intervalo de los errores de [-0.19;0.19] metros al 95% de confianza.



**Figura 112.** Histograma de validación cruzada con 49 datos

**Tabla 43.**

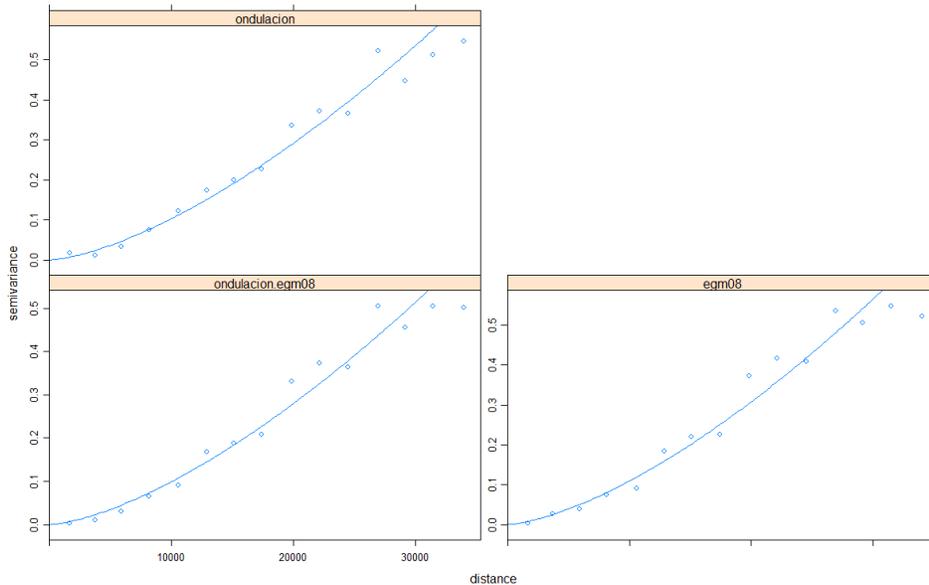
*Estadísticos de la validación cruzada con Cokriging universal al 15% de datos*

Parámetro	ME	RMSE	RSR
Valor	0.0026	0.0991	0.0374

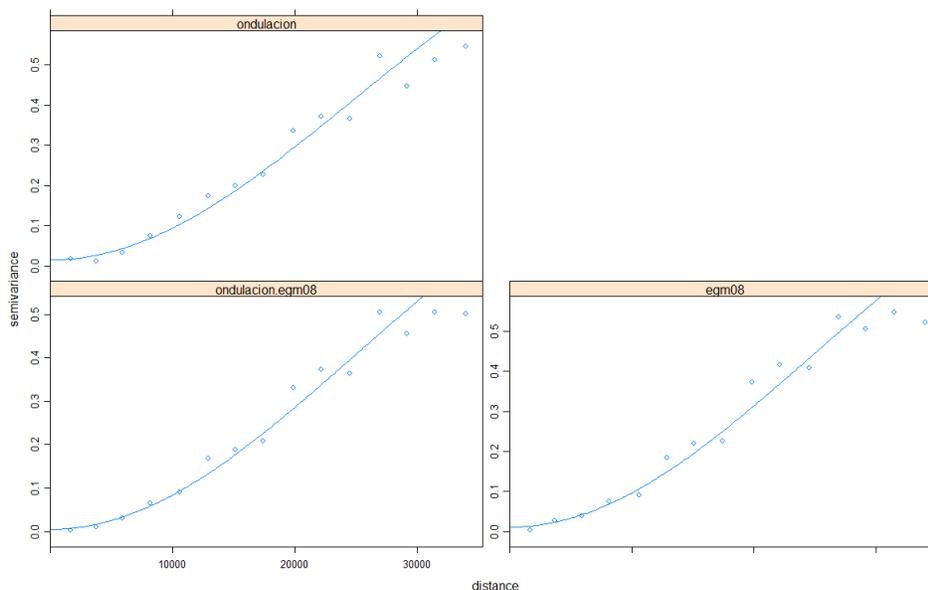
Como se muestra en la Tabla 43, el RMSE es más bajo que los anteriores mostrados, lo cual ya muestra un posible resultado óptimo.

### 4.6.3.3 Cokriging residual

De forma similar a como se trabajó con 16 y 33 datos, se obtuvieron los residuales a partir de la locación geográfica de las muestras, para posteriormente generar el semivariograma cruzado de los residuales y ajustarlos a un modelo teórico autorizado, que para 49 puntos fueron dos modelos lo que cumplieron con el MCL, siendo el potencia y gaussiano. El ajuste de los semivariogramas cruzados se muestra en las siguientes figuras:



**Figura 113.** Cross – semivariograma ajustado al modelo Potencia



**Figura 114.** Cross – semivariograma ajustado al modelo Gaussiano

Los parámetros de los modelos ajustados son los siguientes:

**Tabla 44.**

*Parámetros del modelo Potencia ajustado – Cokriging residual al 15% de datos*

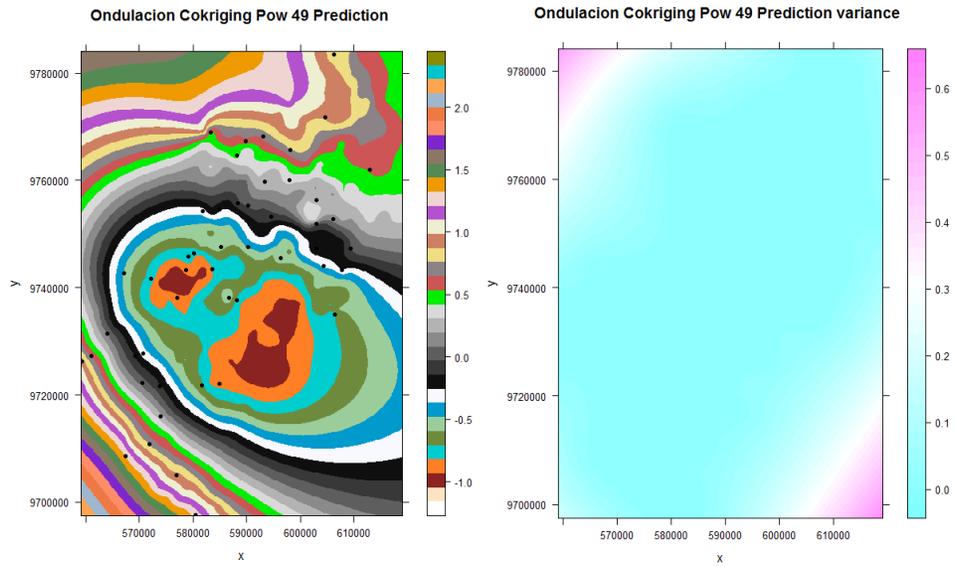
Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	1.031908e-07		
Simple (Ond-egm)	1.090840e-07	35000	1.56
Cruzado	9.891425e-08		

**Tabla 45.**

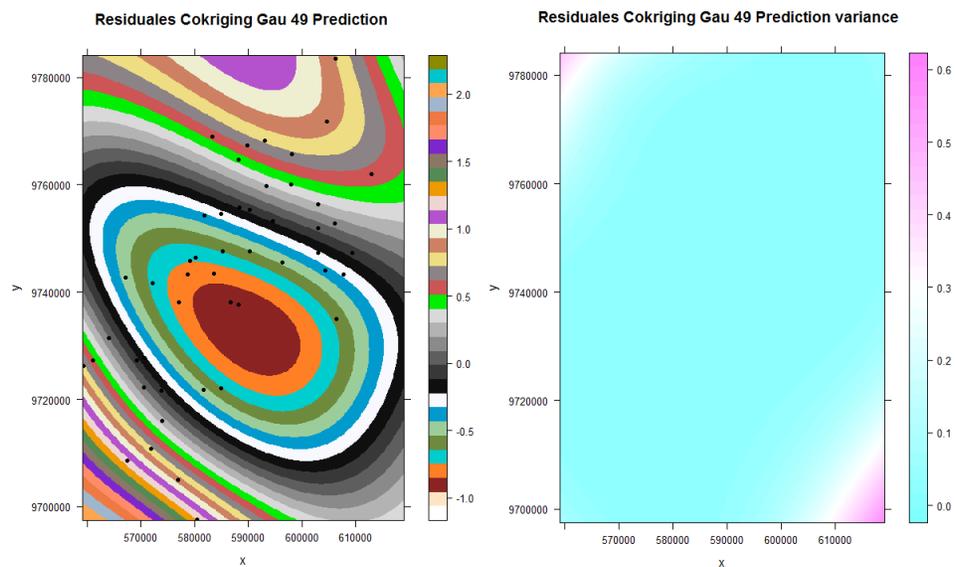
*Parámetros del modelo Gaussiano ajustado – Cokriging residual al 15% de datos*

Semivariograma	Sill	Rango	Nugget
Simple (Ond-real)	1.01182418		0.01470078
Simple (Ond-egm)	1.09029888	35000	0.01077834
Cruzado	1.00970887		0.00477000

Con los modelos ajustados se procedió a predecir los valores de los residuales sobre la grilla de puntos con la resolución calculada, cuyos resultados se pueden previsualizar a continuación:

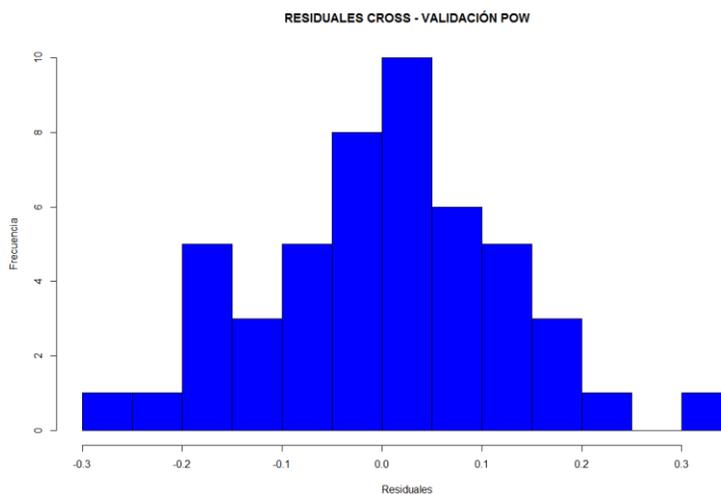


**Figura 115.** Visualización del modelo Potencia de predicción y varianza de residuales

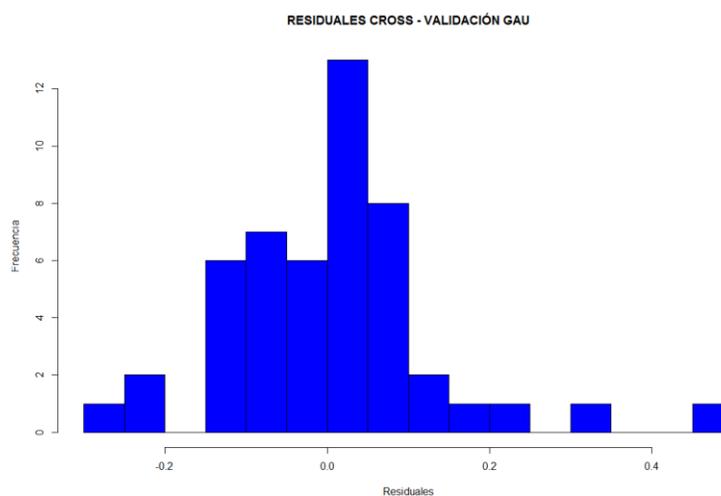


**Figura 116.** Visualización del modelo Gaussiano de predicción y varianza de residuales

Finalmente, se realizó la cross – validación de los modelos cuyos resultados se muestran en los histogramas y tablas con los indicadores estadísticos que se muestran a continuación:



**Figura 117.** Histograma de validación cruzada del modelo Potencia con 49 datos



**Figura 118.** Histograma de validación cruzada del modelo Gaussiano con 49 datos

**Tabla 46.**

*Estadísticos de la validación cruzada con Cokriging residual al 15% de datos*

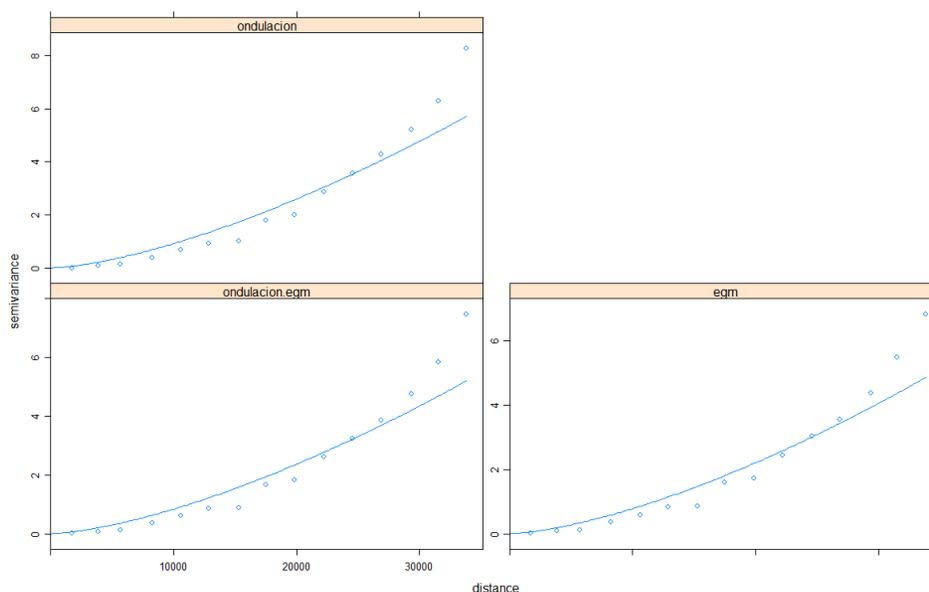
Modelo\Parámetro	ME	RMSE	RSR
Potencia	0.0031	0.1222	0.0462
Gaussiano	0.0020	0.1269	0.0479

De los histogramas se puede observar que el modelo potencia es ligeramente mejor que el gaussiano, cuyo rango de valores se encuentra en  $[-0.24;0.24]$  metros al 95% de confianza, lo cual se verifica con los estadísticos calculados.

#### 4.6.4 Obtención, predicción y validación del modelo con 20% de datos

##### 4.6.4.1 Cokriging ordinario

Con la muestra de 66 puntos, se procedió a generar el semivariograma cruzado empírico, para seguir con el ajuste del modelo autorizado, que al igual que los otros casos de Cokriging ordinario, fue el modelo potencia el único que cumplió con el modelo de correionalización lineal. El semivariograma ajustado se muestra a continuación:



**Figura 119.** Cross – semivariograma ajustado al modelo Potencia

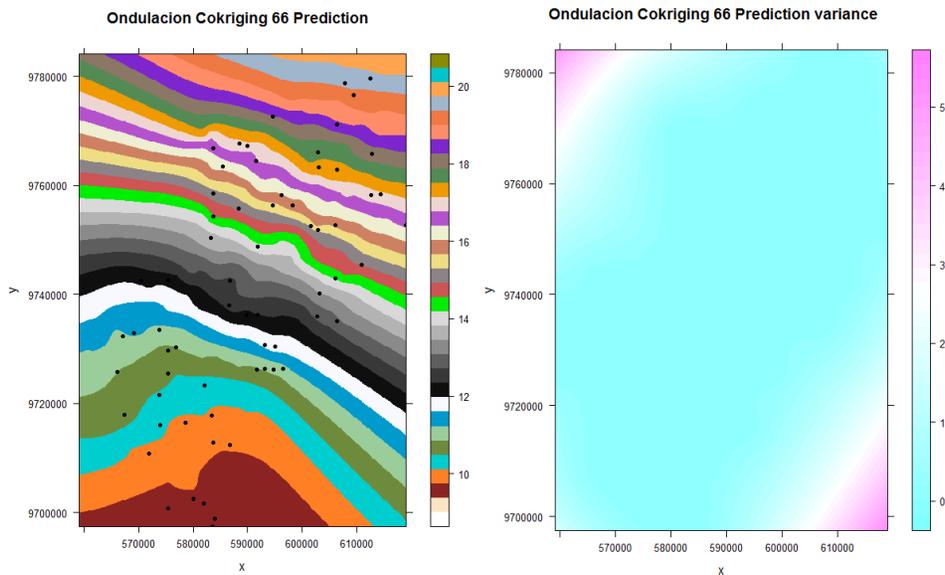
Los parámetros ajustados del modelo se muestran en la Tabla 47.

**Tabla 47.**

*Parámetros del modelo Potencia ajustado – Cokriging ordinario al 20% de datos*

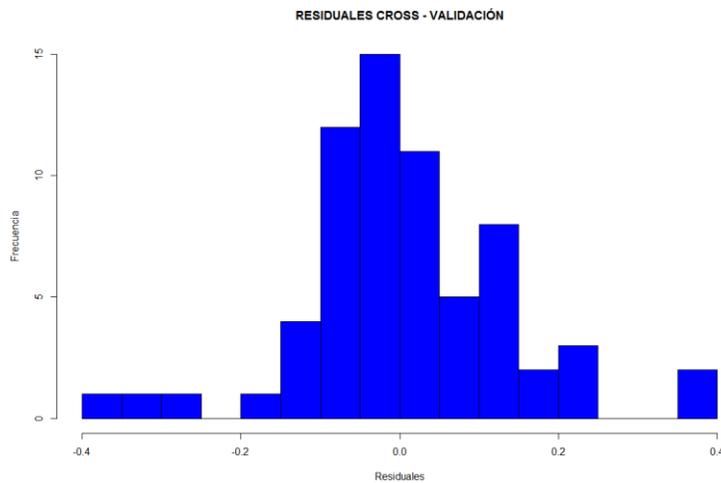
Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	9.196783e-07		
Simple (Ond-egm)	7.791906e-07	35000	1.58
Cruzado	8.375929e-07		

El siguiente paso fue generar el modelo de predicción de la ondulación geoidal sobre la grilla con resolución definida, cuyo resultado se muestra en la Figura 120.



**Figura 120.** Visualización del modelo de predicción y varianza de la ondulación geoidal

A continuación se procedió a realizar la validación cruzada, mostrando el resultado en el histograma con un intervalo de los errores de  $[-0.25;0.25]$  metros al 95% de confianza:



**Figura 121.** Histograma de validación cruzada con 66 datos

Los indicadores estadísticos que fueron calculados de la validación cruzada, se muestran en la Tabla 48:

**Tabla 48.**

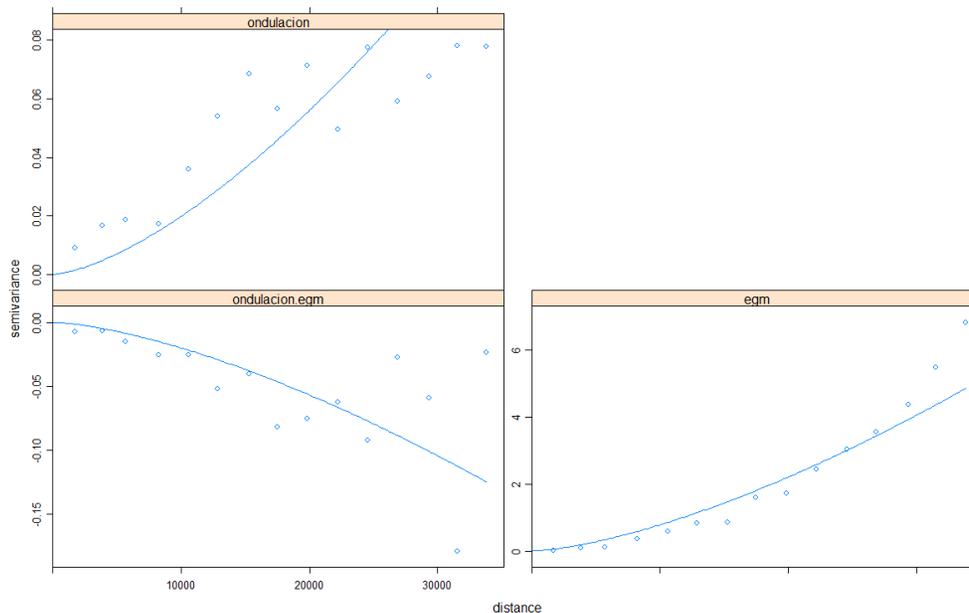
*Estadísticos de la validación cruzada con Cokriging ordinario al 20% de datos*

Parámetro	ME	RMSE	RSR
Valor	0.0061	0.1289	0.0421

#### 4.6.4.2 Cokriging universal

En el caso universal, se procedió a obtener la función de la tendencia al igual que los casos anteriores, para posteriormente generar el semivariograma empírico y ajustarlo a un modelo teórico, que de forma similar a los demás casos de este método, fue únicamente el modelo potencia.

El semivariograma cruzado ajustado, se muestra en la Figura 122:



**Figura 122.** Cross – semivariograma ajustado al modelo Potencia

Los parámetros del modelo autorizado son los siguientes:

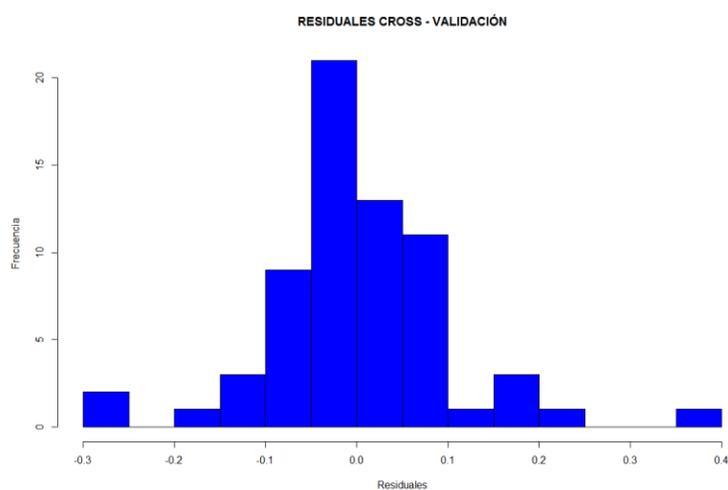
**Tabla 49.**

*Parámetros del modelo Potencia ajustado – Cokriging universal al 20% de datos*

Semivariograma	Nugget	Rango	Parámetro $\theta$
Simple (Ond-real)	1.982494e-08		
Simple (Ond-egm)	7.791906e-07	35000	1.58
Cruzado	-2.011451e-08		

Como se mostró en los casos con 16, 33 y 49 puntos muestrales, el semivariograma cruzado posee valores negativos, indicando un comportamiento indirecto entre variables, es decir, a medida que una aumenta, la otra disminuye.

Con el modelo autorizado ajustado se procedió predecir los valores de ondulación geoidal sobre la grilla de predicción con la resolución calculada, para finalmente realizar la validación cruzada, con sus respectivos residuales del proceso de validación mostrado en el histograma e indicadores estadísticos de los mismos, que son mostrados a continuación:



**Figura 123.** Histograma de validación cruzada con 66 datos

**Tabla 50.**

*Estadísticos de la validación cruzada con Cokriging universal al 20% de datos*

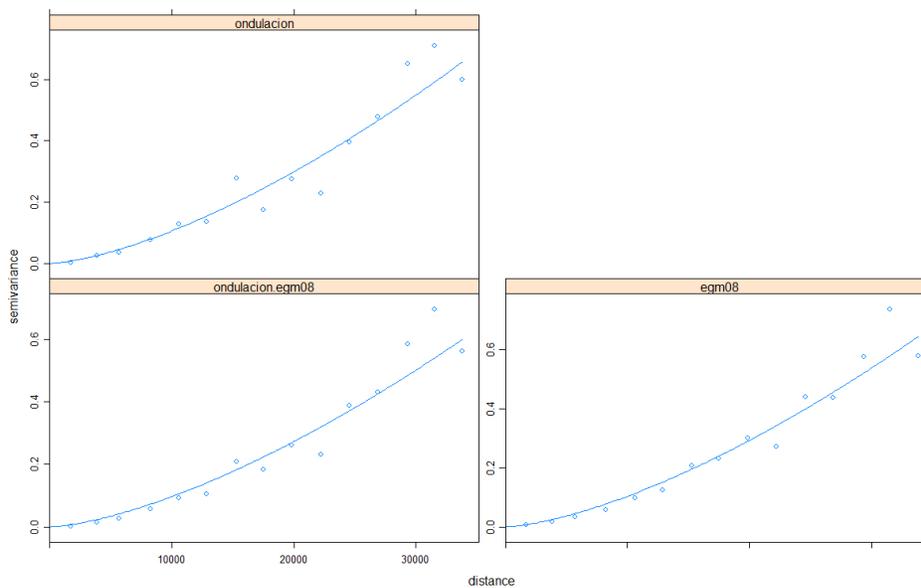
Parámetro	ME	RMSE	RSR
Valor	0.0046	0.0999	0.0326

Como se puede apreciar en el histograma, los errores de validación cruzada se agrupan en el intervalo de  $[-0.19;0.19]$  metros al 95% de confianza, lo cual es un excelente resultado.

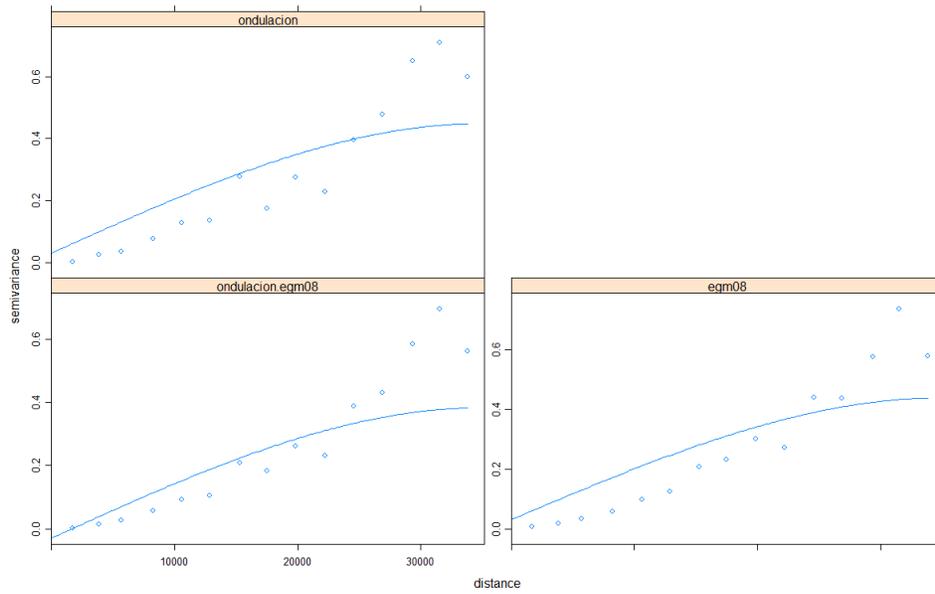
#### 4.6.4.3 Cokriging residual

En primera instancia se aplicó una regresión lineal múltiple al igual que en los casos anteriores, para posteriormente general el semivariograma cruzado empírico, que en este caso, presentó 3 modelos que cumplieron con el MCL, estos son: potencia, gaussiano y esférico.

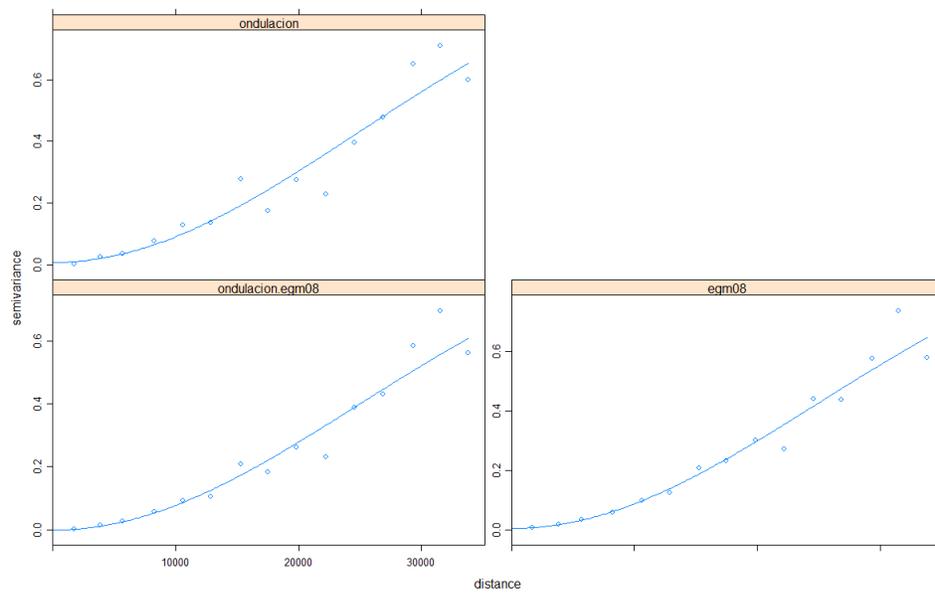
Los semivariogramas cruzados ajustados para cada uno de estos se presentan en las siguientes figuras:



**Figura 124.** Cross – semivariograma ajustado al modelo Potencia



**Figura 125.** Cross – semivariograma ajustado al modelo Esférico



**Figura 126.** Cross – semivariograma ajustado al modelo Gaussiano

Los parámetros de los modelos ajustados se presentan en las Tablas 51, 52 y 53.

**Tabla 51.***Parámetros del modelo Potencia ajustado – Cokriging residual al 20% de datos*

<b>Semivariograma</b>	<b>Nugget</b>	<b>Rango</b>	<b>Parámetro <math>\theta</math></b>
Simple (Ond-real)	1.057688e-07		
Simple (Ond-egm)	1.035710e-07	35000	1.58
Cruzado	9.662044e-08		

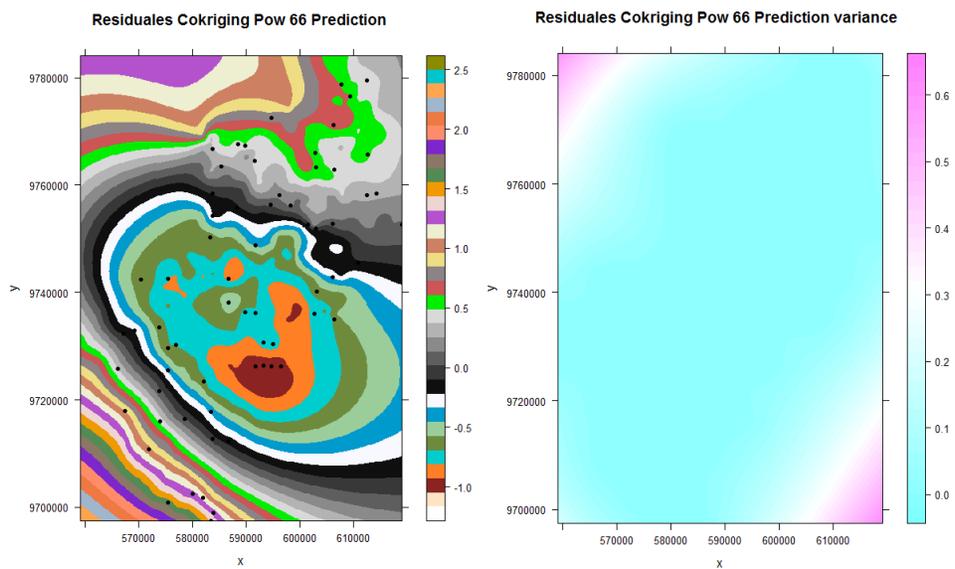
**Tabla 52.***Parámetros del modelo Esférico ajustado – Cokriging residual al 20% de datos*

<b>Semivariograma</b>	<b>Sill</b>	<b>Rango</b>	<b>Nugget</b>
Simple (Ond-real)	0.41900392		0.03004537
Simple (Ond-egm)	0.40845954	35000	0.03004537
Cruzado	0.41369813		-0.03004537

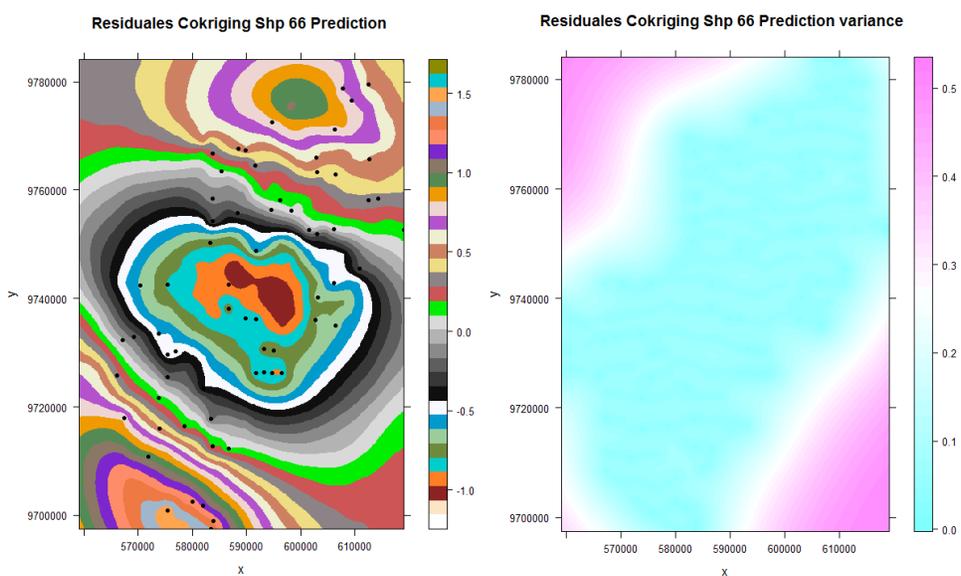
**Tabla 53.***Parámetros del modelo Gaussiano ajustado – Cokriging residual al 20% de datos*

<b>Semivariograma</b>	<b>Sill</b>	<b>Rango</b>	<b>Nugget</b>
Simple (Ond-real)	1.064960415		0.006931560
Simple (Ond-egm)	1.057592850	35000	0.003788503
Cruzado	1.006330097		-0.001668913

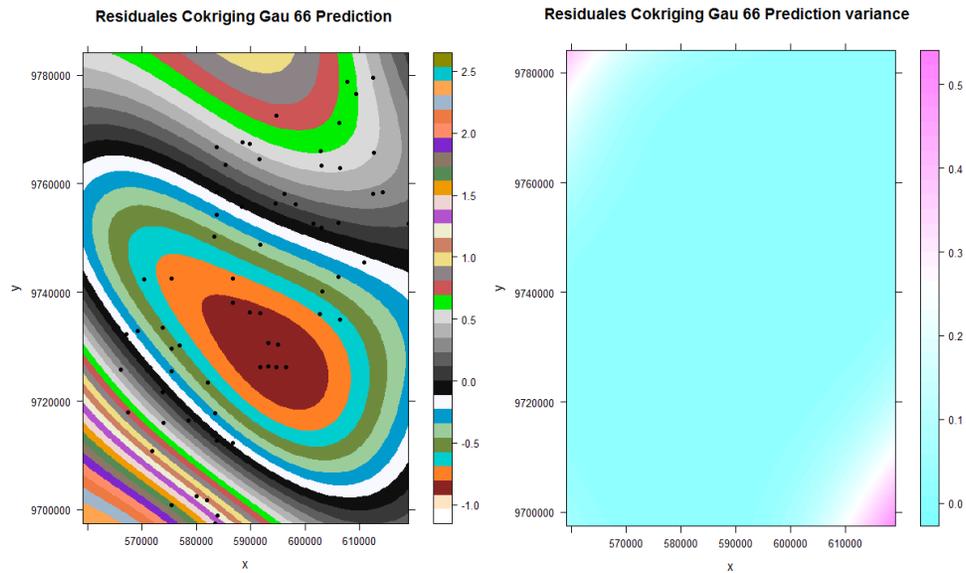
Con los modelos ajustados, se procedió a realizar la estimación de los residuales de la ondulación geoidal sobre la grilla con resolución de 135 metros establecida previamente. La previsualización del modelo de residuales para potencia, esférico y gaussiano se muestran en las Figuras 127, 128 y 129 respectivamente.



*Figura 127.* Visualización del modelo Potencia de predicción y varianza de residuales

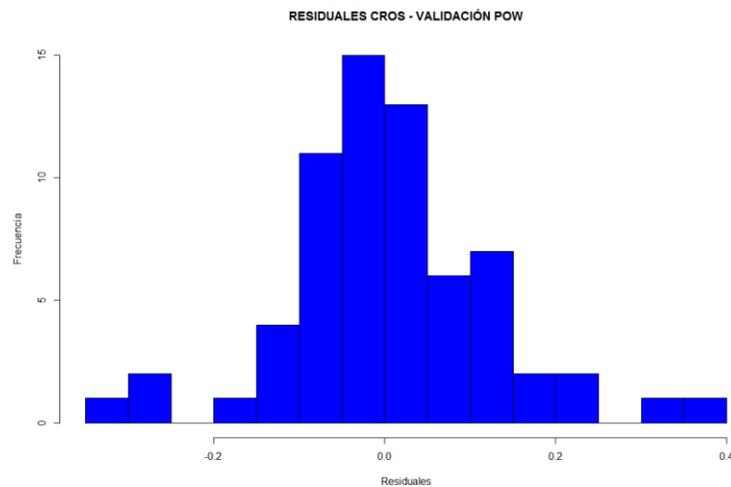


*Figura 128.* Visualización del modelo Esférico de predicción y varianza de residuales

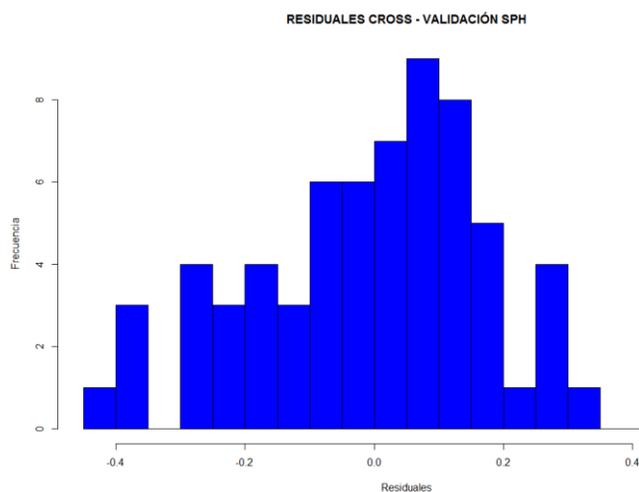


**Figura 129.** Visualización del modelo Gaussiano de predicción y varianza de residuales

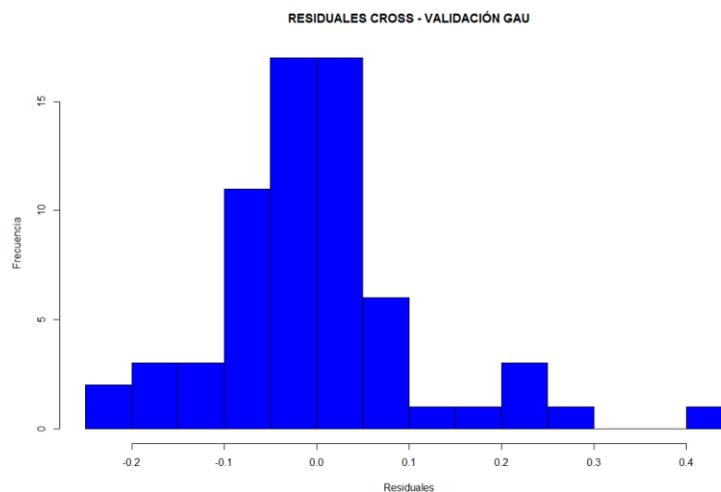
Finalmente, se realizó la validación cruzada con los modelos autorizados, cuyos resultados se presentan en los histogramas siguientes, y en las tablas con los indicadores estadísticos de cada uno de ellos.



**Figura 130.** Histograma de validación cruzada del modelo Potencia con 66 datos



**Figura 131.** Histograma de validación cruzada del modelo Esférico con 66 datos



**Figura 132.** Histograma de validación cruzada del modelo Gaussiano con 66 datos

**Tabla 54.**

*Estadísticos de la validación cruzada con Cokriging residual al 20% de datos*

Modelo\Parámetro	ME	RMSE	RSR
Potencia	0.0056	0.1191	0.0389
Esférico	0.0019	0.1880	0.0614
Gaussiano	0.0002	0.1121	0.0366

En base a los histogramas y los indicadores estadísticos mostrados, se observa que el modelo gaussiano es mejor, siendo el intervalo de valores en los que se encuentra los residuales entre  $[-0.22; 0.22]$  metros al 95% de confianza.

## 4.6.5 Obtención, predicción y validación del modelo con 299 de datos

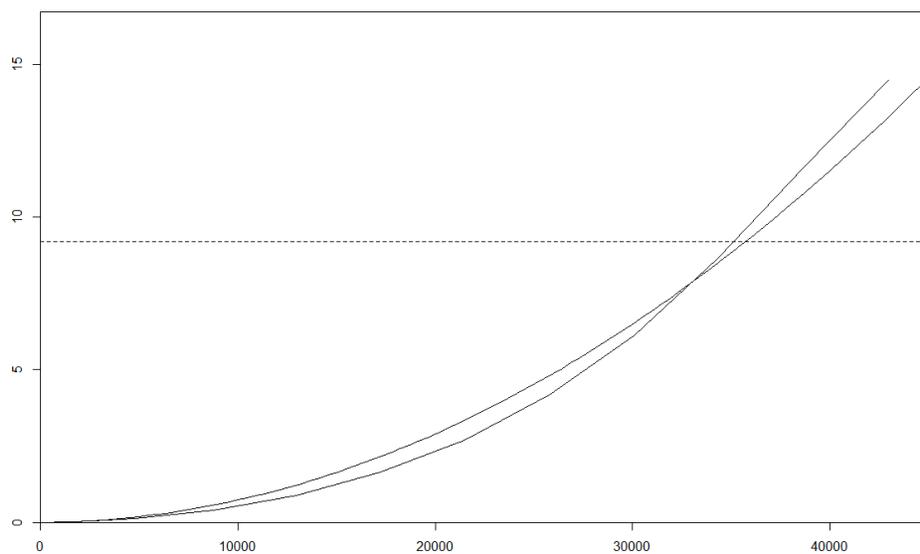
### 4.6.5.1 Kriging ordinario

En el caso de Kriging, se utilizó el mismo software R, con el paquete RGeostats (Renard, Ors, & Desassis, 2017), Rcpp (Eddelbuettel, *et al.*, 2018) y sp (Pebesma, *et al.*, 2018).

Kriging ordinario fue realizado con una muestra de 299 puntos escogidos aleatoriamente, y como primer paso, al igual que en Cokriging, fue generar el semivariograma empírico y ajustarlo a un modelo teórico. Según los datos mostrados en la sección 4.4.6 de análisis exploratorio de datos, los posibles modelos autorizados son no acotados, entre los que se destacan K – Bessel, J – Bessel, potencia. Recordando lo que se mencionó con anterioridad, en el paquete RGeostats existen 28 modelos, por lo cual el análisis del ajuste de modelos teóricos fue más extenso.

A pesar de contar con una opción de análisis de varios modelos a la vez, se decidió realizar un estudio individual con los posibles modelos a ajustar, ya que se dio el caso, que la función de “automodelaje” no siempre realiza el mejor ajuste, sobre todo cuando existe más de tres modelos posibles.

El modelo que se ajustó al semivariograma empírico fue el potencia, coincidiendo con lo realizado en Cokriging ordinario. La Figura 133 muestra el semivariograma ajustado.



**Figura 133.** Semivariograma ajustado con el modelo Potencia

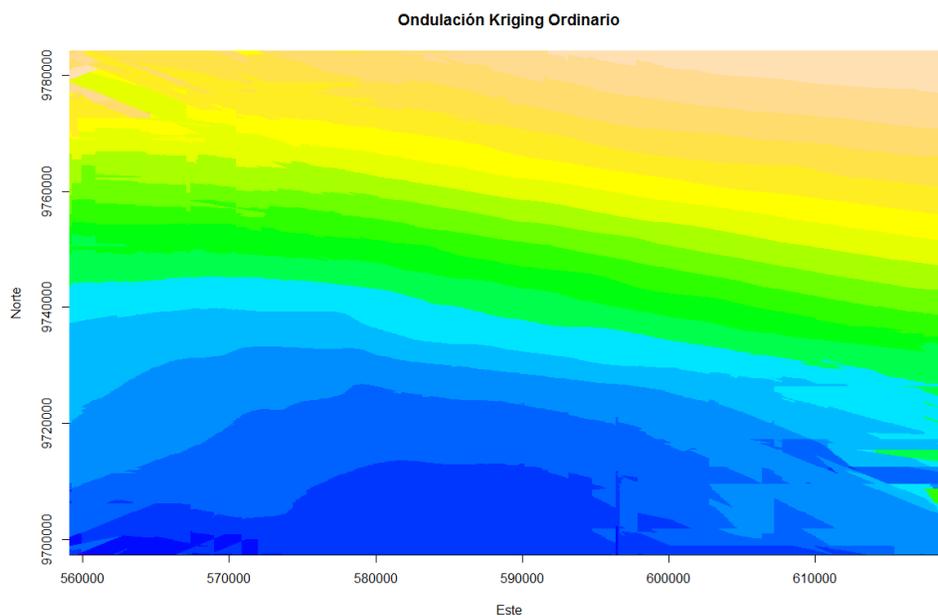
Los parámetros del modelo ajustado se presentan en la Tabla 55 a continuación:

**Tabla 55.**

*Parámetros del modelo Potencia ajustado – Kriging ordinario (299 datos)*

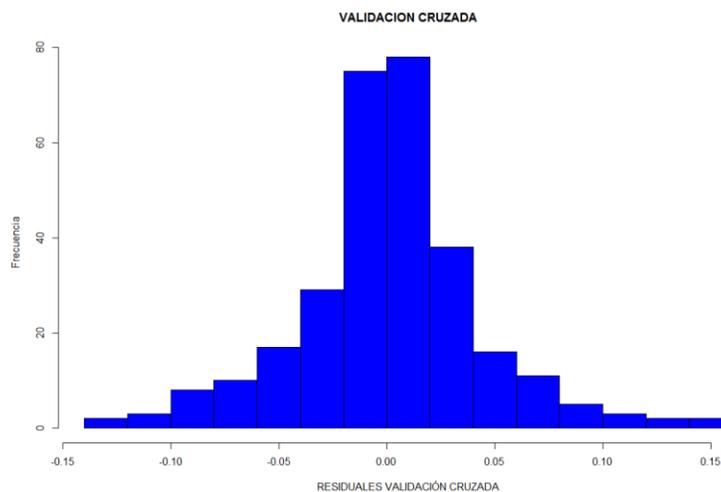
Semivariograma	Sill	Rango	Parámetro $\theta$
Ond-real	13.3	42958	1.99

Con el modelo ajustado se procedió a estimar el valor de ondulación geoidal sobre una grilla con resolución definida para toda la zona de estudio, cuyo resultado se puede previsualizar en la Figura 134:



**Figura 134.** Visualización del modelo de predicción de la ondulación geoidal

La validación cruzada se realizó con los mismos parámetros establecidos para Cokriging; en este caso el histograma correspondiente a los residuales (errores de estimación) es el siguiente:



**Figura 135.** Histograma de validación cruzada del modelo Potencia con 299 datos

Los estadísticos de la validación cruzada se muestran a continuación:

**Tabla 56.**  
*Estadísticos de la validación cruzada con Kriging ordinario (299 datos)*

Parámetro	ME	RMSE	RSR
Valor	0.0011	0.0421	0.0139

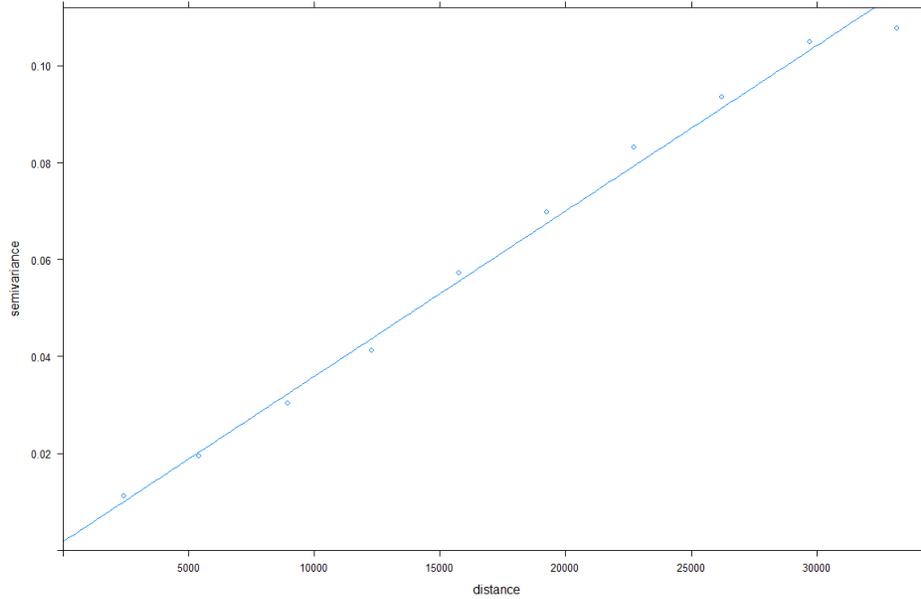
En el histograma mostrado, se tiene que los errores de predicción se encuentran en un rango de [-0.08;0.08] metros al 95% de confianza. Además, sus indicadores estadísticos también son más pequeños que los mostrados en Cokriging, esto se debe claramente a la utilización de más puntos para generar el modelo de predicción, sin embargo, el factor más importante para validar los modelos generados, será la comparación con datos que no fueron usados en la predicción, como se verá más adelante.

#### 4.6.5.2 Kriging universal

En este método se trabajó de forma similar a Cokriging universal, es decir, en primer lugar se obtuvo la función de la ondulación geoidal en base a la variable auxiliar, para posteriormente construir el semivariograma empírico y ajustarlo a un modelo teórico.

Los modelos autorizados para Kriging universal fueron 6: Bessel, circular, exponencial, gaussiano, pentaesférico y esférico; con dichos modelos se procedió a ajustar el semivariograma empírico y realizar la predicción sobre la grilla de puntos a una escala definida. A continuación se aplicó una validación cruzada y fueron calculados los indicadores estadísticos para cada modelo usado.

Como no es de mayor interés Kriging para el presente estudio, se presenta únicamente el modelo teórico que mejores resultados mostró, que fue el circular. El semivariograma ajustado fue el siguiente:



**Figura 136.** Semivariograma ajustado con el modelo Circular

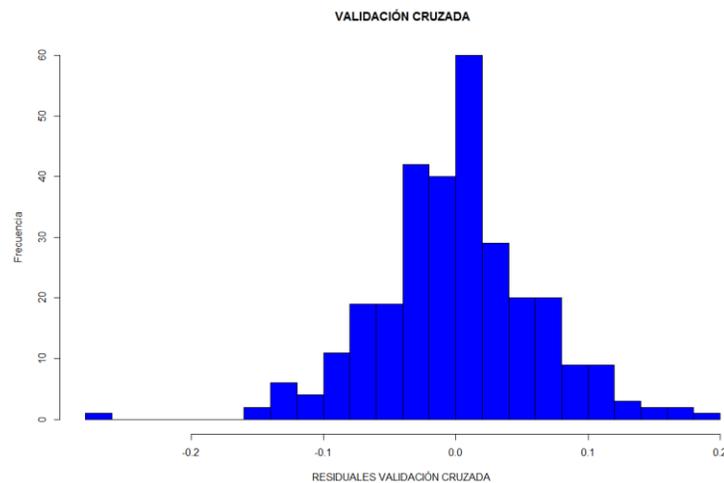
Los parámetros del modelo ajustado fueron los siguientes:

**Tabla 57.**

*Parámetros del modelo Circular ajustado – Kriging universal (299 datos)*

Semivariograma	Sill	Rango	Nugget
Ond-real	5.596234620	2088318	0.001804561

La validación cruzada se muestra en su histograma de residuales en la figura siguiente:



**Figura 137.** Histograma de validación cruzada del modelo Circular con 299 datos

Según el histograma de validación cruzada, los errores de la validación cruzada se encuentran en un intervalo de  $[-0.11;0.11]$  metros al 95% de confianza.

En la Tabla 58 se presentan los estadísticos calculados para todos los modelos autorizados en Kriging universal, siendo el circular el que mejor resultado alcanzó.

**Tabla 58.**

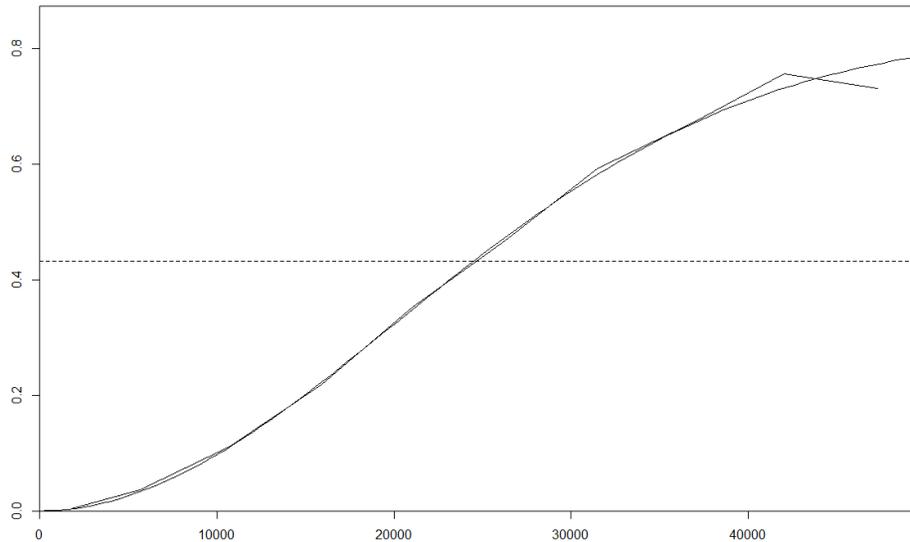
*Estadísticos de la validación cruzada con Kriging universal (299 datos)*

<b>Modelo\Parámetro</b>	<b>ME</b>	<b>RMSE</b>	<b>RSR</b>
Bessel	0.00019	0.07190	0.02368
Circular	0.00061	0.06023	0.01984
Exponencial	0.00053	0.06527	0.02150
Gaussiano	6.28632e-06	0.09381	0.03090
Pentaesférico	0.00061	0.06031	0.01986
Esférico	0.00061	0.06027	0.01985

#### 4.6.5.3 Kriging residual

Al igual que Cokriging residual, se procedió a utilizar una regresión lineal múltiple para obtener los residuales en función de las coordenadas geográficas de los puntos de muestreo y así trabajar con la tendencia subyacente de la variable en cuestión.

Como sucedió en Kriging universal, existió más de un modelo autorizado, que en este caso fueron los siguientes: wendland2, wendland1, estable, gaussiano, exp2dfact, cúbico, cauchy y cardinalSine. De estos, el que mejor se ajustó al semivariograma empírico y por ende el que se muestra en las siguientes figuras y tablas fue el modelo estable.



**Figura 138.** Semivariograma ajustado con el modelo Estable

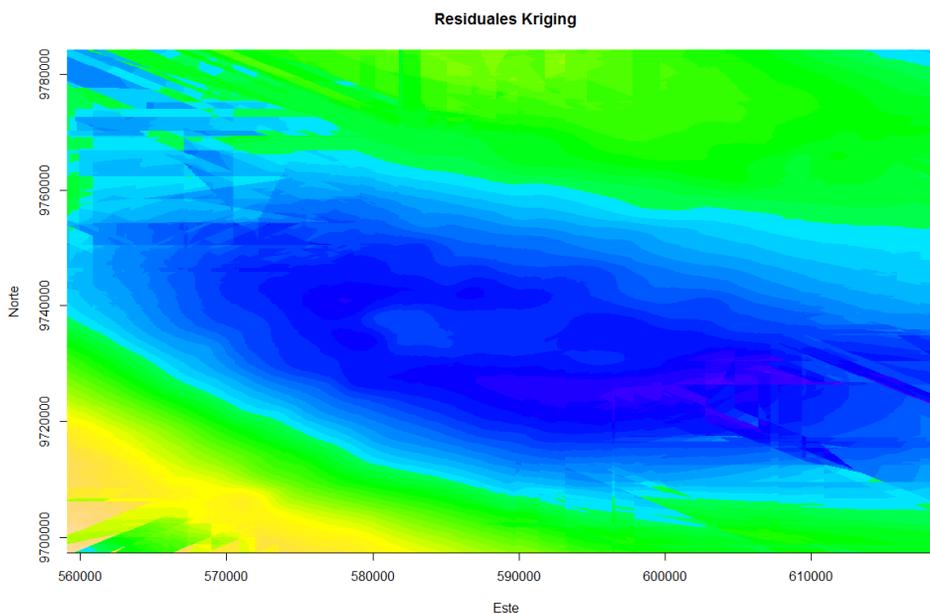
Los parámetros del modelo ajustado se muestran en la Tabla 59:

**Tabla 59.**

*Parámetros del modelo Estable ajustado – Kriging residual (299 datos)*

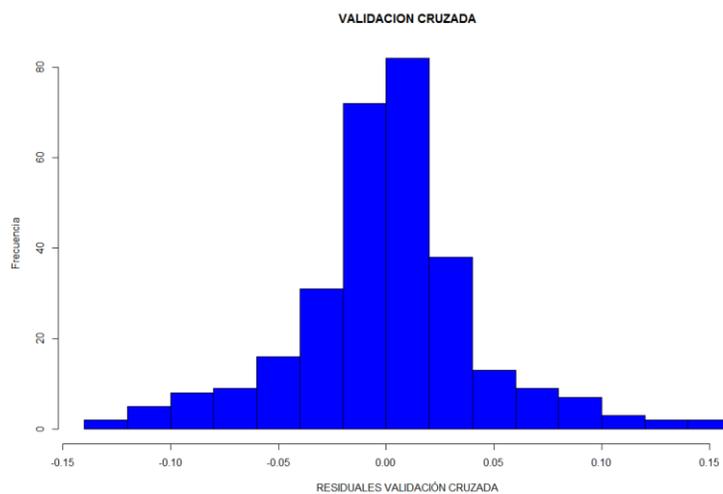
<b>Semivariograma</b>	<b>Sill</b>	<b>Rango</b>	<b>Parámetro</b>
Ond-real	0.827	49696	1.98

La predicción realizada con el modelo estable, se puede previsualizar en la Figura 139:



**Figura 139.** Visualización del modelo de predicción de residuales de la ondulación geoidal

Con el modelo ajustado, se procedió a realizar la validación cruzada con los errores de predicción (residuales), cuyo histograma se muestra a continuación:



**Figura 140.** Histograma de validación cruzada del modelo Estable con 299 datos

En el histograma anterior, se observó que los valores de residuales de la validación se encuentran en un rango entre  $[-0.08;0.08]$  metros al 95% de confianza, algo similar a los dos casos de Kriging analizados (ordinario y universal).

De la misma forma con que se vino trabajando en los demás métodos, se calculó indicadores estadísticos para facilitar la toma de decisiones, cuyos resultados se muestran en la Tabla 60 de todos los modelos autorizados que se analizó, siendo el modelo estable el mejor.

**Tabla 60.**

*Estadísticos de la validación cruzada con Kriging residual (299 datos)*

<b>Modelo\Parámetro</b>	<b>ME</b>	<b>RMSE</b>	<b>RSR</b>
Wendland2	0.00094	0.05628	0.01854
Wendland1	0.00083	0.04642	0.01529
Estable	0.00053	0.04295	0.01415
Gaussiano	0.00358	0.09922	0.03268
Exp2dfact	0.00358	0.09922	0.03268
Cúbico	0.00085	0.04650	0.01532
Cauchy	0.00417	0.09851	0.03245
CardinalSine	0.00534	0.13247	0.04363

#### **4.7 Validación del modelo de ondulación geoidal**

En este apartado, se realizó la comprobación de la precisión del modelo, es decir, la bondad de la predicción realizada con los distintos métodos de Cokriging y Kriging. Para ello, se resume en la siguiente tabla el método y modelo teórico usado para la generación del modelo de predicción de ondulación geoidal en la zona de estudio.

**Tabla 61.**  
*Resumen de métodos y modelos teóricos usados*

<b>Método de predicción</b>	<b>N° de muestras</b>	<b>Modelo teórico</b>	<b>RSR</b>	<b>RMSE</b>
Cokriging ordinario	16	Potencia	0.0438	0.1416
Cokriging residual	16	Potencia	0.0498	0.1609
Cokriging universal	16	Potencia	0.0434	0.1403
Cokriging ordinario	33	Potencia	0.0495	0.1327
Cokriging residual	33	Gaussiano	0.0400	0.1073
Cokriging universal	33	Potencia	0.0460	0.1234
Cokriging ordinario	49	Potencia	0.0514	0.1361
Cokriging residual	49	Potencia	0.0462	0.1222
Cokriging universal	49	Potencia	0.0374	0.0991
Cokriging ordinario	66	Potencia	0.0421	0.1289
Cokriging residual	66	Gaussiano	0.0366	0.1121
Cokriging universal	66	Potencia	0.0326	0.0999
Kriging ordinario	299	Potencia	0.0139	0.0421
Kriging residual	299	Estable	0.0141	0.0602
Kriging universal	299	Circular	0.0198	0.0429

Como se puede observar en la Tabla 61, el modelo teórico que prevaleció para el modelamiento de la ondulación geoidal fue el potencia, y esto indica que el proceso de forma general, corresponde a la hipótesis intrínseca. Además, en Cokriging se puede notar que en todos los casos del predictor ordinario, el modelo ajustado fue únicamente el potencia en concordancia con la no estacionariedad de la variable, mientras que los dos otros predictores usados para analizar tendencia, que son, universal y residual, si presentaron modelos autorizados acotados, aunque el modelo potencia fue el que mejores resultados mostró, menos en dos casos.

Para demostrar que el método Cokriging genera un mejor modelo de predicción de ondulación, se procedió a validar los distintos modelos de la Tabla 61, con puntos muestrales que no fueron utilizados en el modelamiento según cada caso de estudio, es decir, el modelo con 16 muestras

usadas se validará con 316 puntos, el de 33 muestras con 299 puntos, el de 49 muestras con 283 puntos, el de 66 muestras con 266 puntos y el de 299 muestras con 33 puntos.

Cabe indicar que la selección de las submuestras se hizo de forma aleatoria para cada una de estas y de forma individual tomando del total de puntos, en otras palabras, cada submuestra presenta datos distintos (algunos puntos obviamente se repiten), y no considerando los primeros 16 puntos más 17 para completar las 33 muestras y así sucesivamente, sino de forma individual en cada submuestra, evitando así algún sesgo en los datos.

En éste acápite se presentan los resultados de la validación de cada modelo generado según lo indicado en la Tabla 61, para lo cual se calcularon indicadores estadísticos que sirvieron de facilitadores en la selección del mejor o mejores modelos de predicción realizados. A más de los estadísticos usados anteriormente (ME, RMSE, RSR), se calculó el coeficiente de correlación de concordancia de Lin, con el cual se provee de una herramienta más para asegurar lo mejor posible el modelo seleccionado.

Las Tablas 62, 63, 64 ,65 y 66, muestran los resultados de la validación con puntos excluidos de la modelación en cada submuestra, con cada uno de los métodos usados para generar los modelos de predicción:

**Tabla 62.**

*Resultados de la validación del modelo generado con 16 puntos muestrales*

<b>Método</b>	<b>ME</b>	<b>RMSE</b>	<b>RSR</b>	<b>Coef. Lin</b>
Cokriging ordinario	0.0058	0.1644	0.0541	0.9985
Cokriging residual	-0.3667	0.3995	0.1315	0.9912
Cokriging universal	0.0373	0.1621	0.0533	0.9985

**Tabla 63.***Resultados de la validación del modelo generado con 33 puntos muestrales*

<b>Método</b>	<b>ME</b>	<b>RMSE</b>	<b>RSR</b>	<b>Coef. Lin</b>
Cokriging ordinario	0.0100	0.1581	0.0513	0.9986
Cokriging residual	0.6292	0.6442	0.2091	0.9786
Cokriging universal	0.0217	0.1353	0.0439	0.9990

**Tabla 64.***Resultados de la validación del modelo generado con 49 puntos muestrales*

<b>Método</b>	<b>ME</b>	<b>RMSE</b>	<b>RSR</b>	<b>Coef. Lin</b>
Cokriging ordinario	0.0006	0.1292	0.0415	0.9991
Cokriging residual	0.3784	0.3962	0.1274	0.9918
Cokriging universal	0.0258	0.1001	0.0322	0.9994

**Tabla 65.***Resultados de la validación del modelo generado con 66 puntos muestrales*

<b>Método</b>	<b>ME</b>	<b>RMSE</b>	<b>RSR</b>	<b>Coef. Lin</b>
Cokriging ordinario	-0.0191	0.1192	0.0392	0.9992
Cokriging residual	-0.0175	0.0944	0.0310	0.9995
Cokriging universal	0.0063	0.0859	0.0282	0.9996

**Tabla 66.***Resultados de la validación del modelo generado con 299 puntos muestrales*

<b>Método</b>	<b>ME</b>	<b>RMSE</b>	<b>RSR</b>	<b>Coef. Lin</b>
Kriging ordinario	0.0153	0.1278	0.0438	0.9990
Kriging residual	0.4686	0.4855	0.1665	0.9857
Kriging universal	0.0351	0.1302	0.0446	0.9989

## CAPÍTULO V

### ANÁLISIS Y DISCUSIÓN DE RESULTADOS

Como se mostró en las tablas anteriores con los indicadores estadísticos, son tres los parámetros calculados con los que se evaluó la bondad de los modelos generados, aunque cada uno de estos tiene propósitos diferentes. Por ejemplo, si bien el error medio (ME) es un indicador para evaluar el sesgo existente del modelo (debe ser cercano a cero) (Reyes, *et al.*, 2018), no es un buen criterio para identificar la precisión del modelo; la raíz del error cuadrático medio (RMSE) debe ser lo más pequeño posible ya que Kriging (o Cokriging) minimiza los errores y por ende es un indicador de la precisión, pero esto no quiere decir que dicho modelo sea siempre el correcto (Oliver & Webster, 2014); la razón de la desviación estándar del RMSE (RSR) es un criterio más robusto que el RMSE para escoger el modelo, este indicador cuanto más cercano a cero sea, será el modelo correcto a usar (Cruz, Villaseñor, López & Ortiz, 2013; Golmohammadi, Prasher, Madani & Rudra, 2014), razón por la cual se trabajó en base a este último indicador para la selección del modelo final.

De las tablas anteriores, se demostró que según los indicadores estadísticos de RMSE y RSR, los mejores modelos de predicción de la variable ondulación geoidal fueron los generados mediante Cokriging universal con 66 muestras y Cokriging universal con 49 muestras, de sus respectivas submuestras analizadas.

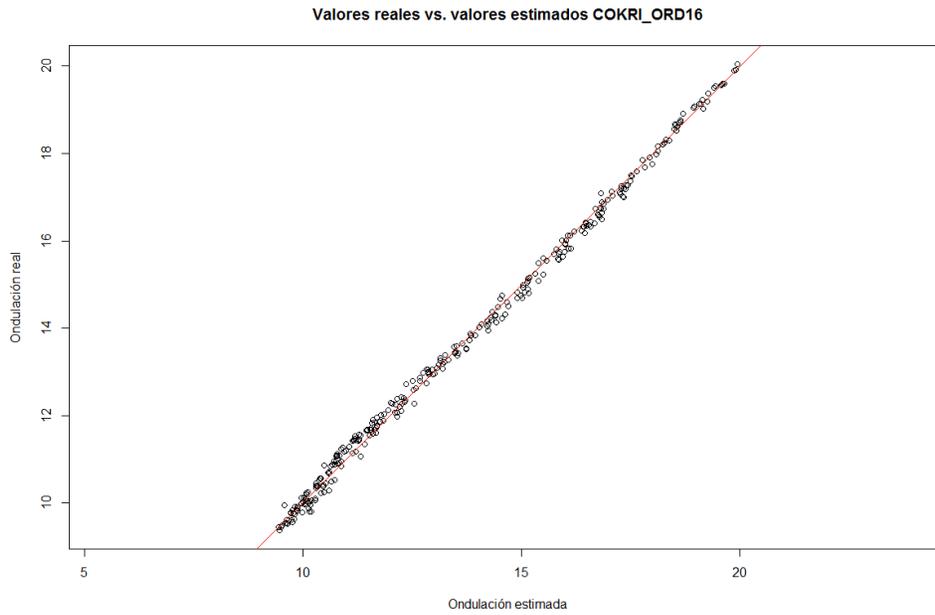
También es importante destacar, que en base a los valores de los estadísticos mostrados, en todos los casos de Cokriging, el predictor universal fue el más preciso tanto con 16, 33, 49 y 66 muestras.

En cuanto a Kriging, se obtuvo que el realizado a través del predictor ordinario fue el más preciso de los tres métodos (Kriging ordinario, residual y universal).

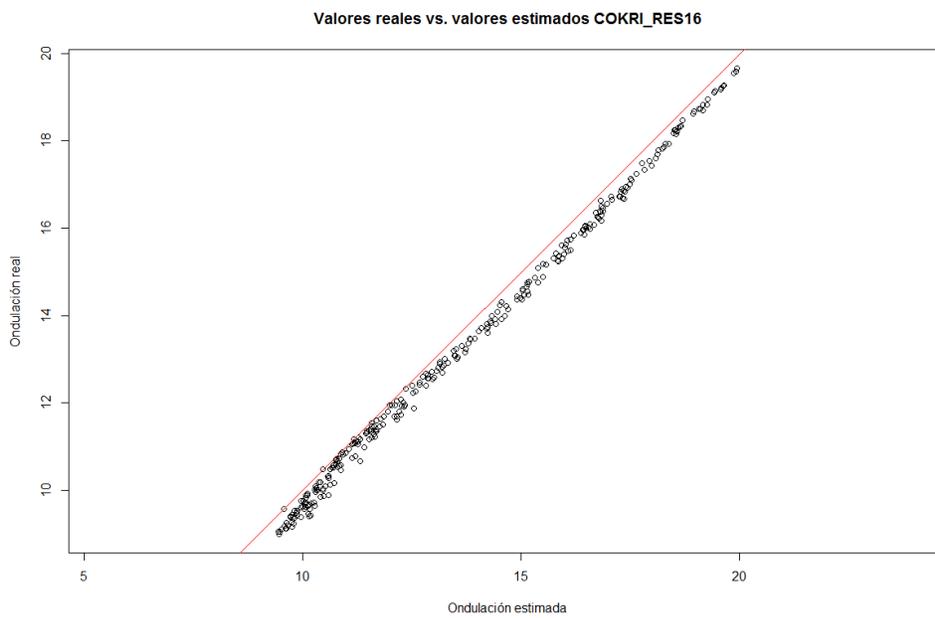
Además de estos tres indicadores estadísticos, también se calculó el coeficiente de concordancia de correlación de Lin, el cual fue muy alto en todos los casos analizados, indicando un grado de pertenencia excelente entre los valores reales y predichos garantizando una correcta modelación de la variable. Cabe mencionar que se utilizó el coeficiente de Lin en vez del coeficiente de Pearson, ya que el segundo no es apropiado cuando se evalúa la concordancia entre las cantidades de una misma variable obtenidas mediante distintas metodologías (Lin, 1989; Lin, 2000), además de que el coeficiente de concordancia se calcula con un nivel de confianza (95% usado en este análisis) (Leal, 2016) y permite conocer si existe algún tipo de error sistemático, situación que el coeficiente de Pearson no analiza y por ende su valor sería erróneamente elevado (Bland & Altman, 2007).

De los 4 casos (submuestras) analizados mediante Cokriging, se afirma que el mejor resultado de predicción fue con un conjunto de 66 puntos, los cuales si comparamos entre los resultados obtenidos con Kriging y Cokriging se puede notar claramente que los indicadores estadísticos son mejores usando una variable auxiliar para la predicción de la ondulación geoidal (ver Tabla 65 y 66), lo cual concuerda con varios estudios de distintas variables de la naturaleza en que sus predicciones son más efectivas utilizando una o varias covariables (Hengl, 2009; Mulla & McBratney, 2002; Waller & Gotway, 2004).

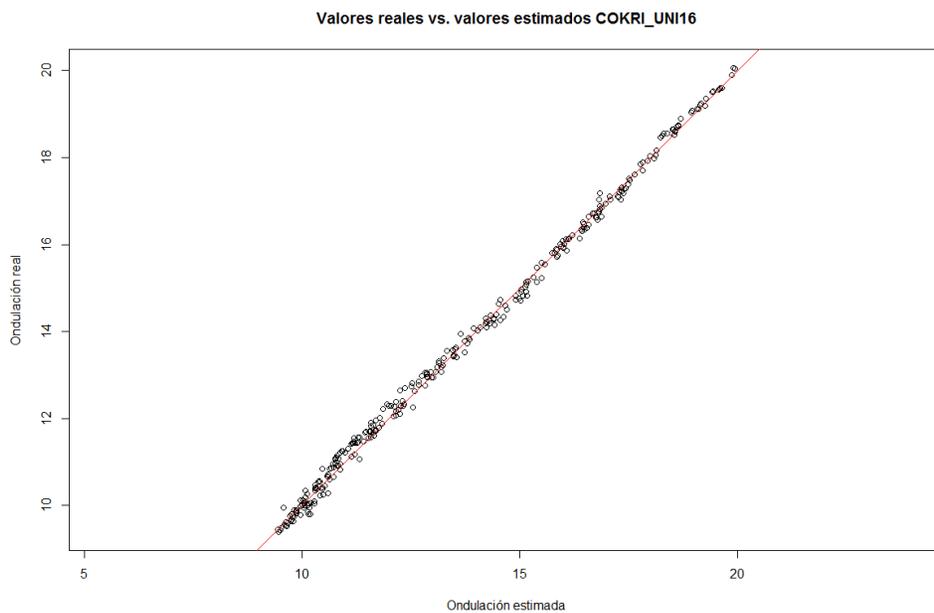
La bondad de los modelos generados se puede verificar en un gráfico de dispersión entre los valores reales de ondulación geoidal y los estimados de la variable, como se muestra en las siguientes figuras:



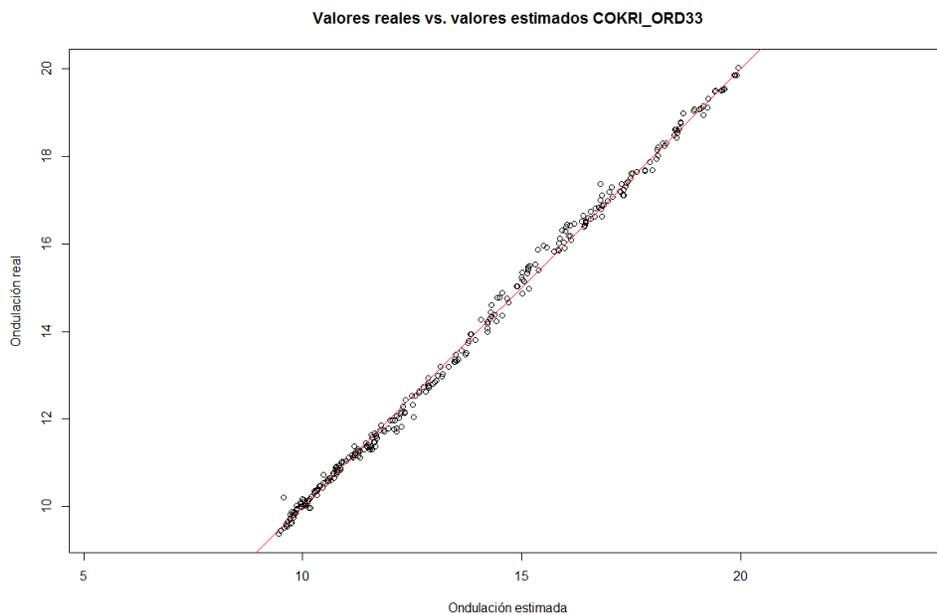
**Figura 141.** Comparación entre valores reales y predichos – Cokriging ordinario 16 puntos



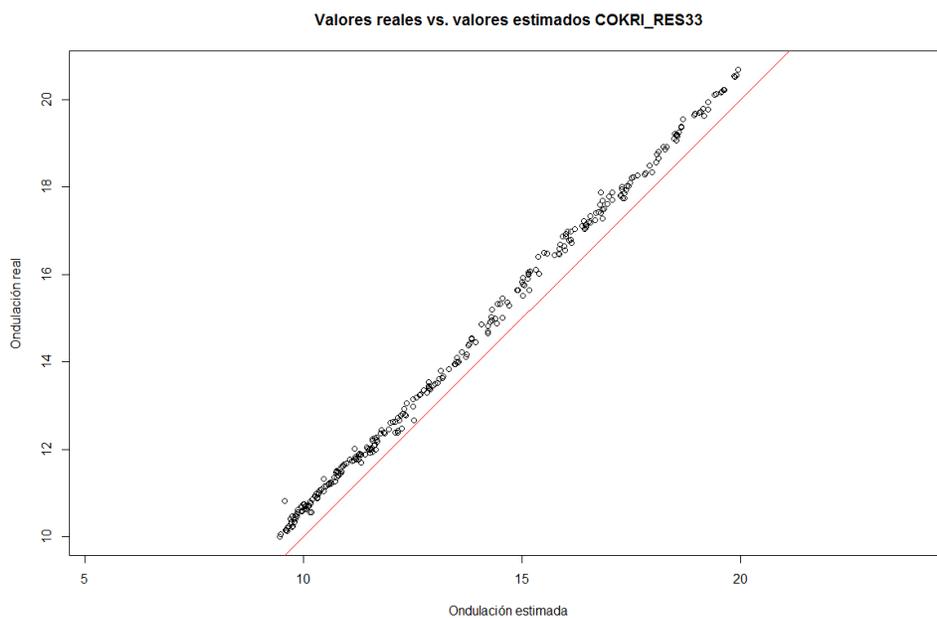
**Figura 142.** Comparación entre valores reales y predichos – Cokriging residual 16 puntos



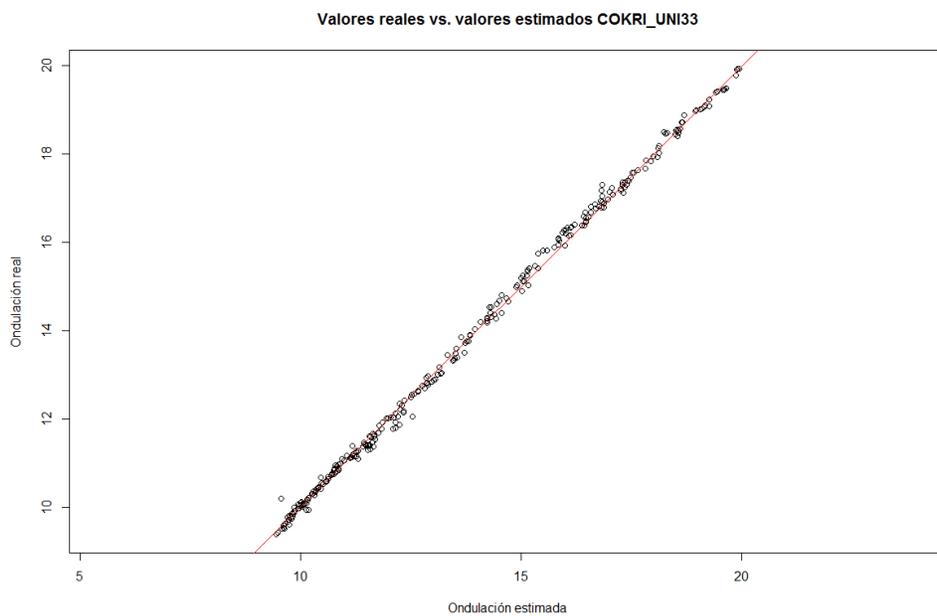
**Figura 143.** Comparación entre valores reales y predichos – Cokriging universal 16 puntos



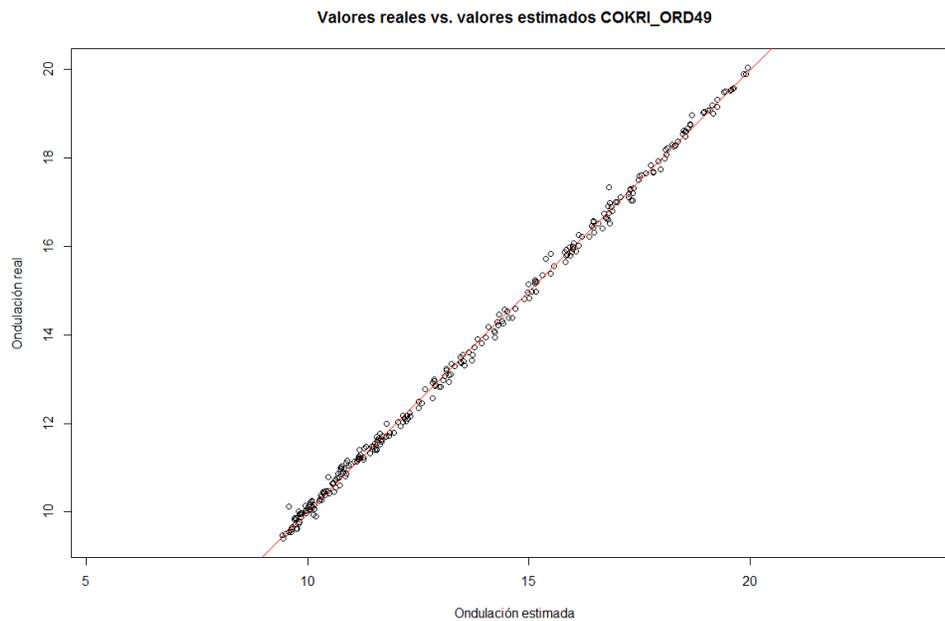
**Figura 144.** Comparación entre valores reales y predichos – Cokriging ordinario 33 puntos



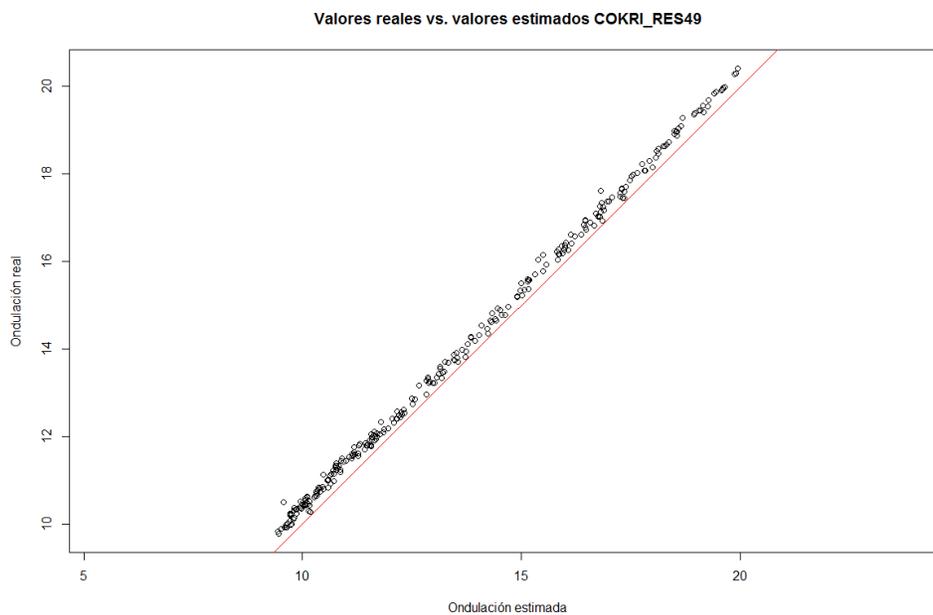
**Figura 145.** Comparación entre valores reales y predichos – Cokriging residual 33 puntos



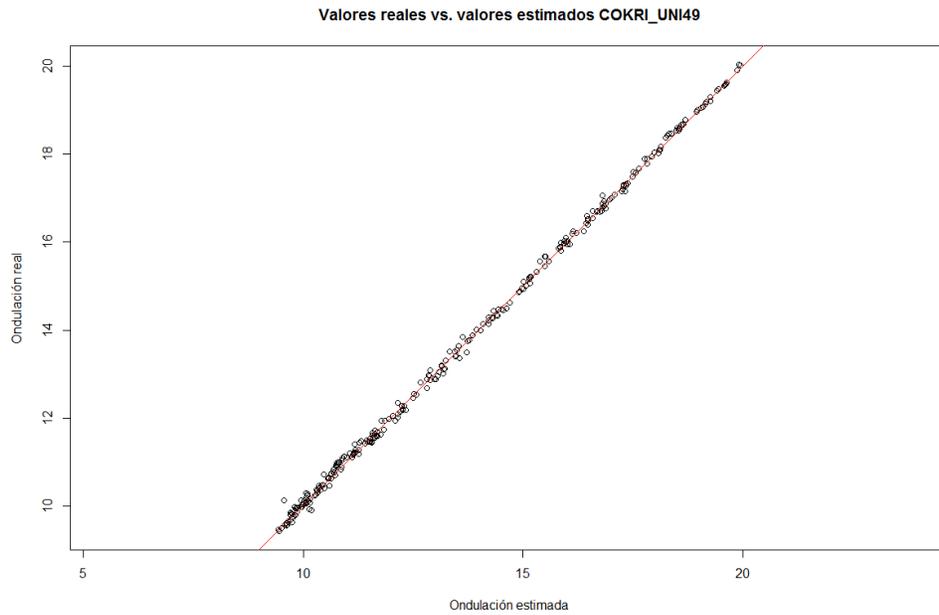
**Figura 146.** Comparación entre valores reales y predichos – Cokriging universal 33 puntos



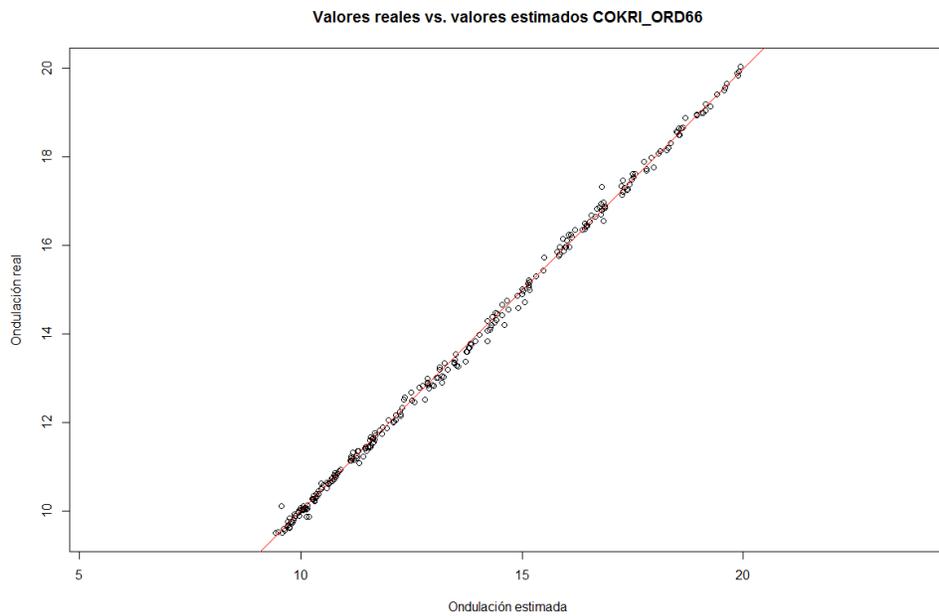
**Figura 147.** Comparación entre valores reales y predichos – Cokriging ordinario 49 puntos



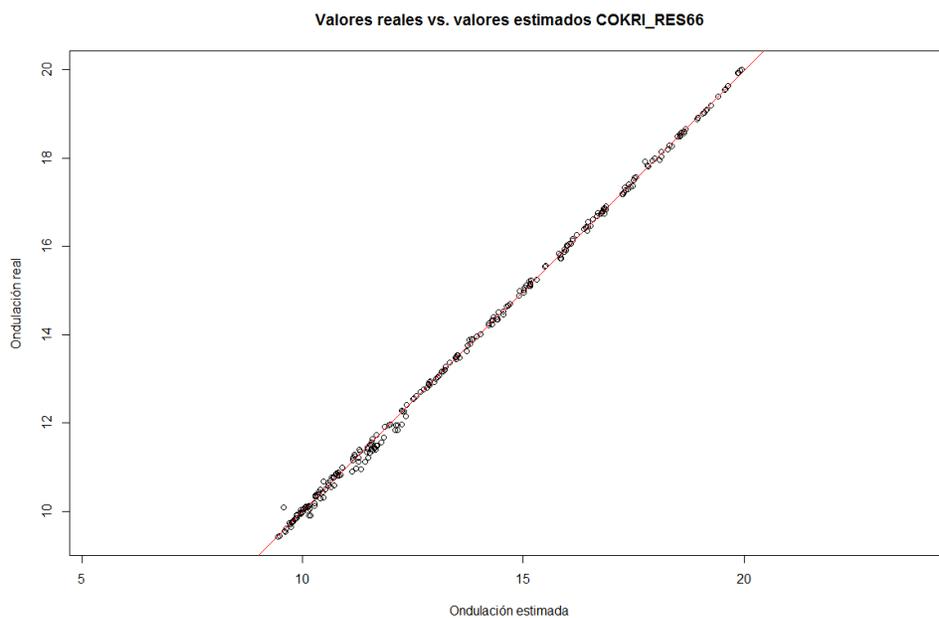
**Figura 148.** Comparación entre valores reales y predichos – Cokriging residual 49 puntos



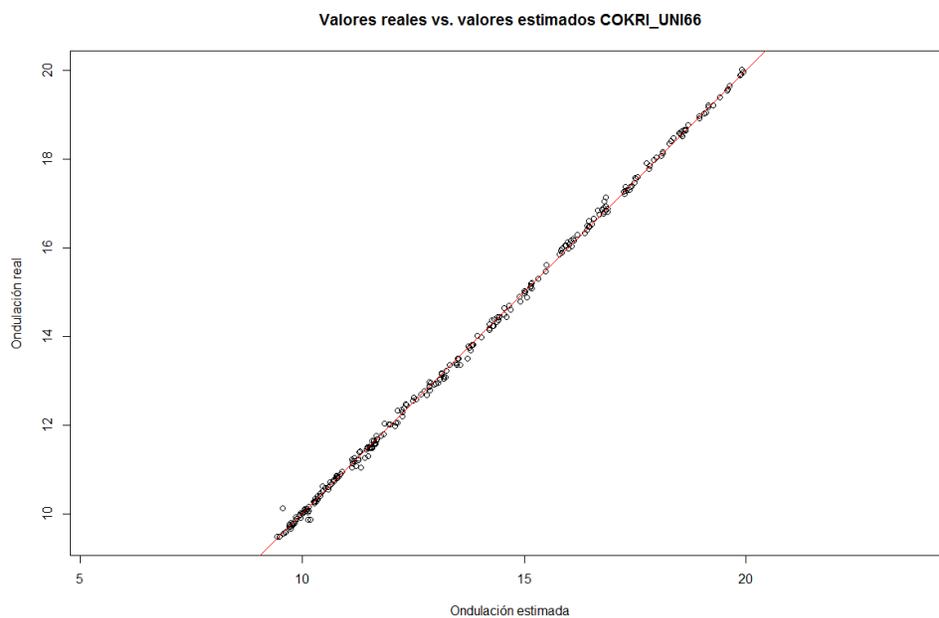
**Figura 149.** Comparación entre valores reales y predichos – Cokriging universal 49 puntos



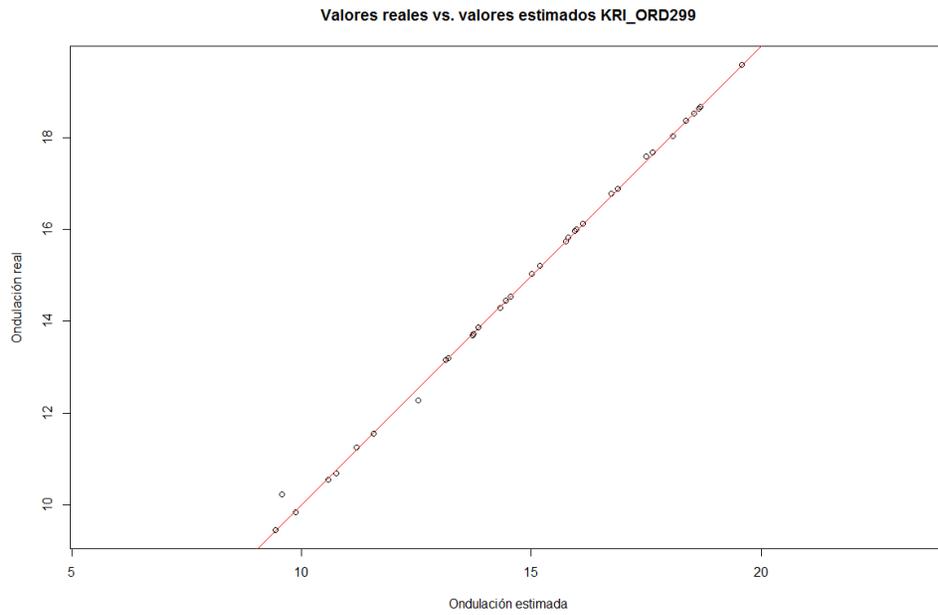
**Figura 150.** Comparación entre valores reales y predichos – Cokriging ordinario 66 puntos



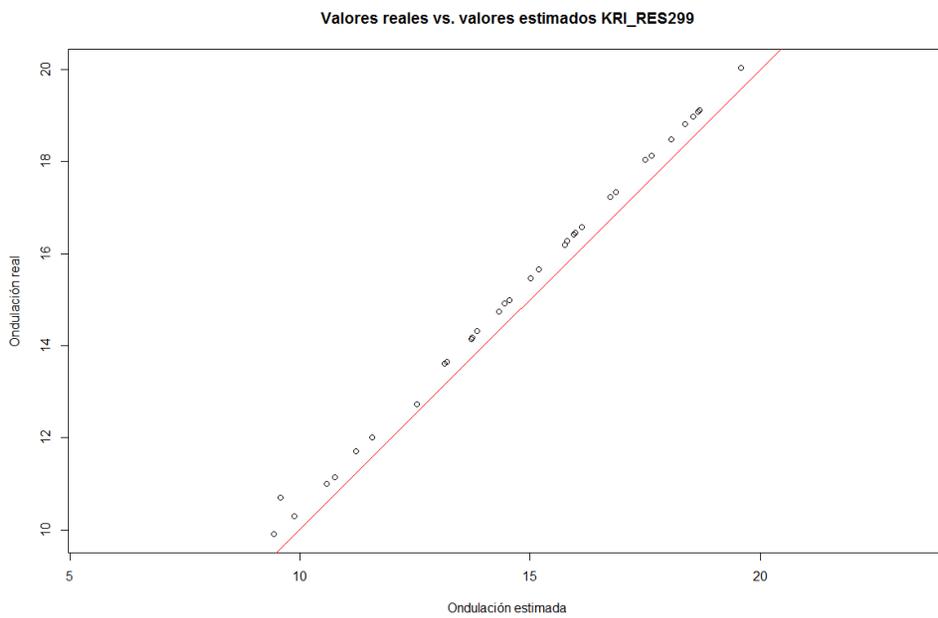
**Figura 151.** Comparación entre valores reales y predichos – Cokriging residual 66 puntos



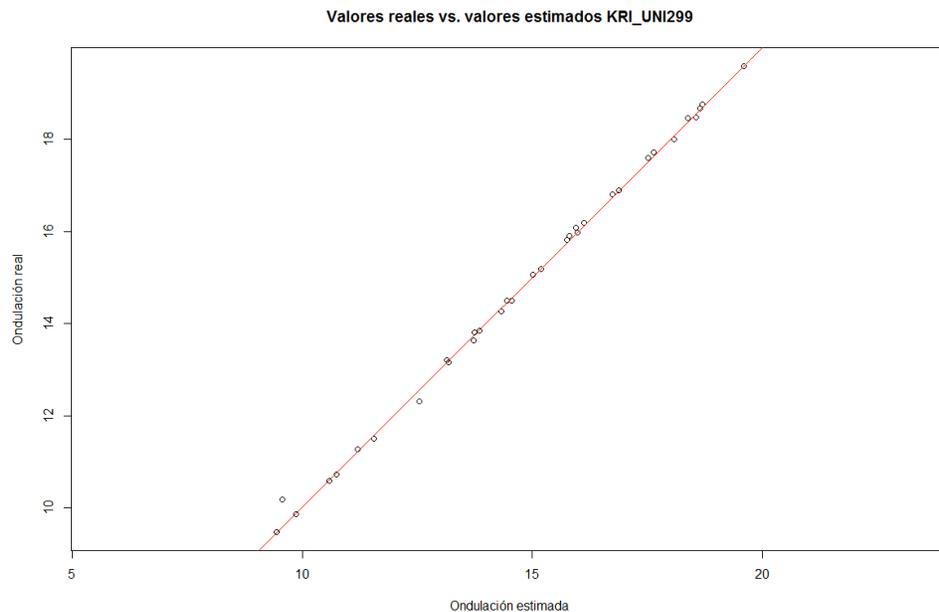
**Figura 152.** Comparación entre valores reales y predichos – Cokriging universal 66 puntos



**Figura 153.** Comparación entre valores reales y predichos – Kriging ordinario 299 puntos



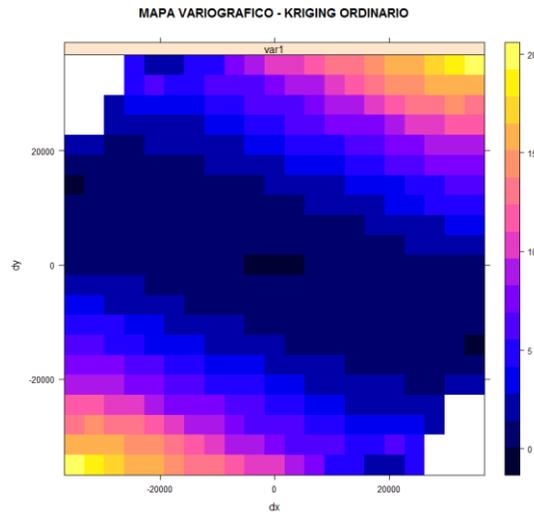
**Figura 154.** Comparación entre valores reales y predichos – Kriging residual 299 puntos



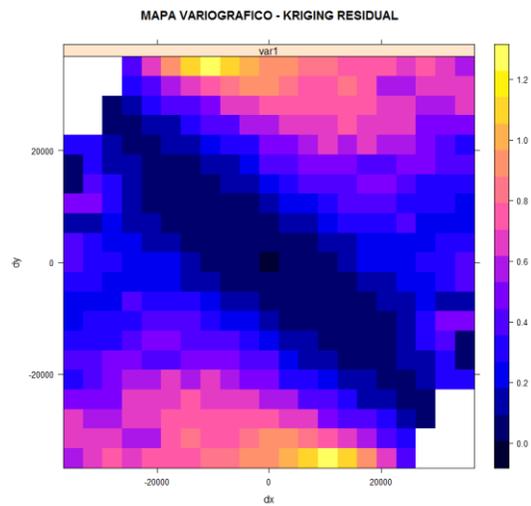
**Figura 155.** Comparación entre valores reales y predichos – Kriging universal 299 puntos

Haciendo una comparación visual de la dispersión de los puntos entre las Figuras 150, 152, 153 (Cokriging con 66 puntos) y las 154, 155, 156 (Kriging con 299), se observa claramente que existe una menor dispersión en los modelos multivariados.

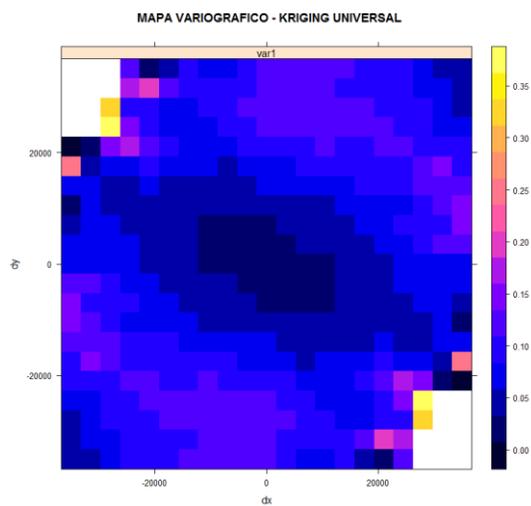
En cuanto a la estacionariedad de la ondulación geoidal en la zona de estudio y en base a los gráficos de dispersión en función de las coordenadas de los puntos muestreados, se dice que el fenómeno no es estacionario sino únicamente intrínseco, lo cual fue comprobado en el ajuste de los modelos autorizados para el MCL, que en su mayoría fue el modelo potencia (no acotado). La presencia de una deriva subyacente en los valores de la ondulación geoidal medidos, fue analizada a través de técnicas como Cokriging (y Kriging) universal y residual; la eliminación o reducción de esta tendencia se relaciona con la eficiencia del método aplicado, lo cual puede verificarse gráficamente mediante un mapa variográfico, que es una herramienta gráfica para la visualización del variograma experimental en cualquier dirección que facilita la identificación de anisotropía (Emery, 2007), como se muestra a continuación:



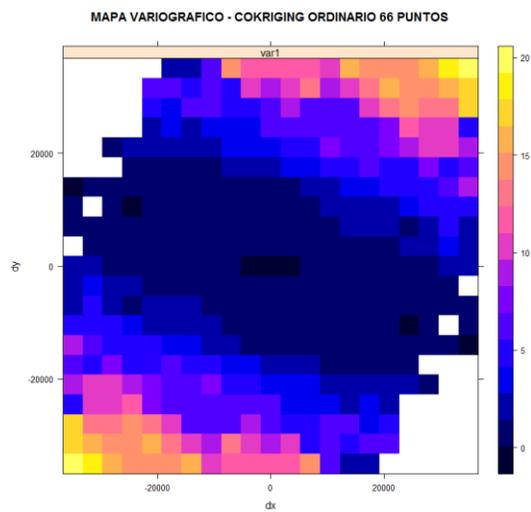
*Figura 156.* Mapa variográfico con Kriging ordinario – 299 puntos



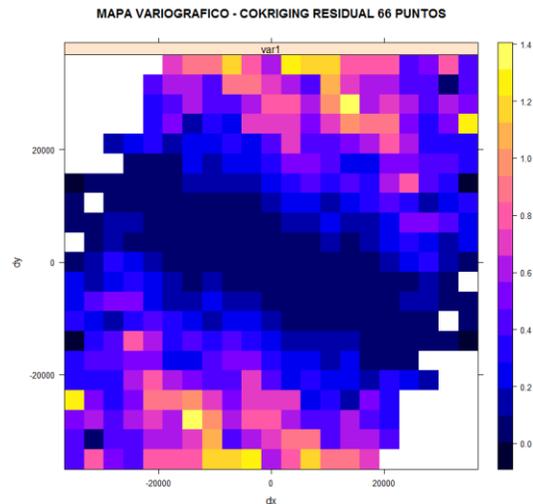
*Figura 157.* Mapa variográfico con Kriging residual – 299 puntos



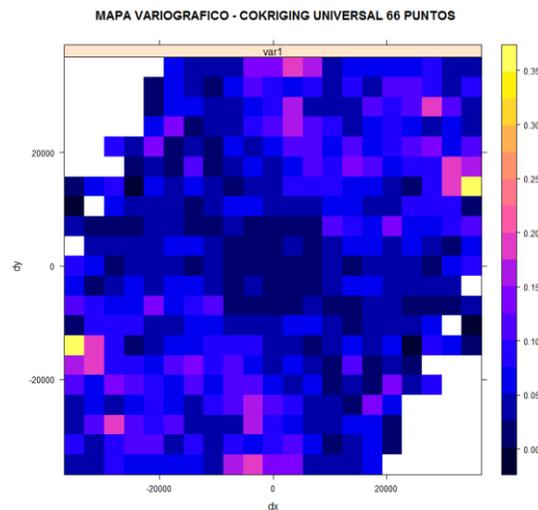
*Figura 158.* Mapa variográfico con Kriging universal – 299 puntos



*Figura 159.* Mapa variográfico con Cokriging ordinario – 66 puntos



**Figura 160.** Mapa variográfico con Cokriging residual – 66 puntos



**Figura 161.** Mapa variográfico con Cokriging universal – 66 puntos

En las figuras anteriores se demuestra la funcionalidad de los métodos de predicción empleados; Kriging ordinario el cual no analiza tendencia, se puede observar que la tendencia es muy marcada y sigue un sentido suroeste – noreste; Kriging residual analiza la tendencia desde las coordenadas de muestreo en el que se encuentra subyacente la deriva de la variable, y se nota una disminución de la misma aunque todavía persiste, y se puede inferir que debido a esto su validación del modelo mostró estadísticos muy elevados; Kriging universal es el que aparentemente modeló de mejor

forma la tendencia, aunque sus estadísticos son ligeramente mayores al Kriging ordinario (ver Tabla 66). En cuanto a Cokriging, el predictor ordinario muestra una clara tendencia marcada en sentido suroeste – noreste (igual que en Kriging); Cokriging residual mostró una mejora en cuanto al tratamiento del drift, sin embargo, el Cokriging universal es que mejor resultado alcanzó en el modelamiento de la tendencia existente (ver Tabla 65).

En base a los estadísticos calculados, y a los gráficos estadísticos utilizados, se estableció que el mejor modelo de predicción de la variable ondulación geoidal para la zona de estudio fue el generado a través de Cokriging universal con 66 puntos muestrales, obteniendo un RMSE de 8 cm de precisión del modelo, RSR de 2 cm y un intervalo de confianza de errores de la validación del modelo entre [-0.16;0.16] metros al 95% de confianza; superando a trabajos anteriormente realizados en la misma zona de estudio pero con diferentes metodologías, como en Leiva (2014), quien utilizó Kriging ordinario y obtuvo un RMSE de 12 cm., y en Chicaiza, Leiva, Arranz & Buenaño (2017), quienes utilizaron una regresión lineal con anomalía de Bouguer consiguiendo 24 cm en la suma de residuales al cuadrado, en ambos casos utilizaron 299 puntos muestrales.

Los dos modelos que presentaron los mejores resultados con 49 y 66 datos (Cokriging universal en ambos casos), fueron comparados con los modelos geopotenciales globales más utilizados, que son el EGM96 y EGM08, y además con el modelo de predicción local generado por el IGM para el país, obtenido a través de una red neuronal artificial de función de base radial entrenada con 833 puntos (IGM, 2013; Tierra & Acurio, 2016). Los resultados de esta comparación se muestran en las Tablas 67 y 68:

**Tabla 67.***Resultados de la comparación entre el modelo de 49 puntos y otros modelos*

Parámetro	Cokri_uni49	EGM96	EGM08	Modelo IGM
RMSE	0.1001	0.8377	0.5924	0.0916

**Tabla 68.***Resultados de la comparación entre el modelo de 66 puntos y otros modelos*

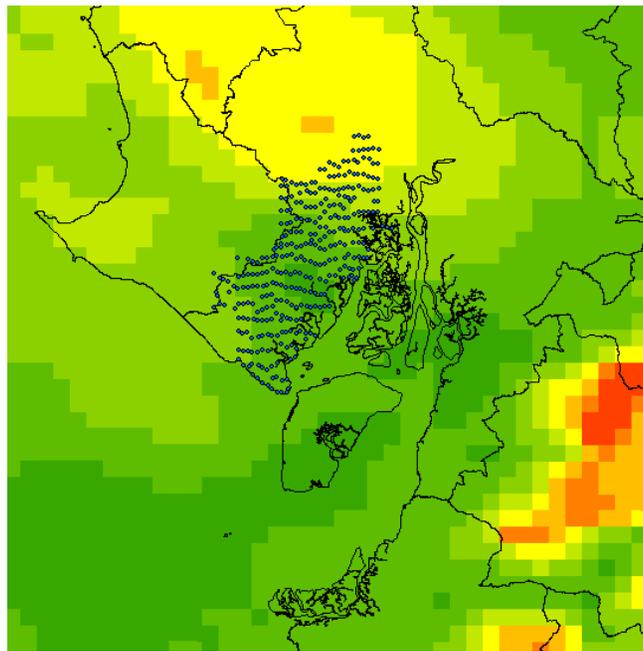
Parámetro	Cokri_uni66	EGM96	EGM08	Modelo IGM
RMSE	0.0859	0.8319	0.5982	0.0894

Los estadísticos calculados muestran una mejora considerable si comparamos los valores entre los modelos geopotenciales globales y los modelos generados en este estudio. En cuanto al modelo del IGM, se observa que en el caso de Cokriging universal con 49 puntos, en modelo IGM es mejor en 8 mm., mientras que con el modelo de Cokriging universal con 66 puntos, dicho modelo es mejor que del IGM en 4 mm., y a pesar de que en términos generales ambos modelos son iguales, se destaca la importancia de que los modelos generados en este estudio utilizan menos puntos para la obtención del modelo de predicción, lo que se traduce en eficiencia de costos, tiempo y personal, es decir, en la optimización del proceso. Cabe mencionar que la densidad de puntos total en el área de estudio es aproximadamente 0.12 puntos/km<sup>2</sup>.

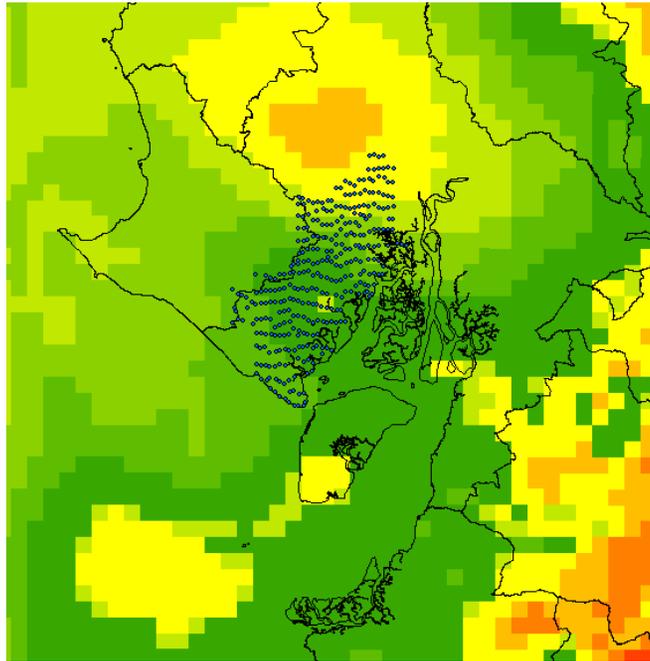
Al observar el crecimiento sistemático de los valores de ondulación geoidal, se trató de explicar este gradiente en dirección suroeste – noreste relacionándolo con la topografía de la zona de estudio, sin embargo no existió algún patrón visible que se corresponda con esta tendencia, lo cual se vio reflejado en la correlación existente con los DEM analizados. Por lo que se infiere que esta deriva es causada por algún proceso físico del campo de gravedad terrestre, que puede estar relacionado con la dinámica de placas tectónicas (Banerjee *et al.*, 1999), por movimientos de convección de la corteza del planeta (Introcaso, 2006), o variación en la densidad de masas de la

Tierra (Del Cogliano, 2006), lo cual se relaciona directamente con las anomalías de gravedad y por ende con la ondulación geoidal.

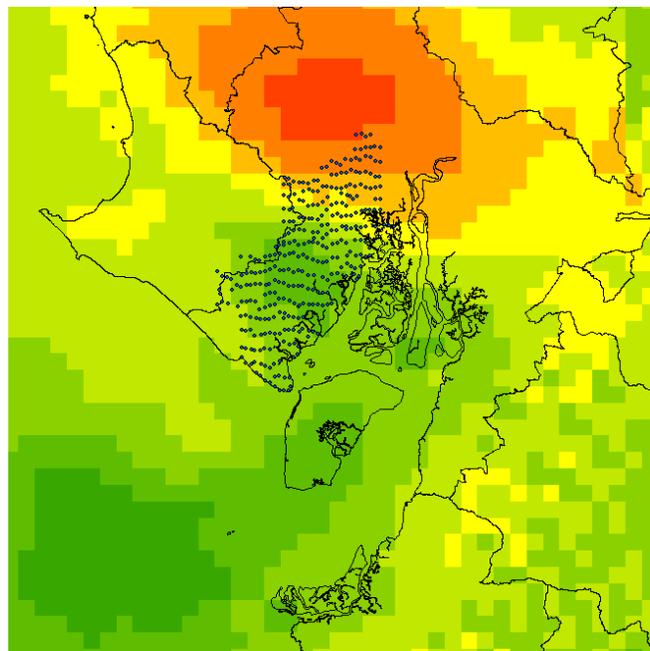
Según Le Meur (1998), (citado en Introcaso, 2006) sugiere que en zonas de subducción existen cambios en anomalías del geoide, lo cual coincide con la ubicación de la zona de estudio, que está entre la fosa oceánica del Pacífico y el comienzo de la Cordillera de la Andes, por lo que se supondría que inciden directamente sobre los valores de anomalías de gravedad y estos en la ondulación geoidal. La variabilidad de las anomalías de gravedad que existe en la zona de estudio se puede observar en las Figuras 162, 163 y 164:



**Figura 162.** Anomalía de Aire libre en la zona de estudio (puntos azules)



**Figura 163.** Anomalía de Bouguer en la zona de estudio (puntos azules)



**Figura 164.** Anomalía Isostática en la zona de estudio (puntos azules)

Como se indica en las figuras anteriores, existe una alta incidencia de las anomalías de gravedad en la parte noreste de la zona de estudio, lo que podría explicar la existencia de tendencia creciente en los datos de ondulación geoidal.

En los resultados de la validación del modelo, se ha demostrado que usando la ondulación geoidal derivada del modelo EGM08 a 1 min de resolución como variable auxiliar en el modelamiento con Cokriging, se alcanzó mejores resultados que en el modelamiento univariante. Además, el enfoque que se dio a la variable auxiliar fue en analogía a la técnica Remove – Restore, en donde la ondulación del modelo geopotencial (gracias a que está construido con expresiones de armónicos esféricos completos de orden y grado 2159, más coeficientes adicionales de grado 2190 y orden 2159), sirvió para remover la tendencia existente en los datos de la variable objetivo, proporcionando información de longitud de onda larga mientras que los datos GPS/nivelación diferencial, incorporaron la información de longitud de onda corta para en conjunto formar el modelo de predicción resultante, de forma similar a los trabajos de Del Cogliano (2006), Ligas y Szombara (2018), Achola, Fukuda y Kuroishi (2012).

La eficiencia del modelo generado en este trabajo, se puede expresar en términos económicos, con el fin de dar una idea del ahorro de presupuesto que se requeriría para llevar a cabo un modelo de predicción de ondulación geoidal a través de Cokriging y compararlo con los costes que serían necesarios usando las técnicas comunes (Kriging). El control geodésico suplementario de un punto, es decir, sus coordenadas horizontales y verticales tienen un costo aproximado de \$400; si se consideran los 299 datos que fueron utilizados en Kriging versus los 66 datos de Cokriging, se traducirían a \$119600 y \$26400 respectivamente, por lo que se estaría reduciendo el presupuesto en \$93200 (77.92%) para proyectos de la escala de trabajo expuesta en este estudio, con lo cual se

optimizan recursos que podrían ser empleados en otras ciudades del país o proyectos relacionados a esta temática.

El modelo de predicción de ondulación geoidal final, se representó en un mapa, el mismo que se puede observar en la Figura 165, y sus errores de predicción se muestran en la Figura 166:

### MODELO DE ONDULACIÓN GEOIDAL

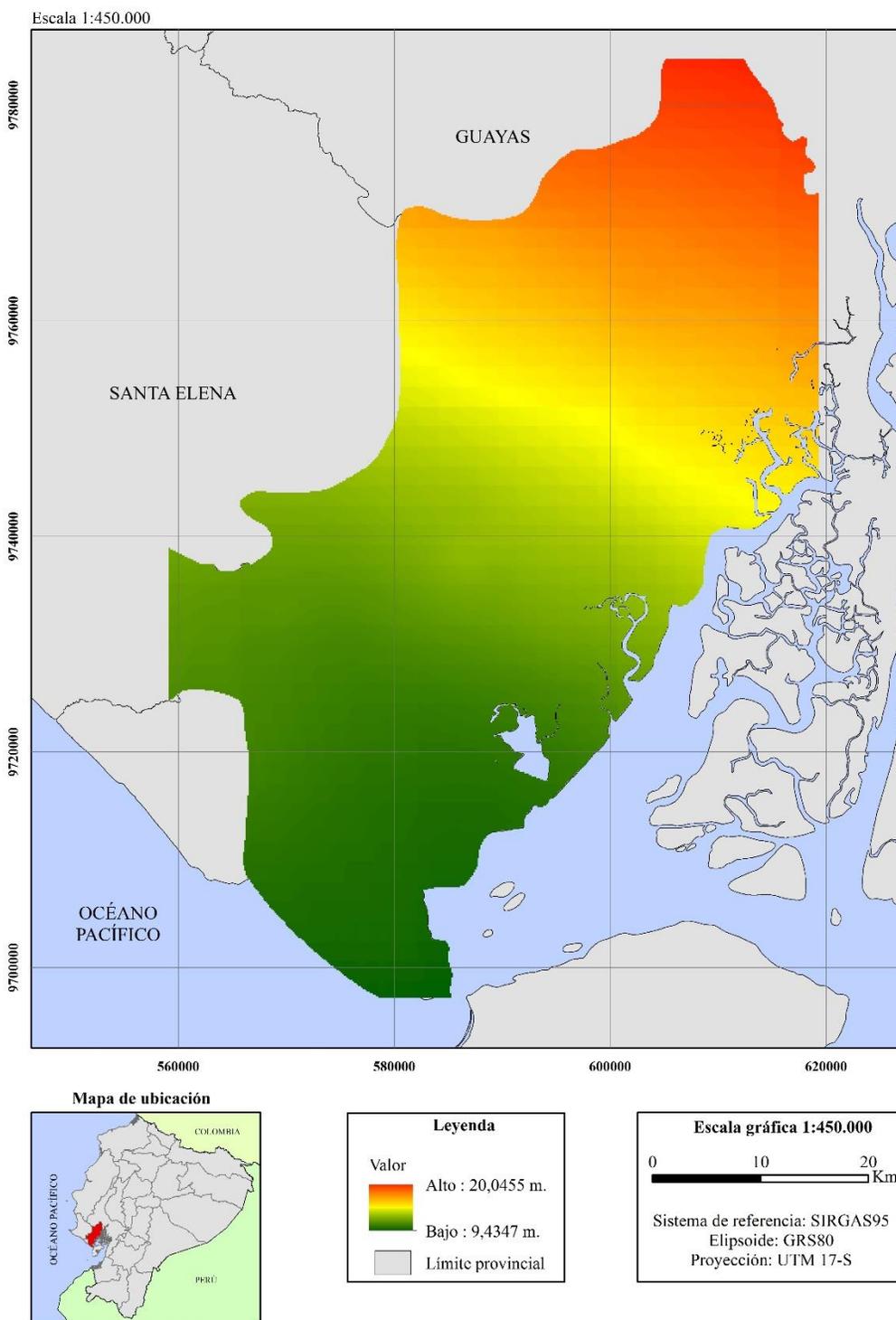
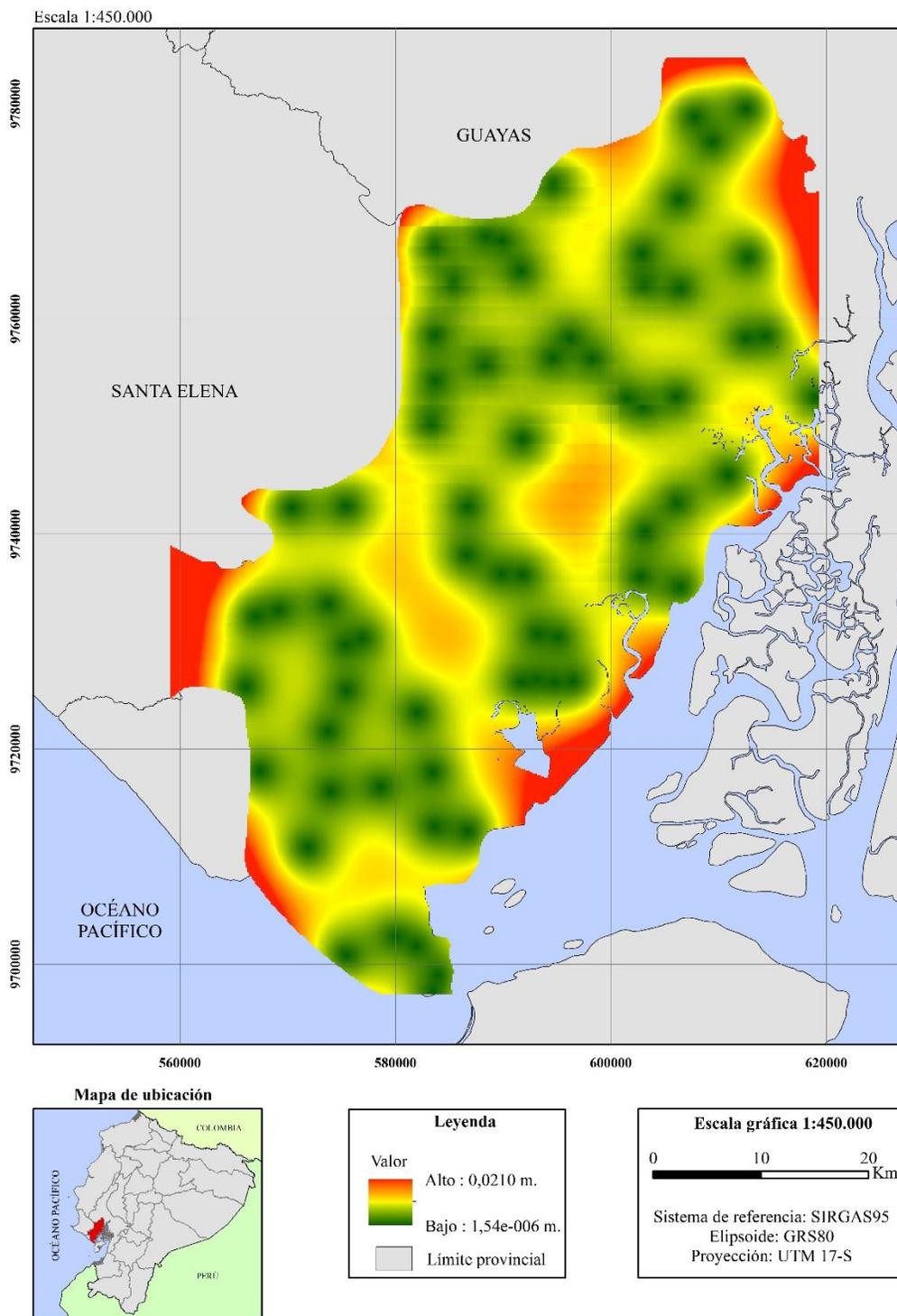


Figura 165. Modelo de ondulación geoidal generado con Cokriging universal con 66 puntos

**MODELO DEL ERROR DE PREDICCIÓN DE ONDULACIÓN GEOIDAL**



**Figura 166.** Modelo de errores de ondulación geoidal (Cokriging universal con 66 puntos)

## CAPÍTULO VI

### CONCLUSIONES Y RECOMENDACIONES

#### 6.1 Conclusiones

Se demostró que a partir de muestras dispersas y en menor cantidad de datos, el modelo de predicción de ondulación geoidal generado mediante Cokriging universal con 66 puntos, fue el más preciso de todos los casos analizados, obteniendo un RMSE de 8 cm. y RSR de 2 cm. en la validación del modelo resultante (Tabla 65), con un error de predicción (desviación estándar) máximo de 2 cm. en el área de estudio (Figura 166).

Mediante los indicadores estadísticos calculados para la validación del modelo de predicción, se comprobó que utilizando una variable auxiliar, en este caso, la ondulación geoidal derivada del EGM08 (1 min), se logró mayor precisión en comparación al realizado con el estimador Kriging, comprobando así, la hipótesis planteada en este trabajo.

De las 7 posibles variables auxiliares estudiadas, se concluye que, el DEM no posee correlación espacial con la ondulación geoidal en la zona de estudio, las anomalías de gravedad (Aire libre, Bouguer e Isostática) tienen una buena correlación por lo que serían una opción para el modelamiento de la variable objetivo, y la ondulación geoidal derivada de las dos resoluciones espaciales disponibles del modelo geopotencial EGM08 cada 2.5' min y 1 min, la segunda fue la que mejor correlación espacial obtuvo, y por consiguiente, seleccionada como la variable auxiliar idónea para este estudio.

La presencia de tendencia en los datos muestreados, fue analizada a través de los predictores residual y universal, y a pesar de que ambos métodos redujeron la deriva (ver Figuras 156, 157,

158, 159, 160 ,161), fue Cokriging universal el que modeló de mejor manera los datos con la tendencia subyacente de la ondulación geoidal.

La tendencia existente en los puntos muestreados de la variable objetivo, puede ser debido a procesos geofísicos que inciden sobre las anomalías de gravedad en la zona de estudio, atribuido a las características específicas del área estudiada, por encontrarse en una región entre la zona de subducción de placas tectónicas y el comienzo de un macizo montañoso como es la Cordillera de los Andes, por lo que se infiere una variación en las densidades de masas internas de la Tierra, que repercuten en los valores de dichas anomalías, especialmente en la dirección noreste (ver Figuras 162, 163, 164).

Al comparar el modelo generado de Cokriging universal con 66 puntos, y los modelos geopotenciales globales más usados en el país (EGM96 y EGM08), el modelo geoestadístico obtenido en este estudio es mucho más preciso que estos dos, y al contrastarlo con el modelo del IGM, posee una precisión mayor en 4 mm. a este último.

El modelo final generado, se ofrece como una herramienta de soporte para todos los trabajos relacionados a las geociencias que requieran alturas referidas al nivel medio del mar, con altas precisiones, de forma rápida y económica.

En términos de eficiencia, la metodología desarrollada en el presente estudio, permite optimizar la obtención alturas niveladas más precisas, potenciando el uso de técnicas GNSS y reduciendo el presupuesto económico necesario en comparación a las técnicas geoestadísticas univariante.

## 6.2 Recomendaciones

Para el caso de la zona de estudio y basado en el valor del error cuadrático medio, es factible utilizar el modelo generado en trabajos donde se requieran precisiones de hasta 8 cm., caso contrario se debería tener en cuenta el mapa de los errores de predicción en los sitios específicos que sean usados.

Se deberían realizar estudios en otras áreas del país para demostrar la aplicabilidad de la metodología empleada, y así ser utilizada como una alternativa eficiente en los distintos trabajos de las geociencias y por profesionales o técnicos del ramo que requieran de esta variable.

Con este método se abre las puertas a nuevos campos de investigación, no solo en la geografía, sino en las demás ciencias que requieran optimizar tiempos, recursos, y calidad en sus modelos de predicción, por lo que se recomendaría seguir con esta línea de investigación.

**REFERENCIAS BIBLIOGRÁFICAS**

- Achola, P., & Fukuda, Y. (2015). Recovery of orthometric heights from ellipsoidal heights using offsets method over Japan. *Earth, Planets and Space*, 67(134), 3–7. doi:10.1186/s40623-015-0306-z
- Achola, P., Fukuda, Y., & Kuroishi, Y. (2012). A high-resolution gravimetric geoid model for Japan from EGM2008 and local gravity data. *Earth Planets Space*, 64, 361–368. doi:10.5047/eps.2011.11.004
- Ahad, N., Yin, T., Othman, A., & Yaacob, C. (2011). Sensitivity of normality tests to non-normal data. *Sains Malaysiana*, 637–641.
- Alaska Satellite Facility. (2018). *About ALOS PALSAR*. Recuperado el 3 de Diciembre de 2018, de <https://www.asf.alaska.edu/sar-data/palsar/about-palsar/>
- Alaska Satellite Facility. (2018). *Terrain-Corrected (RTC)*. Recuperado el 3 de Diciembre de 2018, de <https://www.asf.alaska.edu/sar-data/palsar/terrain-corrected-rtc/>
- Alfaro, M. (2007). *Estimación de Recursos Mineros*. París, Francia: Escuela de Minas de París.
- Al-Krargy, E., Hosny, M., & Dawod, G. (2015). *Investigating the Precision of Recent Global Geoid Models and Global Digital Elevation Models for Geoid Modelling in Egypt*. Sharm El-Sheikh, Egypt: Regional Conference on Surveying & Development.
- Alsamamra, H., Ruiz, J., Pozo, D., & Tovar, J. (2009). A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain.

*Agricultural and Forest Meteorology*, 149(8), 1343–1357.  
doi:10.1016/j.agrformet.2009.03.005

Balmino, G., Vales, N., Bonvalot, S., & Briais, A. (2011). Spherical harmonic modeling to ultra-high degree of Bouguer and isostatic anomalies. *Journal of Geodesy*, 86(7), 499–520.  
doi:10.1007/s00190-011-0533-4

Banerjee, P., Foulger, G., Satyaprakash, & Dabral, C. (1999). Geoid undulation modelling and interpretation at Ladak, NW Himalaya using GPS and levelling data. *Journal of Geodesy*, 73(2), 79–86. doi:10.1007/s001900050221

BGI. (2012). *WGM2012 Earth's gravity anomalies*. Recuperado el 22 de Noviembre de 2018, de <http://bgi.omp.obs-mip.fr/data-products/Grids-and-models/wgm2012>

Bishop, T., McBratney, A., & Whelan, B. (2001). Measuring the quality of digital soil maps using information criteria. *Geoderma*, 103(1), 95–111. doi:[https://doi.org/10.1016/S0016-7061\(01\)00071-4](https://doi.org/10.1016/S0016-7061(01)00071-4)

Bland, J., & Altman, D. (2007). Agreement between methods of measurement with multiple observations per individual. *Journal of Biopharmaceutical Statistics*, 17(4), 571–582.  
doi:10.1080/10543400701329422

Bogaert, P., Mahau, P., & Beckers, F. (1995). *The Spatial Interpolation of Agroclimatic Data. Cokriging Software and Source Code*. Roma, Italia: FAO.

Bonvalot, S., Balmino, G., Briais, A., Kuhn, M., Peyrefitte, A., Vales, N., . . . Sarrailh, M. (2012). *World Gravity Map. Scale 1:50000000* (1era ed.). París, Francia: IGN France.

- Cañizares, C. (2015). *Análisis del Sistema de Alturas para la Red de Control Básico Vertical del Ecuador*. Sangolquí, Ecuador, Tesis de grado: Universidad de las Fuerzas Armadas - ESPE.
- Caro, A. (2012). *Modelización geoestadística para la predicción de actividad de Cs-137 en el suelo*. Madrid, España, Tesis de doctorado: Universidad Autónoma de Madrid.
- Chicaiza, E., Leiva, C., Arranz, J., & Buenaño, X. (2017). Spatial uncertainty of a geoid undulation model in Guayaquil, Ecuador. *Open Geosciences*, 9(1), 255–265. doi:10.1515/geo-2017-0021
- Cruz, G., Villaseñor, J., López, L., & Ortiz, E. (2013). Distribución espacial de la riqueza de especies de plantas vasculares en México. *Revista Mexicana de Biodiversidad*, 84(4), 1189–1199. doi:10.7550/rmb.31811
- De By, R. (2001). *Principles of Geographical Information Systems* (Vol. 1). Enschede, Holanda: ITC Educational Textbook Series.
- Del Cogliano, D. (2006). *Modelado del Geoide con GPS y Gravimetría. Caracterización de la Estructura Geológica de Tandil*. Rosario, Argentina, Tesis de doctorado: Universidad Nacional de Rosario.
- DeMers, M. (2001). *GIS Modeling in Raster. GIS & Remote Sensing*. West Sussex, Inglaterra: Wiley.
- Díaz, M. (2002). *Geoestadística aplicada*. México D.F., México: Instituto de Geofísica, UNAM.
- Drewes, H. (2014). *Sistemas de Referencia*. Quito, Ecuador: Instituto Geográfico Militar.

- Drewes, H., & Sánchez, L. (2002). *Curso de Sistemas de Referencia en Geodesia*. Quito: Instituto Geográfico Militar.
- Drewes, H., Sánchez, L., Blitzkow, D., & Freitas, S. (1998). *SIRGAS - Grupo de Trabajo III: Sistema de Referencia Vertical para América del Sur*. Recuperado el 3 de Diciembre de 2018, de <http://azimuth.univalle.edu.co/docsdownload/archivo3.doc>
- Eddelbuettel, D., Francois, R., Allaire, J., Ushey, K., Kou, Q., Russell, N., . . . Chambers, J. (2018). *Package 'Rcpp'*. Recuperado el 7 de Diciembre de 2018, de <https://cran.r-project.org/web/packages/Rcpp/Rcpp.pdf>
- Emery, E. (2007). *Apuntes de Geoestadística*. Santiago, Chile: Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile.
- ESRI. (2002). *Spatial Analyst Help Documentation. ArcGIS Users' Guide*. Redlands, EE.UU.: ESRI Inc.
- Farjas, M. (2009). *Nivelación Geométrica*. Recuperado el 3 de Diciembre de 2018, de [http://ocw.upm.es/ingenieriacartografica-geodesica-y-fotogrametria/topografia-ii/Teoria\\_NG\\_Tema4.pdf](http://ocw.upm.es/ingenieriacartografica-geodesica-y-fotogrametria/topografia-ii/Teoria_NG_Tema4.pdf)
- Fernández, I. (2001). *Localizaciones Geográficas. Las Coordenadas Geográficas y la Proyección UTM (Universal Transversa Mercator)*. Valladolid, España: Editorial de la Universidad de Valladolid.
- Florinsky, I., & Kuryakova, G. (2000). Determination of grid size for digital terrain modelling in landscape investigations - exemplified by soil moisture distribution at a micro-scale.

*International Journal of Geographical Information Science*, 14(8), 815–832.  
doi:<https://doi.org/10.1080/136588100750022804>

García, J. (2014). Density modeling of the Escollos Alijos Seamount from inversion of its geoid undulation anomaly. *Geofísica Internacional*, 53(3), 333–341. doi:10.1016/S0016-7169(14)71509-0

Giraldo, R. (2001). Propuesta de un indicador como variable auxiliar en el análisis Cokriging. *Revista Colombiana de Estadística*, 24(1), 1–12.

Giraldo, R. H. (2007). *Introducción a la Geostatística. Teoría y aplicación*. Bogotá, Colombia: Departamento de Estadística. Universidad Nacional de Colombia.

Golmohammadi, G., Prasher, S., Madani, A., & Rudra, R. (2014). Evaluating Three Hydrological Distributed Watershed Models: MIKE-SHE, APEX, SWAT. *Hydrology*, 1(1), 20–39. doi:10.3390/hydrology1010020

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation (Applied Geostatistics)*. New York, EE.UU.: Oxford University Press.

Hansen, F. (2008). *Apuntes de Cartografía*. Aguascalientes, México: INEGI.

Heiskanen, W., & Moritz, H. (1985). *Physical Geodesy* (4ta. ed.). Madrid, España: Editorial IGN.

Hengl, T. (2006). Finding the right pixel size. *Computers & Geosciences*, 32(9), 1283–1298. doi:10.1016/j.cageo.2005.11.008

Hengl, T. (2009). *A practical guide to geostatistical mapping* (2da ed.). Amsterdam, Holanda: University of Amsterdam.

- Hengl, T., Heuvelink, G., & Rossiter, D. (2007). About regression-kriging: From equations to case studies. *Computers & Geosciences*, 33(10), 1301–1315. doi:10.1016/j.cageo.2007.05.001
- Hoaglin, D., Mosteller, F., & Tukey, J. (1983). *Understanding Robust and Exploratory Data Analysis*. New York: John Wiley & Sons.
- Hofmann, B., & Moritz, H. (2005). *Physical Geodesy*. Austria: SpringerWienNewYork.
- Hoyer, M., Wildermann, E., Suárez, H., & Hernández, J. (2004). Modelo geoidal combinado para Venezuela (MGCV04). *Interciencia*, 29(12), 667–672.
- IAG. (2012). *The International Association of Geodesy*. Recuperado el 3 de Diciembre de 2018, de [http://www.iag-aig.org/templates\\_img/iagbooklet.pdf](http://www.iag-aig.org/templates_img/iagbooklet.pdf)
- IGM. (1978). *Ley de cartografía nacional*. Quito, Ecuador.
- IGM. (2006). *Memoria Técnica del Control Geodésico Suplementario Horizontal y Vertical de los Fotoíndices 9, 14, 15 y 19 del Proyecto Cantón de Guayaquil*. Quito, Ecuador.
- IGM. (2013). *Cálculo de la altura geoidal del Ecuador continental*. Recuperado el 8 de Diciembre de 2018, de <http://www.geoportaligm.gob.ec/ModeloGeoidal-war/#>
- Introcaso, A. (2006). *Geodesia física* (Volúmen especial). Rosario, Argentina: Boletín del Instituto de Fisiografía y Geología.
- Isaaks, E., & Srivastava, R. (1989). *Applied Geostatistics*. New York, EE.UU.: Oxford University Press.
- Journel, A. G., & Huijbregts, C. J. (1978). *Mining Geostatistics*. New York, EE.UU.: Academic Press.

- Kiamehr, R., & Sjöberg, L. (2005). Effect of the SRTM global DEM on the determination of a high-resolution geoid model: a case study in Iran. *Journal of Geodesy*, 79(9), 540–551. doi:<https://doi.org/10.1007/s00190-005-0006-8>
- Knotters, M., Brus, D., & Oude, J. (1995). A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. *Geoderma*, 67(3), 227–246. doi:[https://doi.org/10.1016/0016-7061\(95\)00011-C](https://doi.org/10.1016/0016-7061(95)00011-C)
- Kopeikin, S., Mazurova, E., & Karpik, A. (2015). Towards an exact relativistic theory of earth's geoid undulation. *Physics Letters A*, 379(26), 1555–1562. doi:<https://doi.org/10.1016/j.physleta.2015.02.046>
- Leal, C. (2016). *Coeficiente de Correlación de Concordancia bajo la distribución t multivariada*. Santiago, Chile, Tesis de doctorado: Universidad Católica de Chile.
- Leiva, C. (2014). *Determinación de modelos de predicción espacial de la variable ondulación geoidal, para la zona urbana del cantón Quito y la zona rural del cantón Guayaquil, utilizando técnicas geoestadísticas*. Quito, Ecuador, Tesis de maestría: Universidad Politécnica Nacional.
- Lemoine, F., Kenyon, S., Factor, J., Trimmer, R., Pavlis, N., Chinn, D., . . . Olson, T. (1998). *The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96*. Greenbelt, EE.UU.: NASA/TP.
- Li, J., & Heap, A. (2008). *A Review of Spatial Interpolation Methods for Environmental Scientists*. Canberra, Australia: GeoCat: Geoscience Australia.

- Ligas, M., & Kulczycki, M. (2010). Simple spatial prediction – least squares prediction, simple kriging, and conditional expectation of normal vector. *Geodesy and Cartography*, 59(2), 69–81. doi:10.2478/v10277-012-0002-0
- Ligas, M., & Szombara, S. (2018). Geostatistical prediction of a local geometric geoid - kriging and cokriging with the use of EGM2008 geopotential model. *Studia Geophysica et Geodaetica*, 62(2), 187–205. doi:10.1007/s11200-017-0713-7
- Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45, 255–268.
- Lin, L. (2000). A note on the concordance correlation coefficient. *Biometrics*, 56, 324–325.
- López, J. (2015). *Historia de las Proyecciones Cartográficas*. Madrid, España: Centro Nacional de Información Geográfica (CNIG). doi:10.7419/162.05.2015
- Luna, M. (2017). *Contribución a la definición de un modelo de velocidades de la corteza terrestre para Ecuador a partir de datos GNSS*. Madrid, España, Tesis de doctorado: Universidad Politécnica de Madrid.
- MAGAP. (2016). *Modelo Digital del Terreno (MDT) de Ecuador*. Recuperado el 3 de Diciembre de 2018, de [http://metadatos.sigtierras.gob.ec:8080/geonetwork/srv/spa/catalog.search#/metadata/MDT\\_SIGTIERRAS\\_16092016](http://metadatos.sigtierras.gob.ec:8080/geonetwork/srv/spa/catalog.search#/metadata/MDT_SIGTIERRAS_16092016)
- Matheron, G. (1969). *Curso de Geoestadística*. París, Francia: Centro de Geoestadística de la Escuela de Minas de París.

- Matheron, G. (1970). *La teoría de las variables regionalizadas y sus aplicaciones*. París, Francia: Centro de Geoestadística de la Escuela de Minas de París.
- Matheron, G. (1989). *Estimating and choosing. An essay on probability in practice*. Berlín, Alemania: Springer-Verlag.
- McCarthy, D. (2000). *IERS Conventions*. IERS Technical Note, International Earth.
- Mena, J. (2008). *Geodesia Superior* (Vol. 1). Madrid, España: Centro Nacional de Información Geográfica.
- Milbert, D. (1991). GEOID90; a high-resolution geoid for the United States. *EOS, Trans Am Geophys Union*, 72(49), 545–554.
- Molnár, D., & Julien, P. (2000). Grid-size effects on surface runoff modeling. *Journal of Hydrologic Engineering*, 5(1), 8–16. doi:[https://doi.org/10.1061/\(ASCE\)1084-0699\(2000\)5:1\(8\)](https://doi.org/10.1061/(ASCE)1084-0699(2000)5:1(8))
- Mulla, D., & McBratney, B. (2002). Soil spatial variability. En *Handbook of Soil Science* (págs. A321-A352). Florida, EE.UU.: CRC Press.
- Myers, D. E. (1987). Optimization of Sampling Locations for Variogram Calculations. *Water Resources Research*, 23(3), 283–293.
- NASA. (2005). *EGM96 General Description*. Recuperado el 22 de Noviembre de 2018, de <https://cdis.nasa.gov/926/egm96/gendesc.html>
- Navarro, S. (2010). *Nivelación. Apuntes de Topografía*. Barranquilla, Colombia: UNI Norte.

- NGA. (2008). *EGM2008 - WGS 84 Version*. Recuperado el 22 de Noviembre de 2018, de [http://earth-info.nga.mil/GandG/wgs84/gravitymod/egm2008/egm08\\_wgs84.html](http://earth-info.nga.mil/GandG/wgs84/gravitymod/egm2008/egm08_wgs84.html)
- Noda, A., & Niino, H. (2003). Critical grid size for simulating convective storms: a case study of the Del City supercell storm. *Geophysical Research Letters*, 30(16), 1–4. doi:<https://doi.org/10.1029/2003GL017498>
- Nornadiah, M., & Yap, B. (2011). Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21–33.
- Nornadiah, M., Norin, R., Nur, N., Az'lina, A., & Adriana, I. (2012). A comparison of normality tests using SPSS, SAS and MINITAB: An application to Health Related Quality of Life data. *2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE)*, (pp. 1-6). Langkawi, Malasia. doi:10.1109/ICSSBE.2012.6396570
- Núñez, A. (2006). *Determinación de un geoide de precisión en áreas de pequeña. Aplicación en el Parque Nacional de Doñana*. Cataluña, España, Tesis de doctorado: Universidad Politécnica de Cataluña.
- Oliver, M. (2010). *Geostatistics Applications for Precision Agriculture*. New York, EE.UU.: Springer.
- Oliver, M., & Webster, R. (2014). A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena*, 113, 56–69. doi:<http://dx.doi.org/10.1016/j.catena.2013.09.006>

- Oliver, M., & Webster, R. (2015). *Basic Steps in Geostatistics: The Variogram and Kriging*. New York, EE.UU.: Springer.
- Pacino, M. (1999). *Geoide y Estructura Litosférica en la Argentina*. Rosario, Argentina, Tesis de doctorado: Universidad Nacional de Rosario.
- Paredes, N. (1986). *Determinación del datum vertical en La Libertad, Ecuador*. Guayaquil, Ecuador: INOCAR.
- Pavlis, N., Holmes, A., Kenyon, S., & Factor, J. (2008). *An Earth Gravitational Model to Degree 2160: EGM08*. EE.UU.: National Geospatial-Intelligence Agency.
- Pebesma, E. (2014). *Gstat user's manual*. Recuperado el 26 de Noviembre de 2018, de <http://www.gstat.org/gstat.pdf>
- Pebesma, E., Bivand, R., Rowlingson, B., Gomez, V., Hijmans, R., Sumner, M., . . . O'Rourke, J. (2018). *Package 'sp'*. Recuperado el 5 de Diciembre de 2018, de <https://cran.r-project.org/web/packages/sp/sp.pdf>
- Rapp, R. (1984). *Geometric Geodesy, Volume I*. Ohio: Department of Geodetic Science and Surveying, The Ohio State University.
- Rapp, R., & Pavlis, N. (1990). The Development and Analisis of Geopotential Coefficient Models to Spherical Harmonic Degree 360. *Journal of Geophysical Research*, 95(B13), 885–911. doi:<https://doi.org/10.1029/JB095iB13p21885>
- Renard, D., Ors, F., & Desassis, N. (2017). *RGeostats Manual*. Recuperado el 26 de Noviembre de 2018, de <http://rgeostats.free.fr/doc/Manual/Manual.pdf>

- Reyes, A., Giraldo, R., & Mateu, J. (2015). Residual Kriging for Functional Spatial Prediction of Salinity Curves. *Communications in Statistics - Theory and Methods*, 44(4), 798–809. doi:10.1080/03610926.2012.753087
- Reyes, J., & Wendroth, O. (2017). *PLS 655 Spatial and Temporal Statistics. Manual for Computations Using R*. Lexington, EE.UU.: University of Kentucky.
- Reyes, J., Wendroth, O., Matocha, C., Zhu, J., Ren, W., & Karathanasis, A. (2018). Reliably Mapping Clay Content Coregionalized with Electrical Conductivity. *Soil Physics & Hydrology*, 82(3), 578–592. doi:10.2136/sssaj2017.09.0327
- Ribeiro, P., & Diggle, P. (2016). *Package 'geoR'*. Recuperado el 5 de Diciembre de 2018, de <https://cran.r-project.org/web/packages/geoR/geoR.pdf>
- Rivoirard, J. (2003). *Course on multivariate geostatistics*. París, Francia: Ecole des Mines de Paris.
- Royston, P. (1995). Remark AS R94: A remark on Algorithm AS 181: The W test for normality. *Applied Statistics*, 44, 547–551.
- Sánchez, L. (2003). *Determinación de la superficie vertical de referencia para Colombia*. Múnich, Alemania, Tesis de doctorado: Technische Universität Dresden.
- Sánchez, L., & Martínez, W. (1997). *Guía metodológica para la obtención de alturas sobre el nivel medio del mar utilizando el sistema GPS*. Bogotá, Colombia: Instituto Geográfico Agustín Codazzi.
- Sarkar, D. (2018). *Package 'lattice'*. Recuperado el 5 de Diciembre de 2018, de <https://cran.r-project.org/web/packages/lattice/lattice.pdf>

- Seeber, G. (2003). *Satellite Geodesy* (2da ed.). New York, EE.UU.: Walter de Gruyter.
- Singh, S., & Srivastava, R. (2018). *Development of Geoid Model - A Case Study on Western India*. Estambul, Turquía: FIG Congress 2018.
- SIRGAS. (2002). *Sistema de Referencia Geocéntrico para las Américas. Boletín Informativo N° 6*. Recuperado el 5 de Diciembre de 2018, de [http://www.sirgas.org/fileadmin/docs/Boletines/bol\\_006.pdf](http://www.sirgas.org/fileadmin/docs/Boletines/bol_006.pdf)
- SIRGAS. (2014). *SIRGAS - Definición*. Munich: Deutsches Geodätisches Forschungsinstitut (DGFI). Recuperado el 5 de Diciembre de 2018, de <http://www.sirgas.org/index.php?id=52>
- Stevenson, M., Nunes, T., Heuer, C., Marshall, J., Sanchez, J., Thornton, R., . . . Jay, M. (2018). *Package 'epiR'*. Recuperado el 5 de Diciembre de 2018, de <https://cran.r-project.org/web/packages/epiR/epiR.pdf>
- Thode, H. (2002). *Testing for Normality*. New York, EE.UU.: Marcel Dekker.
- Tierra, A. (2009). Evaluación del EGM08 y EGM96 en el Ecuador a partir de datos de GPS y nivelación geométrica. *Revista Geoespacial*(6), 72–80.
- Tierra, A., & Acurio, V. (2016). Modelo neuronal para la predicción de la altura geoidal local en el Ecuador. *Revista Geoespacial*(13.1), 59–72.
- Torge, W. (2001). *Geodesy* (3era. ed.). Berlín - New York: Walter de Gruyter.
- Tsoulis, D., Novák, P., & Kadlec, M. (2009). Evaluation of precise terrain effects using high-resolution digital elevation models. *Journal of Geophysical Research*, 114(B2), B02404: 1–14. doi:10.1029/2008JB005639

- Vaníček, P., Janák, J., & Véronneau, M. (2000). *Impact of Digital Elevation Models on Geoid Modelling*. Recuperado el 3 de Diciembre de 2018, de <http://www2.unb.ca/gge/Personnel/Vanicek/ImpactOfDEM.pdf>
- Venables, W., & Ripley, B. (2002). *Modern Applied Statistics with S* (4ta. ed.). New York, EE.UU.: Springer-Verlag.
- Volfová, A., & Šmejkal, M. (2012). *Geostatistical Methods in R*. Praga, República Checa: Faculty of Civil Engineering Czech Technical University in Prague.
- Wagner, H. (2003). Spatial covariance in plant communities: integrating ordination, geostatistics, and variance testing. *Ecology*, *84*, 1045–1057.
- Waller, L., & Gotway, C. (2004). *Applied spatial statistics for public health data*. New Jersey, EE.UU.: Wiley-interscience.
- Wendroth, O. (2018). *Spatial and Temporal Statistics: Crossvariograms and Cokriging*. Quito, Ecuador: Universidad Central del Ecuador - Universidad de Kentucky.
- White, J. (1997). Soil Zinc Map of the USA using Geostatistics and Geographic Information Systems. *Soil Science Society of american journal*, *61*, 185–194.
- Yap, B., & Sim, C. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, *81*(12), 2141–2155. doi:<https://doi.org/10.1080/00949655.2010.520163>
- Yun, H. (1999). Precision geoid determination by spherical FFT in and around the Korean peninsula. *Earth Planets Space*, *51*(1), 13–18. doi:<https://doi.org/10.1186/BF03352204>

Zakatov, P. (1997). *Curso de Geodesia Superior*. Madrid, España: Rubiños-1860, S.A.

Zilkoski, D. (1990). *Minimum steps required when estimating GPS-derived orthometric heights*.

Anaheim, EE.UU.: Proc. GIS/LIS'90 Fall convention.