



# 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: DETERMINACIÓN DE MODELOS DE PREDICCIÓN  
ESPACIAL PARA LA TRANSFORMACIÓN DE  
COORDENADAS ENTRE LOS SISTEMAS PSAD56 E ITRF94,  
ÉPOCA 1995.4, UTILIZANDO TÉCNICAS GEOESTADÍSTICAS

AUTORES: GUAJALA AGILA HENRY GONZALO  
ZAMBRANO SOLÍS MARÍA JOSÉ

DIRECTOR: MSc. LEIVA GONZÁLEZ CÉSAR ALBERTO  
SANGOLQUÍ

2016



# 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**

## CERTIFICACIÓN

Certifico que el trabajo de titulación “**DETERMINACIÓN DE MODELOS DE PREDICCIÓN ESPACIAL PARA LA TRANSFORMACIÓN DE COORDENADAS ENTRE LOS SISTEMAS PSAD56 E ITRF94, ÉPOCA 1995.4, UTILIZANDO TÉCNICAS GEOESTADÍSTICAS**”, realizado por los señores: **AGILA GUAJALA HENRY GONZALO** y **ZAMBRANO SOLÍS MARÍA JOSÉ**, ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar a los señores: **AGILA GUAJALA HENRY GONZALO** y **ZAMBRANO SOLÍS MARÍA JOSÉ** para que lo sustente públicamente.

Sangolquí, 17 de Febrero del 2016

Una firma manuscrita en tinta azul, que parece ser la del director César Alberto Leiva González.

César Alberto Leiva González

**DIRECTOR**



**DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA CONSTRUCCIÓN**  
**CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE**

**AUTORÍA DE RESPONSABILIDAD**

Nosotros, **AGILA GUAJALA HENRY GONZALO**, con cédula de identidad N° 1724422350 y **ZAMBRANO SOLÍS MARÍA JOSÉ**, con cédula de identidad N° 1600517542, declaramos que este trabajo de titulación “**DETERMINACIÓN DE MODELOS DE PREDICCIÓN ESPACIAL PARA LA TRANSFORMACIÓN DE COORDENADAS ENTRE LOS SISTEMAS PSAD56 E ITRF94, ÉPOCA 1995.4, UTILIZANDO TÉCNICAS GEOESTADÍSTICAS**” ha sido desarrollado considerando los métodos de investigación existentes, así como también se ha respetado los derechos intelectuales de terceros considerándose en la citas bibliográficas.

Consecuentemente declaramos que este trabajo es de nuestra autoría, en virtud de ello nos responsabilizamos del contenido, veracidad y alcance de la investigación mencionada.

Sangolquí, 17 de Febrero del 2016



Agila Guajala Henry Gonzalo

1724422350



Zambrano Solís María José

1600517542



**DEPARTAMENTO DE CIENCIAS DE LA TIERRA Y LA CONSTRUCCIÓN**  
**CARRERA DE INGENIERÍA GEOGRÁFICA Y DEL MEDIO AMBIENTE**

**AUTORIZACIÓN**

Nosotros, **AGILA GUAJALA HENRY GONZALO** y **ZAMBRANO SOLÍS MARÍA JOSÉ**, autorizamos a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca Virtual de la institución el presente trabajo de titulación **“DETERMINACIÓN DE MODELOS DE PREDICCIÓN ESPACIAL PARA LA TRANSFORMACIÓN DE COORDENADAS ENTRE LOS SISTEMAS PSAD56 E ITRF94, ÉPOCA 1995.4, UTILIZANDO TÉCNICAS GEOESTADÍSTICAS”** cuyo contenido, ideas y criterios son de nuestra autoría y responsabilidad.

Sangolquí, 17 de Febrero del 2016

Agila Guajala Henry Gonzalo

1724422350

Zambrano Solís María José

1600517542

## DEDICATORIA

Este proyecto está dedicado:

A mi Dios por ser mi guía y mi fortaleza.

A mi familia que ha sido mi apoyo y la fuerza que me ha impulsado a seguir adelante, en especial a mis padres Antonio y Doris, mi abuelita Yoli y a mi hermana María Antonieta, los mismos que han estado pendientes de cada uno de los momentos importantes en mi vida, porque ellos siempre han tenido una palabra de aliento para que pueda cumplir mis metas, siendo para mí el mejor regalo que Dios me pudo dar.

A mis compañeros de clase con quienes compartí tantos momentos inolvidables.

A Henry por ser mi amigo, mi compañero, mi hermano de corazón y por siempre contagiarme de su positivismo.

A José que siempre ha sido incondicional y ha compartido con alegría cada una de mis metas.

*María José*

## DEDICATORIA

Este proyecto está dedicado:

A mi Dios por ser mi guía durante toda mi vida universitaria, por ser el pilar y la fuerza en esta dura etapa.

De manera muy especial a mis padres María y Gonzalo, que día a día con esfuerzo y sacrificio han logrado sacar adelante a mí y a mis hermanos. Aprender a valorar sus malas noches, sus ganas de darnos todo lo necesario, su humildad y su tenacidad, es lo más importante que siempre llevaré en mi corazón y en mi vida profesional. Gracias Infinitas.

A mis hermanos Diego, Paola, Ricardo y Carlitos, por ser el significado de risas y peleas, de felicidad y tristeza, de confianza y apoyo durante esta etapa, sin ustedes mi corazón estaría vacío. Gracias por su cariño y por ayudarme a entender el significado de hermandad.

A mis mejores amigos Juan y Bryan por jamás dejarme solo y siempre estar ahí, para lo que necesite.

A todos los amigos que logré en la universidad, a Jeremy, Stalyn, Harry, Diego, Paul, Raúl, Raquel, Naty, Anita, Guillermo, Fabián y en especial a Maria José por ser una amiga incondicional.

A mi PL-QUITO 131, y en especial a Marthita por sus sabios consejos y por su ejemplo de lucha.

A Alexandra que me ha mostrado su confianza y apoyo incondicional.

*Henry*

## AGRADECIMIENTO

En primer lugar queremos agradecer a Dios por ser nuestro guía durante toda nuestra vida universitaria, por ser el pilar y la fuerza en esta dura etapa. De su mano hemos salido adelante, ya que él nos ha librado de enfermedades y peligros, lo cual nos ha permitido esforzarnos cada día en nuestra meta y así poder terminar nuestros estudios.

Agradecemos de manera muy especial a nuestros padres, por la confianza que han puesto en nosotros, y sobre todo por su valioso ejemplo de lucha, sacrificio, honradez y humildad, ya que sin ello no se consigue nada en la vida.

Agradecemos a la Universidad de las Fuerzas Armadas y en especial a nuestros docentes de la Carrera de Ingeniería Geográfica y del Medio Ambiente, que durante toda esta etapa, nos han impartido conocimientos anécdotas y experiencias.

Agradecemos al Ing. César Leiva por ser el mentor y guía en la elaboración de este proyecto.

Agradecemos al Ing. Marco Luna, Dr. Alfonso Tierra y al Ing. Ricardo Romero por su ayuda y apoyo incondicional, con el único fin de compartir conocimientos y promover la investigación en nosotros.

Agradecemos al Instituto Geográfico Militar por abrimos las puertas, y facilitarnos la información y datos, para llevar a cabo este proyecto.

*Los Autores*

## ÍNDICE DE CONTENIDO

### CARÁTULA

<b>CERTIFICACIÓN</b> .....	ii
<b>AUTORÍA DE RESPONSABILIDAD</b> .....	iii
<b>AUTORIZACIÓN</b> .....	iv
<b>DEDICATORIA</b> .....	v
<b>AGRADECIMIENTO</b> .....	vii
<b>ÍNDICE DE FIGURAS</b> .....	xii
<b>ÍNDICE DE TABLAS</b> .....	xiv
<b>RESUMEN</b> .....	xv
<b>ABSTRACT</b> .....	xvi
<b>GLOSARIO</b> .....	xvii
<b>SIGLAS</b> .....	xxii
<b>CAPÍTULO I</b> .....	1
<b>ASPECTOS GENERALES</b> .....	1
1.1. Antecedentes .....	1
1.2. Planteamiento del problema .....	2
1.3. Justificación e importancia del problema .....	3
1.4. Descripción del área de estudio .....	4
1.5. Objetivos .....	5
1.5.1. Objetivo General .....	5
1.5.2. Objetivos Específicos .....	5
1.6. Metas .....	6
<b>CAPÍTULO II</b> .....	7
<b>FUNDAMENTO TEÓRICO</b> .....	7
2.1. Sistema Geodésico de Referencia .....	7
2.1.1. Sistema de Referencia .....	7
2.1.1.1. Sistema de Referencia Topocéntrico .....	7
2.1.1.2. Sistema de Referencia Geocéntrico .....	8
2.2. Marco de Referencia .....	9
2.3. Superficies de Referencia .....	10



2.3.1.	Geoide.....	10
2.3.2.	Elipsoide .....	10
2.3.3.	Superficie Topográfica .....	11
2.3.4.	Relación entre superficie topográfica, geoide y elipsoide .....	11
2.4.	Dátum Geodésico.....	12
2.4.1.	Dátum Geodesico Clásico.....	12
2.4.2.	Dátum Geodésico Moderno .....	13
2.5.	Sistemas de Coordenadas .....	14
2.5.1.	Coordenadas Cartesianas .....	15
2.5.1.1.	Sistema de Coordenadas Cartesianas Globales .....	15
2.5.1.2.	Sistema de Coordenadas (E, N, U) .....	16
2.5.2.	Coordenadas Geodésicas .....	17
2.5.3.	Coordenadas Planas .....	18
2.5.3.1.	Coordenadas Universal Transversa de Mercator (UTM) .....	18
2.6.	Transformación entre Sistemas de Referencia .....	19
2.6.1.	Modelo de Helmert o de Semejanza.....	19
2.6.1.1.	Método de Bursa-Wolf .....	20
2.6.1.2.	Método de Molodensky Badekas.....	21
2.7.	Parámetros de Transformación entre ITRF'S .....	22
	<b>CAPÍTULO III</b> .....	<b>24</b>
	<b>GEOESTADÍSTICA</b> .....	<b>24</b>
3.1.	Estadística Espacial .....	24
3.1.1.	Definiciones .....	24
3.1.1.1.	Geoestadística .....	24
3.2.	Análisis Exploratorio de Datos .....	25
3.2.1.	Introducción .....	25
3.2.2.	Gráficos Exploratorios .....	25
3.3.	Análisis Estructural.....	26
3.3.1.	Un enfoque estocástico para la variación espacial .....	26
3.3.2.	Variables Regionalizadas.....	27
3.3.3.	Estacionariedad.....	28
3.3.3.1.	Estacionariedad de Segundo Orden .....	28

3.3.3.2.	Estacionariedad Débil o Intrínseca .....	29
3.3.4.	Estimación del Variograma .....	29
3.3.4.1.	Características del Variograma.....	30
3.3.5.	Nube Variográfica .....	33
3.3.6.	Modelamiento del Variograma .....	34
3.3.6.1.	Modelo Potencia .....	35
3.3.6.2.	Modelo Lineal.....	36
3.3.6.3.	Modelo Esférico.....	36
3.3.6.4.	Modelo Exponencial.....	37
3.3.6.5.	Modelo Gaussiano .....	38
3.3.6.6.	Modelo Cuadrático .....	38
3.3.6.7.	Modelo Cúbico .....	39
3.3.6.8.	Modelo Circular.....	39
3.3.6.9.	Modelo Pepita Puro .....	40
3.3.6.10.	Modelo Anidado .....	40
3.4.	Predicción Espacial.....	41
3.4.1.	Predicción Espacial Óptima.....	41
3.4.2.	Características Generales del Kriging .....	41
3.4.3.	Kriging Ordinario .....	43
3.4.3.1.	Kriging Ordinario por bloques.....	44
3.4.3.2.	Pesos en el Kriging .....	46
3.4.3.3.	Vecindad de Kriging.....	46
3.4.3.3.1.	Estrategia de búsqueda .....	47
3.4.3.4.	Criterio de vecindad única y vecindad móvil .....	48
3.4.3.4.1.	Criterio de vecindad única .....	48
3.4.3.4.2.	Criterio de vecindad móvil .....	48
3.4.3.5.	Distancia de vecindad de búsqueda .....	49
3.4.3.6.	Tendencia.....	49
3.4.4.	Validación Cruzada .....	50
3.4.5.	Verificación del Modelo .....	50
	<b>CAPÍTULO IV .....</b>	<b>53</b>
	<b>METODOLOGÍA .....</b>	<b>53</b>

4.1.	Recopilación y validación de los datos.....	53
4.2.	Determinación del tamaño de la muestra.....	57
4.3.	Obtención y validación de Modelos Geoestadísticos .....	58
4.3.1.	Modelo Geoestadístico de predicción de la variable desplazamiento de latitud ( $\Delta\phi$ ).....	58
4.3.1.1.	Análisis Exploratorio de Datos .....	58
4.3.1.2.	Variograma Experimental.....	61
4.3.1.2.1.	Tamaño y número del lag .....	61
4.3.1.2.2.	Tolerancia en distancia ( $\epsilon$ ) y tolerancia angular ( $\theta$ ).....	64
4.3.1.2.3.	Cálculo del variograma.....	64
4.3.1.3.	Nube variográfica .....	65
4.3.1.4.	Ajuste del variograma a un modelo teórico.....	66
4.3.1.5.	Validación Cruzada .....	67
4.3.1.6.	Predicción .....	69
4.3.1.7.	Elaboración del mapa de predicción para el manejo en un SIG .....	73
4.3.1.8.	Verificación del Modelo Geoestadístico .....	73
4.3.1.8.1.	Coefficiente de determinación para la verificación del modelo .....	77
4.3.2.	Modelo Geoestadístico de predicción de la variable desplazamiento de longitud ( $\Delta\lambda$ )....	77
4.3.2.1.	Análisis Exploratorio de Datos .....	77
4.3.2.2.	Variograma Experimental.....	79
4.3.2.2.1.	Cálculo del variograma.....	79
4.3.2.3.	Nube variográfica .....	80
4.3.2.4.	Ajuste del variograma a un modelo teórico.....	81
4.3.2.5.	Validación Cruzada .....	82
4.3.2.6.	Predicción .....	83
4.3.2.7.	Elaboración del mapa de predicción para el manejo en un SIG .....	87
4.3.2.8.	Verificación del modelo geoestadístico.....	87
4.3.2.8.1.	Coefficiente de determinación para la validación del modelo .....	90
	<b>CAPÍTULO V</b> .....	<b>91</b>
	<b>ANÁLISIS DE RESULTADOS</b> .....	<b>91</b>
5.1.	Comparación del Modelo Geoestadístico, con el Modelo Matemático de Helmert.....	91

5.2.	Determinación de la escala. ....	94
<b>CAPÍTULO VI</b> .....		97
<b>CONCLUSIONES Y RECOMENDACIONES</b> .....		97
6.1.	Conclusiones.....	97
6.2.	Recomendaciones .....	98
<b>REFERENCIAS</b> .....		99

## ÍNDICE DE FIGURAS

Figura 1.	Mapa de ubicación. ....	5
Figura 2.	Sistemas de Referencia Geocéntrico y Clásico.....	9
Figura 3.	Geoide y Elipsoide de referencia. ....	11
Figura 4.	Relación entre Geoide y Elipsoide.....	12
Figura 5.	Dátum Geodésico Clásico.....	13
Figura 6.	Dátum Geodésico Moderno. ....	14
Figura 7.	Tipos de coordenadas.....	15
Figura 8.	Sistema Cartesiano Geocéntrico. ....	16
Figura 9.	Sistema de coordenadas ENU.....	17
Figura 10.	Sistema de coordenadas geodésico .....	18
Figura 11.	Geometría del Modelo de Bursa-Wolf.....	21
Figura 12.	Modelo de Molodensky-Badekas. ....	22
Figura 13.	Parámetros de transformación del ITRF2008 a ITRFS anteriores. ....	23
Figura 14.	Parámetros del Variograma. ....	31
Figura 15.	Variograma efecto pepita puro. ....	32
Figura 16.	Variograma No Acotado (Proceso solamente Intrínseco) .....	33
Figura 17.	Nube Variográfica.....	34
Figura 18.	Modelo Potencia. ....	35
Figura 19.	Modelo Lineal.....	36
Figura 20.	Modelo Esférico.....	37
Figura 21.	Modelo Exponencial. ....	37
Figura 22.	Modelo Gaussiano. ....	38
Figura 23.	Modelo Cuadrático. ....	39

Figura 24. Modelo Cúbico. ....	39
Figura 25. Modelo Circular. ....	40
Figura 26. Kriging Ordinario por bloques. ....	45
Figura 27. Octantes .....	47
Figura 28. Elección de vecindario móvil de radio r. ....	48
Figura 29. Variograma con crecimiento sistemático. ....	50
Figura 30. Diagramas de caja preliminares. ....	55
Figura 31. Diagrama de dispersión para la variable desplazamiento de latitud. ....	56
Figura 32. Diagrama de dispersión para la variable desplazamiento de longitud. ....	56
Figura 33. Histograma de la variable desplazamiento de latitud .....	59
Figura 34. Diagrama de caja de la variable desplazamiento de latitud .....	60
Figura 35. Gráfico de dispersión para la variable desplazamiento de latitud .....	60
Figura 36. Puntos de muestreo en el Ecuador continental .....	62
Figura 37. Puntos de muestreo sin agrupaciones (sin clusters). ....	63
Figura 38. Prueba de hipótesis que muestra la dispersión de los puntos .....	63
Figura 39. Variograma Experimental Omnidireccional - Variable $\Delta\phi$ . ....	65
Figura 40. Nube Variográfica de la variable desplazamiento de latitud .....	66
Figura 41. Modelamiento de Variograma Experimental - Variable $\Delta\phi$ .....	67
Figura 42. Histograma de Error de Validación Cruzada -Variable $\Delta\phi$ . ....	69
Figura 43. Modelo geoestadístico para la variable desplazamiento de latitud. ....	71
Figura 44. Modelo de error de predicción para el desplazamiento de latitud .....	72
Figura 45. Mapa de la distribución de los puntos de modelamiento y validación. ....	74
Figura 46. Histograma del error del modelo - Variable $\Delta\phi$ . ....	76
Figura 47. Histograma de la variable desplazamiento de longitud .....	78
Figura 48. Diagrama de caja de la variable desplazamiento de longitud .....	78
Figura 49. Gráfico de dispersión de la variable desplazamiento de longitud .....	79
Figura 50. Variograma Experimental Omnidereccional - Variable $\Delta\lambda$ .....	80
Figura 51. Nube Variográfica de la variable desplazamiento de longitud .....	81
Figura 52. Modelamiento de Variograma Experimental - Variable $\Delta\lambda$ . ....	81
Figura 53. Histograma de Error de Validación Cruzada -Variable $\Delta\lambda$ . ....	83
Figura 54. Modelo geoestadístico para la variable desplazamiento de longitud. ....	85
Figura 55. Modelo de error de predicción para el desplazamiento de longitud .....	86

Figura 56. Histograma del error de verificación del modelo - Variable $\Delta\lambda$ .....	89
Figura 57. Diferencia en la posición en metros (m).....	94
Figura 58. Diferencia en la posición en metros (m).....	94

## ÍNDICE DE TABLAS

Tabla 1 Tipos de predictores Kriging y sus propiedades .....	42
Tabla 2 Resumen estadístico para la variable desplazamiento de latitud .....	58
Tabla 3 Cálculo del tamaño de lag adecuado .....	64
Tabla 4 Parámetros de Estructura Básica - Variable $\Delta\phi$ .....	67
Tabla 5 Error de verificación del modelo ( $\Delta\lambda_{\text{OBSERVADA}} - \Delta\lambda_{\text{PREDICHA}}$ ) .....	75
Tabla 6 Resumen estadístico error de verificación del modelo desplazamiento de latitud.....	76
Tabla 7 Resumen estadístico para la variable desplazamiento de longitud .....	77
Tabla 8 Parámetros de Estructura Básica - Variable $\Delta\phi$ .....	82
Tabla 9 Error de verificación del modelo ( $\Delta\phi_{\text{OBSERVADA}} - \Delta\phi_{\text{PREDICHA}}$ ) .....	88
Tabla 10 Resumen estadístico error de verificación del modelo desplazamiento de longitud.....	89
Tabla 11 Resumen estadístico para la comparación de metodologías .....	92
Tabla 12 Cálculo del RMSE .....	96

## RESUMEN

Los sistemas de referencia son estructuras geométricas utilizadas para referir coordenadas de puntos en el espacio, clasificados a partir de su dátum, en sistemas de referencia topocéntricos y geocéntricos. Actualmente, con el avance de las tecnologías y metodologías de medición debido a la aparición del sistema GPS, la información desarrollada en el sistema de referencia local PSAD56, es incompatible con el sistema de referencia global SIRGAS95. Para lo cual, Leiva (2003) calculó 7 parámetros de transformación, mediante el modelo matemático de Helmert, para compatibilizar la geoinformación entre los dos sistemas de referencia, establecida como metodología estándar del Ecuador, compatible hasta escalas 1:25000. En este estudio se realizó una metodología alternativa usando técnicas geoestadísticas, a través de modelos de predicción espacial para las variables desplazamiento de latitud y longitud ( $\Delta\phi$ ,  $\Delta\lambda$ ), los mismos que fueron aplicados a la transformación de coordenadas geodésicas ( $\phi$ ,  $\lambda$ ), entre ambos sistemas de referencia. Se usó 149 puntos geodésicos, destinando 107 puntos para el modelamiento con kriging ordinario puntual en el software R, paquete Rgeostats y los 42 puntos restantes para la verificación del modelo. Al comparar los resultados obtenidos con el método geoestadístico y el determinístico (Helmert), se comprobó que el método geoestadístico es más eficiente para transformar coordenadas entre estos sistemas de referencia, mejorando los estadísticos descriptivos del error de transformación, dando como resultado un error máximo de 0,567 m. al 95% de confianza, mientras el método de Helmert presentó un error máximo de 1.35 m.; siendo esta metodología aplicable a escalas 1:2000 y menores.

### PALABRAS CLAVES

- **SISTEMAS DE REFERENCIA**
- **MÉTODO DE HELMERT**
- **GEOESTADÍSTICA**
- **KRIGING**
- **PREDICCIÓN**

## ABSTRACT

Reference systems are geometric structures used to refer coordinates of points in space, ranked from its datum systems topocentric and geocentric reference. Actually, with the advancement of technologies and measurement methodologies due to the appearance of the GPS system, the information developed in the local reference system PSAD56, is incompatible with the global reference system SIRGAS95. For that, Leiva (2003) estimated seven transformation parameters by the mathematical model of Helmert, to compatible between the two geo-reference systems established as standard methodology Ecuador, supports up scales of 1: 25,000. In this study an alternative methodology was performed using geostatistical techniques through models of spatial prediction for the variable displacement of latitude and longitude ( $\Delta\phi$ ,  $\Delta\lambda$ ), the same as were applied to the transformation of geodetic coordinates ( $\phi$ ,  $\lambda$ ) between the two reference systems. 149 geodetic points was used, allocating 107 points for the modeling with ordinary kriging point in the R software, Rgeostats package and the remaining 42 points for the verification of the models. When comparing the results obtained with the geostatistical method and deterministic (Helmert), it was found that the geostatistical method is more efficient to transform coordinates between the reference systems, improving the descriptive statistical error processing and resulting in a maximum error of 0.567 m. 95% confidence, while the method of Helmert presented a maximum error of 1.35 m.; so, this methodology applicable at scales of 1: 2000 and minor scales.

## KEYWORDS

- REFERENCE SYSTEM
- HELMERT METHOD
- GEOSTATISTICS
- KRIGING
- PREDICTION



## GLOSARIO

**Anisotropía.-** Una propiedad del proceso espacial o de los datos donde la dependencia espacial (autocorrelación) cambia con la distancia y la dirección entre dos localizaciones (ESRI, 2003).

**Autocorrelación.-** Correlaciones estadísticas entre variables espaciales aleatorias del mismo tipo, atributo, nombre, etc., donde la correlación depende de la distancia y / o la dirección que separa las localizaciones. Define como cambian los valores de la variable a partir de su ubicación (ESRI, 2003).

**Artefactos.-** En los procedimientos de interpolación (kriging), cuando se utiliza una vecindad móvil o un modelo de semivariograma con un efecto pepita importante crea discontinuidades con la superficie interpolada (artefactos), saltando hacia arriba o hacia abajo para tomar cualquier valor de un punto que corresponde a un nodo de la grilla (punto de estimación), se trata de formas artificiales que afectan notablemente la calidad del modelo (Bohling, 2005).

**Cuartil.-** Es una medida de posición no central. A partir de tres valores se divide la distribución en cuatro partes iguales, o cuatro intervalos, dentro de cada uno está: el 25% de los datos de la distribución ( $Q_1$ ), definido como el primer cuartil que representa el 25% de la distribución.  $Q_2$ , es el segundo cuartil que corresponde al 50% de la distribución y  $Q_3$ , el tercer cuartil que representa el 75% de la distribución (INEI, 2006).

**Curtosis.-** Es una medida de forma, conocida como medida de apuntamiento ya que mide si los valores de la distribución se encuentran concentrados alrededor de los valores medios de la muestra. Se definen 3 tipos de distribuciones según su grado de curtosis: distribución leptocúrtica ( $Cr > 0$ ), mesocúrtica ( $Cr = 0$ ) y distribución platicúrtica ( $Cr < 0$ ) (INEI, 2006).

**Coefficiente de asimetría (As).-** Es la medida que determina la simetría de la distribución de una variable respecto a su media, es decir, si hay el mismo número de datos a la izquierda y derecha de la media. Se pueden presentar los siguientes casos:  $As > 0$ , la distribución es asimétrica positiva, la cola de la distribución se alarga para valores superiores a la media.  $As = 0$ , la distribución es simétrica, hay el mismo

número de elementos a la izquierda y derecha de la media (distribución normal).  $As < 0$ , la distribución es asimétrica negativa, la cola de la distribución se alarga para valores inferiores a la media (INEI, 2006).

**Coefficiente de determinación.-** Es un coeficiente que determina la bondad de ajuste del modelo desarrollado. Este valor se obtiene del cuadrado del coeficiente de correlación y se representa por  $R^2$ . Puede tomar valores entre 0 y 1. (INEI, 2006).

**Coefficiente de variabilidad.** Es una medida de dispersión relativa, calculada a partir de la división de la desviación estándar para la media. La ventaja de este coeficiente es que no tiene unidades. Se interpreta como porcentaje y se denota por  $CV$  (INEI, 2006).

**Correlación.-** Determina la relación o dependencia que existe entre las variables (INEI, 2006).

**Desviación Estándar.-** Es una medida de dispersión que se obtiene de la raíz cuadrada de la varianza. Este estadístico se mide en la misma unidad que la variable por lo da una mejor interpretación que la varianza (INEI, 2006).

**Dependencia Espacial .-** La idea de que las cosas cercanas entre sí son más similares que las cosas más separadas (ESRI, 2003).

**Diagrama de Caja.-** Es un gráfico para el análisis de datos, que permite tener una idea visual de la distribución de los muestra. Permite determinar si hay simetría, ver el grado de variabilidad existente y detectar los “outliers” (posibles datos discordantes). Se denota también como BOXPLOT. (INEI, 2006).

**Distribución Normal o Curva Normal.-** Es la distribución de probabilidad más utilizada en estadística y teoría de probabilidad. Conocida también como distribución de Gauss (INEI, 2006).

**Histograma.-** Es un gráfico de barras que representa la distribución de frecuencias de una variable continua. Describe el comportamiento de un conjunto de datos en cuanto a su tendencia central, forma y dispersión. La anchura de las barras muestra el rango de valores en cada grupo, y la altura de la barra indica el número de valores en cada grupo (ESRI, 2003).

**Interpolar.-** La estimación de los valores en los lugares donde no se ha observado datos, utilizando los datos de los lugares donde se ha recopilado dicha información. Por lo general, la interpolación es para estimaciones dentro de la zona en la que se

han recogido los datos, en lugar de extender las estimaciones a áreas fuera de la zona de interés (ESRI, 2003).

**Intervalo de Confianza.-** Es el rango de valores donde se presume se encuentra el valor del parámetro, a partir de una probabilidad determinada. Generalmente se constituye intervalos de confianza con 95% de probabilidad (INEI, 2006).

**Isotropía.-** Una propiedad de un proceso natural o de los datos donde la dependencia espacial (autocorrelación) cambia sólo con la distancia entre dos ubicaciones (la dirección no es importante) (ESRI, 2003).

**Kriging .-** Un método de interpolación estadístico que utiliza un solo tipo de datos (atributo individual) para predecir (interpolarse) los valores de ese mismo conjunto de datos en localizaciones no muestreadas. Kriging también proporciona los errores estándar de las predicciones (ESRI, 2003).

**Kriging Ordinario.-** Predicción espacial (interpolación) utilizando modelos semivariograma o covarianza que se basan en las relaciones espaciales entre los datos. Kriging ordinario hace ciertas suposiciones sobre el modelo, como la estacionariedad intrínseca, y que la media real de los datos es constante, pero desconocida (ESRI, 2003).

**Lag .-** Es la distancia  $h$  que se conoce como paso ("lag" en inglés). Un lag es un vector, tiene distancia y dirección (ESRI, 2003).

**Media.-** La media se obtiene al sumar todos los valores observados, dividido por el número total de observaciones. (INEI, 2006).

**Mediana.-** Es el valor que divide al conjunto de datos ordenados, en aproximadamente dos partes: 50% de valores son inferiores a este valor y otro 50% son superiores al mismo, por lo cual se la conoce como una medida de tendencia central. (INEI, 2006).

**Muestreo aleatorio simple.** Es una modalidad del muestreo probabilístico donde cada elemento de la población tiene la misma probabilidad de ser seleccionado para integrar la muestra (INEI, 2006).

**Pepita.-** Parámetro de un modelo de covarianza o semivariograma que representa el error independiente, el error de medición en la variable o a la escala de la misma. El efecto pepita es visto como una discontinuidad en el origen de cualquier modelo de covarianza o semivariograma (ESRI, 2003).

**Rango.-** Es un valor que mide la amplitud de los valores de un conjunto de datos y se calcula por diferencia entre el valor mayor y el valor menor:  $R = X_{mayor} - X_{menor}$  (INEI, 2006).

**Rango (variograma) .-** Un parámetro de un modelo de covarianza o semivariograma que representa una distancia más allá de la cual hay poco o nada de autocorrelación entre las variables, es decir, define el límite de la dependencia espacial (ESRI, 2003).

**Recorrido Intercuartílico.-** Es una medida de dispersión, donde su valor se obtiene de la diferencia del tercer cuartil ( $Q_3$ ) menos el primer cuartil ( $Q_1$ ) como se expresa a continuación:  $R1 = Q_3 - Q_1$  (INEI, 2006).

**Semivariograma.-** El variograma dividido por dos (ESRI, 2003).

**Sesgo.-** Se denomina así al grado de inclinación que adopta la distribución de los datos respecto a su valor promedio tomado como centro de gravedad. Si la inclinación se encuentra a la izquierda, se determina un sesgo negativo, mientras si la inclinación es a la derecha se determina un sesgo positivo (INEI, 2006).

**Sill.-** Un parámetro de un modelo de variograma o semivariograma que representa el máximo valor que alcanza un semivariograma. A grandes distancias, no hay correlación en las variables, por lo que el sill de la semivariograma es igual a la varianza de la variable aleatoria. Algunos semivariogramas teóricos no tienen un sill. Todos los modelos de semivariogramas utilizados en el Análisis Geostatístico tienen un sill (ESRI, 2003).

**Validación .-** El procedimiento donde se extrae parte de los datos y el resto de los datos se utiliza para predecir la parte extraída de los datos (ESRI, 2003).

**Variable Aleatoria.-** Conocida también como variable estocástica o probabilística. Es la característica considerada en un experimento aleatorio cuyo valor de ocurrencia sólo puede saberse con exactitud una vez observado (INEI, 2006).

**Variable Continua.-** Es una variable cuantitativa, cuyos valores se representan a través del conjunto de números reales. Puede tomar cualquier valor dentro de un intervalo (INEI, 2006).

**Varianza.-** Es una medida de dispersión de la información, obtenida del promedio de los cuadrados de las desviaciones de los valores de la variable respecto a su media aritmética, por lo cual siempre será mayor que cero. Cuando más se acerca a cero, los valores se encuentran concentrados alrededor de la media. Por el contrario, mientras

mayor sea la varianza, más dispersos están los datos. Tiene el inconveniente de ser poco significativo, debido a que se mide en el cuadrado de la unidad de la variable (INEI, 2006).

**Variograma.**-Una función de la distancia y la dirección que separa dos localizaciones, para cuantificar autocorrelación. El variograma se define como la varianza de la diferencia entre dos variables en dos localizaciones. El variograma generalmente aumenta con la distancia, y es descrito por los parámetros de pepita, sill, y rango (ESRI, 2003).

## SIGLAS

- ASPRS:** American Society Photogrammetric and Remote Sensing.
- CNUGGI:** Comité Nacional de la Unión Geodésica y Geofísica Internacional.
- DORIS:** Doppler Orbitography and Ratio-positoning Integrated by Satellite.
- ESRI:** Environmental Systems Research Institute.
- GLONASS:** Russia's Global Navigation Satellite System.
- GNSS:** Global Navigation Satellite System.
- GPS:** Global Positioning System.
- GRS80:** Geodetic Reference System 1980.
- IERS:** International Earth Rotation and Reference Systems Service.
- IGM:** Instituto Geográfico Militar.
- INEI:** Instituto Nacional de Estadística e Informática.
- ITRF:** International Terrestrial Reference Frame.
- ITRS:** International Terrestrial Reference System.
- LLR:** Lunar Laser Ranging.
- NAVSTAR:** Navigation Satellite Timing and Ranging.
- NSSDA:** National Standard for Spatial Data Accuracy.
- RENAGE:** Red Nacional GPS del Ecuador.
- SIRGAS:** Sistema de Referencia Geocéntrico para las Américas.
- SLR:** Satellite Laser Ranging.
- PSAD56:** Provisional South American Datum 1956.
- VEMOS:** Velocity Model for SIRGAS.
- VLBI:** Very Long Baseline Interferometry
- WGS84:** World Geodetic System 1984.

# CAPÍTULO I

## ASPECTOS GENERALES

### 1.1. Antecedentes

Los Sistemas Geodésicos de Referencia son indispensables para referir las observaciones geodésicas y estimar los parámetros a una base global única, la cual se materializa a través de un Marco de Referencia que tenga un dátum estable a largo plazo y un conjunto de puntos que hayan sido posicionados con alta precisión y exactitud (Santacruz, 2010). El Instituto Geográfico Militar (IGM), organismo rector de la cartografía en el Ecuador establece la Red Nacional GPS del Ecuador (RENAGE), como el Marco Geodésico de Referencia Nacional materializado a través de 135 vértices geodésicos y el sustento técnico por el cual el Ecuador cruzó el umbral de la Geodesia Convencional a la era de los Sistemas Satelitales de Navegación Global (Cisneros, 2013).

La RENAGE está enlazada al Sistema de Referencia Geocéntrico para las Américas (SIRGAS) (realización SIRGAS 95, ITRF94, época 1995.4). SIRGAS, es la densificación del Marco de Referencia Terrestre Internacional (ITRF) en América Latina y El Caribe. Las realizaciones o densificaciones de SIRGAS asociadas a diferentes épocas y soluciones del ITRF materializan el mismo sistema de referencia y sus coordenadas, reducidas a la misma época y marco de referencia, son compatibles en el nivel milimétrico (SIRGAS, 2015). Anteriormente en el Ecuador se utilizaba un Sistema de Referencia Local, el Datum Provisional para América del Sur de 1956 (PSAD56), que es un sistema que posee como elipsoide de referencia el Internacional de Hayford y como punto de origen La Canoa ubicado en la República de Venezuela; incompatible con el uso del Sistema de Posicionamiento Global (GPS). Hoy en día se manejan sistemas de referencia geocéntricos que son compatibles con las distintas metodologías y tecnologías, como los sistemas satelitales de navegación global (GNSS).

Para homogenizar las mediciones dentro del territorio se han realizado distintas metodologías para la transformación entre el Sistema oficial PSAD56 y el sistema

SIRGAS95, como la “Determinación de parámetros de transformación entre los sistemas PSAD56 y WGS84 para el País”, que resolvieron la diferencia entre coordenadas con fines cartográficos hasta escalas 1:25000 y menores, con el método de transformación conforme de Helmert (Leiva, 2003). Otro estudio es, la “Transformación entre el Sistema PSAD56 y los Marcos de Referencia ITRF utilizando los modelos de Helmert y de velocidades de placas tectónicas VEMOS” (Santacruz, 2010). Adicionalmente se realizó la investigación de una nueva “Metodología para la Transformación entre Sistemas de Referencia utilizando una Red Neuronal Artificial” (Romero & Tierra, 2014).

## **1.2. Planteamiento del problema**

La RENAGE se encuentra enlazada al Sistema de Referencia Geocéntrico SIRGAS95, el cual se usa actualmente en el Ecuador para el manejo y obtención de información cartográfica. Sin embargo, municipios o niveles de gobierno aún cuentan con cartografía referenciada en el Sistema de Referencia Clásico PSAD56, definido como dátum oficial según la Ley de Cartografía del país, incompatible con el cambio tecnológico y la evolución de las técnicas de medición con la aparición de los sistemas globales de navegación satelital como GPS; por ende, se han desarrollado diversas metodologías que automatizan el proceso de transformación de coordenadas entre los dos sistemas de referencia mencionados.

En el año 2003 se realizó la transformación de coordenadas entre los dos sistemas mediante el cálculo de 7 parámetros de transformación (3 traslaciones, 3 rotaciones y un factor de escala), establecidos de manera oficial por el IGM y desarrollado por Leiva (2003), compatible hasta escalas 1:25000. Posteriormente, Romero y Tierra (2014) realizaron otra metodología para la transformación de coordenadas, mediante el uso de redes neuronales artificiales. Estas investigaciones alcanzan una buena precisión a una escala determinada, pero aún no logran encontrar una eficiente compatibilidad entre ambos sistemas de referencia. Debido a lo expuesto, el proyecto busca mediante técnicas geoestadísticas, modelar la diferencia que existe entre ambos sistemas de referencia, partiendo de 150 puntos geodésicos del IGM, densificados en el Ecuador continental.



De inicio se puede mencionar que esta diferencia matemática tiene un comportamiento determinístico, por lo cual la variable desplazamiento de latitud ( $\Delta\varphi$ ) y longitud ( $\Delta\lambda$ ) son variables determinísticas; pero, según Meyer (2009), la variable a modelar puede presentar aleatoriedad, debido a consideraciones físicas del experimento, evidencia empírica o en algunos casos, simplemente un juicio personal basado en una experiencia previa con una situación similar; de esta manera se adopta un modelo probabilístico para la caracterización de dichas variables sin descartar todas las relaciones determinísticas.

El modelo determinístico planteado inicialmente se convierte en un modelo probabilístico expresado de la siguiente manera:

$$\Delta\varphi = (\varphi_{\text{PSAD56}} - \varphi_{\text{SIRGAS95}}) + \varepsilon \quad (1.1)$$

$$\Delta\lambda = (\lambda_{\text{PSAD56}} - \lambda_{\text{SIRGAS95}}) + \varepsilon \quad (1.2)$$

El ruido estocástico ( $\varepsilon$ ) explica las fuentes de variabilidad con lo cual el modelo determinístico se convierte, como ya se mencionó, en un modelo probabilístico. La elección de adoptar entre un modelo determinístico o probabilístico puede ser difícil de hacer algunas veces. Puede depender de lo intrincado de nuestra técnica de medida y la precisión asociada. Por ejemplo, si las medidas precisas son tan difíciles de obtener que las lecturas repetidas de la misma cantidad produzcan resultados variables, un modelo probabilístico es indudablemente más adecuado para describir la situación (Meyer, 2009). Con este modelo probabilístico se busca transformar con el menor error, toda la información generada en el Sistema PSAD56 al Sistema SIRGAS95 y de ésta manera manejar toda la geoinformación en un solo sistema de referencia.

### **1.3. Justificación e importancia del problema**

Existe la necesidad de estandarizar las mediciones, observaciones y levantamientos georreferenciados dentro del territorio nacional, adoptando un sistema de referencia geocéntrico único que de mejor respuesta a la ubicación y distribución espacial de los datos geográficos asociados a los sistemas de posicionamiento global: GPS, GLONASS, GALILEO, etc. Por esta razón, es

importante e indispensable la transformación de la información que se encuentra en el Sistema de Referencia PSAD56 al SIRGAS95, con la finalidad de actualizar y compatibilizar la geoinformación y la cartografía existente en el país, con el menor error posible. De esta manera toda la información generada en PSAD56 no se desecharía, ya que es la base para los diversos estudios que se realizan en el Ecuador. Tradicionalmente, para la transformación entre sistemas de referencia se utilizan modelos matemáticos como el de Helmert en el cual se calculan parámetros de transformación. Esta investigación pretende dar una nueva metodología a la transformación de coordenadas, mediante el uso de técnicas geoestadísticas que a través de modelos de predicción espacial, determinan el comportamiento de las variables desplazamiento de latitud ( $\Delta\phi$ ) y desplazamiento de longitud ( $\Delta\lambda$ ), entre los sistemas PSAD56 y SIRGAS95.

#### **1.4. Descripción del área de estudio**

El área de estudio se encuentra delimitada por el polígono que se muestra en la Figura 1. Este polígono fue determinado a partir de 150 puntos geodésicos, de los cuales, 84 puntos son observados tanto en el sistema PSAD56 y SIRGAS95, y 66 puntos observados en el sistema PSAD56 y el marco ITRF08 (previamente transformados al marco ITRF94, época 1995.4).



## 1.5. Objetivos

### 1.5.1. Objetivo General

- Determinar dos modelos de predicción espacial para la transformación de coordenadas entre los sistemas de referencia PSAD56 y SIRGAS95, mediante la utilización de técnicas Geoestadísticas.

### 1.5.2. Objetivos Específicos

- Recopilar y validar los datos de las coordenadas de los puntos en los sistemas PSAD56 y SIRGAS95.
- Generar un modelo de predicción para la variable desplazamiento de latitud ( $\Delta\phi$ ), entre los sistemas de referencia PSAD56 y SIRGAS95, utilizando técnicas Geoestadísticas.

- Generar un modelo de predicción para la variable desplazamiento de longitud ( $\Delta\lambda$ ), entre los sistemas de referencia PSAD56 y SIRGAS95, utilizando técnicas Geoestadísticas.
- Validar los modelos de predicción geoestadísticos, a través de puntos de control.
- Comparar los resultados de la transformación de coordenadas entre el método estocástico desarrollado y el método matemático de Helmert utilizado por Leiva (2003).

### **1.6. Metas**

- Una base de datos de las coordenadas de los puntos en los sistemas PSAD56 y SIRGAS95.
- Un modelo de predicción para la variable desplazamiento de latitud ( $\Delta\varphi$ ), entre los sistemas de referencia PSAD56 y SIRGAS95, en formato .grid.
- Un modelo de predicción para la variable desplazamiento de longitud ( $\Delta\lambda$ ), entre los sistemas de referencia PSAD56 y SIRGAS95, en formato .grid.
- Estadísticos de la validación de los modelos obtenidos.
- Una tabla de comparación de los resultados de la transformación de coordenadas entre el método estocástico desarrollado y el método matemático de Helmert utilizado por Leiva (2003).

## **CAPÍTULO II**

### **FUNDAMENTO TEÓRICO**

#### **2.1. Sistema Geodésico de Referencia**

##### **2.1.1. Sistema de Referencia**

Un sistema de referencia es una estructura geométrica para referir coordenadas de puntos en el espacio, el cual define la orientación y ubicación de los ejes coordenados; el sistema se establece a partir de un conjunto de convenciones, normas y conceptos teóricos que definen en cualquier momento, modelos, parámetros y constantes, que son la base necesaria para la representación de cantidades geométricas y físicas, tanto en los sistemas de referencia geodésicos (latitud, longitud, altura), como en los sistemas de referencia cartesianos (X,Y,Z) (Santacruz, 2010).

##### **2.1.1.1. Sistema de Referencia Topocéntrico**

El sistema de referencia topocéntrico o local es aquel en el cual el origen de coordenadas está desplazado del geocentro, adaptando de esta manera un elipsoide de referencia al geoide y buscando un acoplamiento en una zona de interés. El elipsoide elegido y posicionado, se adapta adecuadamente al geoide en las inmediaciones del punto dátum (siempre que la elección haya sido criteriosa), pero a medida que se aleja crece la probabilidad de que esta adaptación disminuya. Por esta razón los sistemas así definidos fueron utilizados por países, permitiendo llevar adelante todos los proyectos geodésicos en sus respectivos territorios (CNUGGI, 1999).

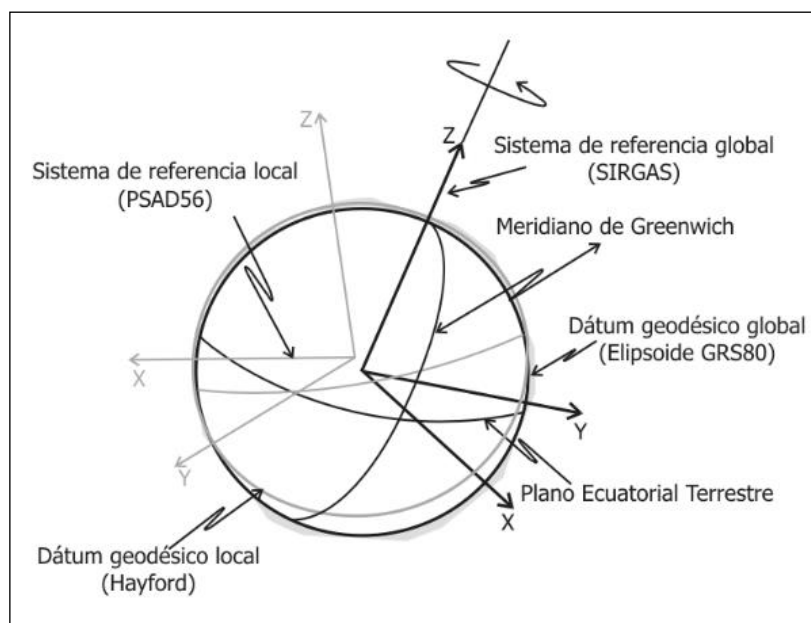
Este es el caso del Sistema PSAD56 utilizado de manera oficial en el Ecuador, regido bajo la Ley de Cartografía Nacional, la cual menciona como elipsoide de referencia el Internacional de Hayford y su punto origen o dátum materializado sobre la superficie terrestre en la Canoa-Venezuela.

### 2.1.1.2. Sistema de Referencia Geocéntrico

Según Leiva (2003), se define un sistema de referencia geocéntrico a aquel en el cual el origen de coordenadas cartesiano (X, Y, Z) del sistema coincide con el centro de masas de la Tierra (ver figura 2). En comparación de los sistemas geodésicos locales, los sistemas geocéntricos son tridimensionales y de alcance global. El concepto de dátum define parámetros que fijan el origen, la orientación de la terna de referencia y la escala del sistema de coordenadas (Drewes, 2014). Entre los sistemas de referencia geocéntricos más conocidos se encuentran:

- **World Geodetic System 1984 (WGS84):** Es un sistema geodésico desarrollado por el Departamento de Defensa de Estados Unidos de alcance mundial, el cual se alinea con los criterios y convenciones establecidos por la IERS (International Earth Rotation and Reference Systems Service). Empleado para definir las órbitas radiotransmitidas por los satélites NAVSTAR GPS (Romero, 2013).
- **International Terrestrial Reference System (ITRS):** Según Santacruz (2010), es un sistema geocéntrico que co-gira con la Tierra en su movimiento diurno en el espacio. Es mantenido por el IERS, y entre sus principales características está la determinación de la orientación del eje de rotación de la Tierra, información sobre la escala del Tiempo Universal Coordinado (UTC) y el mantenimiento del Marco de Referencia Terrestre Internacional (ITRF) (Moirano, 2000).
- **Sistema de Referencia Geocéntrico para las Américas (SIRGAS):** Es un sistema de referencia asociado a las convenciones del IERS, densificado mediante la materialización del ITRF que cuenta con la cooperación de todos los países de las Américas ONU (citado en Romero, 2013). Se estableció en 1993, en la Conferencia Internacional para la Definición de un Sistema de Referencia para América del Sur, celebrada en Asunción, Paraguay; sus mediciones GPS fueron realizadas en mayo/junio de 1995, con errores en las coordenadas de los puntos SIRGAS en el orden de 1cm, siendo el sistema de referencia para la solución definitiva el ITRF 94, época 1995.4 y mediante su origen, orientación y escala se define al dátum geodésico SIRGAS en base al

elipsoide GRS80. Como tal, el proyecto SIRGAS difundió los procedimientos que los países deberán seguir para obtener redes geodésicas GPS con la más alta precisión y compatibilidad en el mismo sistema de referencia y época (CNUGGI, 1999).



**Figura 2. Sistemas de Referencia Geocéntrico y Clásico**

Fuente: (Leiva, 2003)

## 2.2. Marco de Referencia

Según Drewes (2014), el marco de referencia materializa un sistema de referencia físicamente y lo realiza matemáticamente, a través de observaciones, es decir, se trata de un conjunto de puntos ubicados en la superficie terrestre, con coordenadas y velocidades conocidas respecto a ese sistema de referencia. El marco está formado por coordenadas cartesianas tridimensionales (X, Y, Z) dadas en una época fija y su variación respecto al tiempo (velocidades) de una serie de estaciones. Para ello se emplean técnicas de observación espacial como:

- VLBI
- SLR
- LLR
- GPS

- DORIS

El ITRF establecido y mantenido por la IERS constituye el marco de referencia del ITRS (International Earth Reference System). El IERS publica soluciones de estas observaciones, determinando un ITRF al que se le asignan dos dígitos identificativos; por ejemplo, para el año 1994 es el ITRF94 (Benitez & García-Asenjo, 2001).

### 2.3. Superficies de Referencia

#### 2.3.1. Geoide

El Geoide se define como la superficie equipotencial del campo de gravedad de la Tierra que mejor se adapta al nivel medio del mar, libre de perturbaciones exteriores y extendiéndose por debajo de los continentes, océanos y mares (Fernández, 2001). El Geoide además, es la superficie de referencia para las observaciones de nivelación geométrica, con la que se definen las alturas (ver figura 3).

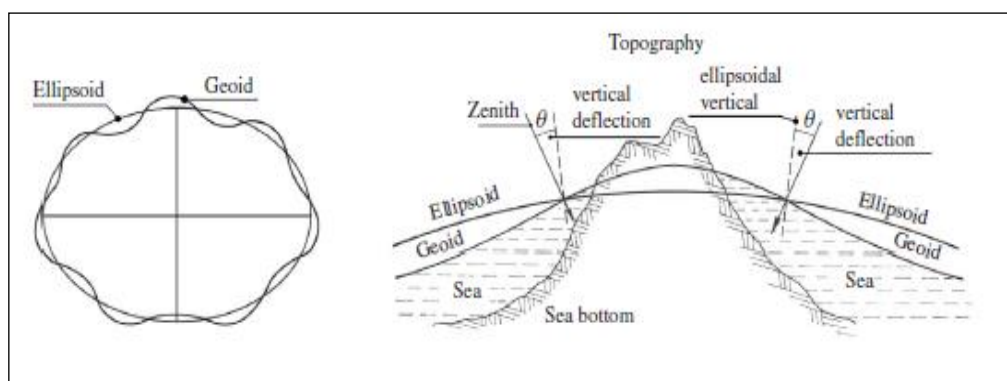
#### 2.3.2. Elipsoide

Debido a que la Tierra no es redonda, y su figura es similar a una esfera achatada en los polos, nace la necesidad de tener una superficie de referencia que represente a la forma de la Tierra. Esta superficie se denomina elipsoide, que resulta de revolucionar una elipse sobre su eje menor (b). Esta elipse de revolución está definida por su semieje mayor (a) y su achatamiento (f).

$$f = \frac{a-b}{a} \quad (2.1)$$

En un sistema de coordenadas geodésico, el centro de la elipse de revolución coincide con el centro del sistema de referencia, y su semieje menor se hace coincidir con el eje de rotación terrestre. (Furones, 2010).





**Figura 3. Geoide y Elipsoide de referencia.**

Fuente: (Seeber, 2003)

### 2.3.3. Superficie Topográfica

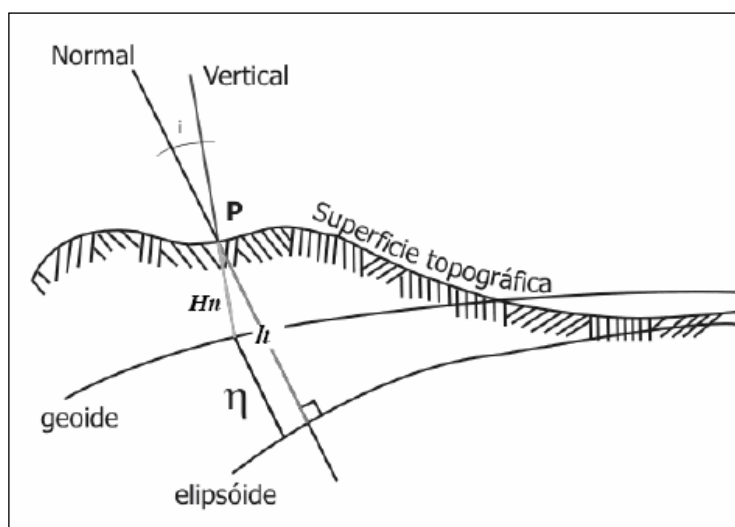
La superficie física o topográfica de la Tierra es la frontera entre la parte sólida y la fluida, donde se realizan las mediciones de distancias geodésicas, ángulos, entre otros. Los fondos oceánicos estarían incluidos en esta definición como la superficie de discontinuidad entre la parte sólida y las masas de agua oceánicas. No es posible representar la superficie topográfica a través de una función analítica, por lo que se obtiene un modelo de la misma mediante la definición de coordenadas establecidas en un sistema de referencia geodésico, y la definición de métodos de interpolación (García-Asenjo & Hernández, 2005).

### 2.3.4. Relación entre superficie topográfica, geoide y elipsoide

La relación entre la superficie topográfica, elipsoide y el geoide se da mediante la ondulación geoidal ( $\eta$ ), a través de la siguiente ecuación:

$$h = H + \eta \quad (2.2)$$

En donde  $H$  es la altura ortométrica, que en ocasiones para fines prácticos, se puede sustituir  $H$  por  $H_n$  que corresponde a la altura nivelada (medida en la superficie topográfica), siendo  $h$  la altura elipsoidal (ver figura 4).



**Figura 4. Relación entre Geoide y Elipsoide.**

Fuente: (Leiva, 2003)

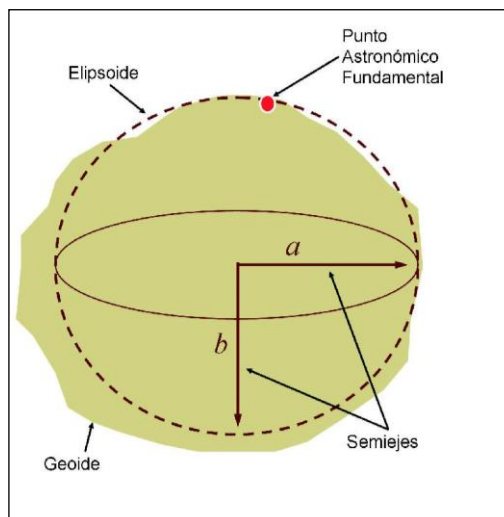
## 2.4. Dátum Geodésico

Se define el dátum geodésico como el conjunto de parámetros que definen la posición, tamaño y orientación de un elipsoide respecto a la Tierra, por ende su conexión con un sistema de referencia. Para definir un datum es importante conocer la geometría del elipsoide  $a$  y  $f$ , su posición respecto al geocentro  $\Delta X$ ,  $\Delta Y$ ,  $\Delta Z$  (coordenadas del centro del elipsoide respecto al geocentro), su orientación  $R1$ ,  $R2$ ,  $R3$  (orientación de los ejes del elipsoide respecto a los terrestres) y la escala  $k$ . Si  $\Delta X = \Delta Y = \Delta Z = 0$ , el dátum Geodésico se conoce como global o absoluto (Seeber, 2003).

### 2.4.1. Dátum Geodesico Clásico

El dátum clásico está asociado a un sistema local cuyo alcance es un espacio geográfico o región determinada (ver figura 5). Es también llamado dátum horizontal, ya que la determinación de la altura se hace independiente de las coordenadas latitud y longitud. El posicionamiento y orientación de un dátum horizontal se define con respecto al sistema astronómico local de un punto cualquiera, el cual, convencionalmente, le proporcionaba el nombre al dátum correspondiente (Leiva, 2003). Un dátum horizontal consta de las siguientes cantidades iniciales: la latitud y la longitud de un punto inicial (origen); el acimut de

una línea (dirección); el semi-eje mayor y el achatamiento del elipsoide seleccionado para los cálculos; y la ondulación geoidal en el origen (Sosa, 1980, citado en Leiva, 2003).

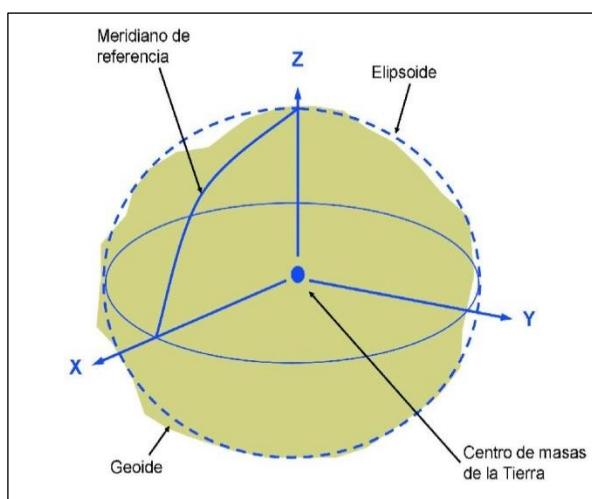


**Figura 5. Datum Geodésico Clásico.**

Fuente: (Atlas Global de la Región de Murcia, 2010)

#### 2.4.2. Datum Geodésico Moderno

Según Leiva (2003), el datum geodésico o geocéntrico está asociado a un sistema de referencia global válido para cualquier punto del planeta, cuyo fin es unificar la plataforma de referencia para la definición de coordenadas a nivel mundial (ver figura 6). Con la implementación del GPS, el Departamento de Defensa de los Estados Unidos de América implementa la serie WGS (World Geodetic System), cuya característica fundamental es que su origen de coordenadas coincide con el centro de masas de la Tierra. Actualmente está en uso a nivel mundial el datum WGS84, que al concebirse con fines militares, la Asociación Internacional de Geodesia (IAG) impulsa la versión civil de los sistemas de referencia global como los GRS (Geodetic Reference System) y en especial el uso del datum GRS80 que en la práctica es similar al WGS84 (Santacruz, 2010).



**Figura 6. Datum Geodésico Moderno.**

Fuente: (Atlas Global de la Región de Murcia, 2010)

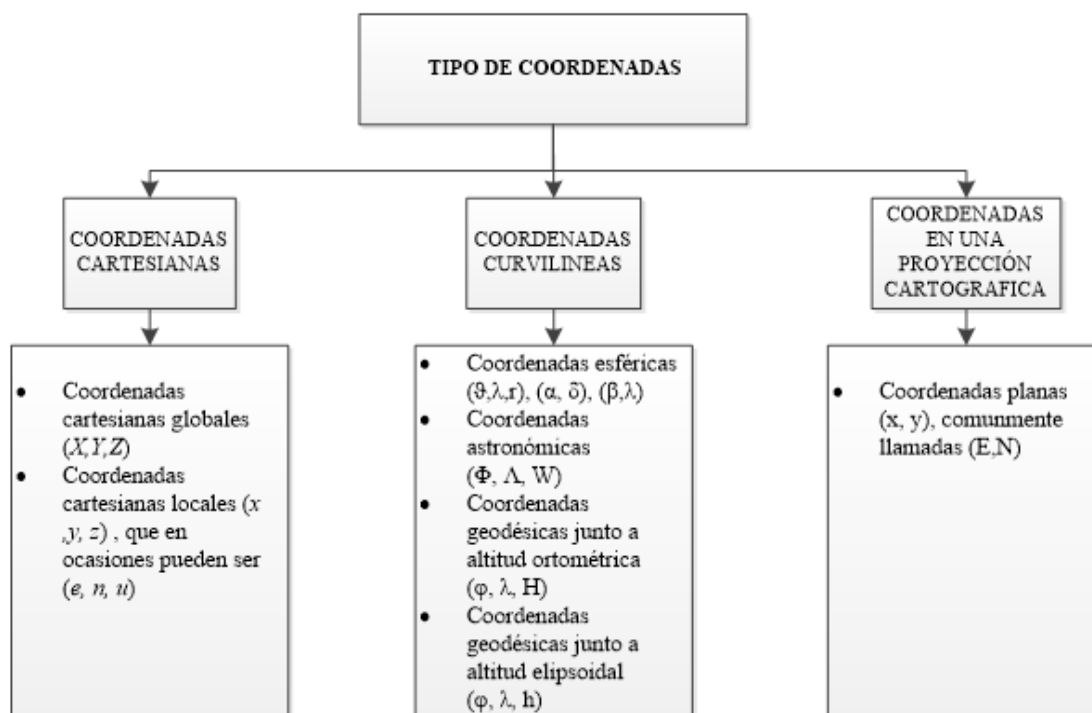
## 2.5. Sistemas de Coordenadas

Un sistema de coordenadas establece la localización espacial de un punto. El utilizar un sistema de coordenadas particular permite que el objeto de estudio se represente en su forma más simple, geoméricamente interpretable y susceptible de ser medida (García-Asenjo & Hernández, 2005).

Drewes (2014), afirma que para definir un sistema de coordenadas se debe especificar:

- El tipo de coordenadas (rectilíneo, curvilíneo, planas)
- La ubicación del origen
- La orientación de los ejes
- La unidad de medida

Es importante conocer que no se pueden “medir” coordenadas, se calculan en un sistema. Los Sistemas de coordenadas terrestres están fijos a la Tierra y giran con ella; son usados para definir las coordenadas de puntos sobre la superficie terrestre. Generalmente, son clasificados en función del tipo de coordenadas (García-Asenjo & Hernández, 2005):



**Figura 7. Tipos de coordenadas.**

Fuente: (García-Asenjo & Hernández, 2005)

Los sistemas de coordenadas más empleados en Geodesia son:

- Coordenadas cartesianas globales  $(X, Y, Z)$
- Coordenadas geográficas (elipsoidales)  $(\phi, \lambda, h)$
- Coordenadas planas  $(E, N)$
- Coordenadas cartesianas locales  $(E, N, U)$

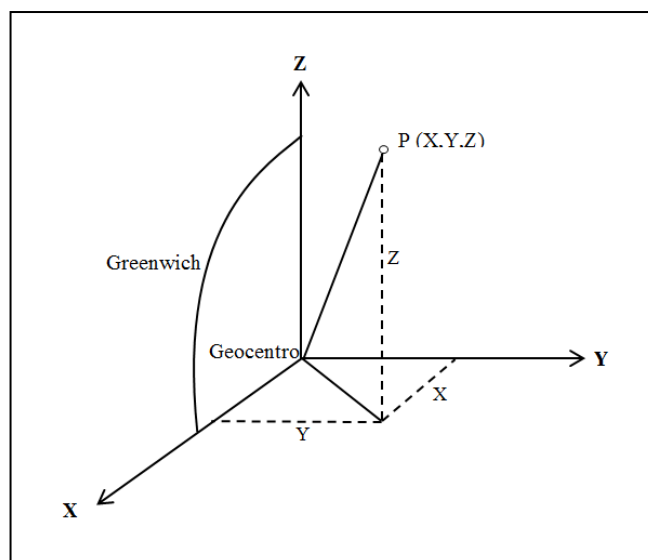
### 2.5.1. Coordenadas Cartesianas

#### 2.5.1.1. Sistema de Coordenadas Cartesianas Globales

Este sistema de coordenadas cartesiano global (ver figura 8) se define a partir de:

- Su origen es el centro de masas terrestre.
- Ejes rectilíneos ortogonales.
- El eje Z coincide con el eje de rotación y perpendicular al plano XY, plano del ecuador.

- El eje X en el plano  $Z=0$ , está orientado en la dirección del meridiano de Greenwich.
- El eje Y en el plano  $Z=0$ , perpendicular a los dos anteriores ejes.



**Figura 8. Sistema Cartesiano Geocéntrico.**

Fuente: (Adaptado de García-Asenjo & Hernández, 2005)

La ortogonalidad de sus ejes constituyen un sistema derecho, es decir, cumple con la ley de la mano derecha. Este sistema de coordenadas es uno de los sistemas más utilizados en Geodesia y su unidad de medida es métrica (Drewes, 2014).

### 2.5.1.2. Sistema de Coordenadas (E, N, U)

Consiste en un sistema de referencia local, ligado al campo de gravedad, es decir, su orientación está en función de la vertical local (cenit, línea de la plomada). Los ejes  $x$ ,  $y$ ,  $z$  equivalen a *neu* o también denominado *enu*, que corresponde a la direcciones este, norte y altura (Drewes, 2014).

- El eje  $z$  apunta en la dirección del cenit, perpendicular al plano  $x$ ,  $y$ .
- El eje  $x$  apunta hacia el norte elipsoidal (meridiano).
- El eje  $y$  apunta hacia el este y forma así un sistema de mano izquierda.

El vector expresado en componentes ENU se obtiene a partir del vector expresado en coordenadas cartesianas geocéntricas (Drewes, 2014) (ver figura 9); de acuerdo a la siguiente expresión  $(\Delta X, \Delta Y, \Delta Z)$ .

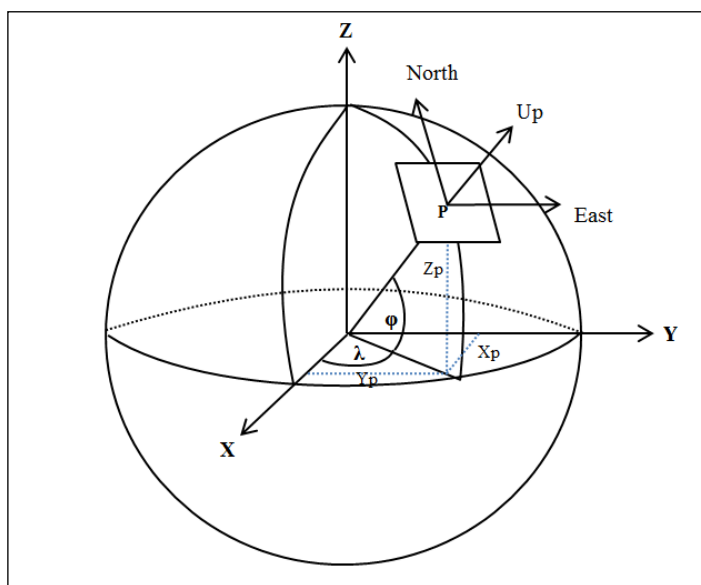
De forma matricial la expresión para el sistema local (E, N, U) es:

$$\begin{pmatrix} E \\ N \\ U \end{pmatrix} = R^{-1}(\varphi, \lambda) \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} \quad (2.3)$$

Donde la matriz inversa de rotación es:

$$R^{-1}(\varphi, \lambda) = \begin{pmatrix} -\text{sen}\lambda & \cos\lambda & 0 \\ -\text{sen}\lambda\cos\lambda & -\text{sen}\varphi\text{sen}\lambda & \cos\varphi \\ \cos\varphi\cos\lambda & \cos\varphi\text{sen}\lambda & \text{sen}\varphi \end{pmatrix} \quad (2.4)$$

Y  $\Delta X, \Delta Y, \Delta Z$ , corresponden al vector entre las coordenadas cartesianas geocéntricas.



**Figura 9. Sistema de coordenadas ENU.**

Fuente: (Adaptado de Drewes, 2014)

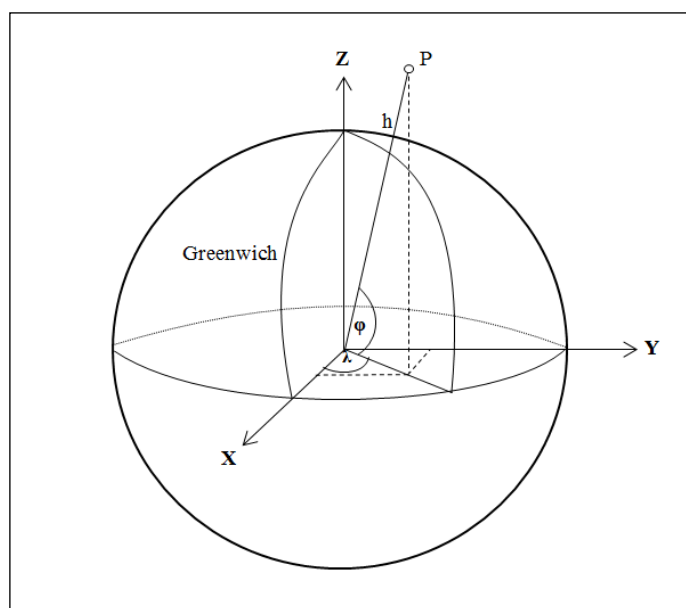
### 2.5.2. Coordenadas Geodésicas

Las coordenadas geográficas o geodésicas representan las coordenadas básicas en los sistemas geodésicos convencionales, por tal su uso práctico en la navegación, la cartografía y la ingeniería (Leiva, 2014). Su origen se encuentra en el centro

geométrico de un elipsoide de revolución adoptado como modelo matemático para la Tierra.

Según Nelsi (2006), en este sistema de coordenadas un punto se determina bidimensionalmente con dos coordenadas longitud ( $\lambda$ ) y latitud ( $\varphi$ ) sobre una superficie de referencia que suele ser un elipsoide. Las coordenadas pueden ir acompañadas de una altura elipsoidal o sobre el nivel medio del mar. Los elementos de las coordenadas geodésicas son definidos y representados en la Figura 10:

- **Latitud geodésica ( $\varphi$ ):** Ángulo medido en el plano meridiano que forman la normal al elipsoide en el punto P y el plano del ecuador.
- **Longitud geodésica ( $\lambda$ ):** Ángulo medido en el plano del ecuador en el sentido directo que forman el plano meridiano que contiene el punto P considerado y el plano meridiano de Greenwich.
- **Altura elipsoidal ( $h$ ):** Distancia entre el punto P y el elipsoide, medida a lo largo de la normal al elipsoide que pasa por dicho punto. Esta altitud no tiene ningún significado físico, solamente carácter geométrico.



**Figura 10. Sistema de coordenadas geodésico**

Fuente: (Adaptado de García-Asenjo & Hernández, 2005)

### 2.5.3. Coordenadas Planas

#### 2.5.3.1. Coordenadas Universal Transversa de Mercator (UTM)



El sistema de proyección UTM, es un sistema en el cual se constituye geoméricamente el mapa, de manera que los meridianos y paralelos se transformen en una red regular rectangular, conservando los ángulos originales. Para lo cual, un elipsoide de referencia se proyecta sobre un cilindro tangente a un meridiano establecido (meridiano central), de forma que al desarrollar el cilindro, el Ecuador se transforma en una recta que se toma como eje de las X, y el meridiano central se transforma en otra recta perpendicular a la anterior que será el eje de las Y. Este tipo de transformación se denomina conforme y toma como base la proyección Mercator. La universalidad de la proyección UTM se da al considerar 60 zonas de 6° de ancho en longitud, y una zona de proyección entre los paralelos 84° N y 80° S en latitud ; el resultado, es la representación completa de la superficie de referencia a excepción de las regiones polares. (Fernández, 2001).

Para evitar que la distorsión de las magnitudes lineales aumente conforme se aumenta la distancia al meridiano central se emplea un factor de escala 0.9996, en el meridiano central, así que las líneas a 1°37' al este y oeste presentarán escala verdadera (Drewes, 2014). El Ecuador utiliza la Proyección Universal Transversa de Mercator (UTM), siendo esta la usada en mapas topográficos por sus especiales características, ya que representa a la Tierra sobre el plano. El Ecuador continental se encuentra en las zonas 17 y 18 y la parte insular, Galápagos, en las zonas 15 y 16.

Para esta investigación la referencia espacial utilizada en la base de datos para asignar una posición a cada registro fue la proyección UTM Zona 17 Sur para mantener la compatibilidad con la cartografía, ya que la mayor parte del Ecuador continental se encuentra en la zona 17.

## **2.6. Transformación entre Sistemas de Referencia**

### **2.6.1. Modelo de Helmert o de Semejanza**

El modelo matemático Helmert o de Semejanza, llamado así por Fiedrich Robert Helmert, quién lo formuló en el año 1880, consiste en un modelo de transformación tridimensional, el cual establece un sistema inicial de coordenadas cartesianas en el espacio, las cuales son transformadas mediante la aplicación de 7 parámetros de transformación (3 traslaciones, 3 rotaciones y un factor de escala) a un nuevo

sistema de coordenadas, donde las diferencias entre sistemas están determinadas por los factores de traslación, rotación y escalamiento (Rebolledo, 2010).

Las tres traslaciones prueban la no coincidencia de los orígenes en los dos sistemas de referencia; las tres rotaciones son necesarias para expresar el no paralelismo y el factor de escala es requerido para homogenizar los dos sistemas. Para el cálculo de parámetros de transformación de un sistema a otro, se necesitan puntos comunes cuyas coordenadas cartesianas son conocidas en ambos sistemas (Leiva, 2003).

Según Ramírez y Ortiz (2003), este modelo entrega una transformación exacta sólo en sistemas geodésicos de referencia matemáticos y perfectamente definidos; por ejemplo: dos sistemas de tipo satelital. No es recomendable el uso en sistemas geodésicos de referencia local, ya que fueron realizados por mediciones terrestres clásicas, sujetas a errores generados por la tecnología de la época; por esta razón, mientras mayor distorsión presente la red clásica peor será la transformación tridimensional de Helmert.

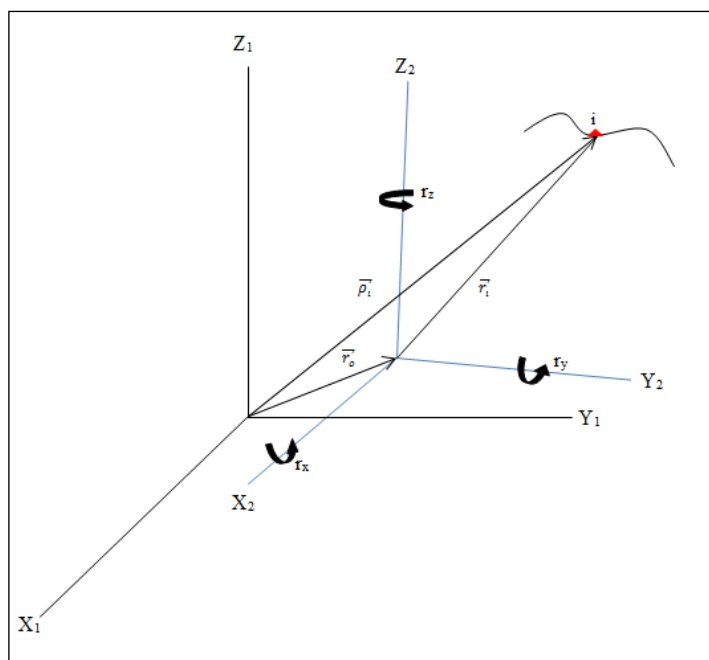
### 2.6.1.1. Método de Bursa-Wolf

El modelo de Bursa-Wolf se basa en una transformación de coordenadas cartesianas tridimensionales entre los dátums de origen y de destino.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_2 = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} + (1 + \delta) * \begin{bmatrix} 1 & \varepsilon_z & -\varepsilon_y \\ -\varepsilon_z & 1 & \varepsilon_x \\ \varepsilon_y & -\varepsilon_x & 1 \end{bmatrix} * \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_1 \quad (2.6)$$

Los siete parámetros de la transformación son tres traslaciones ( $T_x$ ,  $T_y$ ,  $T_z$ ), que van dirigidas desde el centro del sistema de referencia local hacia el centro del sistema de referencia global, tres ángulos de rotación ( $\varepsilon_x$ ,  $\varepsilon_y$ ,  $\varepsilon_z$ ) aplicados a cada uno de los ejes del sistema de coordenadas y el factor de escala ( $1 + \delta$ ), donde  $\delta$  es expresado en partes por millón (ppm). En la ecuación la matriz de rotación es una simplificación, válida siempre y cuando los ángulos de rotación sean pequeños.  $X_1$ ,  $Y_1$ ,  $Z_1$  es el dátum origen y  $X_2$ ,  $Y_2$ ,  $Z_2$  es el dátum destino (Mateu, Valero, & Blanquer, 2006).

Thompson (1976), expresa que el Modelo de Bursa-Wolf es adecuado para la combinación de dos redes de satélites, pero no para un terrestre y una red geodésica satelital. La geometría de la expresión del modelo se muestra en la Figura 11.



**Figura 11. Geometría del Modelo de Bursa-Wolf.**

Fuente: (Adaptado de Thompson, 1976)

### 2.6.1.2. Método de Molodensky Badekas

Este modelo de transformación fue discutido por Molodensky (1962) y Badekas (1969). Es un modelo que relaciona dos sistemas de referencia tridimensionales mediante los 7 parámetros de transformación de Helmert, pero a diferencia del modelo de Bursa-Wolf, este modelo incluye un centroide de coordenadas, es decir realiza la transformación alrededor del baricentro de las coordenadas transformadas en lugar del origen de coordenadas geocéntrico, por lo cual es recomendado para la transformación entre un sistema geodésico de referencia satelital y un sistema geodésico de referencia local (Ramírez y Ortiz, 2003 citado en Rebolledo, 2010).

Para estimar los 7 parámetros de transformación en este modelo se necesita de dos pasos fundamentales; el primero consiste en determinar las tres rotaciones y el factor de escala y el segundo en determinar las tres traslaciones, pero previamente ambos

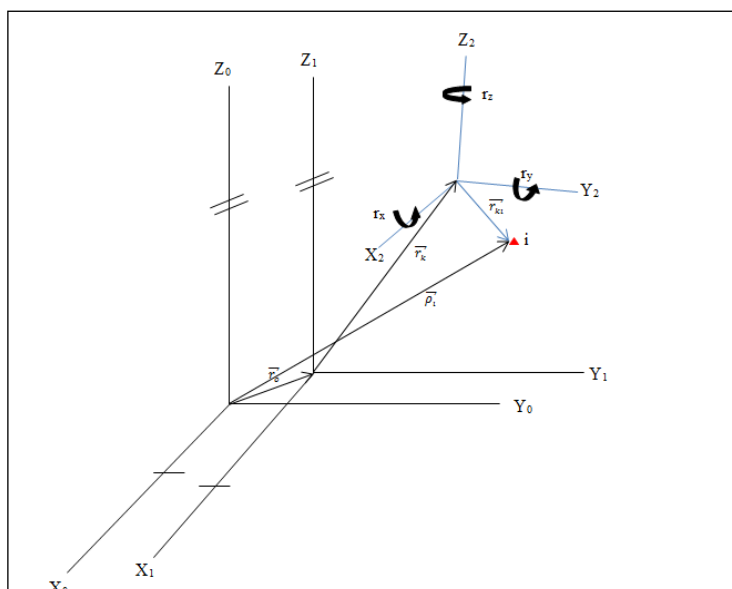
sistemas se deben referir a los baricentros mencionados anteriormente (Cañizares & Leiva, 2013).

La siguiente ecuación es una adaptación del modelo de Bursa Wolf, con la diferencia que este modelo añade un baricentro para la determinación de la coordenada final  $(X_2, Y_2, Z_2)$  como se indica a continuación:

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} + (1 + \delta) * \begin{bmatrix} 1 & \varepsilon_z & -\varepsilon_y \\ -\varepsilon_z & 1 & \varepsilon_x \\ \varepsilon_y & -\varepsilon_x & 1 \end{bmatrix} * \begin{bmatrix} X_1 - X_0 \\ Y_1 - Y_0 \\ Y_1 - Z_0 \end{bmatrix} \quad (2.7)$$

Dónde:

$(X_0, Y_0, Z_0)^t$ : Vector de coordenadas del baricentro.



**Figura 12. Modelo de Molodensky-Badekas.**

Fuente: (Adaptado de Thompson, 1976)

## 2.7. Parámetros de Transformación entre ITRF'S

Los parámetros de transformación entre ITRF's y sus variaciones en el tiempo (rates) se derivan de las publicaciones anuales que realiza el IERS. La relación estándar de transformación entre dos ITRF's es una similitud Euclidiana de siete parámetros: tres componentes de traslación, un factor de escala y tres ángulos de rotación, designados respectivamente:  $T_x, T_y, T_z, D, R_x, R_y, R_z$ , y los rates:  $\dot{T}_x, \dot{T}_y, \dot{T}_z, \dot{D}, \dot{R}_x, \dot{R}_y, \dot{R}_z$  (ITRF, 2015).

Para la transformación entre ITRF's se utiliza los siete parámetros de transformación como lo indica la Figura 13.

Transformation parameters from ITRF2008 to past ITRFs.								
SOLUTION	Tx	Ty	Tz	D	Rx	Ry	Rz	EPOCH
UNITS----->	mm	mm	mm	ppb	.001"	.001"	.001"	
RATES	Tx	Ty	Tz	D	Rx	Ry	Rz	
UNITS----->	mm/y	mm/y	mm/y	ppb/y	.001"/y	.001"/y	.001"/y	
ITRF2005	-2.0	-0.9	-4.7	0.94	0.00	0.00	0.00	2000.0
rates	0.3	0.0	0.0	0.00	0.00	0.00	0.00	
ITRF2000	-1.9	-1.7	-10.5	1.34	0.00	0.00	0.00	2000.0
rates	0.1	0.1	-1.8	0.08	0.00	0.00	0.00	
ITRF97	4.8	2.6	-33.2	2.92	0.00	0.00	0.06	2000.0
rates	0.1	-0.5	-3.2	0.09	0.00	0.00	0.02	
ITRF96	4.8	2.6	-33.2	2.92	0.00	0.00	0.06	2000.0
rates	0.1	-0.5	-3.2	0.09	0.00	0.00	0.02	
ITRF94	4.8	2.6	-33.2	2.92	0.00	0.00	0.06	2000.0
rates	0.1	-0.5	-3.2	0.09	0.00	0.00	0.02	
ITRF93	-24.0	2.4	-38.6	3.41	-1.71	-1.48	-0.30	2000.0
rates	-2.8	-0.1	-2.4	0.09	-0.11	-0.19	0.07	
ITRF92	12.8	4.6	-41.2	2.21	0.00	0.00	0.06	2000.0
rates	0.1	-0.5	-3.2	0.09	0.00	0.00	0.02	
ITRF91	24.8	18.6	-47.2	3.61	0.00	0.00	0.06	2000.0
rates	0.1	-0.5	-3.2	0.09	0.00	0.00	0.02	
ITRF90	22.8	14.6	-63.2	3.91	0.00	0.00	0.06	2000.0
rates	0.1	-0.5	-3.2	0.09	0.00	0.00	0.02	
ITRF89	27.8	38.6	-101.2	7.31	0.00	0.00	0.06	2000.0
rates	0.1	-0.5	-3.2	0.09	0.00	0.00	0.02	
ITRF88	22.8	2.6	-125.2	10.41	0.10	0.00	0.06	2000.0
rates	0.1	-0.5	-3.2	0.09	0.00	0.00	0.02	

**Figura 13. Parámetros de transformación del ITRF2008 a ITRFS anteriores.**

Fuente: (ITRF, 2015)

## **CAPÍTULO III**

### **GEOESTADÍSTICA**

#### **3.1. Estadística Espacial**

La estadística espacial es una disciplina de la estadística general que trata del análisis descriptivo e inferencial de datos geográficos. Un concepto básico de este término es el de dependencia o autocorrelación espacial, que analiza la falta de independencia que se produce entre las observaciones de una variable para sus diferentes localizaciones (Martori, 2008).

Para Giraldo (2005), la estadística espacial es la reunión de un conjunto de metodologías apropiadas para el análisis de datos que corresponden a la medición de variables aleatorias en diversos sitios de una región, de otra forma, se puede decir que trata con el análisis de realizaciones de un proceso estocástico. En estadística espacial se distinguen tres tipos de datos: datos geoestadísticos o georreferenciados (geostatistical data), datos en rejilla o datos en un área (lattice data), y datos de procesos puntuales o patrones espaciales (point processes data). Este estudio se basa en la aplicación de técnicas geoestadísticas, para la obtención de un modelo espacial de predicción para las variables de desplazamiento de latitud y longitud ( $\Delta\varphi$  y  $\Delta\lambda$ ).

##### **3.1.1. Definiciones**

###### **3.1.1.1. Geoestadística**

El término “geoestadística” fue expuesto por G. Matheron en los años 60. La geoestadística es una rama de la estadística aplicada que se especializa en el análisis y la modelación de la variabilidad espacial en ciencias de la tierra. Actualmente, la geoestadística es un conjunto de técnicas usadas para analizar y predecir valores de una propiedad distribuida en espacio o tiempo. A diferencia de la estadística clásica o convencional, tales valores no se consideran independientes, por el contrario se suponen de manera implícita que están correlacionados unos con otros, es decir que existe una dependencia espacial. De una manera más simple, esto quiere decir que

mientras más cercanos estén situados dos puntos están más correlacionados y mientras más separados hay menos relación entre ellos (Díaz, 2002).

Para aplicar técnicas geoestadísticas debe existir continuidad espacial en los datos, con ello se puede conocer la forma en que varía cualquier variable continua en el espacio. Según Díaz (2002) un análisis geoestadístico está compuesto por tres etapas: (a) el análisis exploratorio de los datos, (b) el análisis estructural y (c) las predicciones en sitios no muestreados.

La primera etapa, aplica técnicas estadísticas convencionales que nos permiten obtener todo un conjunto de información, desconocida a priori sobre la muestra bajo estudio. La segunda describe la correlación entre puntos en el espacio y la tercera permite la predicción en sitios de la región no muestreados por medio de la técnica de Kriging.

## **3.2. Análisis Exploratorio de Datos**

### **3.2.1. Introducción**

La aplicación de la Geoestadística está prevista de una serie de pasos que son de suma importancia para el correcto análisis y desarrollo de un modelo de predicción. Inicialmente se realiza el Análisis Exploratorio de Datos (AED), el cual permite estudiar los datos muestrales sin tener en cuenta su distribución geográfica, es decir es una etapa de aplicación de la estadística. Se comprueba la consistencia de los datos, para posteriormente eliminar los erróneos e identificar las distribuciones de las que provienen.

El AED además de examinar la distribución de los datos permite identificar valores extremos, buscar tendencias globales, examinar la variabilidad, correlación espacial y realizar transformaciones si son requeridas, con el fin de entender de una manera diferente el fenómeno a ser investigado y así establecer supuestos necesarios para definir que procedimiento de predicción es el más conveniente (Giraldo, 2005).

### **3.2.2. Gráficos Exploratorios**

Los gráficos exploratorios son construidos para establecer suposiciones sobre el mecanismo de generación de los datos. Según (Chasco, 2003) un buen método

gráfico es aquel capaz de analizar y representar dos características fundamentales en toda distribución espacial: alisado y asperezas. El alisado hace referencia a la tendencia central de la variable (medida a través de la mediana) y la dispersión de los datos, mientras las asperezas se relacionan con aquellos datos localizados a cierta distancia de la mediana, como los llamados atípicos (outliers). Los gráficos que representan el alisado son histogramas, diagramas de tallo y hoja, mientras para las asperezas se tiene los diagramas de caja y de bigote, y gráficos de dispersión. Cabe mencionar que los gráficos de dispersión facilitan la detección de tendencia en los datos, bajo el supuesto de estacionariedad (el nivel promedio de la variable debe ser constante en todos los puntos del área de estudio) que establece la teoría geoestadística. La nube de puntos resultante, incluso el ajuste de una línea de regresión, permite establecer de manera empírica si existe dicha tendencia (Giraldo, 2005).

### **3.3. Análisis Estructural**

El análisis estructural es la parte principal dentro del proceso geoestadístico, ya que radica en estimar y modelar una función que refleje la correlación espacial de la variable regionalizada, adoptando la hipótesis más adecuada acerca de su variabilidad. Dependiendo de las características de estacionariedad del fenómeno, se modelará la función de covarianzas o la de semivarianzas. Por su importancia, generalmente se estudia la estimación y modelación de la función de semivarianzas o semivariograma (Díaz, 2002).

#### **3.3.1. Un enfoque estocástico para la variación espacial**

Para el análisis de la variabilidad espacial de una variable; Matheron encontró un enfoque alternativo al tradicional determinista, que consiste en tratar a la variable de interés como una variable aleatoria, siguiendo una metodología estocástica o probabilística. Esto implica que para cada punto  $x$ , en el espacio hay una propiedad  $Z(x)$  que representa una variable aleatoria con media  $\mu$  y varianza  $\sigma^2$  (Oliver, 2010).

La variabilidad espacial utiliza funciones para modelar la variación de las variables en determinadas regiones, que son utilizadas posteriormente para interpolar en el espacio el valor de la variable en lugares que no han sido muestreadas. Es



importante conocer que la variabilidad espacial se relaciona directamente con la autocorrelación espacial de los datos, ya que permitirá definir como varían los valores de la variable en función de su localización. El primer paso en el análisis de la variabilidad espacial consiste en la visualización de la regularidad o irregularidad de la distribución observada (Castillejo, 2007).

### 3.3.2. Variables Regionalizadas

Para Alfaro (2007), una variable regionalizada es una función que representa la variación en el espacio de una cierta magnitud asociada a un fenómeno natural, de manera que posee una estructura espacial de correlación. Matemáticamente la variable regionalizada se representa como la variable aleatoria  $Z$  definida en un punto en el espacio  $x$  tal como  $Z(x)$ . Puede denotarse como  $Z(x)$ ,  $Z(x_1, x_2)$  y  $Z(x_1, x_2, x_3)$  dependiendo del espacio dimensional del problema (1D, 2D, 3D respectivamente).

La geoestadística moderna considera a la variable de interés como variable aleatoria con una media  $\mu$  y varianza  $\sigma^2$ . El conjunto de variables aleatorias,  $Z(x_1)$ ,  $Z(x_2)$ ,..... es un proceso aleatorio y el actual valor de  $Z$  observado es solo un número cualquiera entre las potenciales realizaciones de este proceso. Por lo tanto, se puede estimar la covarianza espacial para describir esta relación entre pares de puntos; para una variable aleatoria, como:

$$C(x_1, x_2) = E\{[Z(x_1) - \mu(x_1)][Z(x_2) - \mu(x_2)]\} \quad (3.1)$$

Dónde:  $\mu(x_1)$  y  $\mu(x_2)$  son las medias de  $Z$  en  $x_1$  y  $x_2$ , y  $E$  denota el valor esperado. Las predicciones de la variable regionalizada son puntuales, por tal razón la media de la variable es desconocida lo que hace necesario involucrar el concepto de estacionariedad (Oliver, 2010).

Para describir o modelar un fenómeno regionalizado se puede adoptar una posición determinística que conduzca a una evaluación precisa de los valores de la regionalización a partir de un número limitado de observaciones. Sin embargo, ello requiere un conocimiento detallado de la génesis del fenómeno y de las leyes físicas o matemáticas que rigen la evolución de la variable regionalizada.

Además muchos de los fenómenos regionalizados que habitualmente se estudian son lo suficientemente complejos como para que un marco determinista sólo pueda

dar una visión muy parcial de los mismos. Es por ello que nos vemos obligados a renunciar al enfoque determinista y adoptar un enfoque probabilístico que permite formalizar tanto los conocimientos como la incertidumbre que se tenga en cuanto al fenómeno aleatorio regionalizado (Larraz & Montero, 2008).

Así, desde otro punto de vista, el valor regionalizado puede ser visto como el resultado de algún mecanismo aleatorio, mecanismo que recibe el nombre de variable aleatoria. La variable regionalizada se interpreta como una realización de una función aleatoria. La variable regionalizada suele ser muy irregular localmente (lo cual impide su representación por medio de una función matemática determinista).

### 3.3.3. Estacionariedad

La Estacionariedad se refiere a un proceso invariante cuando hay un cambio en su posición espacial. La falta de estacionariedad puede deberse bien a la existencia de anomalías en el espacio, bien a la existencia de una tendencia o gradiente espacial cuya dimensión sea mayor que el área de estudio. La teoría geoestadística se basa en la hipótesis de estacionariedad y puede definirse principalmente por la estacionariedad de segundo orden, pero también existen algunos fenómenos físicos que desarrollan una estacionariedad débil o intrínseca (Castillejo, 2007).

#### 3.3.3.1. Estacionariedad de Segundo Orden

Sea  $\{Z(x): x \in D \subseteq R^d\}$  una variable regionalizada definida en un dominio  $D$  contenido en  $R^d$  (generalmente una variable medida en la superficie de una región). Se dice que  $Z(x)$  es estacionario de segundo orden si cumple:

- El valor esperado de la variable aleatoria es finito y constante para todo  $x$ , en el dominio  $D$ , esto es,  $E[Z(x)] = \mu, \mu \in R, \forall x \in D \subseteq R^d$
- La función de covarianza existe y sólo depende de la distancia entre las localizaciones involucradas, esto es,  $C(Z(x), Z(x+h)) = C(h), \forall x, x+h \in D$ , siendo  $h$ , el vector distancia entre dichas localizaciones. La distancia  $h$  se conoce como paso ("lag" en inglés).

La existencia de la covarianza implica que la varianza existe, es finita y no depende de  $h$ , es decir,  $V(Z(x_i)) = C(0) = \sigma^2$ , donde  $C(0)$  recibe el nombre de varianza a priori del proceso (Giraldo, 2005).

### 3.3.3.2. Estacionariedad Débil o Intrínseca

Según Clark (citado en Giraldo, 2005) existen algunos fenómenos físicos reales en los que la varianza no es finita. En estos casos se trabaja sólo con la hipótesis que pide que los incrementos  $[Z(x) - Z(x + h)]$  sean estacionarios, esto es si:

- $Z(x)$  tiene esperanza finita y constante para todo punto en el dominio. Lo que implica que la esperanza de los incrementos es cero.

$$E[Z(x) - Z(x + h)] = 0 \quad (3.2)$$

- Para cualquier vector  $h$ , la varianza del incremento está definida y es una función única de la distancia.

$$E[\{Z(x) - Z(x + h)\}^2] = Var [Z(x) - Z(x + h)] = 2\gamma (h) \quad (3.3)$$

La cantidad  $\gamma (h)$  es conocida como la semivarianza en el paso  $h$ , o la varianza por punto cuando los puntos se consideran en pares. Como en la covarianza, la semivarianza depende solamente del paso  $h$  y no de las posiciones absolutas de los datos. A la función  $\gamma (h)$ , usualmente se la conoce como variograma (Leiva, 2014).

### 3.3.4. Estimación del Variograma

El método más usual de cálculo de las semivarianzas empíricas de los datos  $z(x_1), z(x_2), \dots$ , es el estimador definido por Matheron, conocido como estimador de momentos. Su ecuación es:

$$\gamma(h) = \frac{1}{2m(h)} \sum_{i=1}^{m(h)} \{z(x_i) - z(x_i + h)\}^2 \quad (3.4)$$

Donde,  $\gamma(h)$  es la semivarianza para todas las muestras localizadas en el espacio separado por el intervalo de distancia  $h$ .  $z(x_i)$  y  $z(x_i + h)$  son los actuales valores de

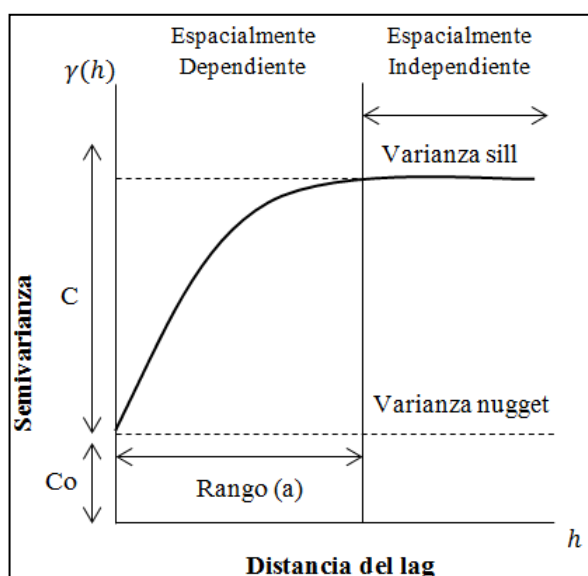
$Z$  en los sitios  $x_i$  y  $x_i + h$  y  $m(h)$  es el número de comparaciones en pares en el paso. El variograma experimental o muestral se obtiene cambiando  $h$  (Oliver, 2010).

Webster y Oliver (citado en Oliver, 2010) mencionan que para estimar el variograma por este método es necesario tener al menos 100 puntos de muestreo para obtener resultados confiables, mientras que con pocos datos los resultados no serán buenos. Para el cálculo del semivariograma es importante tener en cuenta algunos parámetros y reglas, que permiten elevar la eficiencia y la calidad de la estimación, como lo expone Díaz (2002):

- Para la estimación del semivariograma, los pares de las observaciones se agrupan en función de la distancia dentro de un intervalo  $h = |h|$  con una tolerancia  $\pm \Delta h/2$  y dentro de una dirección  $\theta$  con una tolerancia  $\pm \Delta \theta/2$ , con lo cual el semivariograma estimado es considerado suavizado o regularizado.
- El semivariograma muestral habitualmente, se estima para valores de  $|h|$  menores que la mitad de la distancia máxima ( $|h| < d_{\max}/2$ ).
- Se considera un número máximo de 25 lags para cualquier propósito, y un mínimo de 10 lags debe ser usado para determinar con precisión el rango y la meseta del semivariograma.

### 3.3.4.1. Características del Variograma

El variograma representa el índice del cambio que una variable muestra con la distancia, es así que valores observados en localizaciones  $Z(x)$  próximas son muy similares que, conforme aumenta la distancia de paso (lag), esta similitud va disminuyendo. Su forma describe el patrón de variación espacial en términos de su magnitud y forma general. Según Oliver (2010) el variograma puede variar de una manera espacialmente continua y discontinua, dependiendo de varias consideraciones importantes, y de parámetros del variograma que se analizan siempre para un estudio. Estos parámetros son la meseta o sill, el rango o alcance y efecto nugget o pepita (ver figura 14).



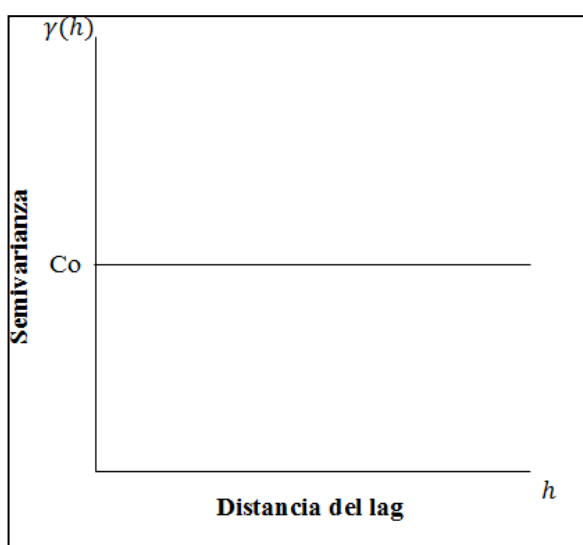
**Figura 14. Parámetros del Variograma.**

Fuente: (Adaptado de Oliver, 2010)

- **Sill o meseta:** Se define como el límite superior del semivariograma cuando la distancia  $h$  tiende al infinito, considerando que la meseta puede ser finita o no. Los semivariogramas con una meseta finita cumplen con la hipótesis de estacionariedad fuerte (de segundo orden); mientras que cuando no es finita, el semivariograma define un fenómeno natural que cumple únicamente la hipótesis intrínseca. La meseta se denota por  $C$  o por  $(Co + C)$  cuando la pepita es diferente de cero (Giraldo, 2005).
- **Rango o alcance:** Corresponde la distancia a partir de la cual dos observaciones son independientes, interpretada como la zona de influencia del semivariograma. Algunos modelos de semivariograma no presentan una distancia finita para la cual dos observaciones sean independientes, es así que se conoce como rango efectivo a la distancia para la cual el semivariograma alcanza el 95% de la meseta. Mientras más pequeño sea el rango, más cerca se está del modelo de independencia espacial; sin embargo el rango no siempre aparece de manera explícita en la fórmula del semivariograma (Giraldo, 2005).
- **Efecto pepita (Nugget effect):** Se expresa por  $Co$  y describe una discontinuidad puntual del semivariograma en el origen, que generalmente se debe a errores de medición en la variable o a la escala de la misma. En

algunas ocasiones puede ser indicativo de que parte de la estructura espacial se concentra a distancias inferiores a las observadas (Giraldo, 2005).

Cuando un semivariograma tiende a la horizontalidad, con valores próximos a la varianza muestral, indica la ausencia de una estructura espacial. La varianza muestral representa la variabilidad de los datos cuando ocurre esta circunstancia, determinando que el intervalo de muestreo es demasiado grande para resolver la variación presente (Moral, 2004). A estos variogramas se les conoce como efecto pepita puro, como se observa en la Figura 15.

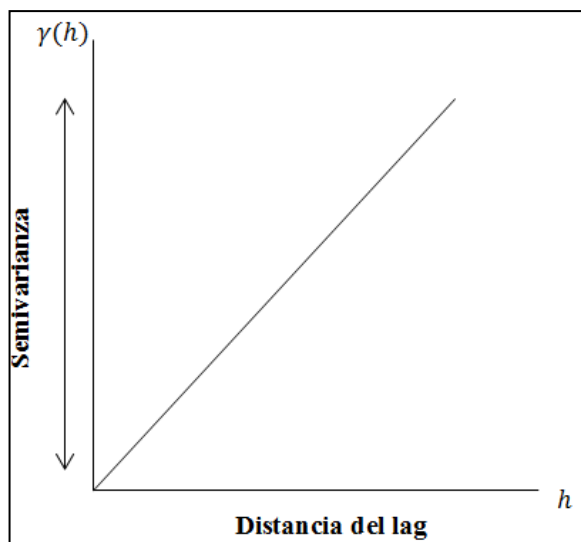


**Figura 15. Variograma efecto pepita puro.**

Fuente: (Adaptado de Oliver, 2010)

Como menciona Oliver (2010, citado en Leiva, 2014) si el variograma aumenta indefinidamente (variograma no acotado) con el incremento del paso  $h$ , como en la Figura 16, el proceso es solamente intrínseco. Un variograma que fluctúa de forma periódica con el aumento de distancia de seguimiento indica repetición regular de la variación. Si la variación difiere de acuerdo con la dirección, es anisotrópico y el variograma lo será también. La anisotropía es geométrica si el gradiente inicial o rango del variograma cambia con la dirección y una simple transformación de las coordenadas eliminará el efecto. La anisotropía zonal no puede ser tratada con la misma facilidad; está presente si la meseta fluctúa con los cambios de dirección, lo que podría indicar la presencia de zonas preferentemente orientadas con diferentes medias. La variación en el medio ambiente puede ocurrir en varias escalas espaciales

al mismo tiempo, y los patrones de la variación puede ser anidados uno dentro del otro. El variograma experimental aparecerá más complejo si más de una escala espacial está presente.

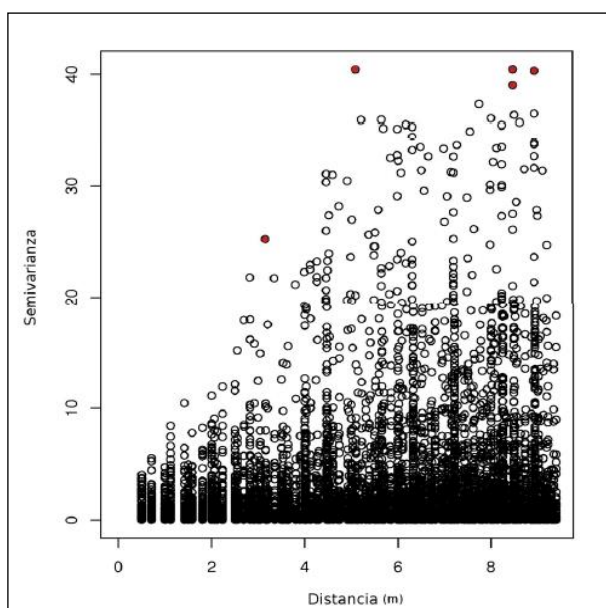


**Figura 16. Variograma No Acotado (Proceso solamente Intrínseco)**

Fuente: (Adaptado de Oliver, 2010)

### 3.3.5. Nube Variográfica

Según Alfaro (2007), la nube variográfica consiste en graficar el valor de todas las diferencias cuadráticas  $\frac{1}{2} [z(x_i) - z(x_j)]^2$  en función de la distancia entre los puntos  $x_i$  y  $x_j$ . Es una herramienta que permite localizar los pares de datos responsables de los valores altos en el cálculo del variograma experimental, como indica la Figura 17.



**Figura 17. Nube Variográfica.**

Fuente: (Gallardo, 2008)

### 3.3.6. Modelamiento del Variograma

El método experimental de variograma de momentos estima el variograma subyacente (el cual es una función continua), considerando un conjunto de puntos discretos a intervalos particulares de paso. Estas estimaciones a menudo varían de punto a punto, ya que están sujetas a errores que surgen en gran parte de la toma de muestras. Para describir la variación espacial, se ajusta un modelo a los valores experimentales (Leiva, 2014).

Se deben tomar en cuenta ciertas características en el modelamiento del variograma; la primera que el modelo debe ser condicional negativo semidefinido (CNSD) de tal manera que no dé lugar a varianzas negativas cuando variables aleatorias son combinadas, y segundo, la función debe ser también capaz de representar las características del variograma descritas en la sección (3.3.4.1). Existen pocas funciones simples que satisfacen las condiciones antes mencionadas. Estas incluyen funciones acotadas (delimitadas), como los modelos gaussiano, esférico y exponencial, que son modelos transicionales (porque en ellos se puede estimar el sill, sea verdadero o no) los cuales representan procesos que son estacionarios de segundo orden, y funciones no delimitadas (no acotadas), como los



modelos logarítmico, potencia y lineal, que representa procesos únicamente intrínsecos (Oliver, 2010).

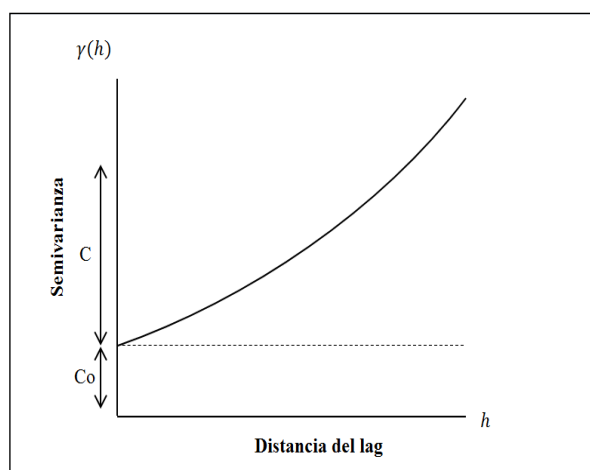
Por otro lado Alfaro (2007) manifiesta que el variograma teórico debe respetar al variograma experimental, sobre todo en los primeros puntos, que son más confiables; así como considerar que el ajuste de variogramas constituye un punto crucial en un estudio geoestadístico, porque todos los cálculos posteriores se harán utilizando exclusivamente el modelo teórico; y, finalmente que, para tener un buen ajuste se debe considerar que uno de los objetivos finales es la predicción puntual, o por bloques dentro de una vecindad restringida, de manera que no se consideren demasiadas muestras en la predicción.

### 3.3.6.1. Modelo Potencia

Este modelo es usado cuando existen casos en que la varianza aparenta incrementarse indefinidamente. Su ecuación es:

$$\gamma(h) = C_0 + h^\alpha \quad h \geq 0 \text{ y } 0 < \alpha \leq 2 \quad (3.5)$$

Donde  $C_0$  es el “nugget”,  $h$  es la distancia y  $\alpha$  describe la curvatura del modelo; si  $\alpha = 1$  tenemos un modelo lineal, si  $\alpha < 1$  la curva es convexa hacia arriba y si  $\alpha > 1$  la curva es cóncava (Gallardo, 2008).



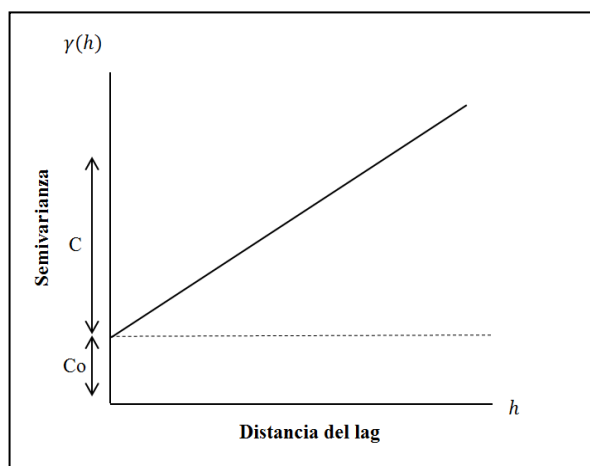
**Figura 18. Modelo Potencia.**

Fuente: (Adaptado de Alfaro, 2007)

### 3.3.6.2. Modelo Lineal

$$\gamma(h) = C_0 + bh \quad (3.6)$$

Donde  $C_0$  es el “nugget”,  $h$  es la distancia y  $b$  es la pendiente de la recta (Gallardo, 2008).



**Figura 19. Modelo Lineal.**

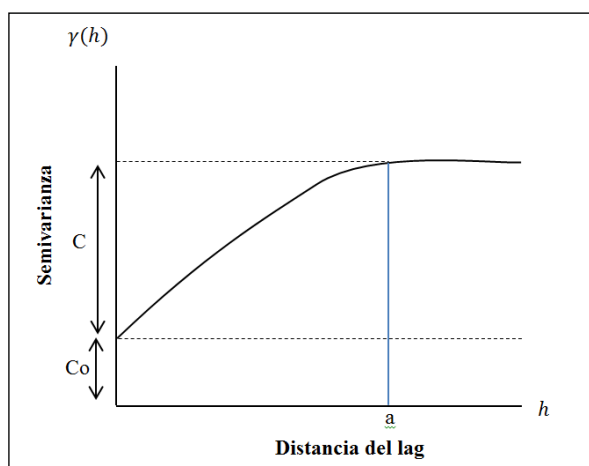
Fuente: (Adaptado de Gallardo, 2008)

### 3.3.6.3. Modelo Esférico

El modelo esférico presenta un crecimiento rápido cerca al origen, pero va decreciendo a medida que la distancia aumenta (Giraldo, 2005). Su expresión matemática es la siguiente:

$$\gamma(h) = \begin{cases} C_0 + C \left( \frac{3}{2} \left( \frac{h}{a} \right) - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right) & 0 < h \leq a \\ C_0 + C & h > a \end{cases} \quad (3.7)$$

En donde  $(C_0 + C)$  representa la meseta,  $a$  el rango y  $h$  la distancia.



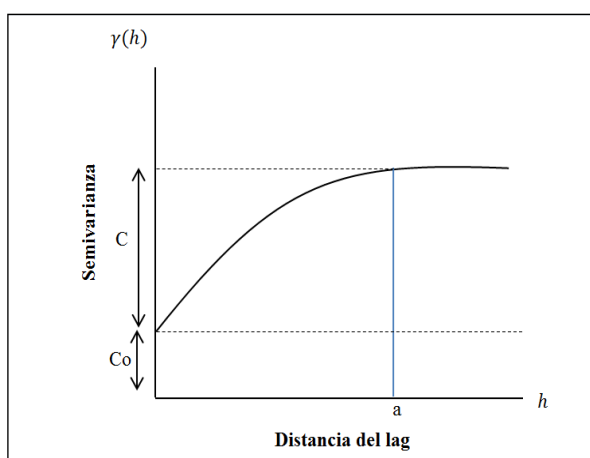
**Figura 20. Modelo Esférico.**

Fuente: (Adaptado de Gallardo, 2008)

#### 3.3.6.4. Modelo Exponencial

Este modelo tiene los mismos parámetros que el esférico, pero crece más lentamente que dicho modelo. Es usado cuando la dependencia espacial tiene un crecimiento exponencial respecto a la distancia entre las observaciones. El valor de  $a$  hace referencia a la distancia para la cual el semivariograma alcanza un valor igual al 95% de la meseta (sill) (Giraldo, 2005). Su ecuación es la siguiente:

$$\gamma(h) = C_0 + C \left( 1 - \exp\left(-\frac{3h}{a}\right) \right) \quad (3.8)$$



**Figura 21. Modelo Exponencial.**

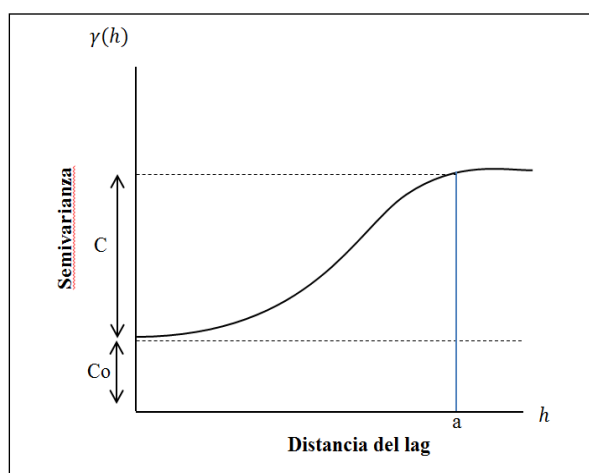
Fuente: (Adaptado de Gallardo, 2008)

### 3.3.6.5. Modelo Gaussiano

Al igual que el modelo exponencial, el semivariograma alcanza el 95% de la meseta. El principal distintivo de este modelo es su forma parabólica cerca al origen (Gallardo, 2008). Matemáticamente se expresa como:

$$\gamma(h) = C_0 + C \left( 1 - \exp\left(-\frac{h^2}{a^2}\right) \right) \quad (3.9)$$

Esta ecuación presenta los mismos parámetros de los modelos anteriores, meseta o sill ( $C_0 + C$ ), distancia ( $h$ ) y rango ( $a$ ).



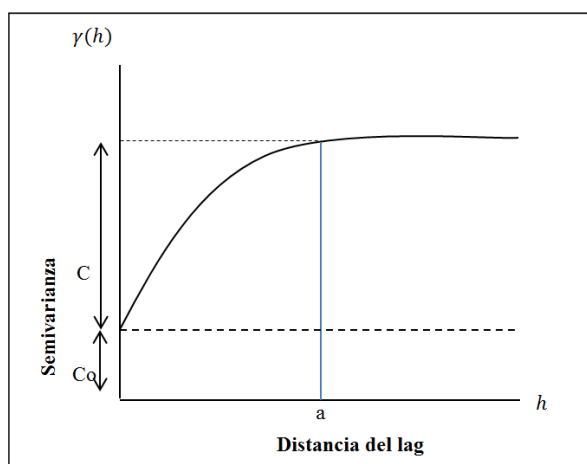
**Figura 22. Modelo Gaussiano.**

Fuente: (Adaptado de Alfaro, 2007)

### 3.3.6.6. Modelo Cuadrático

Es similar a un modelo esférico pero más simple (Alfaro, 2007). Su expresión matemática es la siguiente:

$$\gamma(h) = \begin{cases} C_0 + C \left( 2 \left( \frac{h}{a} \right) - \left( \frac{h}{a} \right)^2 \right) & \text{si } h \leq a \\ C_0 + C & \text{si } h > a \end{cases} \quad (3.10)$$



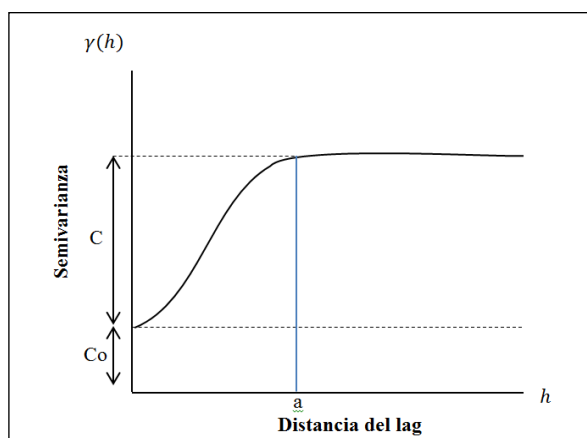
**Figura 23. Modelo Cuadrático.**

Fuente: (Adaptado de Alfaro, 2007)

### 3.3.6.7. Modelo Cúbico

Tiene un comportamiento parabólico en el origen pero su rango es finito e igual a  $a$  (Alfaro, 2007); su ecuación es:

$$\gamma(h) = \begin{cases} C_0 + C \left( 7 \left( \frac{h}{a} \right)^2 - 8.75 \left( \frac{h}{a} \right)^3 + 3.5 \left( \frac{h}{a} \right)^5 - 0.75 \left( \frac{h}{a} \right)^7 \right) & 0 < h \leq a \\ C_0 + C & h > a \end{cases} \quad (3.11)$$



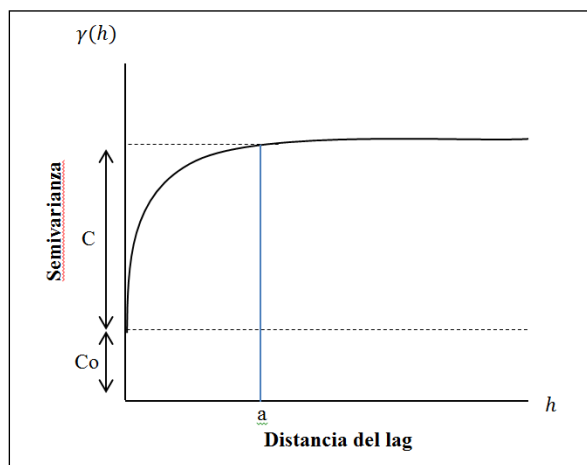
**Figura 24. Modelo Cúbico.**

Fuente: (Adaptado de Alfaro, 2007)

### 3.3.6.8. Modelo Circular

Según Gallardo (2008), este modelo no es muy utilizado, se curva notablemente conforme se aproxima al rango. Su ecuación es:

$$\gamma(h) = \begin{cases} C_0 + C \left( 1 - \frac{2}{\pi} \cos^{-1} \left( \frac{h}{a} \right) + \frac{2h}{\pi a} \sqrt{1 - \left( \frac{h}{a} \right)^2} \right) & 0 < h \leq a \\ C_0 + C & h > a \end{cases} \quad (3.12)$$



**Figura 25. Modelo Circular.**

Fuente: (Adaptado de Gallardo, 2008)

### 3.3.6.9. Modelo Pepita Puro

Este modelo indica que no hay una adecuada correlación espacial entre las observaciones de una variable (Giraldo, 2005). El variograma no depende de la distancia  $h$  (ver figura 15). Su expresión matemática es:

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C_0 & h > 0 \end{cases} \quad (3.13)$$

### 3.3.6.10. Modelo Anidado

En ocasiones un modelo puede ajustarse a partir de la suma de dos o más modelos tales como nugget, esférico, etc. Agregando el componente efecto pepita a uno de los modelos teóricos, o a su vez la suma de dos modelos de variograma teórico, dan como resultado el modelo anidado (Alfaro, 2007). Este modelo se representa de la siguiente manera:

$$\gamma(h) = \gamma_1(h) + \gamma_2(h) \quad (3.14)$$

### **3.4. Predicción Espacial**

#### **3.4.1. Predicción Espacial Óptima**

El método de predicción espacial más estudiado es el kriging, (por Daniel Krige, quien en 1951 manifestó por primera vez esta metodología, mientras realizaba trabajos en la predicción de reservas de oro, y que suelen considerarse como pioneros en los métodos de interpolación espacial) consiste en la predicción lineal insesgada óptima (BLUE: Best Linear Unbiased Estimator), lo cual lo determina como un método de estimación superior a los tradicionales. Este es el método de interpolación de datos dispersos para procesos espaciales aleatorios (Oliver, 2010).

#### **3.4.2. Características Generales del Kriging**

El principal objetivo de todo proyecto geoestadístico es la caracterización de la variable investigada en todas las localizaciones, partiendo de la información que se tiene en puntos muestrales (Moral, 2004). El método de estimación geoestadística conocido como “Kriging”, o krigeado, determina una estimación superior a los métodos tradicionales, como son: el inverso ponderado de la distancia, triangulación, mínimos cuadrados polinomiales, polígonos de thiesen, splines, etc; considerando una serie de características descritas por Laslett et al. (citado en Oliver, 2010):

- Los métodos tradicionales usan el concepto de distancia euclidiana y en ocasiones la arbitrariedad para el cálculo de los pesos que se aplica para cada variable muestral, a diferencia del kriging que analiza la distancia y la geometría de la localización de las muestras.
- El kriging minimiza la varianza del error esperado (diferencia entre el valor real y el predicho). Debido a que el valor real de un punto no muestral no se conoce, el kriging emplea un modelo teórico con una función aleatoria asociada a los valores reales.
- En la interpolación, mediante el uso de métodos geoestadísticos, se puede estimar valores puntuales o en bloques, así como incorporar información secundaria que esté relacionada con la variable principal (ecuaciones de transformación de la variable, vecindad de búsqueda).

- Los métodos de interpolación tradicionales no estiman la varianza en todos los puntos de modelamiento, a diferencia del método kriging que si lo hace.
- El método kriging no asegura la mejor predicción posible, sino mas bien permite evitar un error sistemático.

Según Giraldo (2005), el kriging posee un conjunto de métodos de predicción espacial que se fundamentan en la minimización del error cuadrático medio de predicción. La Tabla 1 muestra el tipo de predictores y sus propiedades.

**Tabla 1**

**Tipos de predictores Kriging y sus propiedades**

TIPO DE PREDICTOR	NOMBRE	PROPIEDADES
<b>LINEAL</b>	<ul style="list-style-type: none"> <li>• Simple</li> <li>• Ordinario</li> <li>• Universal</li> </ul>	<ul style="list-style-type: none"> <li>• Son óptimos si hay normalidad multivariada.</li> <li>• Independiente de la distribución son los mejores predictores linealmente insesgados.</li> </ul>
<b>NO LINEAL</b>	<ul style="list-style-type: none"> <li>• Indicador</li> <li>• Probabilístico</li> <li>• Lognormal, TransGaussiano, Disyuntivo</li> </ul>	<ul style="list-style-type: none"> <li>• Son predictores óptimos.</li> </ul>

Fuente: (Giraldo,2005)

Otra de las clasificaciones generales es según el soporte de la medición de los datos, estos pueden ser: puntual y en bloques. Para Oliver (2010) el Kriging ordinario es el método mas robusto que se analiza con frecuencia.



### 3.4.3. Kriging Ordinario

El Kriging ordinario se basa en el supuesto de que la media es desconocida. Se caracteriza por considerar que se producen fluctuaciones locales de la media, limitando el dominio de estacionariedad de la misma a un ámbito local:  $m(x) =$  constante, pero desconocida. Suponiendo que se hacen mediciones de la variable de interés  $Z$  en los puntos  $x_i$ ,  $i=1, 2, \dots, n$ , de la región de estudio, es decir de la forma  $Z(x_1), Z(x_2), \dots, Z(x_n)$ , y se desea predecir  $Z(x_0)$ , en el punto  $x_0$  donde no hubo una medición. Bajo esta circunstancia, el kriging ordinario propone que la variable puede predecirse como una combinación lineal de las  $n$  variables aleatorias de la siguiente manera:

$$\hat{Z}(X_0) = \lambda_1 Z(x_1) + \lambda_2 Z(x_2) + \lambda_3 Z(x_3) + \dots + \lambda_n Z(x_n) \quad (3.15)$$

$$\hat{Z}(X_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (3.16)$$

Donde  $n$  usualmente representa los puntos de dato dentro de la vecindad local, y  $\lambda_i$  representa los pesos o ponderaciones de los valores originales. Dichos pesos se calculan en función de la distancia entre los puntos muestreados y el punto donde se va a hacer la correspondiente predicción. Para que se cumpla el requisito de insesgamiento la suma de los pesos debe ser igual a 1, con lo cual la esperanza del predictor es igual a la esperanza de la variable (Giraldo, 2005).

Si:

$$\sum_{i=1}^n \lambda_i = 1 \quad (3.17)$$

entonces,

$$E(\hat{Z}(x_0)) = E(Z(x_0)) \quad (3.18)$$

En este caso  $\hat{Z}(X_0)$  es el mejor predictor lineal, porque los pesos se obtienen de tal manera, que minimizan la varianza del error de predicción que se representa de la siguiente forma:

$$V(\hat{Z}(x_0) - Z(x_0)) \quad (3.19)$$

Según Samper y Carrera (1990) (citado en Giraldo, 2005), esta es una característica propia de los métodos Kriging, ya que otros métodos de interpolación como el inverso de la distancia ponderada o el poligonal, no garantizan la varianza mínima de predicción.

Dicha varianza de predicción esta determinada por la ecuación:

$$V(\hat{Z}(x_0)) = E[(\hat{Z}(x_0) - Z(x_0))^2] \quad (3.20)$$

$$V(\hat{Z}(x_0)) = 2 \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(x_i, x_j) \quad (3.21)$$

Donde  $\gamma(x_i, x_j)$  es la semivarianza de Z entre los puntos  $x_i$  y  $x_j$ ,  $\gamma(x_i, x_0)$ , es la semivarianza entre el i-ésimo punto de muestreo y el punto objetivo  $x_0$ . Las semivarianzas se derivan del modelo de variograma porque las semivarianzas experimentales son discretas y a distancias limitadas. La ecuación conduce a un conjunto de  $n+1$  ecuaciones con  $n+1$  incógnitas.

$$\sum_{i=1}^n \lambda_i \gamma(x_i, x_j) + \psi(x_0) = \gamma(x_i, x_0) \quad \text{Para todo } j, \quad (3.22)$$

$$\sum_{i=1}^n \lambda_i = 1$$

El multiplicador de Lagrange  $\psi(x_0)$ , se introduce para alcanzar la minimización. Los pesos,  $\lambda_i$ , se insertan en la ecuación (3.22) para dar la predicción de Z en  $x_0$ . La varianza de predicción de Kriging es luego obtenida como:

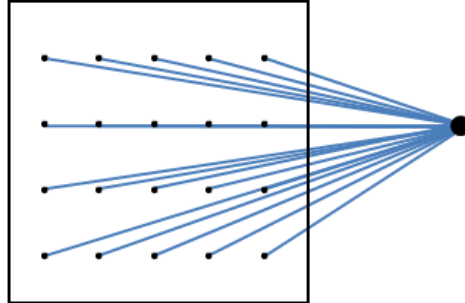
$$\sigma^2(x_0) = \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \psi(x_0) \quad (3.23)$$

Según Oliver (2010), en la interpolación con Kriging los valores de la propiedad por lo general se estiman en los nodos de una malla fina (grid), y la variación se determina mediante líneas de contorno (isolineas), o por modelos digitales. Sus varianzas también se asignan de manera similar y son una guía de la fiabilidad de las predicciones.

### 3.4.3.1. Kriging Ordinario por bloques

En el método descrito anteriormente, el objetivo se centra en la predicción puntual pero, en diversos casos se requiere la estimación del valor medio de una variable

dentro de una superficie o un volumen de cualquier forma o tamaño (bloque) (ver figura 26).



**Figura 26. Kriging Ordinario por bloques.**

Fuente: (Giraldo, 2005).

El valor promedio dentro de un bloque es estimado por:

$$\hat{Z}(B) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (3.24)$$

La varianza del error de predicción del kriging en bloques está dada por:

$$V(\hat{Z}(B)) = E[(\hat{Z}(B) - Z(B))^2] \quad (3.25)$$

$$V(\hat{Z}(B)) = 2 \sum_{i=1}^n \lambda_i \bar{\gamma}(x_i, B) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(x_i, x_j) - \bar{\gamma}(B, B) \quad (3.26)$$

Donde  $\bar{\gamma}(x_i, B)$  es la semivarianza promedio entre el punto de dato  $x_i$  y el bloque objetivo  $B$ , y  $\bar{\gamma}(B, B)$  es la semivarianza promedio dentro de  $B$ , la semivarianza dentro del bloque.

El sistema kriging equivalente por bloques es:

$$\sum_{i=1}^n \lambda_i \gamma(x_i, x_j) + \psi(B) = \bar{\gamma}(x_i, B) \quad \text{Para todo } j, \quad (3.27)$$

$$\sum_{i=1}^n \lambda_i = 1$$

Y su varianza kriging para el bloque se determina como:

$$\sigma^2(B) = \sum_{i=1}^n \lambda_i \bar{\gamma}(x_i, B) + \psi(B) - \bar{\gamma}(B, B) \quad (3.28)$$

Según Isaaks y Srivastava (citado en Giraldo, 2005), el kriging en bloques coincide con el promedio de predicciones hechas con kriging ordinario sobre cada uno de los puntos del enmallado dentro del bloque. En la práctica, establecen como suficiente un enmallado cuadrado de 6x6, para obtener estimaciones estables en los bloques. Para conocer sobre los diferentes tipos de kriging ver Anexo 1.

#### **3.4.3.2. Pesos en el Kriging**

Los pesos en kriging se asignan dependiendo del variograma y la configuración de la toma de muestras. Kriging distribuye la asignación de pesos dentro de un radio de búsqueda, haciéndolo diferente a los interpoladores tradicionales o determinísticos, donde los pesos se asignan arbitrariamente. Los pesos varían de acuerdo a cambios en la proporción entre el efecto pepita y la meseta, el rango del variograma, el tipo de modelo, la configuración del muestreo y el efecto de la anisotropía; siendo los más sensibles la configuración del muestreo y la anisotropía (Webster & Oliver ,2007, citado en Oliver, 2010).

Los pesos cerca de un punto (o un bloque), tienen mayor ponderación que los que están alejados lo cual demuestra que kriging es un predictor local. Cuando la proporción entre efecto pepita y meseta se incrementa los pesos cerca del objetivo decrecen y los que están más lejos aumentan. Para un variograma pepita puro, los pesos kriging son todos iguales y la estimación es simplemente la media de los valores en la vecindad. Para los datos que se distribuyen de manera irregular, los puntos que se agrupan llevan menos peso individual de los que están aislados. El hecho de que los puntos más cercanos al objetivo general llevan el mayor peso tiene implicaciones prácticas. Esto significa que la zona de búsqueda necesita contener no más de 16-20 puntos de datos (Oliver, 2010).

#### **3.4.3.3. Vecindad de Kriging**

Cuando se analiza el kriging, se deben considerar todos los datos disponibles, lo que implica cálculos muy extensos, sin considerar que las muestras alejadas tendrían un peso casi nulo. Por tal razón Alfaro (2007) recomienda restringirse a una vecindad de estimación, que puede ser un círculo o una esfera (dependiendo del espacio

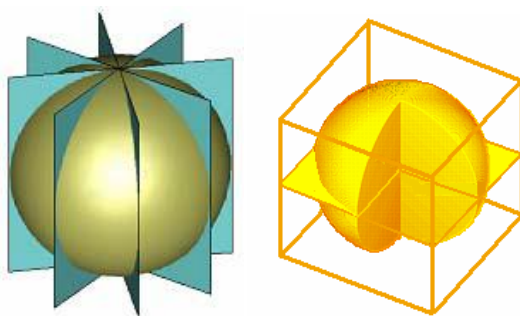
dimensional del estudio), donde el radio de búsqueda en una cierta dirección no debe ser inferior al alcance en esa dirección.

En un espacio de dos dimensiones, con una vecindad que contenga un promedio de 8 muestras, los resultados son buenos. En un espacio de 3 dimensiones la situación se torna más compleja y se debe analizar cada caso particular (Alfaro, 2007).

#### 3.4.3.3.1. Estrategia de búsqueda

Según Alfaro (2007), en una estrategia de búsqueda se establecen varias consideraciones para la estimación puntual; las cuales son:

- **Radios de búsqueda:** En primera aproximación se usa los alcances del variograma en varias direcciones.
- **Mínimo de muestras para krigear:** Controla el caso en que solo una muestra cae en la vecindad
- **Máximo de muestras para krigear:** Esta parámetro se usa para mayor velocidad para los cálculos, estimando para el máximo de muestras cercanas.
- **Muestras por octantes** (ver figura 27): El objetivo del octante de búsqueda es desagrupar, pero hay que tener cierta precaución en su uso, debido a que valores pequeños pueden introducir artefactos.



**Figura 27. Octantes**

Fuente: (Alfaro, 2007)

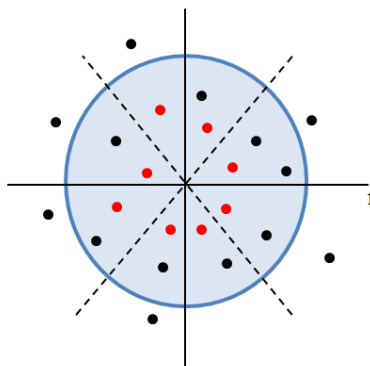
### 3.4.3.4. Criterio de vecindad única y vecindad móvil

#### 3.4.3.4.1. Criterio de vecindad única

Según Larraz y Montero (2008), el criterio de vecindad única utiliza en su totalidad los datos muestrales existentes en el área de estudio para estimar el valor de la función aleatoria en el punto deseado. Cuando el proceso presenta deriva (tendencia) este criterio no puede ser utilizado, excepto si la misma es estimada previamente y eliminada de los datos. Otra dificultad que presenta este criterio, es el excesivo tiempo de cálculo necesario para cada estimación dado el caso que se disponga de una gran cantidad de datos.

#### 3.4.3.4.2. Criterio de vecindad móvil

Para este criterio se utilizan puntos muestrales cercanos a la localización en la que se quiere realizar la estimación, definiendo previamente la forma y el tamaño del vecindario para una adecuada estimación. Un criterio de decisión sobre los puntos que forman parte del vecindario móvil y cuales no, atienden a los conceptos de distancia máxima, distribución angular y proximidad. De esta manera, formarán parte del vecindario aquellos puntos que se hallen a una distancia menor que un radio  $r$ , distribuidos de forma que se acepte un mínimo de  $p$  localizaciones por octante o cuadrante (ver figura 28) y además que resulten ser las  $n$  localizaciones más próximas (Larraz & Montero, 2008).



**Figura 28. Elección de vecindario móvil de radio  $r$ .**

Fuente: (Adaptado de Larraz & Montero, 2008)

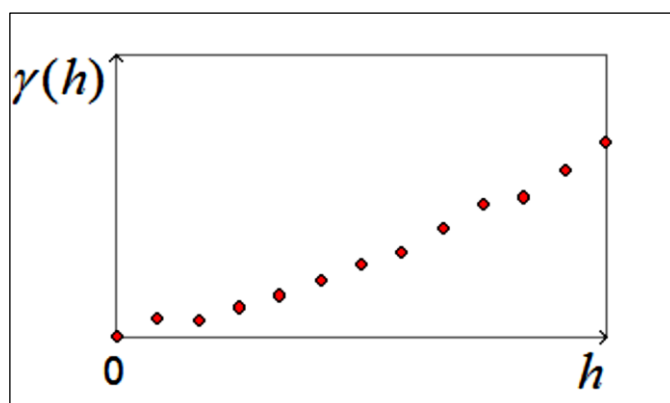
### 3.4.3.5. Distancia de vecindad de búsqueda

Considerando que la región de vecindad de proceso estocástico es estacionario de segundo orden (así se elimina la posibilidad de existencia de deriva local), se debe elegir una distancia adecuada de búsqueda de vecindad. Si la distancia es muy grande, conforme van aumentando el semivariograma experimental va perdiendo fiabilidad debido a la escasez de parejas implicadas en su estimación. Si la distancia es muy pequeña el semivariograma capta bien la estructura de dependencia espacial de los datos. Sin embargo, no es prudente elegir una distancia de búsqueda muy pequeña debido a que la estimación se vuelve sensible a los datos vecinos que intervienen en el cálculo, lo cual puede ocasionar artefactos en diferentes localizaciones del modelo de predicción obtenido. Hay que tener en cuenta la distribución del muestro, es decir que el número de datos por vecindario sea suficiente. Para encontrar el equilibrio en lo que respecta al tamaño del vecindario resulta útil recurrir a la validación cruzada, pudiendo cambiar el vecindario y elegir el que aporte resultados más favorables (Larraz & Montero, 2008).

### 3.4.3.6. Tendencia

Según Alfaro (2007), cuando el semivariograma empírico no alcanza la meseta y tiende a seguir subiendo conforme aumenta la distancia  $h$ , y de forma constante más allá del valor global de la varianza, a menudo es un indicativo de una tendencia espacial significativa en la variable, lo que resulta en una correlación negativa entre los valores de variables separadas por grandes retardos (ver figura 29). Hay tres opciones que se consideran al existir tendencia en los datos (Bohling, 2005):

1. Colocar una superficie de tendencia y trabajar con residuos de la tendencia.
2. Tratar de encontrar una dirección "libre de tendencia" y utilizar el variograma en esa dirección como el variograma para la componente "aleatoria" de la variable (anisotropía).
3. Ignorar el problema y utilizar un variograma lineal o de potencia.



**Figura 29. Variograma con crecimiento sistemático.**

Fuente: (Alfaro, 2007)

### 3.4.4. Validación Cruzada

Para validar el modelo elegido de acuerdo al comportamiento de variable estudiada y por ende de las predicciones hechas con kriging, el método de la validación cruzada es un caso particular del método conocido como Jackknife. El cual consiste en sacar un punto  $n$  de la muestra (asociado a un vecindario) y estimar el valor en ese punto usando Kriging con los  $n - 1$  valores restantes y el modelo de variograma obtenido. De la misma manera se procede con cada uno de los puntos de la muestra y con ello se genera un conjunto de  $n$  errores de predicción. Como resultado se obtiene un mapa o histograma de las diferencias entre el valor observado  $z(x)$  y el estimado  $\hat{z}(x)$ . Es así que, si el modelo del semivariograma muestra adecuadamente la estructura espacial de autocorrelación de los datos, dicha diferencia debe ser pequeña, debido a la cercanía de los valores estimados a los valores observados (Giraldo, 2005).

### 3.4.5. Verificación del Modelo

Para la verificación de un modelo obtenido a partir de un interpolador, Oliver (2010) manifiesta el uso de dos medidas de precisión como el Error Medio (ME) y el Error Cuadrático Medio (MSE) mediante la ecuación 3.29 y 3.30 respectivamente:

$$ME = \frac{1}{N} \sum_{i=1}^N \{z(x_i) - \hat{z}(x_i)\} \quad (3.29)$$



$$MSE = \frac{1}{N} \sum_{i=1}^N \{z(x_i) - \hat{z}(x_i)\}^2 \quad (3.30)$$

El ME es el promedio de la suma de la diferencia entre el valor observado  $z(x_i)$  y el valor estimado  $\hat{z}(x_i)$ , que es definido también como el sesgo o error de la predicción.  $N$  es el número de puntos utilizados. El valor de ME debe ser pequeño (cercano a cero), lo que indica que el método presenta pocos errores. Pero esta medida aun así no describe la magnitud de los errores que pueden ocurrir. Por ello, se utiliza el valor MSE que es la suma de los residuos al cuadrado (varianza de los residuos) en donde valores pequeños indican predicciones más precisas punto por punto (Mueller et al. 2001 y Schloeder et al. 2001, citado en Villatoro et al., 2008).

Por otro lado, existe un coeficiente que permite determinar cuan bueno es el ajuste de un modelo y la medida de la proporción de la variabilidad explicada por el modelo ajustado. Este coeficiente es conocido como coeficiente de determinación ( $R^2$ ) y puede tomar valores entre 0 y 1, mientras cercano este de 1, el modelo presenta mejor ajuste (Walpole, Myers, Myers, & Ye, 2007).

El coeficiente de determinación esta dado por:

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{Z}(x_i) - z(x_i))^2}{\sum_{i=1}^n (z(x_i) - \bar{Z})^2} \quad (3.31)$$

Donde  $\bar{Z} = \sum_{i=1}^n z(x_i)/n$ .

La medida del coeficiente de determinación, generalmente se representa en porcentaje, es decir el valor obtenido se lo multiplica por 100, con lo cual se analiza la efectividad del modelo (Villatoro, Henríquez, & Sancho, 2008)

Frecuentemente para la verificación de un modelo el uso del  $R^2$  es indispensable por su simplicidad, sin embargo, hay errores en su interpretación. La confiabilidad del  $R^2$  está en función del tamaño del conjunto de datos y del tipo de aplicación. Si bien es claro que al obtener un coeficiente de determinación cercano a 1 se puede decir que un modelo tiene un buen ajuste; cuando se conoce el objeto de estudio y la variabilidad que el mismo puede producir, muchas veces se habla de un sobreajuste, debido al que el coeficiente de determinación no puede tener un valor alto. Al agregar más datos a la muestra de estudio se puede deducir la presencia de un

sobreajuste dado que el  $R^2$  aumenta o se mantiene, pero no disminuye. Por tal razón es importante mencionar que para seleccionar un modelo no se debe suscribir un proceso de selección que únicamente incluya la consideración de  $R^2$  (Walpole et al., 2007).

## **CAPÍTULO IV**

### **METODOLOGÍA**

La siguiente metodología pretende proporcionar una herramienta para la transformación de coordenadas entre el Sistema de Referencia PSAD56 (Sistema topocéntrico) y el Sistema SIRGAS95 (Sistema geocéntrico) mediante técnicas geoestadísticas, con el fin de compatibilizar la cartografía existente en municipios u otras dependencias de gestión territorial, con las nuevas tecnologías que usan un dátum geocéntrico para sus mediciones. Además, el desarrollo de un modelo geoestadístico busca disminuir el error de transformación entre ambos sistemas y comparar los resultados, con los obtenidos al transformar coordenadas empleando el modelo matemático de Helmert aplicado por Leiva (2003).

#### **4.1. Recopilación y validación de los datos**

La recopilación de los datos se realizó en el IGM, a través del Departamento de Investigación y Desarrollo, los cuales disponen de una base de datos geodésica de las diferentes comisiones que realiza el Departamento de Geodesia por todo el Ecuador. Esta base de datos consta de 150 coordenadas geodésicas, latitud y longitud ( $\varphi$ ,  $\lambda$ ), conformada en dos grupos:

- 84 Coordenadas en el Sistema de Referencia PSAD56 y Sistema SIRGAS95 (ITRF94, época 1995.4)
- 66 Coordenadas en el Sistema de Referencia PSAD56 e ITRF08 (época 2013.x)

Las 66 coordenadas geodésicas en ITRF08 fueron previamente transformadas al Sistema SIRGAS95 mediante el cambio de marco y su respectiva época, como establece Santacruz (2010), con la finalidad de obtener coordenadas en los dos Sistemas de Referencia de interés (PSAD56 y SIRGAS95).

Para la realización de un modelo geoestadístico es importante contar con una cantidad considerable de datos, Oliver (2010) menciona que se necesitan más de 100 muestras para obtener un modelo de predicción adecuado; por este motivo se recopilieron 150 puntos geodésicos. Esta base de datos está conformada por: el

nombre de cada punto geodésico, las coordenadas Este y Norte en el Sistema SIRGAS95 y las variables diferencia de latitud y diferencia de longitud ( $\Delta\phi$ ,  $\Delta\lambda$ ) (ver Anexo 2). Con las siguientes consideraciones:

- Las variables desplazamiento de latitud y longitud, se obtienen de la diferencia entre la coordenada latitud, longitud del Sistema PSAD56 y la coordenada latitud, longitud del Sistema SIRGAS95 respectivamente; transformada a segundos de arco posteriormente. A continuación se describe la ecuación para una mayor comprensión:

$$\Delta\phi = (\phi_{PSAD56} - \phi_{SIRGAS95})/3600'' \quad (4.1)$$

$$\Delta\lambda = (\lambda_{PSAD56} - \lambda_{SIRGAS95})/3600'' \quad (4.2)$$

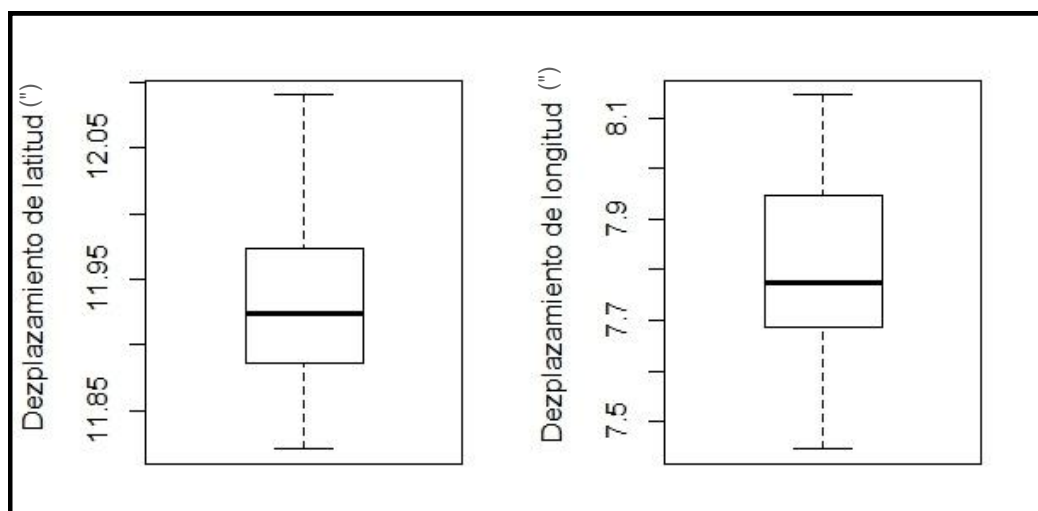
- Las coordenadas UTM, se encuentran en el Sistema de referencia SIRGAS95, elipsoide de referencia GRS80, zona 17 Sur, en dónde se encuentra ubicado espacialmente, la mayor parte del territorio del Ecuador Continental. La razón principal para transformar las coordenadas geodésicas a UTM es debido a que el proceso geoestadístico llevado a cabo en el Software R, paquete RGeostat<sup>1</sup>, necesita una distancia en metros para construir el variograma experimental, parte fundamental para la elección de un modelo de predicción adecuado.

El paso posterior a la recopilación de datos y el más importante en un proceso geoestadístico, fue validar la información con la cual se va a trabajar. Principalmente se determinó si existen valores atípicos que pueden afectar a los modelos de predicción, mediante el uso de gráficos exploratorios de datos entre los cuales están los diagramas de caja y gráficos de dispersión.

Con los 150 puntos se elaboró un diagrama de caja y un gráfico de dispersión “Coordenada vs Variable de estudio”, para cada variable ( $\Delta\phi$ ,  $\Delta\lambda$ ), con la finalidad de identificar valores extremos o anómalos, que de existir es necesario hacer un riguroso análisis para considerar si son o no eliminados. En la Figura 30 se muestra el diagrama de caja para la variable desplazamiento de latitud y desplazamiento de longitud, donde se corrobora que no presentaron valores atípicos.

---

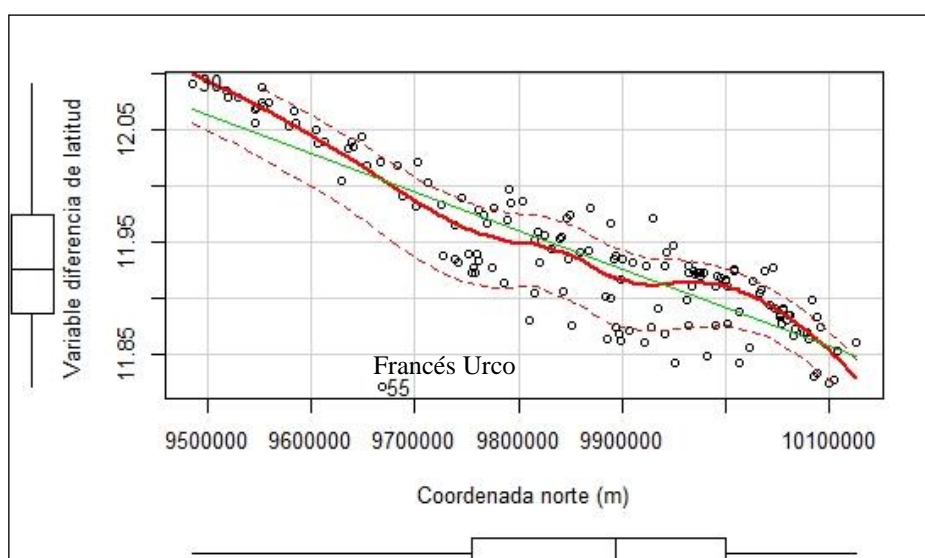
<sup>1</sup> Renard D., Bez N., Desassis N., Beucher H., Ors F., Laporte F.  
RGeostats: The Geostatistical package [10.0.1].  
MINES ParisTech.  
Free download from: <http://cg.ensmp.fr/rgeostats>



**Figura 30. Diagramas de caja preliminares.**

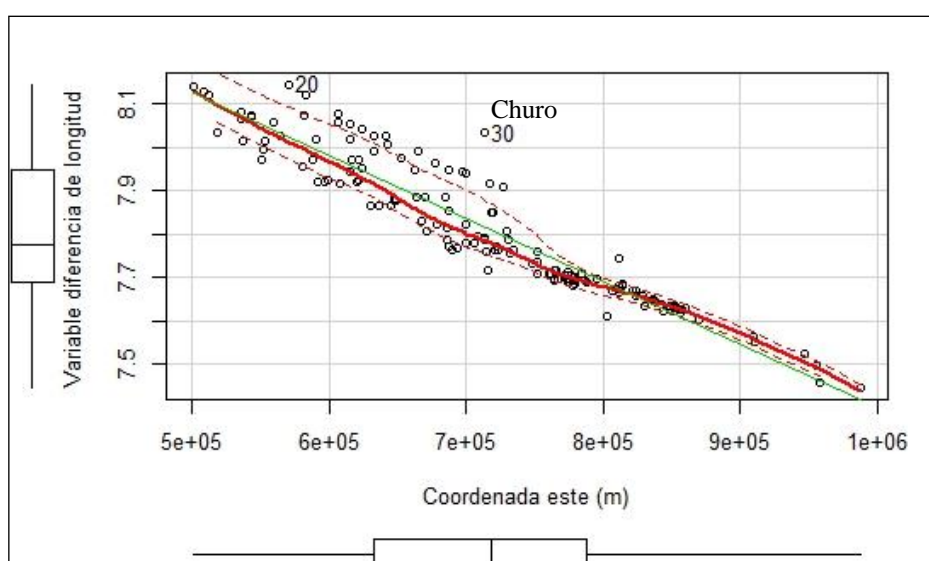
Las Figuras 31 y 32 representan los diagramas de dispersión para la variable desplazamiento de latitud y desplazamiento de longitud.

La Figura 31 para la variable desplazamiento de latitud, determina la línea de tendencia que siguen los datos. Se observó en este gráfico un valor con un comportamiento extraño, alejado de los demás datos, posiblemente debido a un error en su medición o algún tipo de anomalía; el nombre de este dato anómalo corresponde al punto “Francés Urco”. Para este punto geodésico se revisó la investigación desarrollada por Leiva (2003) sobre el cálculo de 7 parámetros de transformación, donde se corrobora que el procesamiento de datos GPS para “Francés Urco” determinó ambigüedades reales (flotantes) que disminuye la precisión, a pesar de haber tenido un tiempo de rastreo de 3 horas; por lo que se decidió sacar este punto del análisis geoestadístico.



**Figura 31. Diagrama de dispersión para la variable desplazamiento de latitud**

De igual manera se analizó la variable desplazamiento de longitud como se indica en la Figura 32. El diagrama de dispersión presentó un punto con un comportamiento diferente a los demás datos. Se comprobó que corresponde al punto geodésico “Churo”, el cual después de realizar una consulta en el IGM, acerca de toma y recepción del punto con GPS de precisión, no se encontró ninguna anomalía, novedad u observación, por lo que se decidió no sacar este punto del estudio geoestadístico.



**Figura 32. Diagrama de dispersión para la variable desplazamiento de longitud**

La base de datos final para el desarrollo de los modelos de predicción constó de 149 puntos geodésicos, como se muestra en el Anexo 3.

#### 4.2. Determinación del tamaño de la muestra

Como menciona Díaz (2002), al disponer de una cantidad de datos adecuada, espacialmente distribuidos de forma homogénea para la variable de estudio (en este caso 149 datos), es posible dividir en dos submuestras, la primera se usa para estimar el variograma y el modelo de predicción, y la segunda evaluará las diferencias entre los valores observados y las predicciones obtenidas mediante el modelo ajustado con el primer subconjunto. Es necesario obtener un tamaño de submuestra adecuado, que provea de calidad y validez en los resultados de la investigación.

Con el fin de conseguir una muestra representativa se partió de un muestreo aleatorio simple sin reposición, que frecuentemente se utiliza según lo mencionado por Capa (2007). La ecuación 4.3 y 4.4 se usaron para el cálculo de la muestra; tomado de Nel (2010)

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \quad (4.3)$$

$$n_0 = \frac{Z^2 * p * q}{E^2} \quad (4.4)$$

Dónde:

$n_0$ = Tamaño de muestra aproximado. Este valor debe corregirse cuando existen poblaciones pequeñas ya que la tasa ( $f=n/N$ ), no debe ser  $\geq 0.05$ .

$n$ = Tamaño de la muestra corregida.

$N$ = Tamaño de la población (149 datos considerados para la investigación)

$Z$ = Valores correspondientes al nivel de significancia.

$p$ = Proporción de datos que poseen la característica de estudio y que se debe estimar. Generalmente  $p=0.5$ , suponiendo que todos los puntos geodésicos tienen la misma posibilidad de ser considerados o no, para el modelamiento.

$q$ = Proporción de datos que no poseen esa característica ( $q=1-p$ )

$E$ = Error de tolerancia de la estimación (5%)

A un nivel de confianza del 95% el valor de  $n$  para el muestreo es de 107 datos, los cuales se usaron para realizar los modelos geoestadísticos, y los 42 puntos restantes para la verificación de los mismos.

En el Software R mediante el comando `sample2` se realizó una elección randómica entre la muestra de 149 puntos, obteniendo 107 puntos para estimar el modelo de predicción.

### 4.3. Obtención y validación de Modelos Geoestadísticos

#### 4.3.1. Modelo Geoestadístico de predicción de la variable desplazamiento de latitud ( $\Delta\phi$ )

##### 4.3.1.1. Análisis Exploratorio de Datos

El objetivo principal del análisis exploratorio de datos es proveer el resumen de la estadística descriptiva para cada variable e identificar valores atípicos o anómalos (outliers) que pudieron no ser identificados en la recopilación y validación de toda la información que se realizó en la sección (4.1). En la Tabla 2 se indica el resumen estadístico para la variable desplazamiento de latitud.

**Tabla 2**

#### Resumen estadístico para la variable desplazamiento de latitud

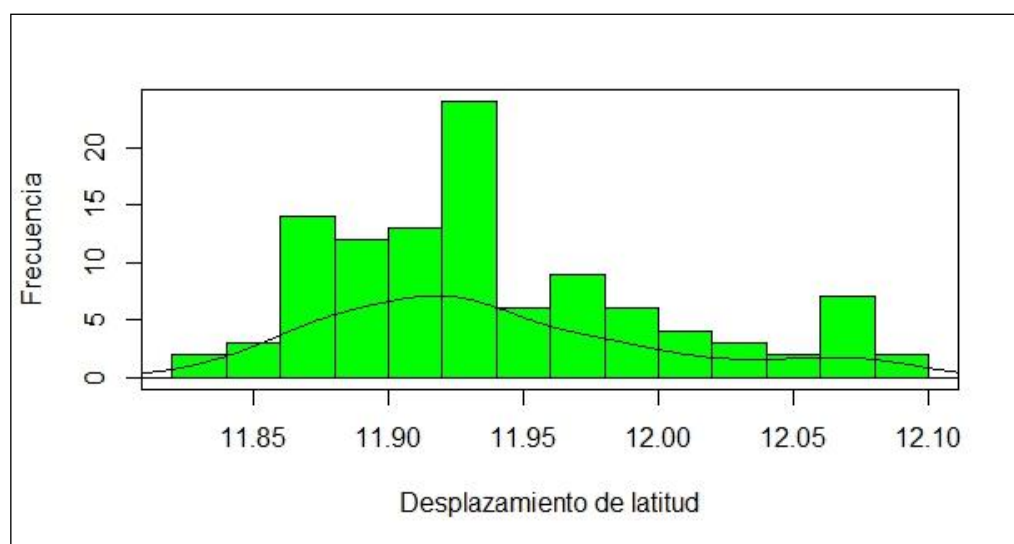
<b>N</b>	107
<b>Media (")</b>	11.939
<b>Desviación estándar (")</b>	0.063
<b>Rango (")</b>	0.263
<b>Coficiente de variación</b>	0.005
<b>Coficiente de asimetría</b>	0.744
<b>Curtosis</b>	-0.091
<b>Mínimo (")</b>	11.827
<b>1er Cuartil (")</b>	11.889
<b>Mediana (")</b>	11.924
<b>3er cuartil (")</b>	11.975
<b>Máximo (")</b>	12.090

<sup>2</sup> Toma una muestra de tamaño especificado a partir de los elementos utilizados.



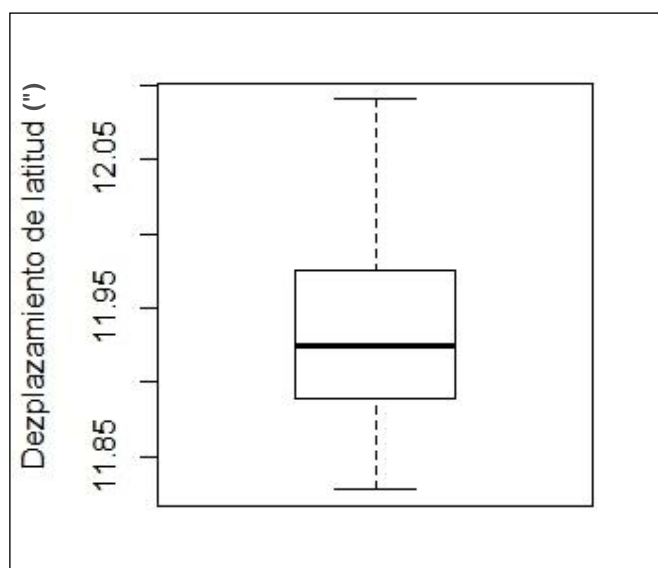
La media al tener un valor de 11.939" mayor que el de la mediana de 11.924", determina que la distribución de los datos es asimétrica sesgada a la derecha, lo cual también es analizado con el coeficiente de asimetría de 0.744 mayor que cero, describiendo la existencia de asimetría positiva, es decir la cola más dispersa está sobre los valores altos de la variable, con un rango de datos de 0.263".

Un coeficiente de variación de 0.5% indica homogeneidad en los datos. Con una curtosis de -0.091 se determina que la distribución se encuentra más plana, lo que representa datos no concentrados cerca de la media. El histograma (ver figura 33), comprueba lo mencionado anteriormente, donde la distribución es platicúrtica y sesgada a la derecha.



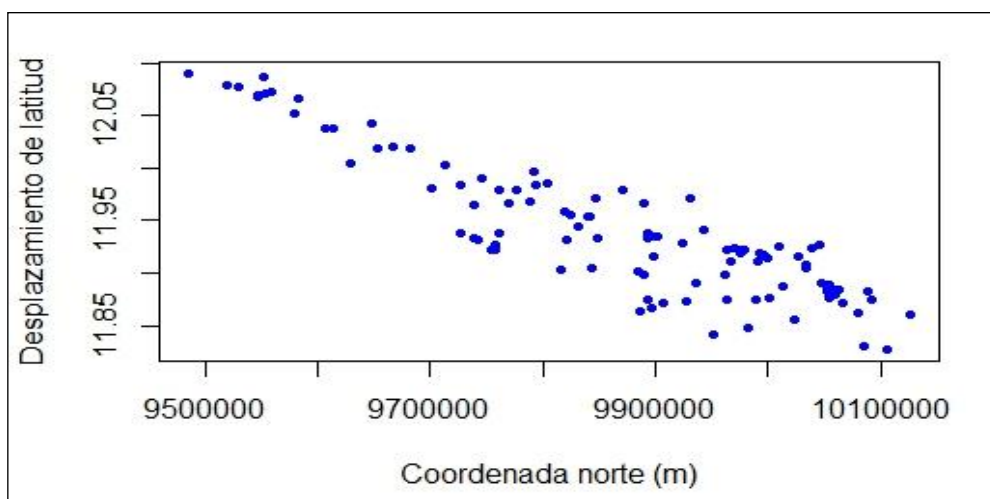
**Figura 33. Histograma de la variable desplazamiento de latitud**

El diagrama de caja de la variable desplazamiento de latitud (ver figura 34), no identificó valores u observaciones atípicas para la variable.



**Figura 34. Diagrama de caja de la variable desplazamiento de latitud**

El Diagrama de dispersión (ver figura 35) no identificó valores anómalos a considerar. Sin embargo, debido al comportamiento de los datos se presentó tendencia en la muestra, la cual fue analizada a partir de lo mencionado en la sección (3.4.3.6.)



**Figura 35. Gráfico de dispersión para la variable desplazamiento de latitud**

### 4.3.1.2. Variograma Experimental

Con los 107 datos del muestreo, se construyó el variograma experimental, bajo el supuesto de isotropía. Previamente es importante realizar la selección de los valores de los parámetros: la distancia de paso (lag), número de pasos (número de lags) y tolerancia en distancia.

Tanto el lag como el número de lags influyen en el resultado de manera significativa; el lag a distancias pequeñas posibilita el análisis de variaciones locales y el comportamiento de la autocorrelación espacial a pequeñas escalas, pero puede presentar un sobreajuste con un modelo teórico; a distancias grandes puede producirse “efecto pepita”, debido a la variabilidad a una distancia más pequeña que el  $h$  considerado. Entonces, considerando que el variograma es un promedio, este es bueno cuando el número  $N$  de parejas es grande. Sin embargo a medida que  $h$  crece,  $N$  decrece, por lo tanto Alfaro (2007) establece la siguiente regla:

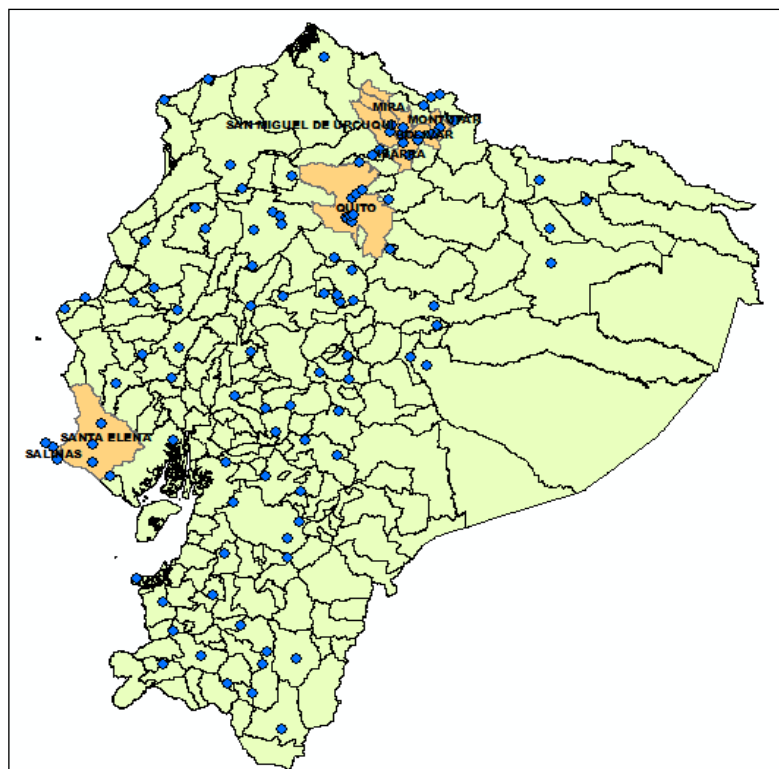
"Un variograma  $\gamma(h)$  es significativo hasta una distancia  $d$  igual a la mitad de la dimensión del campo en la dirección de  $h$ ". Bajo esta condición se considera la siguiente proposición para la solución:

$$lag_{\text{tamaño}} * \# lag = \frac{1}{2} d_{\text{máxima}} \quad (4.5)$$

#### 4.3.1.2.1. Tamaño y número del lag

Para el tamaño del lag, debido a que los datos se encuentran en una malla irregular, es importante realizar un análisis acerca de la distribución de las muestras en toda el área de estudio. Para ello se realizó un análisis estadístico espacial mediante la herramienta “Average Distance Mean” del software ARCGIS que calcula un índice de vecino más cercano en base a la distancia promedio desde cada entidad hasta la entidad vecina más cercana (ESRI, 2015). Además, la herramienta realiza una prueba de hipótesis en la cual se exhibe si la muestra se encuentra dispersa o agrupada. En la Figura 36 se observa los 107 puntos de muestreo, donde se identificó claramente la agrupación de puntos en Quito, Salinas, Santa Elena y en la parte centro norte del país, lo cual genera un sesgo en la muestra a la hora de determinar una distancia de lag adecuada.

Para determinar una distancia de lag promedio se procedió a sacar puntos muy cercanos entre sí, con el fin de obtener puntos dispersos y una distancia de lag que sea representativa para el desarrollo del variograma experimental (ver figuras 37). La Figura 38 muestra que con esta herramienta se obtuvo una primera aproximación de la distancia promedio de 31217.05 metros y una dispersión en los datos con el 1% de significancia.



**Figura 36. Puntos de muestreo en el Ecuador continental**

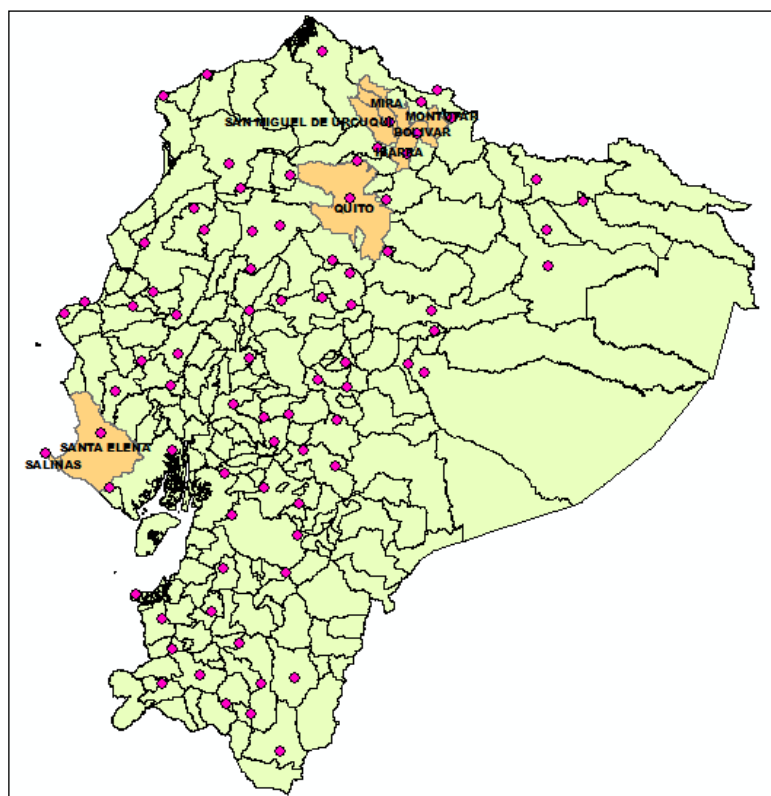


Figura 37. Puntos de muestreo sin agrupaciones (sin clusters)

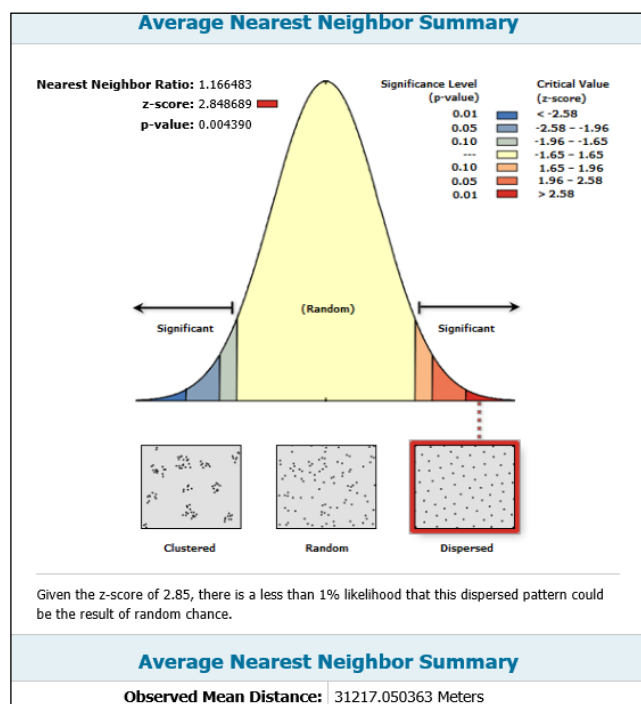


Figura 38. Prueba de hipótesis que muestra la dispersión de los puntos

Al obtener una distancia media entre los puntos de 31217.05 metros, se definió una distancia de lag de 32000 metros, y un número de lags de 10 (Díaz, 2002), justificando así la regla mencionada por Alfaro (2007) citada en la sección (4.3.1.2). De esta manera con los parámetros determinados, se cubrió la mitad de la distancia máxima del campo de estudio de 321561.5 metros. En la Tabla 3 se muestra la distancia, el número de lag, y su distancia correspondiente al multiplicar estos valores.

**Tabla 3**

**Cálculo del tamaño de lag adecuado**

Distancia de lag (m)	Número de lags	lag * n lag (m)
32000	10	320000

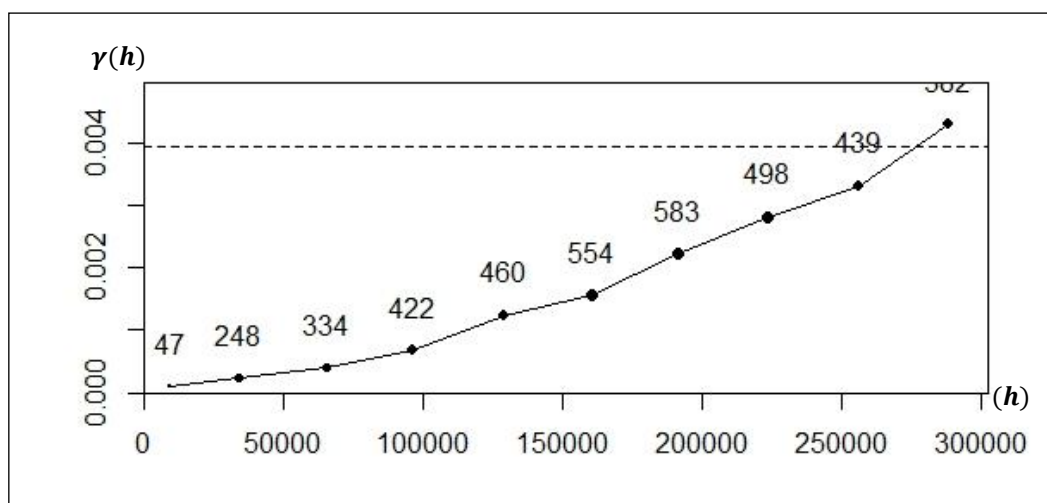
**4.3.1.2.2. Tolerancia en distancia ( $\epsilon$ ) y tolerancia angular ( $\theta$ )**

Según Alfaro (2007), la elección de  $\epsilon$  y  $\theta$  depende de la distribución espacial de los datos y de la práctica. En algunos casos la práctica recomienda utilizar  $\epsilon = 0.5b$ , en que  $b$  es la distancia mínima, llamada paso, para el cálculo de  $\gamma(h)$ .

**4.3.1.2.3. Cálculo del variograma**

Para el cálculo del variograma experimental omnidireccional se utilizó el software R, paquete RGeostats, con la función vario.calc<sup>3</sup>. A continuación la Figura 39 muestra el variograma experimental generado.

<sup>3</sup> Calcula varios tipos de variogramas experimentales para muestreos irregulares, para una o varias variables contenidas en una estructura de base de datos. Si está presente (definido en la entrada Db), se toma un conjunto de peso en cuenta en los cálculos. El variogramas se puede calcular con una distancia regular o no de lag.



**Figura 39. Variograma Experimental Omnidireccional - Variable  $\Delta\phi$**

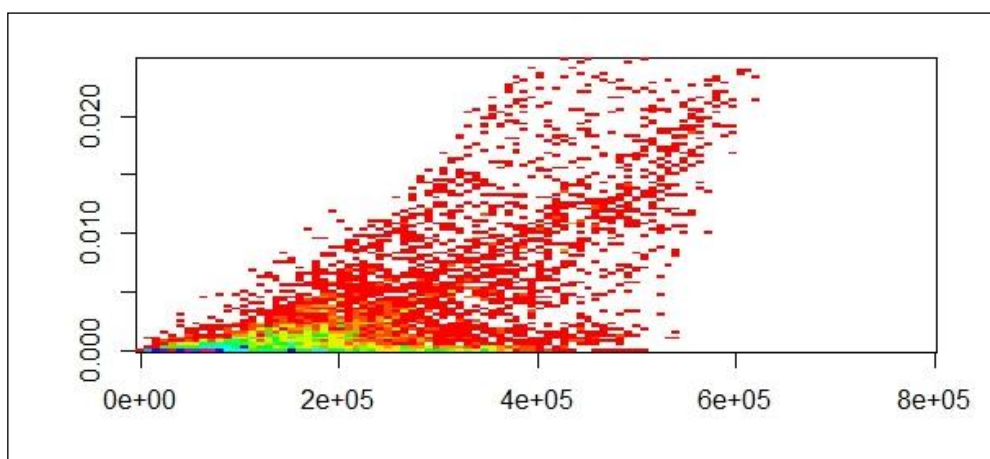
Al analizar el variograma experimental, se observó que los datos presentan ciertas características importantes:

- No hay efecto pepita (discontinuidad en el origen)
- El alcance no se identifica claramente
- La meseta tiende al infinito, con lo cual se evidencia un posible ajuste a un modelo lineal o potencia.

#### 4.3.1.3. Nube variográfica

La nube variográfica para la variable desplazamiento de latitud, se desarrolló con la herramienta cloud.calc<sup>4</sup>, del paquete RGeostats. El resultado de la Figura 40, es una nube variográfica omnidireccional que no presenta valores anómalos, lo que permite un buen ajuste del variograma experimental.

<sup>4</sup> Calcula la nube variográfica



**Figura 40. Nube Variográfica de la variable desplazamiento de latitud**

#### 4.3.1.4. Ajuste del variograma a un modelo teórico

El ajuste de variogramas constituye un punto crucial en un estudio geoestadístico porque todos los cálculos posteriores se harán utilizando exclusivamente el modelo teórico. Por ello es importante distinguir dos variogramas:

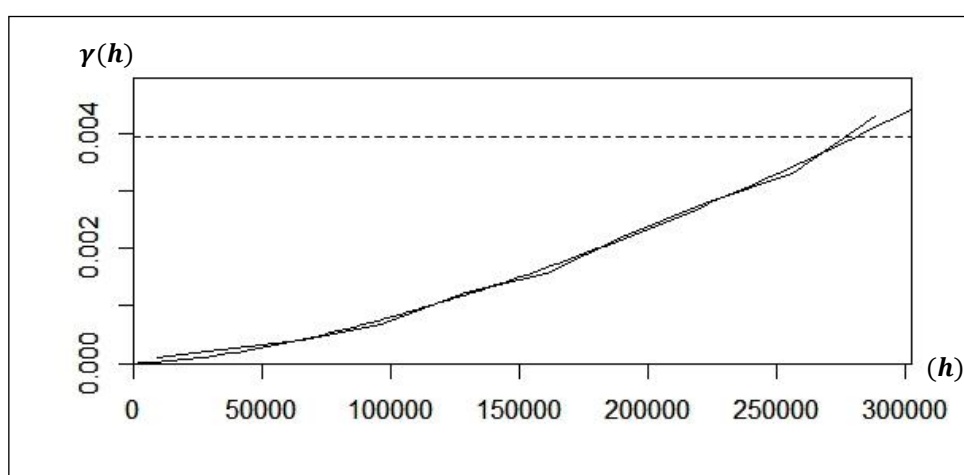
- El variograma experimental, calculado a partir de los datos.
- El variograma teórico, que corresponde a una ecuación que se ajusta al variograma experimental.

El comando para su cálculo es la función `model.auto`<sup>5</sup> del paquete RGeostats, que ajusta automáticamente el variograma experimental a partir de varias opciones de variogramas teóricos (estructuras básicas autorizadas, ver Anexo 4). Un algoritmo numérico se utiliza para calcular los parámetros, que minimizan la distancia entre el modelo y el variograma experimental. Los valores iniciales son elegidos de forma automática y el algoritmo es iterativo. Después de este primer paso, los parámetros con poca influencia se descartan del modelo y el modelo más afín se calcula utilizando nuevamente el algoritmo numérico. Este proceso se itera hasta que más parámetros no puedan ser descartados. Entre las opciones de ajuste también se considera un modelo anidado, que será la que mejor represente el variograma experimental (Desassis & Renard, 2012).

<sup>5</sup> Cálculo del variograma teórico. Ver en Anexo 4 las diferentes estructuras básicas que posee este comando.



Las estructuras básicas consideradas para el ajuste fueron: el efecto pepita, gaussiano, cubico, j-bessel, y de potencia. Se eligieron estas estructuras debido a que gráficamente se observa un posible ajuste con el variograma experimental obtenido y la forma que este presenta. Sin embargo, hay cierta preferencia con el modelo potencia ya que como menciona Bohling (2005) cuando hay presencia de tendencia en los datos, se ignora el problema y se usa este modelo para ajustar el variograma. El resultado que se obtuvo automáticamente fue un modelo potencia, con lo cual se corroboró la mencionado anteriormente (ver figura 41).



**Figura 41. Modelamiento de Variograma Experimental - Variable  $\Delta\phi$**

En la Tabla 4 se muestran los parámetros que se obtuvo para el modelo potencia ajustado al variograma experimental:

**Tabla 4**

**Parámetros de Estructura Básica - Variable  $\Delta\phi$**

Estructura	Parámetros	
	Sill /meseta ( $''^2$ )	Range/ alcance (m)
Potencia	C = 0.00454	305959

#### 4.3.1.5. Validación Cruzada

Para realizar la validación cruzada del modelo se requiere definir previamente la vecindad de búsqueda, que indica la estrategia de selección de muestras para predecir

el valor del dato del objetivo. Con lo mencionado en la sección (3.4.4.) para la validación cruzada se usaron las siguientes características:

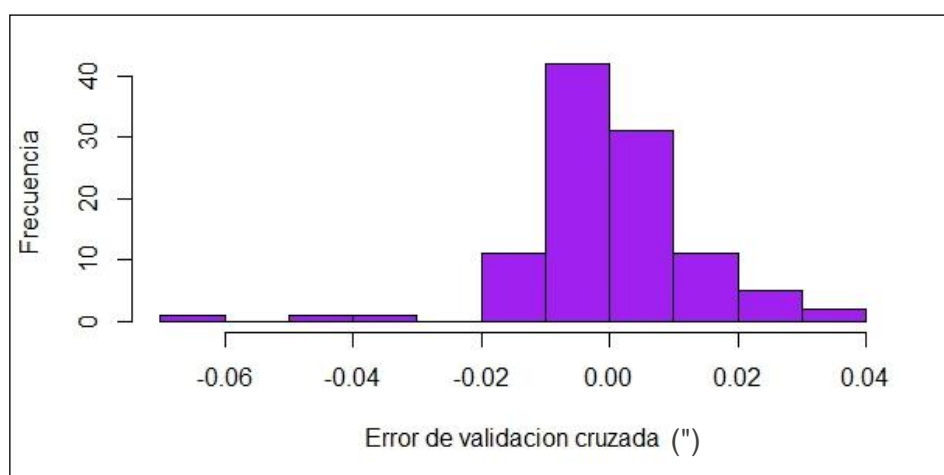
- **Vecindad:** móvil (al conocer la variable de estudio, se presume que un mejor ajuste se obtiene a nivel local, por tal el uso de un vecindario único no es el más adecuado)
- **Mínimo de muestras en vecindad:** 8 muestras
- **Máximo de muestras en vecindad:** 16 muestras
- **Definir sectores:** no (debido a la presencia de tendencia en la variable, no se usan sectores, para evitar que se generen artefactos, que afectan al modelo geoestadístico)
- **Máxima distancia de búsqueda:** la máxima distancia es de 155000 metros debido a que el muestreo está en una malla irregular, por lo tanto hay puntos muy alejados, y si se considera una distancia de búsqueda pequeña, es posible que un cierto número de puntos no interpolen por completo toda el área de estudio. Además bajo el criterio del ajuste con el variograma teórico, en la Figura 41 se constata que a esa distancia se presenció un ajuste casi perfecto del variograma teórico al variograma experimental.

Con el RGeostats, el comando para definir las características de vecindad de búsqueda es el *neigh.input*<sup>6</sup>. Para ejecutar la validación cruzada se usa el comando *x.valid*<sup>7</sup>, que genera dos variables: el error de la validación cruzada para el modelo, y el error normalizado de la validación cruzada, como muestra el siguiente histograma (ver figura 42).

---

<sup>6</sup> Dependiendo la vecindad única o móvil, describe la forma en que las muestras se seleccionan en torno a una ubicación de destino, cuando se realiza una estimación o una simulación.

<sup>7</sup> Procedimiento de validación cruzada



**Figura 42. Histograma de Error de Validación Cruzada -Variable  $\Delta\phi$**

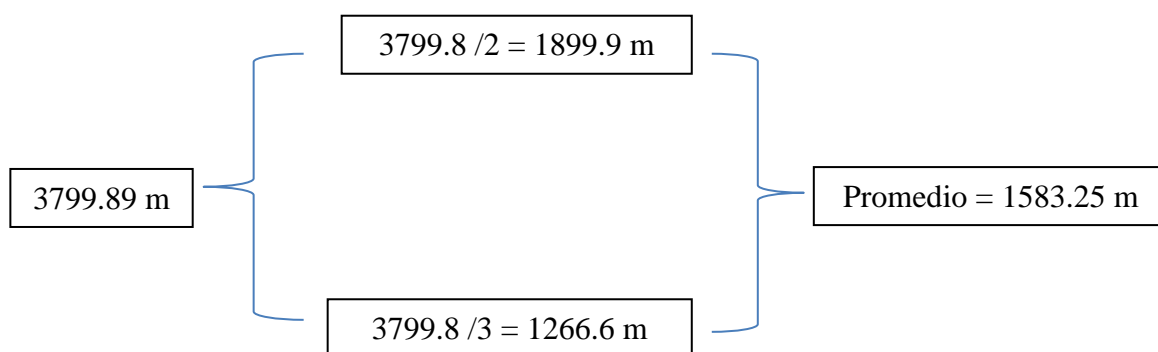
El histograma muestra que la mayor cantidad de valores del error de validación cruzada está en el intervalo de  $-0.02''$  a  $0.02''$ , que aproximadamente<sup>8</sup> representa  $[-0.6 \text{ m.}; 0.6 \text{ m.}]$ , lo cual es un indicativo de buena correspondencia de los datos con el modelo ajustado y con las características de vecindad elegidas.

#### 4.3.1.6. Predicción

Para la predicción espacial de los puntos no muestrales y, en definitiva, para la elaboración de mapas donde se muestre predicción de la variable desplazamiento de latitud, en todos los puntos del área de estudio, se empleó el kriging ordinario puntual sobre un enmallado (grilla) regular con un tamaño de pixel, para Este y Norte, de 1500 m. Para la obtención del tamaño de pixel adecuado se emplea la regla mencionada por Meijerink *et.al* (1994) donde “El tamaño de pixel de la grilla debe estar relacionado de alguna manera a la densidad de los puntos de los datos. Como regla general, la separación de los nodos de la red debe ser 1/2 a 1/3 de la distancia media de los pares de puntos más cercanos”.

Para la variable desplazamiento de latitud, la mínima distancia entre dos puntos es: 3799.89 metros.

<sup>8</sup> Es una aproximación debido a que la variable de estudio y el error del mismo se encuentran en arcos de segundo ( $''$ ), por lo tanto, es difícil interpretar su magnitud, por lo cual se hace la relación en donde 1 arco de segundo = 30 metros en el Ecuador

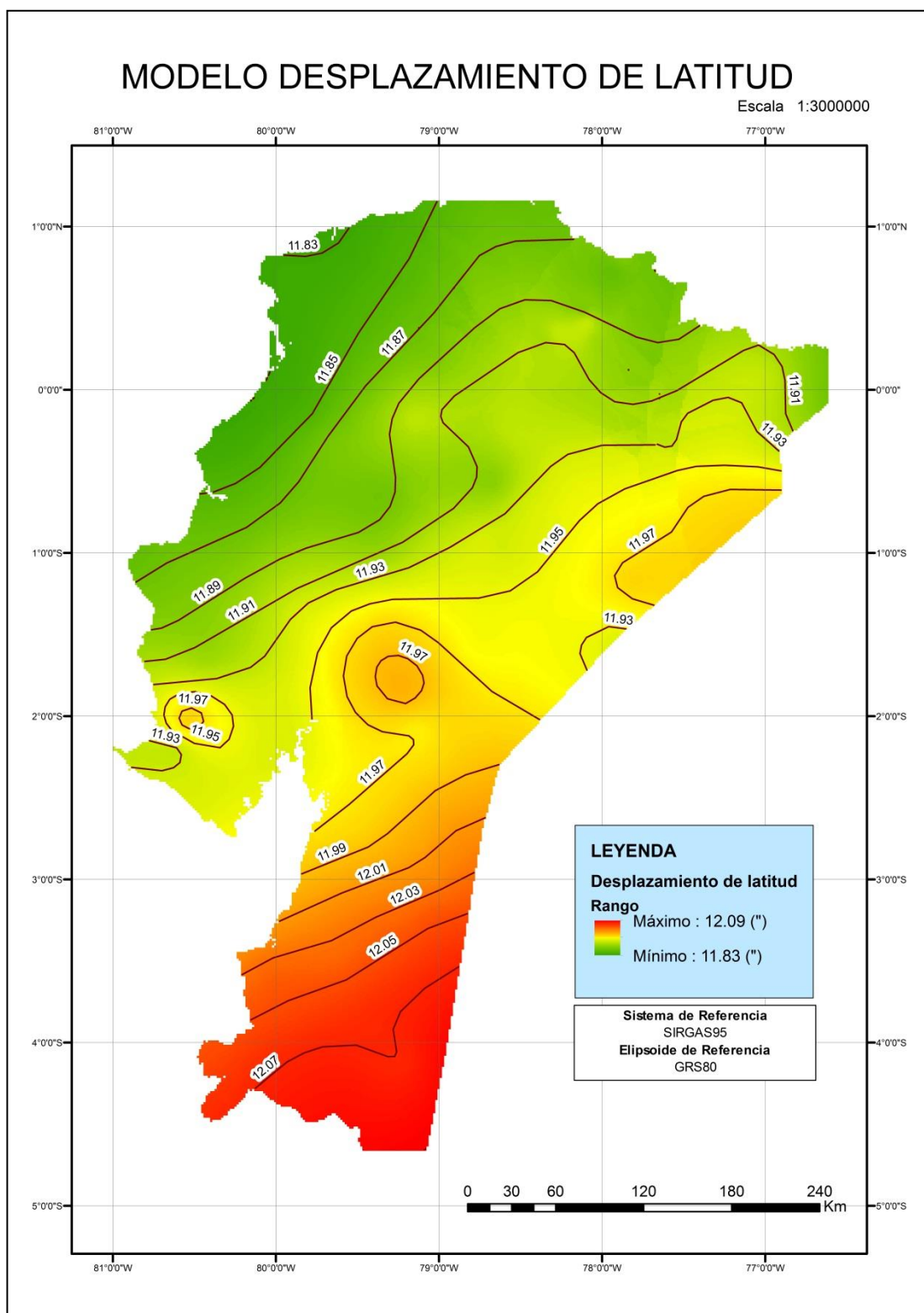


Por lo tanto el tamaño del píxel seleccionado para la predicción fue de 1500 m x 1500 m.

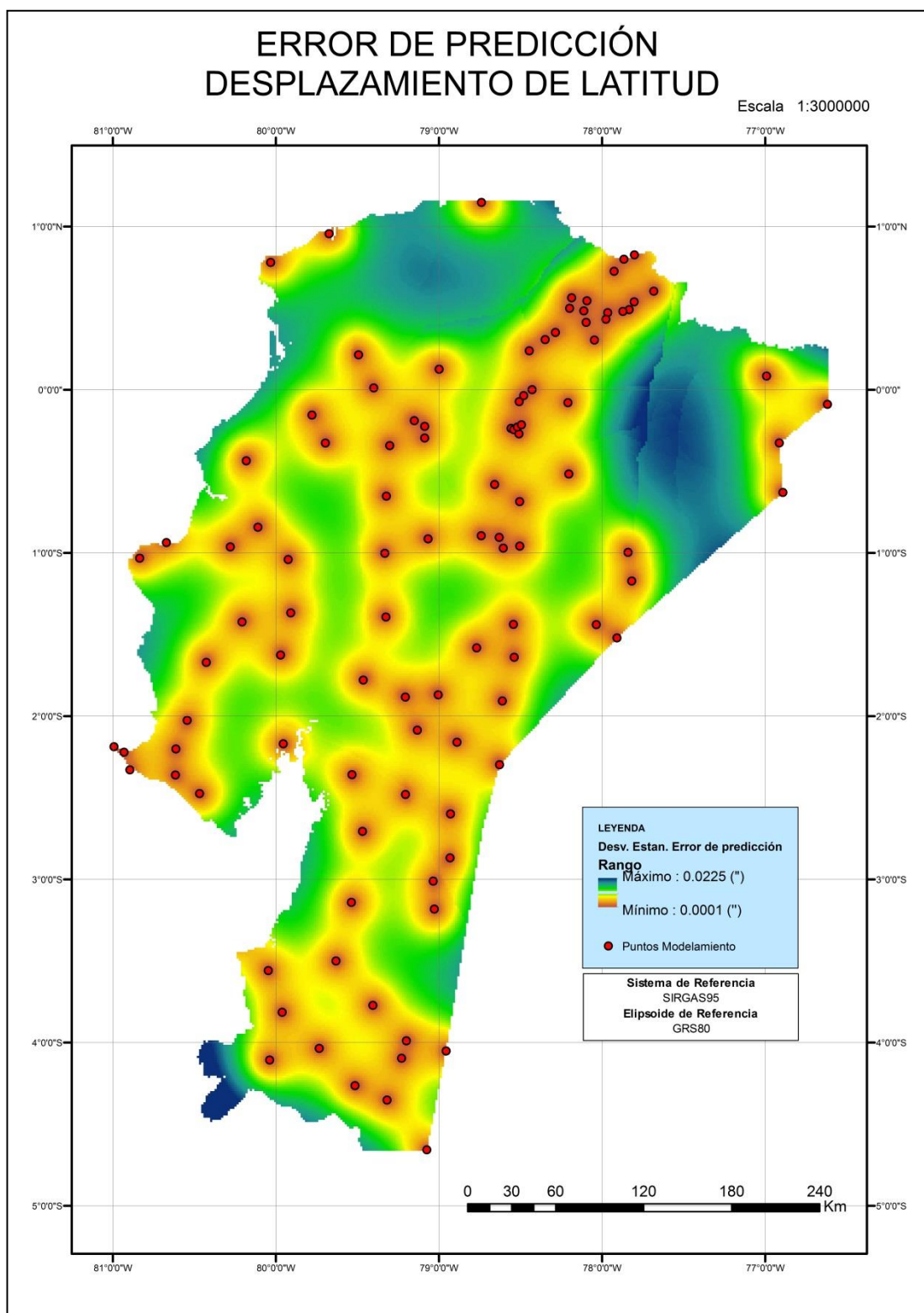
La predicción espacial se realizó a través de la función Kriging del paquete RGeostats, con las características establecidas, para una vecindad móvil. El resultado que se obtiene es una malla regular, con los valores de predicción para cada nodo y su respectiva desviación estándar del error de predicción.

La Figura 43 muestra el modelo de predicción espacial para la variable desplazamiento de latitud, con una resolución espacial de 1500 m. y un rango de variación en el modelo entre 11,83" y 12,09", con líneas de contorno cada 0.02". Como se observa, el desplazamiento de latitud se incrementa de norte a sur en el área de estudio sin presentar discontinuidades (artefactos) a considerar, lo cual da fe de una adecuada distribución de los puntos que se usaron para desarrollar la predicción (Ver Anexo 5, script para el modelo geoestadístico en el software R).

La Figura 44 indica el mapa con la desviación estándar del error de predicción para la variable desplazamiento de latitud; se observó que el error de predicción es de 0.0001" en los puntos de interpolación, y aumenta conforme se aleja de dichos puntos hasta llegar a un máximo de 0.0225", lo cual generó pequeños artefactos en el modelo. Por lo tanto, el intervalo de confianza, al 95%, del error de predicción es de [0.0002"; 0.0441"], equivalente en metros a [0.0006 m; 1.323 m] aproximadamente.



**Figura 43. Modelo geostatístico para la variable desplazamiento de latitud**



**Figura 44. Modelo de error de predicción para el desplazamiento de latitud**

#### **4.3.1.7. Elaboración del mapa de predicción para el manejo en un SIG**

Para la elaboración del mapa de predicción y el mapa de desviación estándar del error de predicción se utilizó un SIG, dada las facilidades que presta en el manejo de información raster y vector. Como paso previo, en el software R, mediante la función `db.write`<sup>9</sup> se exportó un archivo ASCII, que se estructura de la siguiente manera: coordenada este, coordenada norte, valores de predicción y errores de predicción.

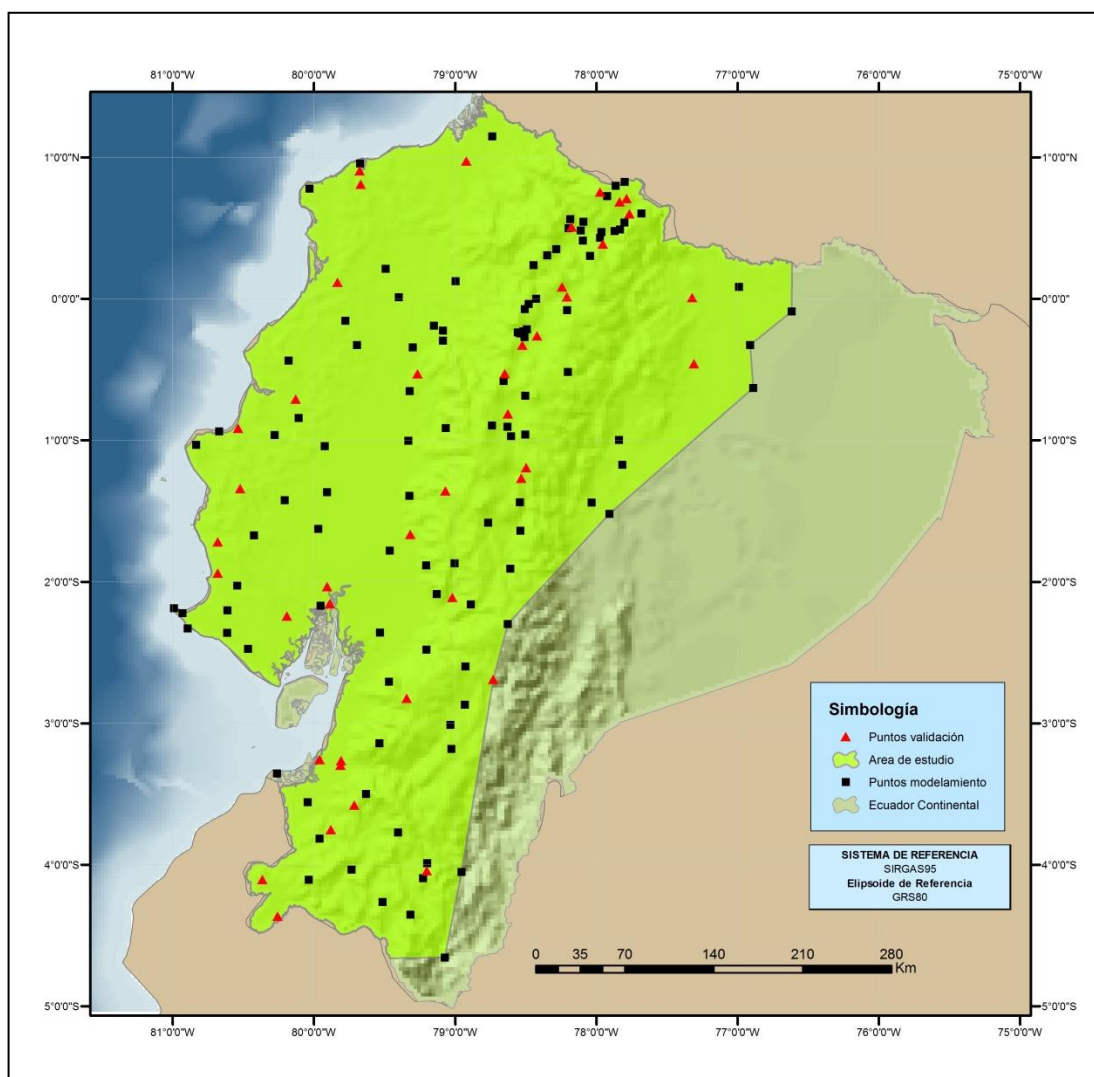
El software ARCGIS importa archivos ASCII, y permite transformarlo a un archivo grid, y posteriormente a un tif. Es importante conservar la resolución espacial de 1500 m del modelo para evitar errores en la visualización y en la predicción. Por último este modelo raster fue recortado a partir del polígono de estudio, obteniendo el modelo de predicción. De igual manera se realizó el modelo de desviación estándar del error de predicción, utilizando la variable errores de predicción del formato ASCII (Ver Anexo 6).

#### **4.3.1.8. Verificación del Modelo Geoestadístico**

La verificación del modelo geoestadístico se realizó mediante el uso de las muestras que fueron extraídas inicialmente, y que no son consideradas en el modelamiento. El total de muestras extraídas para la verificación es de 42 puntos geodésicos distribuidos en todo el polígono de estudio como se observa en la Figura 45.

---

<sup>9</sup> Función que transforma un archivo Db (RGeostats), a un archivo ASCII.



**Figura 45. Mapa de la distribución de los puntos de modelamiento y validación.**

El procedimiento de verificación se basa en comparar las 42 muestras del valor observado de la variable desplazamiento de latitud, con el valor predicho por el modelo geoestadístico. Esta diferencia entre los dos valores se conoce como error de verificación del modelo (ver tabla 5), con lo cual se realiza un análisis estadístico, como se indica en la Tabla 6.



Tabla 5

Error de verificación del modelo ( $\Delta\phi_{\text{OBSERVADA}} - \Delta\phi_{\text{PREDICHA}}$ )

N°	PUNTO GEODÉSICO	DIFERENCIAS OBSERVADAS	DIFERENCIAS PREDICHAS	ERROR VERIFICACIÓN MODELO
		$\Delta\phi$ (")	$\Delta\phi$ (")	$\Delta\phi$ (")
1	ALTO DE LAS CUESTAS	12.0843	12.0671	0.0172
2	ALUBURO PIMAMPIRO	11.8933	11.8940	-0.0007
3	BARBONA	11.8624	11.8681	-0.0057
4	BARCELONA	11.9125	11.9505	-0.0380
5	CAÑA QUEMADA	12.0334	12.0189	0.0145
6	CERRO NEGRO	12.0553	12.0543	0.0010
7	CHAGUARPATA	11.9740	11.9720	0.0020
8	CHILCHILES	12.0494	12.0424	0.0070
9	CHILES	11.8975	11.8670	0.0305
10	CHIQUICHA	11.9411	11.9497	-0.0086
11	CORAZON	11.9289	11.8925	0.0364
12	COREA	11.9280	11.9205	0.0075
13	DAMISA	11.8676	11.8835	-0.0159
14	EL CERRO	11.8526	11.8572	-0.0046
15	EL TIGRE	12.0560	12.0515	0.0045
16	ENVIDIA	11.8237	11.8285	-0.0047
17	FARO-GUA YAQUIL	11.9322	11.9415	-0.0093
18	GATAZO	11.8336	11.8314	0.0022
19	GUANTUGLOMA	11.9097	11.9105	-0.0008
20	HORQUETA	11.8688	11.8715	-0.0026
21	HUICOTANGO	11.9427	11.9472	-0.0045
22	ILALO	11.9206	11.9233	-0.0028
23	LATACUNGA bn	11.9316	11.9283	0.0032
24	LUMBAQUI	11.9163	11.9261	-0.0098
25	MACHALA	12.0341	12.0099	0.0242
26	MACHINES	11.8694	11.8731	-0.0037
27	MASHASHINGO	11.9736	11.9608	0.0128
28	MATAREDONDA	11.8663	11.8757	-0.0094
29	MESOPOTAMIA	11.8905	11.8857	0.0048
30	MULA POTRERO	11.9240	11.9124	0.0116
31	OLIMPO	11.8603	11.8613	-0.0010
32	PAYAMINO	11.9465	11.9500	-0.0035
33	PICHICONA	11.9393	11.9441	-0.0048
34	PUNTÓN	11.8804	11.9207	-0.0404
35	PUNZARA	12.0733	12.0711	0.0022
36	QUAQUE	11.8415	11.8373	0.0041
37	RAYA	11.8761	11.8852	-0.0091
38	SAN ENRIQUE	11.9271	11.9417	-0.0146
39	VERGEL	12.0383	12.0166	0.0217
40	YANTAHUAYCO	11.9904	11.9952	-0.0048
41	YAUSAY	12.0204	12.0143	0.0061
42	ZAMAMA	12.0023	11.9917	0.0105

media	0.0006
desviación es.	0.0145

Tabla 6

## Resumen estadístico error de verificación del modelo desplazamiento de latitud

n	42
Media (")	0.0006
Desviación estándar (")	0.0145
Rango (")	0.0768
Coficiente de asimetría	-0.2762
Curtosis	2.1238
Mínimo (")	-0.0404
1er Cuartil (")	-0.0048
Mediana (")	-0.0007
3er cuartil (")	0.0068
Máximo (")	0.0364

En la Figura 46 se muestra el histograma que representa el error de verificación

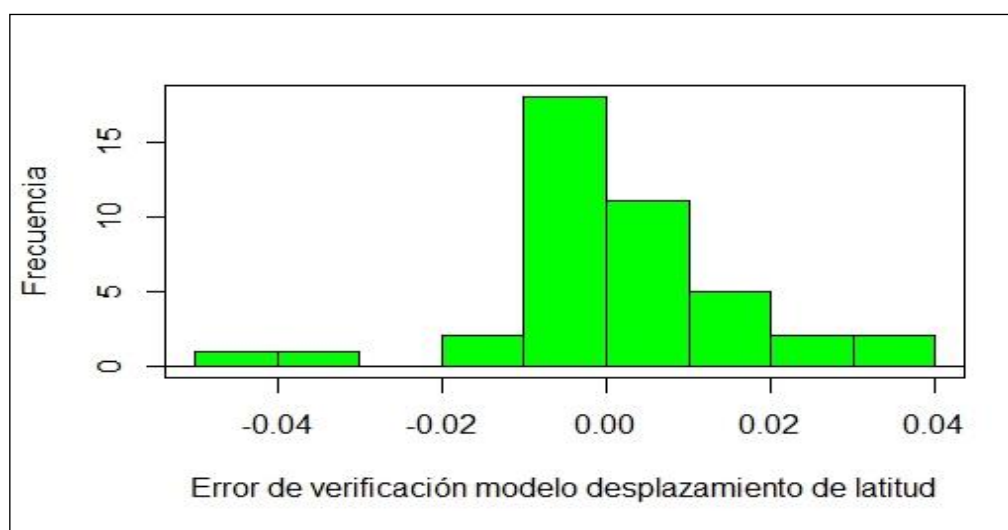


Figura 46. Histograma del error del modelo - Variable  $\Delta\phi$

Como se observó en el gráfico, el error en la verificación del modelo tiende a la normalidad y se agrupa en su mayoría en un intervalo de confianza de  $[-0.004''$  a  $0.005''$ ] al 95%, que a razón de metros representa  $[-0.11$  m. a  $0.15$  m.] aproximadamente.

#### 4.3.1.8.1. Coeficiente de determinación para la verificación del modelo

El coeficiente de determinación  $R^2$ , permite determinar la bondad de ajuste de un modelo. Para la variable desplazamiento de latitud el valor obtenido en el coeficiente de determinación, mediante el uso de la ecuación (3.31) fue de 0.9594, es decir una efectividad del 95.94%.

#### 4.3.2. Modelo Geoestadístico de predicción de la variable desplazamiento de longitud ( $\Delta\lambda$ )

##### 4.3.2.1. Análisis Exploratorio de Datos

La tabla 7 muestra el resumen estadístico para la variable desplazamiento de longitud:

**Tabla 7**

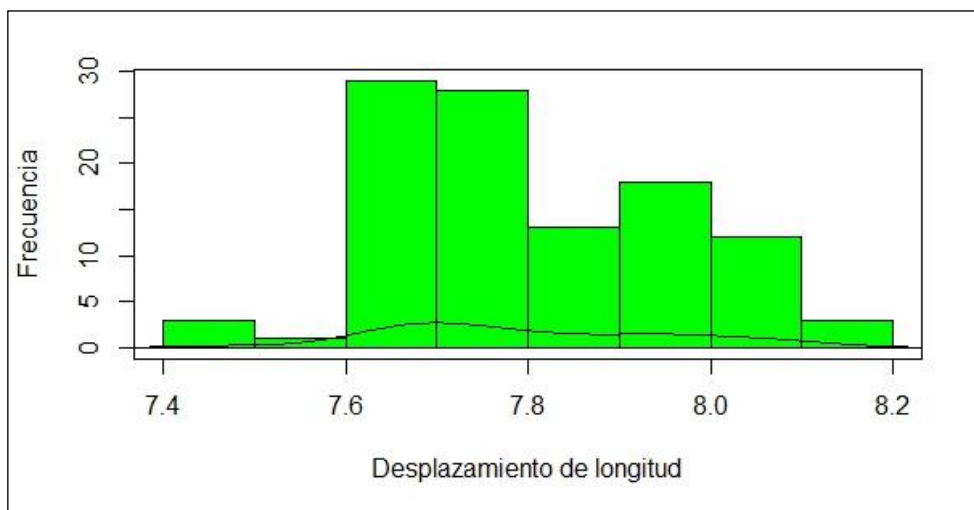
#### Resumen estadístico para la variable desplazamiento de longitud

<b>n</b>	107
<b>Media (")</b>	7.802
<b>Desviación estándar (")</b>	0.159
<b>Rango (")</b>	0.696
<b>Coefficiente de variación</b>	0.020
<b>Coefficiente de asimetría</b>	0.265
<b>Curtosis</b>	-0.672
<b>Mínimo (")</b>	7.445
<b>1er Cuartil (")</b>	7.693
<b>Mediana (")</b>	7.773
<b>3er cuartil (")</b>	7.933
<b>Máximo (")</b>	8.141

La media al tener un valor de 7.802" mayor que el de la mediana de 7.773", determina que la distribución de los datos es asimétrica sesgada a la derecha, lo cual también es analizado con el coeficiente de asimetría de 0.265, mayor que cero, estableciendo que existe asimetría positiva, es decir la cola más dispersa esta sobre los valores altos de la variable. Al tener valores en la media y mediana casi similares indica que no se presenta un sesgado muy marcado.

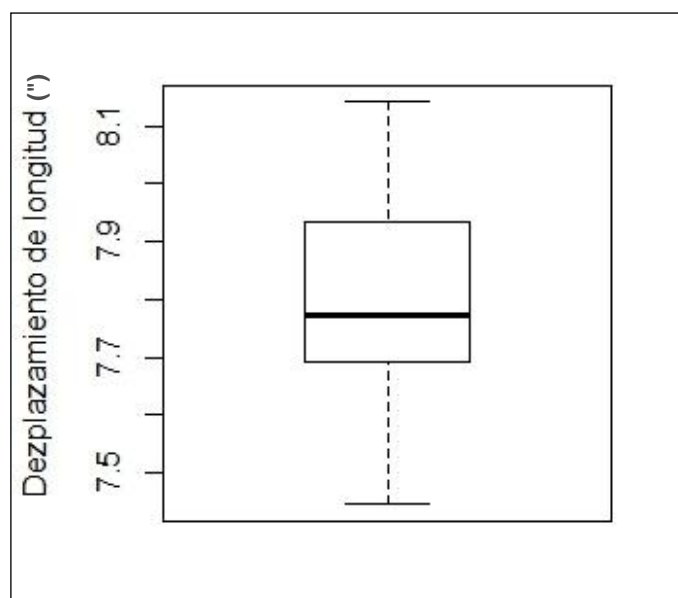
El coeficiente de variación es de 2%, lo cual significa que hay homogeneidad en los datos. Un rango de 0.696", indica que la variabilidad de los datos es mayor en la

variable desplazamiento de latitud. La curtosis de  $-0.672$  determina que la distribución se encuentra más plana, y representa datos no concentrados cerca de la media. El histograma (ver figura 47), comprueba lo mencionado anteriormente, donde la distribución es platicúrtica y no tan sesgada a la derecha.



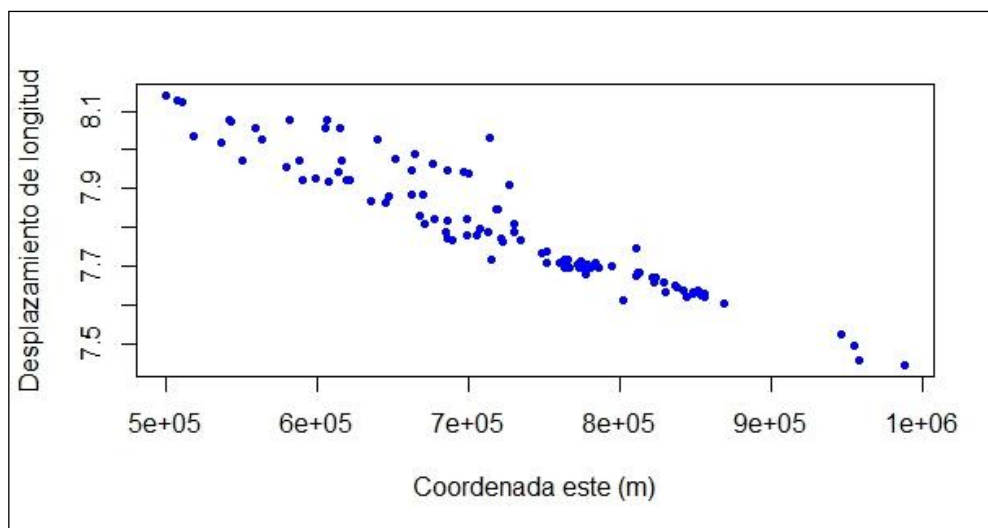
**Figura 47. Histograma de la variable desplazamiento de longitud**

El diagrama de caja (ver figura 48), no muestra valores u observaciones atípicas para la variable.



**Figura 48. Diagrama de caja de la variable desplazamiento de longitud**

El Diagrama de dispersión (ver figura 49) no identificó valores anómalos a considerar. De igual manera, al haber presenciado tendencia en los datos, se la analizó a partir de lo mencionado en la sección (3.4.3.6.).



**Figura 49. Gráfico de dispersión de la variable desplazamiento de longitud**

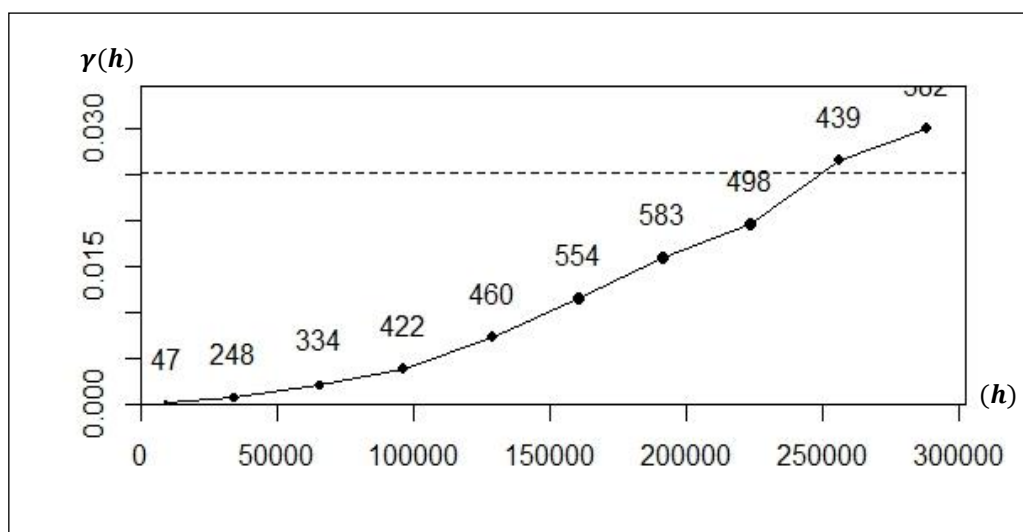
#### 4.3.2.2. Variograma Experimental

Con los 107 datos del muestreo, se construyó el variograma experimental, bajo el supuesto de isotropía. Los valores para el número de lag, y tamaño del lag, son los mismos establecidos para la variable desplazamiento de latitud, mencionados en la sección (4.2.1.2.)

- Tamaño del lag: 32000 m
- Número de lag: 10
- Tolerancia en distancia: 16000 m

##### 4.3.2.2.1. Cálculo del variograma

Para el cálculo del variograma experimental omnidireccional se utilizó el software R, paquete RGeostats, con la función vario.calc. A continuación la Figura 50 muestra el variograma experimental generado, con las 107 muestras.



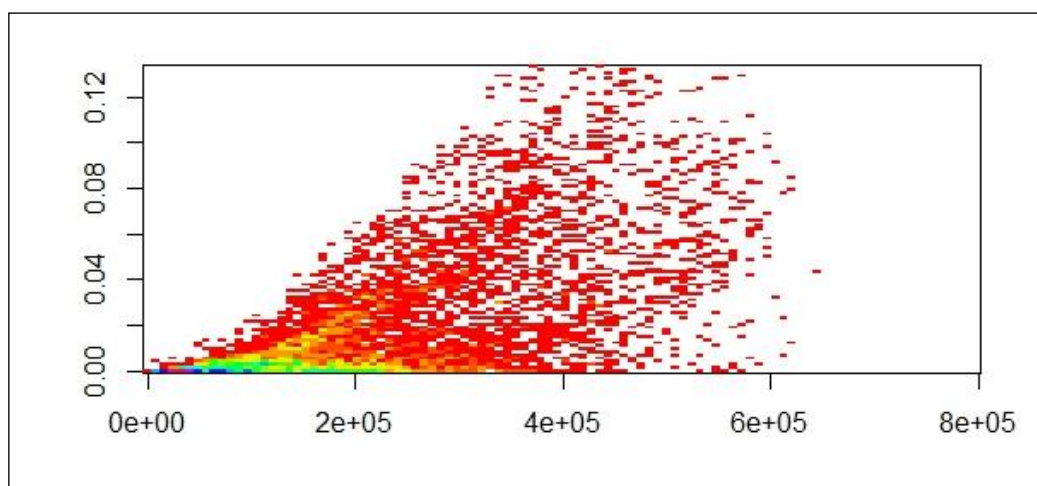
**Figura 50. Variograma Experimental Omnidereccional - Variable  $\Delta\lambda$**

Al analizar el variograma experimental, se observó que el variograma sigue una tendencia y presenta características importantes:

- No hubo presencia de efecto pepita.
- No se identifica claramente el alcance en el semivariograma
- La meseta o sill, al igual que en la variable  $\Delta\phi$ , tienden al infinito.
- Por la forma del variograma experimental, posiblemente se ajusta a un modelo gaussiano, cúbico, potencia y J-bessel.

#### 4.3.2.3. Nube variográfica

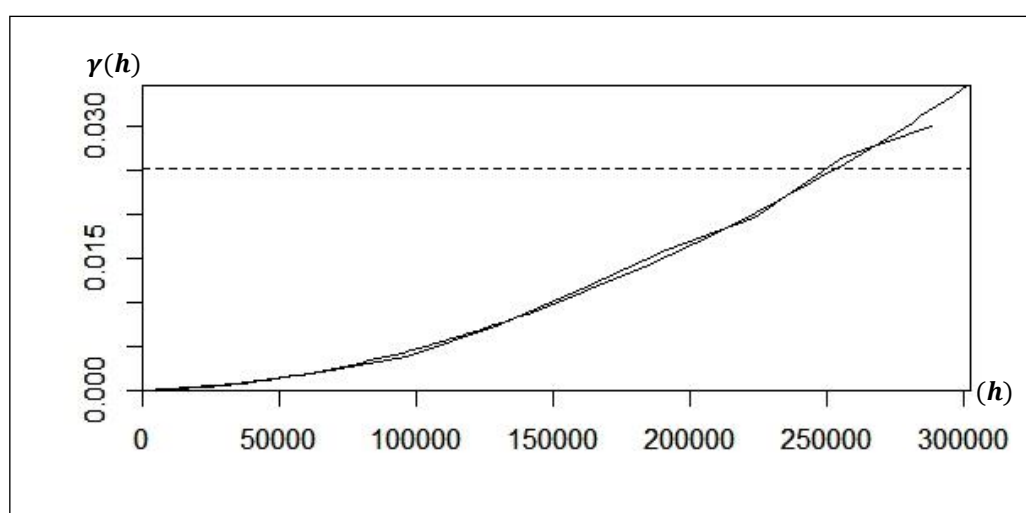
Se generó la nube variográfica con los datos para la variable desplazamiento de longitud. Los resultados que muestra la Figura 51 indican datos agrupados, y no existió la presencia de observaciones atípicas.



**Figura 51. Nube Variográfica de la variable desplazamiento de longitud**

#### 4.3.2.4. Ajuste del variograma a un modelo teórico

De la misma manera, que para el desplazamiento de latitud, se ajustó el variograma experimental con la función `model.auto` del paquete RGeostats. Las estructuras básicas que se consideraron para el ajuste son: el efecto pepita, gaussiano, cubico, j-bessel, y de potencia, a partir del mismo supuesto que se estableció para la variable desplazamiento de latitud. El resultado que se obtuvo automáticamente es un modelo potencia (ver figura 52).



**Figura 52. Modelamiento de Variograma Experimental - Variable  $\Delta\lambda$**

En la Tabla 8 se muestran los parámetros que se obtuvo para el modelo potencia ajustado al variograma experimental:

**Tabla 8**

**Parámetros de Estructura Básica - Variable  $\Delta\lambda$**

Estructura	Parámetros	
	Sill /meseta (") <sup>2</sup>	Range/ alcance (m)
Potencia	C = 0.0317	288044

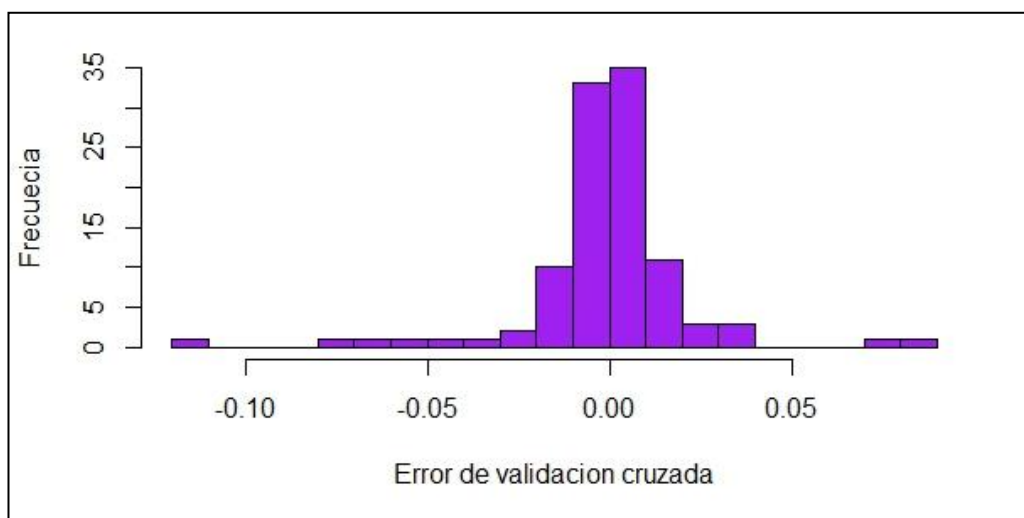
#### 4.3.2.5. Validación Cruzada

Para la validación cruzada del modelo se definió previamente la vecindad de búsqueda, que indica la estrategia de selección de muestras para predecir el valor del dato del objetivo. Los parámetros para la variable desplazamiento de longitud se definen a continuación:

- **Vecindad:** móvil (al conocer la variable de estudio, se presume que un mejor ajuste se obtiene a nivel local, por tal el uso de un vecindario único no es el más adecuado)
- **Mínimo de muestras en vecindad:** 8 muestras
- **Máximo de muestras en vecindad:** 16 muestras
- **Definir sectores:** no
- **Máxima distancia de búsqueda:** la máxima distancia es de 155000 metros debido a que el muestreo está en una malla irregular por lo tanto hay datos muy alejados, y si se considera una distancia de búsqueda pequeña, es posible que un cierto número de puntos no interpolen por completo toda el área de estudio. Con el mismo criterio que en la variable desplazamiento de latitud, se observó en la Figura 52 que a la distancia establecida se presencié un ajuste casi perfecto del variograma teórico al variograma experimental.

Con el RGeostats, el comando para definir las características de vecindad de búsqueda es el *neigh.input*. Para ejecutar la validación cruzada se usa el comando *x.valid*, que genera dos variables: el error de la validación cruzada para el modelo, y el error normalizado de la validación cruzada, como muestra el siguiente histograma (ver figura 53).





**Figura 53. Histograma de Error de Validación Cruzada -Variable  $\Delta\lambda$**

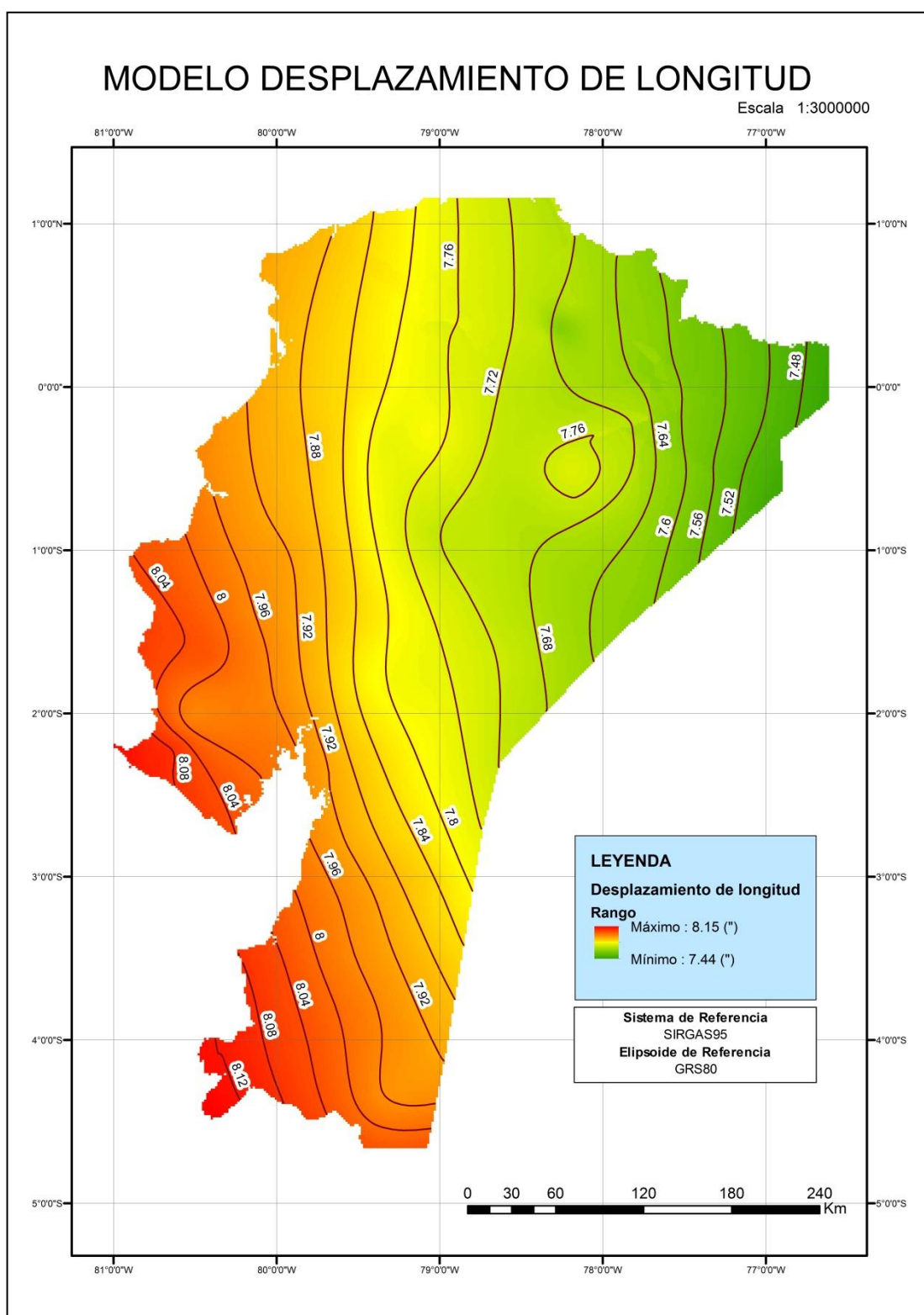
El histograma muestra el agrupamiento del error de validación cruzada entre el intervalo de  $[-0.05''$  a  $0.05''$ ], que aproximadamente representa  $[-1.5$  m ;  $1.5$  m.], lo cual es un indicativo de buena correspondencia de los datos con el modelo ajustado y con las características de vecindad elegidas. Además, se observa que el error en la validación cruzada alcanza un máximo de  $-0.10''$  (3 metros aprox.).

#### **4.3.2.6. Predicción**

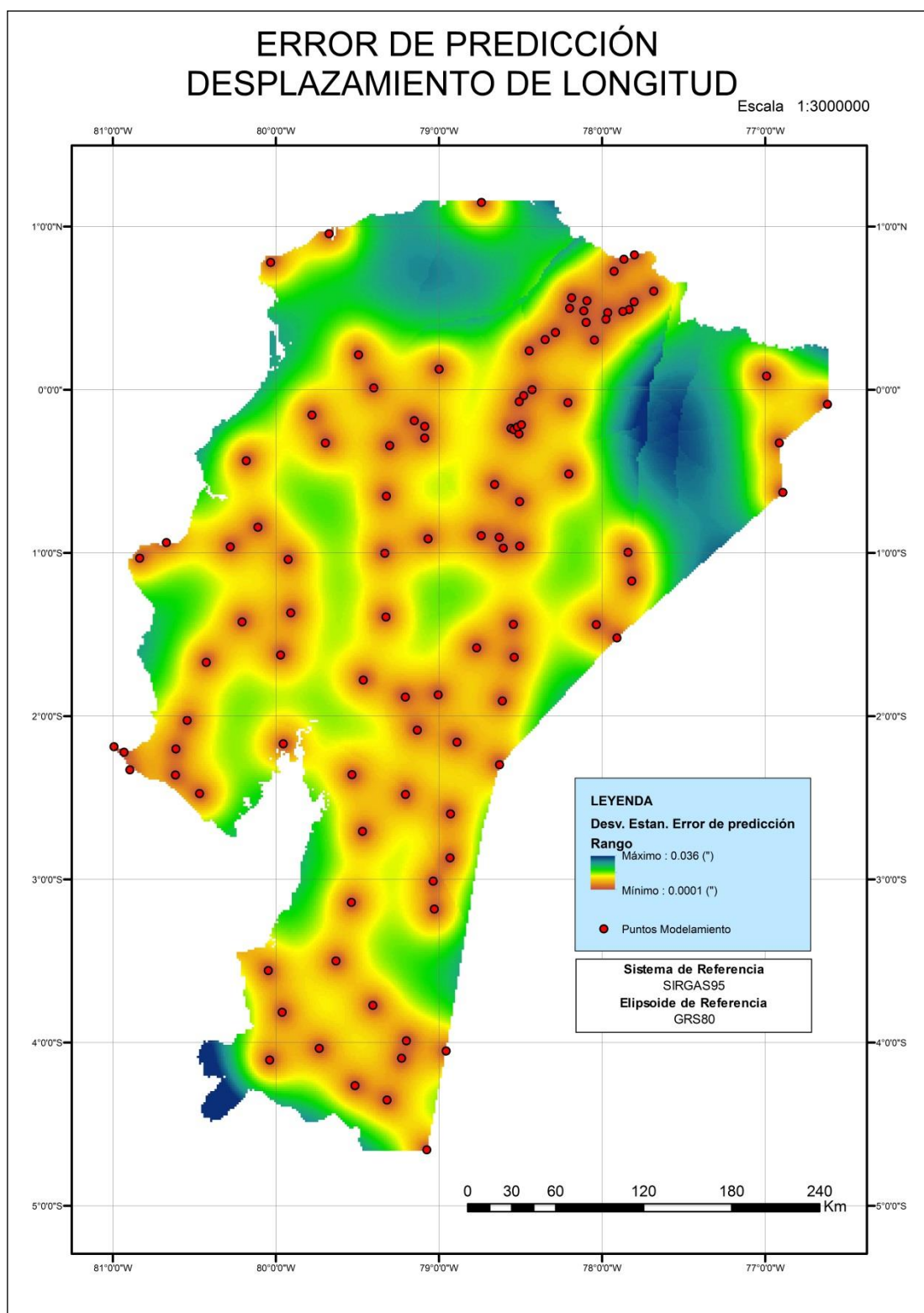
Para la elaboración del mapa de predicción de la variable desplazamiento de longitud, en todos los puntos del área de estudio, se empleó el kriging ordinario puntual sobre un enmallado (grilla) regular con un tamaño de pixel, para Este y Norte, de 1500 m. La predicción espacial se realizó a través de la función Kriging del paquete RGeostats, con las características establecidas, para una vecindad móvil. El resultado que se obtuvo es una malla regular, con los valores de predicción para cada nodo y su respectiva desviación estándar del error de predicción.

La Figura 54 muestra el modelo de predicción espacial para la variable desplazamiento de longitud, con una resolución espacial de 1500 m. y un rango de variación en el modelo entre  $7.44''$  y  $8.15''$ , con líneas de contorno cada  $0.04''$ . En el modelo de predicción, el desplazamiento de longitud se incrementa de este a oeste en el área de estudio y no presenta artefactos, como en el modelo de predicción para el desplazamiento de latitud, Ver Anexo 7, script para el modelo geostatístico en el software R).

La Figura 55 muestra el mapa con la desviación estándar del error de predicción para la variable desplazamiento de longitud; se observó como el error de predicción aumenta conforme se aleja de los puntos de interpolación donde el valor es de 0.0001", y llega a su máximo de 0.036" en zonas donde no se encuentran estos puntos. Además, la presencia de pequeños artefactos en sectores donde no se ubican dichos puntos, incrementa el error hasta su máximo valor. Por lo tanto, el intervalo de confianza, al 95%, del error de predicción fue de [0.0002"; 0.0706"], equivalente en metros a [0.0006 m; 2.1168 m] aproximadamente.



**Figura 54. Modelo geostatístico para la variable desplazamiento de longitud**



**Figura 55. Modelo de error de predicción para el desplazamiento de longitud**

#### **4.3.2.7. Elaboración del mapa de predicción para el manejo en un SIG**

Para la elaboración del mapa de predicción y el mapa de desviación estándar del error de predicción se utilizó un SIG, dada las facilidades que presta en el manejo de información raster y vector. Como paso previo, en el software R, mediante la función `db.write` se exporta un archivo ASCII, que se estructura de la siguiente manera: coordenada este, coordenada norte, valores de predicción y errores de predicción.

El software ARCGIS importa archivos ASCII, y permite transformarlo a un archivo grid, y posteriormente a un tif. Es importante conservar la resolución espacial de 1500 m del modelo para evitar errores en la visualización y en la predicción. Por último este modelo raster se recorta a partir del polígono de estudio, y se obtiene el modelo de predicción. De igual manera se realizó para obtener el modelo de desviación estándar del error de predicción, utilizando la variable errores de predicción del formato ASCII (Ver Anexo 8).

#### **4.3.2.8. Verificación del modelo geoestadístico**

Para la verificación del modelo geoestadístico de la variable desplazamiento de longitud, se usan los 42 puntos que no fueron parte del modelamiento, y que se obtuvieron a partir de un muestreo aleatorio simple, realizado en la base de datos total. El procedimiento de validación se basa en comparar los 42 muestras del valor observado del desplazamiento de longitud, con el valor predicho por el modelo geoestadístico (ver tabla 9). Esta diferencia entre los dos valores genera un error del modelo, con el cual se realiza un análisis estadístico de la nueva variable, como se indica en la Tabla 10.

Tabla 9

Error de verificación del modelo ( $\Delta\lambda_{\text{OBSERVADA}} - \Delta\lambda_{\text{PREDICHA}}$ )

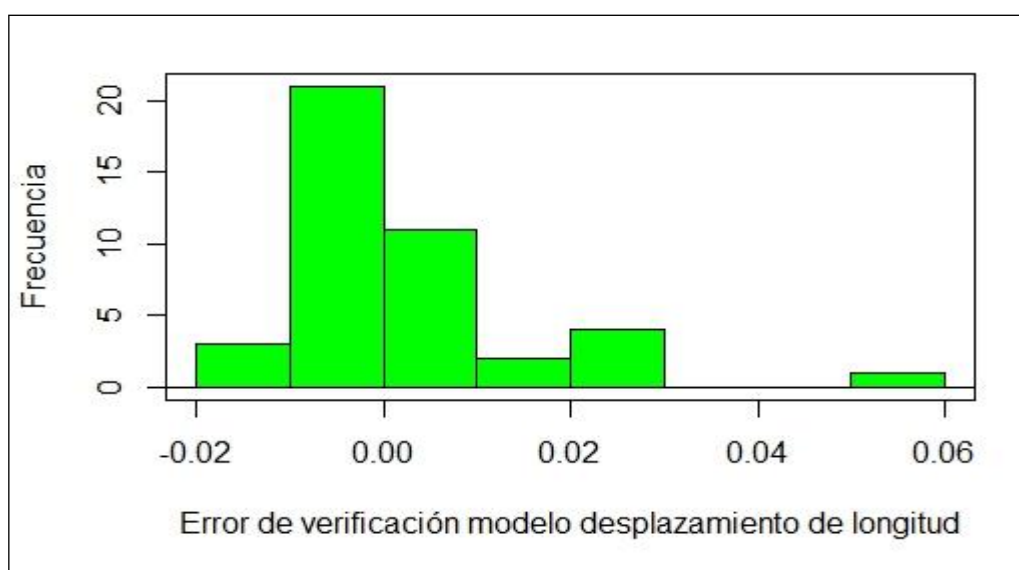
N°	PUNTO GEODÉSICO	DIFERENCIAS OBSERVADAS	DIFERENCIAS PREDICHAS	ERROR VERIFICACIÓN MODELO
		$\Delta\lambda$ (")	$\Delta\lambda$ (")	$\Delta\lambda$ (")
1	ALTO DE LAS CUESTAS	8.1218	8.1249	-0.0031
2	ALUBURO PIMAMPIRO	7.6428	7.6472	-0.0044
3	BARBONA	7.9975	7.9966	0.0009
4	BARCELONA	8.0803	8.0207	0.0596
5	CAÑA QUEMADA	8.0283	8.0004	0.0279
6	CERRO NEGRO	8.1450	8.1214	0.0236
7	CHAGUARPATA	7.7652	7.7825	-0.0174
8	CHILCHILES	8.0059	7.9977	0.0081
9	CHILES	7.6464	7.6468	-0.0004
10	CHIQUICHA	7.6879	7.6880	-0.0001
11	CORAZON	7.7081	7.7087	-0.0006
12	COREA	7.6884	7.7029	-0.0145
13	DAMISA	7.7666	7.7652	0.0014
14	EL CERRO	7.7566	7.7645	-0.0079
15	EL TIGRE	8.0430	8.0387	0.0043
16	ENVIDIA	7.8784	7.8814	-0.0029
17	FARO-GUA YAQUIL	7.9510	7.9561	-0.0051
18	GATAZO	7.8775	7.8756	0.0019
19	GUANTUGLOMA	7.6674	7.6727	-0.0053
20	HORQUETA	7.6239	7.6271	-0.0032
21	HUICOTANGO	7.6840	7.6828	0.0012
22	ILALO	7.6886	7.7015	-0.0130
23	LATACUNGA bn	7.6944	7.6957	-0.0013
24	LUMBAQUI	7.5648	7.5696	-0.0049
25	MACHALA	8.0183	8.0239	-0.0056
26	MACHINES	7.6269	7.6186	0.0083
27	MASHASHINGO	7.7609	7.7665	-0.0056
28	MATAREDONDA	7.6290	7.6197	0.0093
29	MESOPOTAMIA	7.6849	7.6833	0.0016
30	MULA POTRERO	7.6686	7.6704	-0.0018
31	OLIMPO	7.9223	7.9229	-0.0006
32	PA YAMINO	7.5504	7.5558	-0.0054
33	PICHICONA	8.0184	7.9975	0.0209
34	PUNTÓN	8.0654	8.0391	0.0263
35	PUNZARA	7.9396	7.9414	-0.0018
36	QUAQUE	7.8652	7.8745	-0.0093
37	RAYA	8.0171	8.0228	-0.0057
38	SAN ENRIQUE	7.9703	7.9509	0.0194
39	VERGEL	7.9932	7.9992	-0.0060
40	YANTAHUAYCO	7.8839	7.8776	0.0063
41	YAUSAY	7.7608	7.7568	0.0040
42	ZAMAMA	7.8098	7.7933	0.0165

media	0.0028
desviación es.	0.0138

**Tabla 10****Resumen estadístico error de verificación del modelo desplazamiento de longitud**

<b>n</b>	42
<b>Media (")</b>	0.0028
<b>Desviación estándar (")</b>	0.0138
<b>Rango (")</b>	0.0769
<b>Coficiente de asimetría</b>	2.0557
<b>Curtosis</b>	6.1561
<b>Mínimo (")</b>	-0.0174
<b>1er Cuartil (")</b>	-0.0052
<b>Mediana (")</b>	-0.0006
<b>3er cuartil (")</b>	0.0058
<b>Máximo (")</b>	0.0596

En la Figura 56 se muestra el histograma que representa el error de verificación y la distribución que presenta.

**Figura 56. Histograma del error de verificación del modelo - Variable  $\Delta\lambda$** 

Como se observó en el gráfico, el error en la verificación del modelo no posee una normalidad bien definida, pero se concentra la mayoría de los valores en un intervalo de confianza de [-0.001" a 0.007"] al 95%, que en metros representa [-0.04 m.; 0.21 m.] aproximadamente.

#### **4.3.2.8.1. Coeficiente de determinación para la validación del modelo**

El coeficiente de determinación  $R^2$ , permite determinar la bondad de ajuste de un modelo. Para la variable desplazamiento de longitud el valor obtenido en el coeficiente de determinación, mediante el uso de la ecuación (3.31) es de 0.9932, es decir, una efectividad del 99,32%.



## CAPÍTULO V

### ANÁLISIS DE RESULTADOS

#### 5.1. Comparación del Modelo Geoestadístico, con el Modelo Matemático de Helmert.

Para la transformación de coordenadas aplicando los modelos geoestadísticos, se emplearon los 42 puntos de verificación que no se consideraron en el modelamiento. Así también, estos puntos geodésicos se utilizaron para aplicar el modelo matemático de Helmert, mediante el uso de siete parámetros de transformación, propuesto por Leiva (2003) y posteriormente comparar los resultados obtenidos entre los 2 métodos. El objetivo principal es determinar que modelo proporciona un menor error en la transformación de coordenadas entre los sistemas de referencia.

Debido a que los modelos geoestadísticos de las variables desplazamiento de latitud y longitud predicen valores que están en arcos de segundo, para obtener la coordenada transformada en SIRGAS95 latitud y longitud, previamente se calcula la diferencia entre la coordenada PSAD56 ( $\varphi$ ,  $\lambda$ ), y el valor del desplazamiento de latitud y longitud ( $\Delta\varphi$ ,  $\Delta\lambda$ ), dividido para 3600", respectivamente como se indica a continuación:

$$\varphi_{SIRGAS95} = \varphi_{PASAD56} - (\Delta\varphi/3600") \quad (5.1)$$

$$\lambda_{SIRGAS95} = \lambda_{PASAD56} - (\Delta\lambda/3600") \quad (5.2)$$

En el caso del uso del modelo matemático de Helmert, las coordenadas en el sistema PSAD56, son transformadas con siete parámetros de transformación para obtener coordenadas en el sistema geocéntrico SIRGAS95.

Las coordenadas transformadas obtenidas con los dos métodos se analizan y se comparan con las coordenadas observadas en el sistema SIRGAS95 para comprobar como varía el error en la transformación. Es importante mencionar que no es posible interpretar de manera clara el error obtenido en grados, tanto para latitud y longitud en los dos sistemas, por lo tanto se usa el sistema ENU descrito en la sección (2.5.1.2.), que permite analizar el desplazamiento en metros, desde una coordenada inicial (Coordenada SIRGAS95 observada) tanto para Norte y Este (el

desplazamiento en altura “Up” no es considerado en este estudio), hasta una coordenada final (Coordenada SIRGAS95 transformada). En el Anexo 9 y 10 se indica la tabla con los cambios de posición obtenidos mediante el uso del sistema ENU, para los 42 puntos de verificación, tanto para el modelo estocástico como para el modelo determinístico.

Para determinar cuál es el modelo que da mejores resultados y minimiza el error en la transformación de coordenadas, se usan estadísticos entre los cuales los más importantes son la media, la desviación estándar, intervalo de confianza, ME y MSE. Según Oliver (2010), para determinar el error medio y el error medio cuadrático, se usan las formulas descritas en la sección (3.4.5.). La Tabla 11, muestra el resultado obtenido y la comparación tanto para el modelo geoestadístico, como para el modelo matemático de Helmert.

**Tabla 11**

**Resumen estadístico para la comparación de metodologías**

	<b>MODELO DETERMINÍSTICO Modelo Matemático de Helmert (7 parámetros)</b>	<b>MODELO GEOSTADÍSTICO</b>
<b>Muestra</b>	42 puntos	
<b>Media</b>	1.193 m	0.459 m
<b>Mediana</b>	1.173 m	0.330 m
<b>Desviación estándar</b>	0.521 m	0.413 m
<b>Varianza de la muestra</b>	0.271 m <sup>2</sup>	0.171 m <sup>2</sup>
<b>Curtosis</b>	-0.572	6.842
<b>Coefficiente de asimetría</b>	0.160	2.285
<b>Rango</b>	2.081 m	2.146 m
<b>Mínimo</b>	0.164 m	0.035 m
<b>Máximo</b>	2.244 m	2,181 m
<b>Intervalo de confianza</b>	1.035m – 1.350 m	0.334-0.584m
<b>ME</b>	1.193 m	0.459 m
<b>MSE</b>	1.393 m <sup>2</sup>	0.319 m <sup>2</sup>

La media, como una medida de tendencia central muestra en primera instancia el promedio del error en una muestra determinada, en este caso los 42 puntos de verificación. Como se observa, para el modelo determinístico la media es de 1.193 m., mientras que en el modelo geoestadístico la media es de 0.459 m. lo cual indica un menor promedio del error en el modelo geoestadístico.

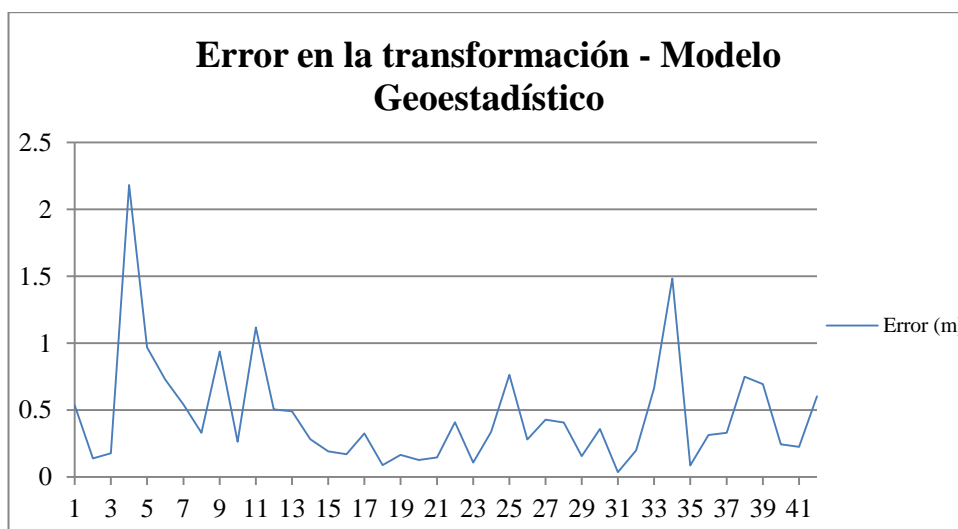
El análisis de la varianza y la desviación estándar determina el grado en que están dispersos los datos respecto a la media; a mayor medida mayor dispersión. Es así, que en el modelo determinístico el error se encuentra más disperso con una varianza de  $0.271 \text{ m}^2$  y una desviación estándar de  $0.521 \text{ m}$ ., en comparación con el modelo estocástico, que concentra más el error en la transformación de coordenadas, con un valor en la varianza de  $0.171 \text{ m}^2$  y una desviación estándar de  $0.413 \text{ m}$ .

El rango del error en los dos modelos no tienen mayor diferencia; pero, el cálculo del intervalo de confianza (I.C) al 95%, presenta una menor variabilidad en el modelo geoestadístico, siendo más estricto con valores entre  $0.327 \text{ m}$  a  $0.567 \text{ m}$ , con lo cual se identificó que el error obtenido en la transformación de coordenadas usando este modelo es menor, en comparación a lo obtenido al usar el modelo matemático de Helmert.

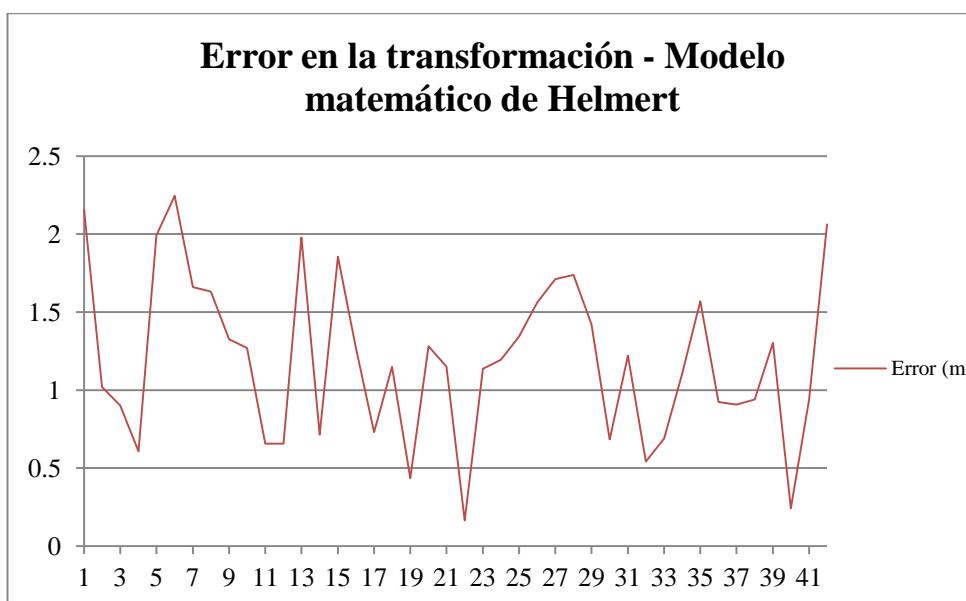
El uso del ME y el MSE permite identificar que modelo me da un mejor ajuste y cual es un mejor estimador en la transformación de coordenadas. Para Mueller y Schloeder (2001) cuando el ME es pequeño, se asocia a un método con pocos errores; en este caso el ME que más se acerca a cero es del modelo geoestadístico con un valor de  $0.459 \text{ m}$ . Esta medida aun así no refleja la magnitud de los errores que pueden ocurrir, para tal efecto se utiliza el valor MSE que es la suma de los residuos al cuadrado (varianza de los residuos), en donde, valores pequeños indican predicciones más precisas punto por punto. Para esta investigación, el modelo geoestadístico describe un MSE pequeño de  $0.319 \text{ m}^2$ , a diferencia del MSE del modelo determinístico con un valor de  $1.393 \text{ m}^2$ . Una vez analizada la estadística descriptiva para la verificación, se determina que el modelo geoestadístico usando kriging ordinario puntual, disminuye el error en la transformación de coordenadas.

Como se observó en la Tabla 11, el uso de un método geoestadístico mejora la estadística descriptiva del error que se genera al transformar coordenadas entre los dos sistemas de referencia, por ende, la metodología desarrollada en esta investigación mejora la transformación con respecto al modelo matemático de Helmert; así se comprueba en la Figura 57 y 58, donde se indica el error en metros obtenido punto por punto para cada método, observando que en el modelo estocástico solo se describe un error mayor a 2 metros correspondiente al punto geodésico

“Barcelona”, mientras en el método determinístico existen 3 puntos geodésicos mayores a este valor.



**Figura 57. Diferencia en la posición en metros (m).**



**Figura 58. Diferencia en la posición en metros (m).**

## 5.2. Determinación de la escala.

Una vez realizado el análisis del error obtenido, es importante determinar una escala adecuada hasta la cual el uso de este método de transformación sea eficiente;

para fines cartográficos a partir de las normativa establecida por el IGM, la cual menciona que la precisión en planimetría no debe diferir en más de  $0.3 \text{ mm}$ . por el denominador de la escala, se sugiere usar este método de transformación, a escalas 1:2000 y menores, con un intervalo de confianza de [0.334 m ;0.584m] al 95 %.

$$\frac{0.584}{0.0003} = 1946.7 \approx 2000$$

$$0.3 \text{ mm} * 2000 = 600 \text{ mm} = \pm 60 \text{ cm}$$

$$58 \text{ cm (máximo valor obtenido en el I.C)} \approx 60 \text{ cm}$$

Sin embargo, a pesar de estar vigente ésta metodología en el Ecuador, no es la única para determinar una escala de trabajo adecuada. Actualmente a nivel internacional se manejan nuevos estándares de precisión para información digital geoespacial de acuerdo a la National Standard for Spatial Data Accuracy (NSSDA) y la American Society for Photogrammetry and Remote Sensing (ASPRS), que establecen nuevos lineamientos para datos planimétricos y mapas a gran escala.

La ASPRS (2013) determina un valor de 0.0125 que se obtiene a partir del 1,25% del factor de escala, que es el recíproco de la relación utilizada para determinar la escala del mapa, es decir, de este valor se obtiene la precisión o la escala de las diferentes clases de datos (dependiendo de la precisión que se requiere y del nivel de detalle de los productos) en control horizontal (ASPRS, 2013). Para obtener la escala, el RMSE (raíz del error medio cuadrático) en  $x$  y  $y$ , es calculado en centímetros, y dividido para 0.0125, mientras para obtener la precisión se multiplica el 0.0125 por el factor de escala establecido, con lo cual se obtiene la escala y precisión que se puede alcanzar. Las ecuaciones del RMSE para  $x$ ,  $y$  y el RMSE total, se expresan a continuación:

$$RMSE_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{i(real)} - x_{i(predicho)})^2} \quad (5.3)$$

$$RMSE_y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{i(real)} - y_{i(predicho)})^2} \quad (5.4)$$

$$RMSE_r = \sqrt{RMSE_x^2 + RMSE_y^2} \quad (5.5)$$

Para la presente investigación se calculó el RMSE de los 42 puntos de verificación transformados con el modelo geoestadístico, tanto para Este ( $x$ ) y Norte ( $y$ ) y su respectivo RMSE total; los resultados se indican en la Tabla 12:

**Tabla 12**  
**Cálculo del RMSE**

	<b>metros</b>	<b>centímetros</b>
<b>RSME<sub>Este</sub></b>	0.4288	42.8811
<b>RSME<sub>Norte</sub></b>	0.4399	43.9907
<b>RSME<sub>Total</sub></b>	0.6143	61.4327

De acuerdo con el RMSE obtenido en  $x$  y  $y$  se calcula la escala de trabajo:

$$\left(\frac{RMSE_x (cm)}{0.0125}\right) = \frac{42.8811}{0.0125} = 3430.49$$

$$\left(\frac{RMSE_y (cm)}{0.0125}\right) = \frac{43.9907}{0.0125} = 3519.26$$

Se determina una escala de trabajo de 1:4000 y menores, para usar el método geoestadístico al transformar coordenadas entre Sistemas de Referencia.

Para los 42 puntos geodésicos, y un factor establecido de 2.4477 (APRS, 2013), la precisión en control horizontal al 95 % de confianza es igual

$$2.4477 * \left(\frac{RMSE_{total}}{1.4142}\right) = 1.7308(RMSE_{total})$$

$$1.7308(61.4327) = 106.3287 \text{ cm} = \pm 1.06 \text{ m}$$

Esta metodología sin embargo no se encuentra establecida oficialmente por el IGM, órgano rector de la cartografía en el país, pero se debería considerar para posteriores estudios.

## CAPÍTULO VI

### CONCLUSIONES Y RECOMENDACIONES

#### 6.1. Conclusiones

- La determinación de los modelos geoestadísticos se realizó mediante un total de 149 puntos geodésicos distribuidos en todo el Ecuador Continental, destinando 107 puntos para la predicción realizada con el interpolador Kriging ordinario puntual en el software R, paquete RGeostats y los 42 puntos restantes se usaron para la verificación de los modelos obtenidos.
- Para la predicción de la variable desplazamiento de latitud, el variograma teórico que mejor se ajustó al variograma experimental fue un modelo potencia, que al elegir características de vecindad móvil para el predictor de kriging, se obtuvo que la mayor parte de errores de predicción en la validación cruzada se encuentra en el intervalo de [-0.6 m.; 0.6 m.]. Además, el modelo obtenido para la desviación estándar del error de predicción de la variable ( $\Delta\phi$ ) determinó un error máximo de predicción de 1.32 m., al 95% de confianza. Adicionalmente, a partir de la verificación del modelo obtenido, se analizó la bondad de ajuste mediante el coeficiente de determinación  $R^2$ , presentando una efectividad del 95.94%, lo cual demuestra que la predicción realizada es buena.
- Para la predicción de la variable desplazamiento de longitud, el variograma teórico que mejor se ajustó al variograma experimental fue un modelo potencia, que al elegir características de vecindad móvil para el predictor kriging, se obtuvo que la mayor parte de errores de predicción en la validación cruzada se encuentra en el intervalo de [-1.5 m.; 1.5 m.]. Además, el modelo obtenido para la desviación estándar del error de predicción de la variable ( $\Delta\lambda$ ) determinó un error máximo de predicción de 2.12 m, al 95% de confianza. Adicionalmente, a partir de la verificación del modelo obtenido, se analizó la bondad de ajuste mediante el coeficiente de determinación  $R^2$ , presentando una efectividad del 99.32%, lo cual demuestra que la predicción realizada es buena.

- Se realizaron mapas en el software ArcGIS, a partir de los modelos de predicción obtenidos para las variables desplazamiento de latitud y longitud, a través de una malla regular con una resolución espacial de 1500 m x 1500 m. Los productos resultantes son archivos digitales en formato .grid y .tif (modelos digitales), para cada variable, delimitados por el polígono de estudio.
- Al analizar los estadísticos descriptivos del error de transformación (mostrados en la Tabla 11), tanto para el método geoestadístico como para el método determinístico (Helmert), se comprobó que el método geoestadístico es más eficiente para transformar coordenadas entre sistemas de referencia, lo que se verificó al analizar los intervalos de confianza al 95% del error de transformación: método geoestadístico [0.327 m.; 0,567 m.]; método de Helmert [1.035 m.; 1.350 m].
- El método geoestadístico para la transformación de coordenadas entre el Sistema de Referencia PSAD56 y SIRGAS95 es aplicable a escalas 1:2000 y menores, basado en la normativa para precisión horizontal utilizada por el IGM. Sin embargo, a partir de los estándares internacionales de precisión para datos horizontales (ASPRS y NSSDA) la escala para usar este método es de 1:4000 y menores.

## 6.2. Recomendaciones

- En base a los resultados obtenidos, en este estudio, se recomienda el uso de los modelos geoestadísticos para la transformación de coordenadas, al método matemático de Helmert, que actualmente se usa en el país, especialmente para escalas grandes.
- Para futuros estudios enmarcados en esta problemática, se recomienda analizar la anisotropía de las variables  $\Delta\phi$ ,  $\Delta\lambda$ , en vista que en el presente estudio se usó la hipótesis de isotropía.
- De igual manera, se recomienda analizar el tratamiento de tendencia de las variables  $\Delta\phi$ ,  $\Delta\lambda$ , ya que los gráficos exploratorios mostraron indicios de presencia de deriva.



## REFERENCIAS

- Alfaro, M. (2007). *Estimacion de Recursos Mineros*. Recuperado el 22 de Junio de 2015, de [http://cg.ensmp.fr/bibliotheque/cgi-bin/public/bibli\\_index.cgi](http://cg.ensmp.fr/bibliotheque/cgi-bin/public/bibli_index.cgi)
- ASPRS. (2013). Accuracy Standards for Digital Geospatial Data. *Photogrammetric Engineering and Remote Sensing*, 1073-1085.
- Atlas Global de la Región de Murcia. (2010). *La cartografía actual: topográfica, ortofotos e imágenes de satélite*. Recuperado el 19 de Noviembre de 2015, de <http://www.atlasdemurcia.com/index.php/secciones/3/cartografia-actual-topografica-ortofotos-satelite/2/>
- Benitez, E., & García-Asenjo, L. (2001). *Sistemas de Referencias*. Recuperado el 21 de Abril de 2015, de [http://www.catastro.minhap.gob.es/documentos/publicaciones/ct/ct42/ct42\\_4.pdf](http://www.catastro.minhap.gob.es/documentos/publicaciones/ct/ct42/ct42_4.pdf)
- Bohling, G. (2005). *Kansas Geological Survey*. Recuperado el 15 de Octubre de 2015, de <http://people.ku.edu/~gbohling/cpe940>
- Bohling, G. (2005). *Resources for C&PE940, Data Analysis in Engineering and Natural Science*. Recuperado el 5 de Julio de 2015, de <http://people.ku.edu/~gbohling/cpe940/Kriging.pdf>
- Cañizares, C., & Leiva, C. (2013). *Determinación de parámetros de transformación entre los sistemas de referencia PSAD56 y SIRGAS95, para el Ecuador, mediante el modelo Molodensky-Badekas*. Quito: Universidad de las Fuerzas Armadas-ESPE.
- Capa, H. (2007). *Elementos y diseños básicos para muestreo*. Quito: EPN.
- Castillejo, A. (2007). *Teoría geostadística aplicada al análisis de la variabilidad espacial arqueológica INTRA-SITE*. Recuperado el 22 de Agosto de 2015, de Universidad Autónoma de Barcelona: <http://www.tdx.cat/bitstream/handle/10803/5526/amc1de1.pdf?sequence=1>
- Chasco, C. (2003). *Metodos Graficos del Analisis Exploratorio de Datos Espaciales*. Recuperado el 24 de Agosto de 2015, de Universidad Autonoma de Madrid: <http://www.asepelt.org/ficheros/File/Anales/2003%20-%20Almeria/asepeltPDF/93.PDF>
- Cisneros, D. (2013). *Campo de velocidades del Ecuador obtenidaa través de mediciones de campanas*. Recuperado el 15 de Junio de 2015, de [http://www.geoportaligm.gob.ec/wordpress/?wpfb\\_dl=65](http://www.geoportaligm.gob.ec/wordpress/?wpfb_dl=65)

- CNUGGI. (1999). *Sistemas Geodésicos*. Recuperado el 20 de Agosto de 2015, de <http://www.ign.gob.ar/images/cnuggi/SistemasGeodesicos.pdf>
- Desassis, N., & Renard, D. (2012). Automatic variograms modeling by iterative least squares. Univariate and multivariate cases. *Mathematical Geosciences*, 453-470.
- Díaz, M. (2002). *Geoestadística Aplicada*. Cuba: Instituto de Geofísica, UNAM. Instituto de Geofísica y Astronomía, CITMA.
- Drewes, H. (2014). *Sistemas de Referencia*. Recuperado el 5 de Agosto de 2015, de [http://www.igm.gob.ec/work/files/adela.camacho/Comunicacion2014/Presentacion.\\_Sistema\\_Referencia.pdf](http://www.igm.gob.ec/work/files/adela.camacho/Comunicacion2014/Presentacion._Sistema_Referencia.pdf)
- ESRI. (2003). *Using ArcGIS® Geostatistical Analyst*. Recuperado el 5 de Diciembre de 2015, de [http://dusk2.geo.orst.edu/gis/geostat\\_analyst.pdf](http://dusk2.geo.orst.edu/gis/geostat_analyst.pdf)
- ESRI. (2015). *ARCGIS Pro*. Recuperado el 18 de Noviembre de 2015, de Promedio de distancia de vecinos cercanos: <https://pro.arcgis.com/es/pro-app/tool-reference/spatial-statistics/average-nearest-neighbor.htm>
- Fernández, I. (2001). *Localizaciones Geográficas. Las Coordenadas Geográficas y la Proyección UTM*. Recuperado el 22 de Junio de 2015, de <http://www.cartesia.org/data/apuntes/cartografia/cartografia-geograficas-utm-datum.pdf>
- Furones, Á. M. (2010). Recuperado el Agosto de 2015, de Universidad Politécnica de Valencia: <http://www.upv.es/unigeo/index/docencia/etsigct/astronomia/teoria/astronomia210.pdf>
- Gallardo, A. (2008). *Métodos geoestadísticos para el análisis de datos ecológicos espacialmente explícitos*. Recuperado el 24 de Septiembre de 2015, de [http://www.escet.urjc.es/biodiversos/espainvestigacion/intercambio/doc/publicaciones/Ecologia\\_espacial\\_book\\_Cap6.pdf](http://www.escet.urjc.es/biodiversos/espainvestigacion/intercambio/doc/publicaciones/Ecologia_espacial_book_Cap6.pdf)
- García-Asenjo, L., & Hernández, D. (2005). *Geodeisa*. Recuperado el 23 de Junio de 2015, de [http://sitopcar.es/modulos/descargas/manuales/Geodesia\\_Redex\\_Cart\\_Matematica.pdf](http://sitopcar.es/modulos/descargas/manuales/Geodesia_Redex_Cart_Matematica.pdf)
- Ginzo, M. J. (2011). *Análisis Geostadístico de datos funcionales*. Recuperado el 17 de Septiembre de 2015, de [http://eio.usc.es/pub/mte/descargas/proyectosfinmaster/proyecto\\_388.pdf](http://eio.usc.es/pub/mte/descargas/proyectosfinmaster/proyecto_388.pdf)

- Giraldo, R. (2005). *Introducción a la Geoestadística. Teoría y Aplicación*. Bogotá: Universidad Nacional de Colombia.
- INEI. (2006). *Glosario de términos estadísticos*. Recuperado el 5 de Diciembre de 2015, de [https://www.inei.gob.pe/media/MenuRecursivo/publicaciones\\_digitales/Est/Lib0900/Libro.pdf](https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib0900/Libro.pdf)
- Internacional, C. N. (1999). *Sistemas Geodésicos*. Bogotá: Subcomité de Geodesia.
- ITRF, I. T. (2015). *Transformation parameters*. Recuperado el 20 de Agosto de 2015, de [http://itrf.ensg.ign.fr/trans\\_para.php](http://itrf.ensg.ign.fr/trans_para.php)
- Larraz, B., & Montero, J. M. (2008). *Introducción a la geoestadística lineal*. La Coruna: Netbiblo.
- Leiva, C. (2003). *Determinación de parámetros de transformación entre los sistemas PSAD56 y WGS84 para el país*. Sangolquí: ESPE.
- 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: Escuela Politécnica Nacional.
- Martori, J. (2008). Nuevas técnicas de estadística espacial para la detección de clusters residenciales de población inmigrante. *Scripta Nova: Revista electrónica de geografía y ciencias sociales*, XII(263).
- Mateu, M., Valero, B., & Blanquer, G. (2006). Transformaciones de datum en el contexto de un sistema de información geográfica. 966-967.
- Meijerink, A., de Brouwer, H., Mannaerts, C., & Valenzuela, C. (1994). *Introduction to the use of geographic information systems for practical hydrology*. Twente: ITC.
- Meyer, P. (2009). *Probabilidad y Aplicaciones Estadísticas*. Washington: Addison-Wesley Iberoamericana.
- Moirano, J. (2000). *Universidad de La Plata*. Recuperado el 15 de Agosto de 2015, de [http://sedici.unlp.edu.ar/bitstream/handle/10915/1998/Documento\\_completo.pdf?sequence=13](http://sedici.unlp.edu.ar/bitstream/handle/10915/1998/Documento_completo.pdf?sequence=13)
- Moral, F. (2004). Aplicación de la Geoestadística en las ciencias ambientales. *Ecosistemas*, 78-86.
- Nel, L. (2010). *Estadística para ingenieros*. MACRO.

- Nelsi, C. d. (2006). *Elementos de Geodésia*. Sao Paulo: Universidad de Sao Paulo.
- Oliver, M. (2010). *Geostatistical applications for precision agriculture*. London: Springer.
- Ramírez, M., & Ortiz, D. (2003). *Estimación de los Parámetros de Transformación Entre los Sistemas de Referencia WGS-84 y PSAD-56 Para una Zona de Calama*. Santiago: Universidad de Santiago de Chile.
- Rebolledo, S. (2010). *Análisis de la significancia de modelos de transformación bidimensional y modelo de similaridad 3D para la compatibilización de bases cartográficas a SIRGAS*. Chile.
- Ribeiro, P., Christensen, O., & Diggle, P. (2003). *Distributed Statistical Computing*. Recuperado el 15 de Octubre de 2015, de <http://www.R-project.org/conferences/DSC-2003/>
- Romero, R. (2013). *Informe de la determinación de la metodología de transformación entre sistemas de referencia PSAD56 y distintas realizaciones del ITRF*. Quito: Instituto Geográfico Militar.
- Romero, R., & Tierra, A. (2014). *Metodología para la transformación entre sistemas de referencia utilizando una red neuronal artificial*. Recuperado el 16 de Junio de 2015, de [http://www.sirgas.org/fileadmin/docs/Boletines/Bol19/54\\_Romero\\_et\\_al\\_2014\\_Transformacion\\_redes\\_neuronales.pdf](http://www.sirgas.org/fileadmin/docs/Boletines/Bol19/54_Romero_et_al_2014_Transformacion_redes_neuronales.pdf)
- Santacruz, A. (2010). *Transformación entre el Sistema PSAD56 y los Marcos de Referencia ITRF utilizando los modelos de Helmert y de velocidades de placas tectónicas VEMOS*. Sangolqui: ESPE.
- Seeber, G. (2003). *Satellite Geodesy*. Berlin: Walter de Gruyter .
- SIRGAS. (2015). *SIRGAS*. Recuperado el 15 de Junio de 2015, de <http://www.sirgas.org/index.php?id=52>
- Thompson, D. (1976). *Combination of Geodetic Networks*. Fredericton: University of New Brunswick.
- Torge, W. (1991). *Geodesy*. New York: Walter de Gruyter.
- Villatoro, M., Henríquez, C., & Sancho, F. (2008). *Comparación de los interpoladores IDW y Kriging en la variación espacial de pH, Ca, CICE y P del SUELO*. Recuperado el 23 de Septiembre de 2015, de [http://www.mag.go.cr/rev\\_agr/v32n01-095.pdf](http://www.mag.go.cr/rev_agr/v32n01-095.pdf)

Walpole, R., Myers, R., Myers, S., & Ye, K. (2007). *Probabilidad y Estadística para ingeniería y ciencias* (Octava ed.). México.