

# UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERÍA ECONÓMICA ESTADÍSTICA Y CIENCIAS SOCIALES



## "PREDICCIÓN DE AMBIENTES DE DEPOSICIÓN SEDIMENTARIA EN BASE A CARACTERÍSTICAS BATIMÉTRICAS, MORFOLÓGICAS Y SEDIMENTARIAS EN EL MARGEN CONTINENTAL PERUANO MEDIANTE MODELOS DE ESTADÍSTICA ESPACIAL"

TESIS

PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO ESTADÍSTICO  
POR LA MODALIDAD DE TESIS

ELABORADO POR

**ERICK ALBACHARRO CHACON MONTALVAN**

Lima - Perú  
2014

**Digitalizado por:**

Consortio Digital del  
Conocimiento MebLatam,  
Hemisferio y Dalse

*A quien ha sido mi motivación, mi inspiración,  
mi soporte, mis ganas de ser mejor, con quien  
siempre estaré en deuda por todo lo que me ha  
dado, mi madre, Marilú.*

*Y a quien me ha brindado su apoyo, me ha  
enseñado mucho de la vida, y me acercó a la  
ciencia, mi padre y profesor, Alfredo.*

*Los amo con todas mis fuerzas, Marilú y Alfredo.*

# Agradecimientos

Mi sincero agradecimiento al IMARPE (Instituto del Mar del Perú) y al IRD (Institut de Recherche pour le Développement) por brindarme el apoyo para realizar mi tesis. A Dimitri Gutierrez y Arnould Bertrand, representantes de ambos institutos en la cooperación DISCOH (Dinámica del Sistema de la Corriente de Humboldt) en el periodo de la realización de este estudio, por su respaldo y colaboración.

Un agradecimiento particular a mi asesor Yann Tremblay; en primer lugar, por darme la oportunidad de trabajar con él; segundo, por el apoyo y orientación para el planteamiento de este proyecto; tercero, por su apoyo en la programación y sus incontables enseñanzas, y porque he aprendido mucho de él. A mi asesor Federico Velazco, por su orientación en los temas relacionados a la sedimentología, por los datos brindados, y por sus valiosas sugerencias. A Daniel Grados y Nicolas Bez por su apoyo en la teoría geoestadística y técnicas espaciales.

Los datos de batimetría fueron de suma utilidad para la elaboración del estudio, gracias por estos datos brindados a Laurence Audin, Rossio Garcia, Daniel Price y Gary Vargas. A investigadores como Paulo Justiniano, Liu Qiliang con quienes pude intercambiar correos para sugerencias sobre la estimación de parámetros para gran cantidad de datos y clustering espacial respectivamente. Y aquellas personas que hicieron un ambiente de trabajo ameno como Vilma, Rocío, Daniel, Giannina, Omar, Andrea, Gabin, Katisha, Marie, Cori y Remi.

Por otro lado, este trabajo no hubiera sido posible sin el apoyo de las maravillosas personas que están a mi alrededor. Mi familia, quienes siempre han confiado en mí y a quienes nunca pienso defraudar. Gracias madre mía, Marita, por haberme amado tanto y porque siempre puedo confiar en ti, cada partícula y molécula mía son muestra de tu cariño y amor. Gracias papá, Alfredo, porque he aprendido mucho de ti, por los videos matemáticos que nos hacías escuchar como el de las “cuchocientas” formas de demostrar el teorema de pitágoras y la música clásica, sabemos que aunque

el camino se ponga difícil, siempre hay soluciones.

Gracias a mis hermanas Edith y Sheyla que toda mi vida han sido grandes ejemplos. A Edith, que desde niño me hizo soñar, gracias a ti tengo grandes aspiraciones porque siempre mostrabas que lo que más deseabas era un mejor futuro para nosotros, tus hermanos menores. A mi hermana Sheyla, con quien comparto muchas cosas aparte del signo zodiacal y del club favorito, que con sus actos y coraje me da el aliento para seguir por mis sueños.

También, gracias a mis compañeros de todas mis locuras de niño y con quienes todavía compartimos nuestro día a día. A ustedes Luis y Álvaro, su preocupación infinita por mí me hace sentir respaldado y que nada me puede hacer caer. Luis, has sido mi ídolo desde que era niño por tu inteligencia, simpatía, responsabilidad; has sido un gran hermano por todo ese tiempo. A mi hermano que más me ha comprendido, probablemente por la cercanía de edades, Álvaro, te agradezco por ser mi hermano y amigo, por compartir tus sueños, por escuchar los míos y por tu alegría y carisma, que me contagia y sin la cual no podría tener un equilibrio en mi vida.

Hay personas que llegan a nuestras vidas en el momento adecuado, gracias Vilma por complementarme y hacer de mí una mejor persona. Quiero que cumplamos nuestros sueños y felicitarte por haber dado el primer paso. Solo nosotros sabemos lo mucho que nos ha costado. Así que, a seguir avanzando porque la vida es cuesta arriba.

Finalmente, hay una persona que me comprendió cuando nadie lo hacía, que me hizo despertar de mi mundo interior y me mostró lo que antes parecía no entender. Me motivó y guió en mi vida estudiantil, y aunque no todo nos salió bien, se que cumpliremos mucho de lo que nos planteamos. Gracias por todo ello José Carlos, eres el mejor amigo.

## Abstract

Biotic and abiotic organisms in the marine habitat erode particle fragments that are accumulated in strategic locations, called environments of deposition, characterized by the structure of the topography and the ocean dynamics. Thin sediments, useful to obtain high-quality sedimentary information, are generally accumulated in areas favoring the continuous deposition of sediments. These environments are characterized by the morphology of the marine topography and the behavior of certain external agents such as hydrodynamic energy, winds, marine currents, and so on. Knowledge of these areas offers significant contributions for the development of *in situ* geological work and it is the basis for research studies in the northern Humboldt Current System (HCS) in Peru, which is the marine ecosystem with the most extreme conditions and unknowns. Therefore, in order to predict the location of continuous deposition environments in Peruvian Continental Margin, the presence of continuous deposition environments are modeled spatially based on the morphology and bathymetry characteristics. In the study, high-resolution grids of bathymetry and likely continuous deposition environments were obtained. The logistic kriging model confirms the association between continuous deposition environments and morphological characteristics of the seabed such as depth, slope, distance to the coast, among others.

**Keywords:** Bathymetry, Continuous sedimentation environments, Geostatistics, Kriging, Morphology, Spatial Correlation.

## Resumen

Los organismos bióticos y abióticos del hábitat marino erosionan fragmentos de partículas que se van acumulando en lugares estratégicos (ambientes de deposición) caracterizados por la estructura del relieve marino y la dinámica existente en este. Los sedimentos finos, útiles para obtener información sedimentaria de calidad, generalmente se acumulan en zonas que favorecen la deposición continua de sedimentos. Estos ambientes son caracterizados por la morfología del relieve marino y el comportamiento de ciertos agentes externos como la energía hidrodinámica, los vientos, las corrientes marinas, entre otros. El conocimiento de estas zonas brinda aportes significativos para el desarrollo de trabajos geológicos *in situ* y como base para estudios de investigación en el Sistema de la Corriente de Humboldt (SCH) del Perú, el cual es el ecosistema marino con mayor cantidad de condiciones extremas e incógnitas. Por ello, con el objetivo de predecir la ubicación de probables ambientes de deposición continua en el margen continental peruano, se modela espacialmente la presencia de ambientes de deposición continua en base a las características de morfología y batimetría. En el estudio se obtuvo grillas de alta resolución de batimetría y de probables ambientes de deposición continua. El modelo del kriging logístico corrobora la asociación entre los ambientes de deposición continua y las características morfológicas del fondo marino tales como la profundidad, pendientes, lejanía a las costas entre otras.

**Palabras clave:** Ambientes de sedimentación continua, Batimetría, Correlación Espacial, Geostatística, Morfología, Kriging.

# Índice general

Carátula . . . . .	I
Dedicatoria . . . . .	II
Agradecimientos . . . . .	III
Resumen . . . . .	v
Lista de Cuadros . . . . .	IX
Lista de Figuras . . . . .	x
Introducción . . . . .	1
<b>1. Planteamiento del Problema . . . . .</b>	<b>5</b>
1.1. Antecedentes . . . . .	5
1.2. Descripción del Problema . . . . .	7
1.3. Formulación del Problema de Investigación . . . . .	8
1.4. Objetivos e Hipótesis . . . . .	9
1.4.1. Objetivos . . . . .	9
1.4.2. Hipótesis . . . . .	10
1.5. Limitaciones y Delimitaciones . . . . .	10
1.5.1. Limitaciones . . . . .	10
1.5.2. Delimitaciones . . . . .	10
1.6. Importancia y Justificación . . . . .	11
<b>2. Marco Teórico . . . . .</b>	<b>12</b>
2.1. El Ecosistema Marino . . . . .	12
2.1.1. Clasificación Elemental del Ambiente Marino . . . . .	13
2.1.2. Importancia del Ambiente Marino . . . . .	15
2.2. El Fondo Marino y sus Sedimentos . . . . .	16
2.2.1. El Fondo Marino . . . . .	17
2.2.2. Los Sedimentos del Fondo Marino . . . . .	18
2.2.3. La Escala de Análisis en el Fondo Marino . . . . .	22
2.3. Los Ambientes de Deposición Sedimentaria . . . . .	24
2.4. La Batimetría y Morfometría . . . . .	24
2.4.1. Batimetría del Fondo Marino . . . . .	24
2.4.2. Morfometría del Fondo Marino . . . . .	26
<b>3. Metodología . . . . .</b>	<b>27</b>
3.1. Diseño de Investigación . . . . .	27

3.1.1.	Tipo de Estudio . . . . .	27
3.1.2.	Área de Estudio . . . . .	27
3.1.3.	Muestra . . . . .	29
3.2.	Operacionalización de las Variables . . . . .	33
3.2.1.	Factor Batimétrico . . . . .	36
3.2.2.	Factor Morfológico . . . . .	36
3.2.3.	Factor Sedimentario . . . . .	39
3.3.	Recolección y Tratamiento de los datos . . . . .	41
3.3.1.	Recolección de Datos . . . . .	41
3.3.2.	Pre-procesamiento de los Datos . . . . .	42
3.3.3.	Procesamiento de los Datos . . . . .	42
3.4.	Metodología del Modelamiento . . . . .	48
3.4.1.	Modelamiento de la Batimetría . . . . .	49
3.4.2.	Modelamiento de la Detección de Ambientes Depositionales . . . . .	52
<b>4.</b>	<b>Técnicas de Análisis . . . . .</b>	<b>54</b>
4.1.	Geoestadística . . . . .	54
4.1.1.	Introducción Histórica . . . . .	54
4.1.2.	Descripción General . . . . .	55
4.1.3.	Terminología y Notaciones . . . . .	57
4.1.4.	Correlación Espacial . . . . .	61
4.1.5.	Estimación Gaussiana de Parámetros . . . . .	66
4.1.6.	Predicción Espacial . . . . .	68
4.2.	Geoestadística para un Procesos de Bernoulli: Modelo Logístico Espacial . . . . .	70
4.2.1.	Modelo Logístico: . . . . .	71
4.2.2.	Modelo Logístico Espacial: . . . . .	71
4.3.	Geoestadística para Datos de Fuentes con Distinta Confiabilidad: Una extensión de la función de correlación . . . . .	72
4.3.1.	Correlación Espacial . . . . .	75
4.3.2.	Estimación de Parámetros . . . . .	79
<b>5.</b>	<b>Resultados . . . . .</b>	<b>81</b>
5.1.	Predicción de la Batimetría del Mar Peruano . . . . .	81
5.1.1.	Características Generales del Mapa de Batimetría Obtenido . . . . .	81
5.1.2.	Comparación del Mapa de Batimetría Obtenido con la Batimetría del Proyecto SRTM30 . . . . .	85
5.1.3.	Comparación entre un Modelo que no Considera la Diferencia entre Fuentes de Datos y el Modelo Empleado . . . . .	90
5.2.	Identificación de Clusters Sedimentarios . . . . .	92
5.2.1.	Descripción General de la Distribución de las Características Sedimentarias en la Plataforma Continental . . . . .	93
5.2.2.	Clusters Encontrados Según los Perfiles de Sedimentología . . . . .	95
5.3.	Modelamiento de Presencia de Depocentros Sedimentarios . . . . .	97
5.3.1.	Variables Consideradas y Factores Incorporados al Modelo . . . . .	97
5.3.2.	Comportamiento Espacial . . . . .	106
5.4.	Predicción de la Ubicación de Depocentros Sedimentarios . . . . .	108
5.4.1.	Características Generales del Mapa de Probabilidades sobre la Presencia de Depocentros . . . . .	108
<b>6.</b>	<b>Discusión . . . . .</b>	<b>113</b>
6.1.	Acerca de la Batimetría . . . . .	113

6.1.1.	¿Qué se encontró? . . . . .	113
6.1.2.	Problemática del Trabajo de Modelamiento . . . . .	114
6.1.3.	Aporte del Mapa Obtenido y Posibles Mejoras en Adelante . . . . .	115
6.2.	Acerca de los Posibles Ambientes de Deposición Continua . . . . .	115
6.2.1.	¿Qué se encontró? . . . . .	116
6.2.2.	Problemática del Trabajo de Modelamiento . . . . .	116
6.2.3.	Aporte del Nuevo Mapa Obtenido . . . . .	117
<b>7.</b>	<b>Conclusiones . . . . .</b>	<b>118</b>
7.1.	De la Batimetría . . . . .	118
7.2.	De los Posibles Ambientes de Deposición Sedimentaria . . . . .	119
7.3.	De los Modelos Espaciales . . . . .	120
<b>Bibliografía . . . . .</b>		<b>121</b>

# Índice de cuadros

3.1. Matriz de consistencia (i) . . . . .	34
3.2. Matriz de consistencia . . . . .	35
3.3. Operacionalización de Variables de Batimetría . . . . .	36
3.4. Operacionalización de Variables de Morfología . . . . .	38
3.5. Operacionalización de Variables de Sedimentología . . . . .	40
5.1. Comparación del Mapa de Batimetría Obtenido con la Batimetría del Proyecto SRTM30	88
5.2. Variables incorporadas al modelo logit sin interacciones y su significancia . . . . .	98
5.3. Variables incorporadas al modelo probit sin interacciones y su significancia . . . . .	99
5.4. Variables incorporadas al modelo comloglog sin interacciones y su significancia . . . . .	100
5.5. Indicadores estadísticos de la bondad de ajuste . . . . .	100
5.6. Capacidad de predicción de los modelos . . . . .	101
5.7. Variables incorporadas al modelo logit con interacciones y su significancia . . . . .	102
5.8. Variables incorporadas al modelo probit con interacciones y su significancia . . . . .	103
5.9. Variables incorporadas al modelo comloglog con interacciones y su significancia . . . . .	104
5.10. Indicadores estadísticos de la bondad de ajuste de los modelos con interacciones . . . . .	105
5.11. Capacidad predictiva del modelo . . . . .	105

# Índice de figuras

2.1. Principales divisiones del ambiente marino. Fuente: Elements of Marine Ecology [Tait and Dipper, 1998] . . . . .	15
2.2. El Ecosistema Marino proporciona servicios ecológicos y económicas. Fuente: Living in the Environment [Miller and Spoolman, 2011] . . . . .	16
2.3. Clasificación del tamaño UddenWentworth de los sedimentos. Fuente: Exploring the World Ocean [Chamberlin and Dickey, 2007] . . . . .	19
2.4. Clasificación de los sedimentos según su ubicación. Fuente: The Sea Floor and its Sediments [Sverdrup et al., 2005] . . . . .	20
2.5. Ciclo del sedimento y procesos involucrados. Fuente: Ocean Sediments [Chamberlin and Dickey, 2007] . . . . .	22
2.6. Batimetría y Topografía de la Tierra. Fuente: NOAA . . . . .	25
2.7. Morfología del Fondo Marino. Fuente: Universidad de Puerto Rico - Departamento de Geología . . . . .	26
3.1. Área de Estudio para la Batimetría (Celeste) . . . . .	28
3.2. Área de Estudio para la Detección de Depocentros (Color beige) . . . . .	29
3.3. Distribución de las fuentes de datos batimétricos. . . . .	30
3.4. Distribución espacial de la muestra de batimetría según número de fuentes . . . . .	32
3.5. Distribución espacial de los puntos muestreados (color rojo) para obtener características sedimentarias (Composición de la textura y coeficiente de asimetría del tamaño de granos). . . . .	33
3.6. Escala de tamaño de grano de las partículas Fuente: A Scale of Grade and Class Terms for Clastic Sediments [Wenworth, 1922]. . . . .	39
3.7. Cantidad de Datos según Fuente: 185,849,891 registros . . . . .	43
3.8. Esquema del modelamiento de la batimetría . . . . .	51
3.9. Esquema del modelamiento general del estudio . . . . .	53
4.1. Ejemplo de Aplicación de Geoestadística. El signo de interrogación representa la ubicación en donde se desea estimar el valor de la profundidad ( $y_0$ ); los $y_i$ , para $i = 1, 2, 3, \dots, 10$ , representan los datos obtenidos del trabajo de campo. La degradación del color del fondo significa que a mayor distancia, existe menor similaridad entre el dato $y_i$ y el valor de la profundidad de interés ( $y_0$ ). . . . .	55
4.2. Correlación de Matern. . . . .	57
4.3. Proceso isotrópico (izquierda) y proceso anisotrópico (derecha). . . . .	59
4.4. Correlación para un proceso isotrópico (izquierda) y un proceso anisotrópico con $\theta = 135$ (derecha). Las flechas negras indican las direcciones principales. . . . .	60
4.5. Proceso anisotrópico (izquierda) y Proceso con coordenadas transformadas - Isotrópico (derecha) . . . . .	60
4.6. Comparación entre los datos espaciales de las fuentes DHN, SMTR30_PLUS e IMARPE. (a) Distribución espacial de las fuentes. (b-c) Profundidad de las superficies según latitud y longitud por fuente. (d) Profundidad de las superficies según longitud por fuente	74

4.7. Variograma Empírico por fuente de IMARPE y SRTM30_PLUS, y entre las dos fuentes. El tamaño de los círculos es proporcional a la cantidad de pares encontrados a la distancia $h$ . . . . .	76
4.8. Correlación Teórica por fuente de IMARPE y SRTM30_PLUS, y entre las dos fuentes. La escala de la distancia $h$ es referencial. (a) Correlación teórica cuando las varianzas, alcance y los efectos pepita de cada función es diferente. (b) Correlación teórica cuando los efectos pepita son distintos pero con alcances y varianzas iguales. . . . .	78
5.1. Predicción de la batimetría a alta resolución (100 metros, exageración vertical x 20).	82
5.2. Error de predicción de la batimetría a alta resolución (100 metros, exageración vertical x 20). . . . .	83
5.3. Predicción de la batimetría a baja resolución (900 metros, exageración vertical x 20).	84
5.4. Error de predicción de la batimetría a baja resolución (900 metros, exageración vertical x 20). . . . .	85
5.5. Comparación a alta resolución entre el mapa de batimetría de SRTM30_3PLUS y el mapa obtenido ( exageración vertical x 20). . . . .	87
5.6. Comparación a baja resolución entre el mapa de batimetría de SRTM30_PLUS y el mapa obtenido (exageración vertical x 20). . . . .	89
5.7. Variograma Empírico para la fuente de IMARPE, SRTM30_PLUS y DHN, y entre IMARPE con cada una de las dos fuentes. El tamaño de los círculos es proporcional a la cantidad de pares encontrados a una distancia $h$ . . . . .	91
5.8. Kriging sin considerar nuestro modelo de correlación. . . . .	92
5.9. Kriging considerando nuestro modelo de correlación. . . . .	92
5.10. Distribución espacial del tamaño de grano según el coeficiente de asimetría. Valores negativos indican granos finos y valores positivos representan granos más gruesos. . .	94
5.11. Cluster de los perfiles sedimentarios . . . . .	95
5.12. Distribución espacial de los clusters de perfiles sedimentarios . . . . .	96
5.13. Características sedimentarias según cada cluster. . . . .	97
5.14. Variograma de los residuales del modelo logístico junto a los modelos ajustados. $\tau^2$ representa el efecto pepita de cada modelo . . . . .	107
5.15. Predicción de probables zonas de deposición continua en la Plataforma Continental y la zona superior del Talud Continental. . . . .	109
5.16. Predicción de probables zonas de deposición continua en la zona norte de la Plataforma Continental y la zona superior del Talud Continental. . . . .	110
5.17. Predicción de probables zonas de deposición continua en la zona sur de la Plataforma Continental y la zona superior del Talud Continental. . . . .	111
5.18. Predicción de probables zonas de deposición continua en la zona sur de la Plataforma Continental y la zona superior del Talud Continental a mayor resolución . . . . .	112

# Introducción

El ecosistema marino está conformado por una parte abiótica constituida por el sustrato, la topografía del fondo marino, dirección de las olas, salinidad, corrientes marinas, entre otras propiedades del agua y una parte biótica que abarca a los macrobentos, hiperbentos, epibentos, peces demersales, peces pelágicos y depredadores superiores [Verfaillie et al., 2006]. El sustrato está conformado por formaciones rocosas de diferente extensión expuestas sobre el fondo marino o por sedimentos consolidados o no consolidados (partículas de diferentes origen, composición y tamaño).

La dinámica de interacción entre los componentes del ecosistema deja rastros de lo ocurrido en forma de fragmentos de diferente origen. Estos fragmentos de diferente origen pueden ser: partículas de minerales erosionadas a rocas o sedimentos expuestos en el fondo marino o provenientes del continente e ingresadas a través de ríos, vientos, entre otros, al océano. Entre ellas; partículas de origen *biogénico* como fragmentos de conchas de moluscos, esqueletos de peces, de algas, etc. ó finalmente, de origen *autigénico* que se producen en el mismo océano como la fosforita que se ha formado por precipitación concéntrica a partir de otras partículas, los nódulos polimetálicos, el carbonato de calcio que ha precipitado y posteriormente transportadas.

Los fragmentos erosionados no permanecen fijos en el lugar donde se dio el evento, necesariamente, sino que son transportados por agentes externos que gobiernan la dinámica del ecosistema; entre ellos identificamos al viento, olas y corrientes de agua. En este contexto, la morfología (expresada por la variabilidad topográfica del fondo marino) juega un rol importante porque este puede favorecer o no a que los fragmentos se acumulen en ciertas zonas del fondo marino. Por lo tanto, ya sea al inicio de la erosión, en el proceso de transporte o cuando la energía de los agentes externos se agote, estos fragmentos se acumulan en el fondo marino y luego se compactan por el efecto de la alta presión por millones de años formando sedimentos que almacenan la información histórica

a diferentes escalas del ecosistema y clima pasado. Por tal motivo, los sedimentos son calificados como la memoria del océano. La sedimentación ocurre en ciertos lugares que presentan condiciones favorables a este proceso, estos lugares son denominados ambientes de deposición sedimentaria.

Existen distintos tipos de ambientes de deposición según las características de los procesos físicos, químicos y biológicos ocurridos en el ecosistema marino, los cuales pueden verse explicados, en parte, por la dinámica de los agentes externos y los rasgos morfométricos de la topografía marina. En el presente estudio, estamos enfocados en los **ambientes actuales/recientes de deposición sedimentaria continua o depocentros sedimentarios** porque en ellos se encuentran sedimentos finos de alta calidad para el análisis geológico, paleoceanográfico y estratigráfico así como de caracterización de cierto tipo de sustrato que habitan determinados organismos bentónicos. Estas zonas se caracterizan por poseer sedimentos con tamaño granulométrico pequeño y por estar compuesto principalmente de arcilla y limo.

La recolección de información sedimentaria en estos ambientes de deposición es muy útil porque revelan los procesos que forman playas, el perfil de las líneas costeras, cañones submarinos y controlan la abundancia de los organismos marinos. De hecho, los sedimentos revelan una historia de los océanos sobre un amplio rango de escalas temporales y espaciales [Chamberlin and Dickey, 2007]. Aún más útil, en la Plataforma Continental Peruana que forma parte del Sistema de la Corriente de Humboldt (SCH), el cual es un ecosistema de embergadura por poseer características valiables como la riqueza de nutrientes debido al fenómeno de surgencia (upwelling<sup>1</sup>), la alta productividad por espacio marítimo, la posesión de zonas de mínimo oxígeno que modifican el comportamiento de sus seres vivos, **la acumulación de sedimentos marinos que crea un registro histórico del ecosistema** y el hecho de que su comportamiento está sujeto a fluctuaciones multidecadales o interanuales debido a su enlace con la Dinámica del Pacífico Ecuatorial [Chavez et al., 2008]. Sin embargo existe una limitante para la ubicación de estas zonas que no ha sido abordada directamente, solo parcialmente a escalas de ciertas regiones del margen continental [Delgado and Gomero, 1995, Herbozo et al., 2006, Velazco, 2012] o muy locales como el caso de bahías [Velazco, 2001, Herbozo et al., 2006]. Por lo cual el presente estudio es de sumo valor para los trabajos sedimentarios “in situ” en términos de costos logísticos y de resultados científicos. Además ello tiene impacto en los resultados de estudios analíticos del bioecosistema marino del SCH debido a los datos de entrada.

---

<sup>1</sup>Fenómeno provocado por vientos que desplazan las aguas de la superficie marina hacia zonas más profundas. Ello provoca la surgencia de aguas frías ricas en nutrientes.

Diversos estudios se han realizado con el fin de identificar depocentros o cuencas sedimentarias en los cuales se muestra el empleo de modelos gravitatorios y técnicas sismográficas para la identificación de cuencas sedimentarias [Tatham, 1975, Bell et al., 2006]. Por medio de estas técnicas se han identificado cuencas sedimentarias a gran escala. Generalmente las técnicas empleadas tienen la finalidad de estimar la profundidad del fondo marino ya que es parte importante de identificar las cuencas sedimentarias. Sin embargo, estos métodos presentan limitaciones bajo pequeñas escalas ya que son sensibles a ciertos materiales.

Por otro lado, las variables involucradas en la caracterización de ambientes de deposición respecto a los factores batimétricos, morfológicos y sedimentarios tienen un comportamiento espacial que no se puede dejar de lado; por lo que, procedimientos y modelos estadísticos espaciales están teniendo gran acogida en investigaciones de sedimentología y en general, en estudios de ecología. Recientes estudios sobre la caracterización de los sedimentos del fondo marino en relación a su morfología y batimetría muestran resultados satisfactorios gracias al empleo de modelos estadísticos espaciales [Lark et al., 2012, Meilianda, 2012, Meilianda et al., 2011, Becker et al., 2009, Verfaillie et al., 2006].

Por lo tanto, el presente estudio se realiza con el objetivo de detectar zonas altamente probables de ser ambientes de deposición continua mediante modelos geoestadísticos. Para ello, primero, pre-procesamos, procesamos y modelamos datos batimétricos de distintas fuentes de datos obteniendo un mapa batimétrico de alta resolución en la área cubierta por el rango de latitud [-20.7,-1.3] y el rango de longitud [-86.5,-69.8], y un mapa del error de predicción. Luego de obtener el mapa de batimetría, identificamos clusters de zonas de sedimentación en base a resultados de análisis sedimentológicos de muestras obtenidas en campo (mediante dragado). Tercero, cuantificamos las características morfológicas del relieve marino. Finalmente, modelamos espacialmente la presencia de zonas de deposición continua según las variables de morfometría (variables independientes) y las variables sedimentarias (variables dependientes).

La organización de la presente tesis es de la siguiente manera: en el capítulo I, se plantea los componentes que conforman el planteamiento del problema. El marco teórico acerca del ecosistema marino, sus sedimentos y los ambientes de deposición se muestra en el capítulo II. La metodología de todo el estudio desde la fuentes de los datos hasta la predicción de los ambientes de deposición continua es presentada en el capítulo III. Las técnicas estadísticas espaciales para el análisis y modelamiento se establece en el capítulo IV. Los resultados pueden ser vistos en el capítulo V y la

discusión se plasma en el capítulo VI. Finalmente la conclusión de los principales aportes del estudio se presentan en el capítulo VII.

# Capítulo 1

## Planteamiento del Problema

Las partículas de distinto origen que se incorporan al fondo marino a través de procesos de sedimentación pueden conformar depósitos cuya acumulación continua es controlada tanto por las características topográficas del área como de la dinámica marina existente, entre otras variables. Luego de un largo proceso de sedimentación se conforman los sedimentos, de los cuales se puede obtener información útil acerca de las características del ecosistema y clima pasado [Chamberlin and Dickey, 2007]. El conocimiento de estas zonas brinda aportes significativos para el desarrollo trabajos geológicos *in situ* y como base para estudios de investigación en el Sistema de la Corriente de Humboldt (HCS) del Perú, el cual es el ecosistema marino con mayor cantidad de condiciones extremas e incógnitas. Sin embargo, no se tiene un conocimiento muy preciso de la ubicación y distribución de estas zonas. En este capítulo se define y delimita el problema bajo estudio.

### 1.1. Antecedentes

El interés de identificar depocentros sedimentarios ha sido común a través de los años. Tatham inicialmente propuso analizar la dispersión de las ondas superficiales para detectar depocentros sedimentarios debido a que las velocidades de la onda sísmica superficial es afectada por la estructura de la corteza ya que hay un fuerte contraste entre las propiedades físicas de los sedimentos clásticos y los materiales de base adyacentes [Tatham, 1975]. Esta técnica tiene la ventaja de proporcionar un modo conveniente, rápido y relativamente barato para detectar las cuencas sedimentarias y estimar su espesor. Sin embargo, sus principales desventajas son la relativa baja resolución,

la distribución de las estaciones sísmicas alrededor del mundo y la distribución de los epicentros de terremotos, por ello Tatham lo propuso como una herramienta útil para examinar variaciones en velocidad respecto a la profundidad y en estimar la profundidad de la base en cuencas sedimentarias sospechosas. Aunque los inicios de esta técnica se remota hace varios años atrás, es una de las más empleadas hasta hoy en la actualidad.

Una alternativa muy conveniente es analizar los registros sísmicos de ecosondas de muy baja frecuencia (baja penetración y muy alta resolución) pero su realización es muy costosa y no suelen ser disponibles en investigaciones científicas convencionales en nuestro medio. Usualmente se aplican en zonas ya definidas como objetivo con el fin de afinar la búsqueda de depocentros con buena calidad de registros.

Otra posible solución es integrar bases de datos acústicas en las que se identifiquen y clasifiquen los tipos de fondo teniendo como punto de partida los análisis sedimentológicos y un software de clasificación acústica de fondos marinos de alto costo a los que hay que integrar además las condiciones hidrodinámicas topográficas y geoquímicas; una primera aproximación ha sido realizada por [Herbozo et al., 2006] para la Plataforma y Talud Continental superior frente a Callao y Pisco, empleando datos de retrodispersión acústica de ecosondas científicas (fuerza del eco absorbida y reflejada por los sedimentos), relacionadas con la topografía; otra aproximación la constituyen la integración de bases de datos sedimentológicas y geoquímicas [Velazco, 2012], ambos casos requieren la integración de la topografía del fondo y de Sistemas de información Geográfica (SIG).

Los modelos gravimétricos también han sido empleados con el mismo interés. Bell identificó tres principales cuencas sedimentarias debajo de la capa de hielo de la Antártida occidental empleando técnicas de aerogravedad y deconvolución de Werner en datos magnéticos [Bell et al., 2006]. Parte importante de este tipos de estudios es tratar de medir la profundidad del fondo marino en base a la velocidad superficial o a los datos de gravimetría como Becker et al., quienes realizaron un mapa de batimetría/topografía de todo el mundo [Becker et al., 2009].

Para el presente estudio a una escala de amplia cobertura (todo el margen continental Peruano) se generará un mapa de batimetría en base a datos de cruceros científicos. Una gran variedad de métodos pueden ser empleados para generar un mapa batimétrico con datos irregularmente espaciados, no obstante se ha demostrado que la Geoestadística es uno de los mejores para este caso en especial [Pineda and Stefanoni, 2007]. Además la aplicación de la Geoestadística ha tenido un gran crecimiento y es usado en una amplia gamma de ramas científicas. En el análisis de sedimentología,

hay grandes avances para entender y estimar la estructura espacial de las variables sedimentarias tales como la composición de la textura sedimentaria del fondo marino y el tamaño del grano de los sedimentos empleando el kriging ordinario y variaciones de este método como el cokriging, el kriging con variables externas o el kriging para una variable respuesta categórica.

Lark et al. realizaron una predicción espacial de las clases de textura del fondo marino por medio del método cokriging desde una base de datos de observaciones puntuales [Lark et al., 2012]. En ella se empleó el cokriging para la variable respuesta categórica clase de textura (arena, fango y grava) en la Plataforma Continental del Reino Unido y se obtuvo resultados consistentes con los mapeos convencionales de esta variable. En otros estudios se analizó la distribución del tamaño del grano [Verfaillie et al., 2006, Meilianda et al., 2011, Meilianda, 2012] encontrándose asociaciones interesantes entre el tamaño del grano y la profundidad del fondo marino. Además que se mostró la mejora de la predicción del tamaño del grano al incorporar la batimetría como variable predictora. Verfaillie et al. emplearon estadística multivariada para el modelamiento predictivo de la distribución de arena superficial en la costa marina a través del tamaño del grano [Verfaillie et al., 2006]. Meilianda et al. analizaron la distribución del tamaño de grano de sedimento en la Plataforma arenosa poco profunda del Mar del Norte usando geoestadística multivariada [Meilianda et al., 2011]. Meilianda aplicó geoestadística multivariada para investigar los sedimentos de la superficie de la Plataforma poco profunda y de alta energía en la isla de la barrera Spiekeroog en la ensenada alemana del Mar del Norte meridional [Meilianda, 2012]. En estos tres estudios se mostró la gran ventaja de utilizar a la batimetría como covariable.

En los tres estudios mencionados anteriormente también se muestra que la caracterización de estas variables es variante a escalas pequeñas y por lo tanto, su asociación con la batimetría también se da a este nivel [Verfaillie et al., 2006, Meilianda et al., 2011, Meilianda, 2012]. Ello involucra la necesidad de obtener datos batimétricos de alta resolución para su análisis conjunto con características sedimentarias. Por lo tanto, en el presente estudio se requiere obtener una grilla batimétrica de alta resolución para utilizarla en la predicción de ambientes de deposición continua.

## 1.2. Descripción del Problema

Para fortalecer y aumentar los conocimientos de la Dinámica del HCS, se necesita conocer la ubicación de ambientes de deposición continua, en donde existe mayor sedimentación (depocentros

sedimentarios) para obtener información geológica que pueden brindar aportes sustanciales en estudios de geología, paleoceanografía, estratigrafía y del impacto de los sedimentos en el Bioecosistema de la Corriente de Humboldt. Actualmente se puede identificar estas zonas con datos de trabajos de campo y con relativa subjetividad basada en la experiencia. Sin embargo, ello puede conducir a obtener resultados no esperados en los trabajos *in situ* en relación al tiempo, costo, área barrida y calidad de los datos; además de limitar el alcance de los estudios relacionados a esta área. Por lo tanto, se requiere de un estudio que brinde estimación de estas zonas de mayor sedimentación continua para que apoyen directamente a la mejora de trabajos de campo, al conocimiento científico y a estudios que necesitan de esta información como base.

### 1.3. Formulación del Problema de Investigación

Los ambientes de deposición sedimentaria están definidos por un conjunto de características del hábitat del fondo marino que surgen de los procesos físicos, químicos y biológicos que ocurren en él. Entre ellas podemos encontrar la batimetría, rasgos morfológicos (Pendiente, rugosidad, localización, dimensión, concavidad, variabilidad, etc), la hidrodinámica (Dinámica de las masas de agua y la energía producida) y el comportamiento de los vientos. Por ello la identificación de estas zonas debe involucrar estos factores, en su mayoría. Por otro lado, la variable respuesta es un evento bernoulli, con espacio muestral {sí, no}, que debe ser creada a partir de las características sedimentarias, identificando el perfil de sedimentología deseado.

Además, el modelamiento espacial de la presencia del depocentro considerando este conjunto de variables no es de manera directa, sino que se debe estimar espacialmente cada una de ellas desde la batimetría hasta los rasgos de morfometría del fondo marino. Por tal motivo, ha de considerarse, en primer lugar, la estimación espacial de estas variables para luego realizar un análisis conjunto con la variable dependiente (probabilidad de ser depocentro). De este modo, el estudio abarca los siguientes problemas específicos:

1. No se tiene información a alta resolución que describa las características del fondo marino peruano. Lo cual es necesario para identificar características morfológicas relacionadas a los ambientes de sedimentación.
2. Se necesita conocer las características sedimentarias de los ambientes de deposición continua.

3. No se cuenta con índices morfológicos del fondo marino, así que es necesario definir indicadores de morfología y calcularlos de manera cuantitativa en base a un mapa de batimetría de alta resolución.
4. Teóricamente, se conoce que hay asociaciones entre los rasgos de morfología y los ambientes de deposición. Sin embargo, al momento, no se ha cuantificado y definido la significancia de estos rasgos morfológicos en el área del margen continental peruano.
5. Finalmente, el problema específico final es que existen grandes áreas en las que no conocemos si hay presencia de ambientes de deposición continua, los cuales son importantes para obtener información de calidad de los procesos sedimentarios.

## 1.4. Objetivos e Hipótesis

### 1.4.1. Objetivos

#### 1. Objetivo Principal

- Predecir la ubicación de probables ambientes de deposición continua en el margen continental peruano.

#### 2. Objetivos Específicos

- Modelar la distribución espacial de la profundidad del fondo marino a una alta resolución para obtener rasgos del fondo marino más detallados.
- Identificar el perfil sedimentario del ambiente de deposición continua mediante clusters.
- Determinar índices morfológicos (Pendiente, rugosidad, localización, dimensión, concavidad, entre otros) del fondo marino que sirvan como variables independientes para la detección de depocentros en el relieve marino.
- Determinar un modelo espacial para la detección de ambientes de deposición continua en el relieve marino considerando las características morfológicas como covariables.
- Detectar probables ambientes de deposición continua con el perfil de sedimentología identificado.

## 1.4.2. Hipótesis

### 1. Hipótesis Principal:

- La ubicación de los ambientes de deposición continua puede ser identificada mediante el tamaño de grano y la composición de la textura del sedimento, y modelada a través de factores batimétricos y morfológicos.

### 2. Hipótesis Secundarias:

- Existen datos de fuentes locales (Perú) que son más confiables en comparación a grillas batimétricas de baja resolución que se emplean a nivel internacional.
- Los tipos de ambientes de deposición son caracterizados según el tamaño de grano y la composición de la textura del sedimento.
- Se puede obtener indicadores morfológicos a partir del mapa de batimetría.
- La presencia de depocentros es explicada por una tendencia en función a los índices de batimetría y morfometría y por un componente espacial.
- Se puede detectar probables ambientes de deposición en base al modelo espacial obtenido, incluso en zonas donde se tiene menor información de los perfiles sedimentarios.

## 1.5. Limitaciones y Delimitaciones

### 1.5.1. Limitaciones

Las características sedimentarias bajo estudio no solo dependen de la geometría morfológica del fondo marino sino también de otros factores que influyen en los procesos de erosión, transporte y sedimentación tales como el viento, las masas de agua, entre otros. Sin embargo, en el presente estudio se tiene información limitada de la hidrodinámica, así que solo se considera dichas características por la disponibilidad de los datos y porque en otros trabajos se obtuvieron resultados satisfactorios [Meilianda et al., 2011, Meilianda, 2012].

### 1.5.2. Delimitaciones

Los resultados eventualmente pueden verse afectados (limitados en precisión) por ciertas características de los datos de sedimentología: (i) La existencia de datos de sedimentos antiguos no

discriminados (Ejm. sedimentos finos no muy consolidados que están siendo actualmente erosionados y que han sido muestreados y analizados).(ii) Por presencia de sedimentos autigénicos cuya formación es en ambientes anóxicos de Plataforma y Talud en zonas de alta productividad (como en el Margen Continental Peruano) pero su tamaño no refleja necesariamente procesos hidrodinámicos, salvo cuando han sido posteriormente transportados y seleccionados desde su lugar de deformación (estos minerales crecen por precipitación concéntrica en procesos de miles a millones de años formándose costras, arenas o gravas), lo cual puede enmascarar procesos.

## 1.6. Importancia y Justificación

El conocimiento de zonas de depocentro de acumulación de sedimento continuo permite mejorar los estudios geológicos y en todo bioecosistema marino porque las características geológicas influyen al comportamiento de todos los elementos de este sistema, tanto de manera directa como indirecta, empezando con los organismos bentónicos hasta el comportamiento de los animales y el esfuerzo de pesca de los humanos.

Particularmente, contribuye en mejorar conocimiento aplicado a evaluación de recursos no vivos del fondo marino y para la evaluación de efectos ambientales de su extracción. Además, contribuye en mejorar conocimiento de tipos de fondos para el desarrollo de la investigación de la ecología de organismos bentónicos y demersales así como a sus actividades de extracción.

Del mismo modo la visualización y obtención de información cuantitativa respecto al hábitat del fondo marino, tales como características de sedimentación y batimetría, en mapas de alta resolución implica un aporte sustancial para el desarrollo de estudios en el bioecosistema marino. Las características sedimentarias nos permiten analizar la paleoceanografía del relieve marino para comprender los bioecosistemas y climas pasados y estimar los recursos orgánicos e inorgánicos [Blondel, 2012], mientras que la batimetría nos permite incorporar valiosa información a una variedad de estudios que se realizan en el marco del HCS tales como la predicción de biomasa, análisis de la sedimentación, análisis de los recursos bentónicos, etc. e incluso puede aportar a estudios de análisis de eventos como tsunamis. [Mofjeld et al., 2004]

De esta manera, se reafirma y muestra, como en la Introducción, que la información espacial resultante respecto a la batimetría y sedimentología serán de ayuda significativa para el conocimiento científico, gestión y preservación de los recursos del mar peruano.

# Capítulo 2

## Marco Teórico

### 2.1. El Ecosistema Marino

Los organismos existen solamente como partes de una entidad compleja hecha de interacciones de elementos inorgánicos y bióticos, al cual denominamos ecosistema. Ellos no pueden vivir como unidades aisladas y las actividades que comprenden sus vidas son dependientes y controladas por las circunstancias externas según las condiciones físicas o químicas, en los cuales ellos viven y las poblaciones de otros organismos con los que interactúan [Tait and Dipper, 1998]. En la tierra, el movimiento de la materia y energía en los ecosistemas operan localmente; pero en el mar, la naturaleza móvil de los ecosistemas marinos, y la gran escala del volumen del océano, hacen mayor la distancia de transporte de los componentes del ecosistema. Los hábitats en la tierra son fáciles de graficar, porque se puede caminar entre ellos y tocarlos. En el mar, la mayoría son microscópicos, y todo recae fuera de la vista, debajo de las olas, en un espacio 3D tan grande que la apreciación es mucho más difícil [Cundy and Kershaw, 2013].

Los ecosistemas marinos están entre los ecosistemas acuáticos más grandes de la Tierra. Ellos incluyen los océanos, la marisma salina y la ecología intermareal, estuarios y lagunas, manglares y arrecifes de coral, el mar profundo y el fondo marino. Ellos pueden ser contrastados con los ecosistemas de agua fresca, los cuales tienen menor contenido de sal [Cunningham and Cunningham, 2013]. En el océano, la parte abiótica de este ecosistema está constituida por el sustrato, la topografía del fondo marino, dirección de las olas, salinidad, corrientes marinas, entre otras propiedades del agua y la parte biótica abarca a los macrobentos, hiperbentos, epibentos, peces demersales, peces

pelágicos y depredadores superiores [Verfaillie et al., 2006].

Particularmente, el océano es el ecosistema más grande de la Tierra. Cubre más del 70 % de la superficie terrestre. Contiene casi el 99 % del espacio de vida en la Tierra, y casi toda el agua del planeta. Es hogar del animal más grande del mundo, el volcán mas grande, el cañón más profundo y la más grande cascada - todo bajo el agua. Y aún así, es el ambiente menos explorado del planeta [Berkenkamp, 2009]. Aunque el océano es un sistema de agua enorme que fluye sobre todo el globo terrestre, se dividen en 5: Pacífico, Atlántico, Índico, Ártico y Antártico. El océano Pacífico es el más grande del mundo. Cubre más de la tercera parte del planeta entero, y es tan grande que todos los continentes y casi todos los otros océanos podrían caber dentro de este [Berkenkamp, 2009]. Además el Océano Pacífico acoge al Sistema de la Corriente de Humboldt (SCH), el cual es un ecosistema de embergadura por poseer características valiables como la riqueza de nutrientes debido al fenómeno de surgencia (upwelling<sup>1</sup>), la alta productividad por espacio marítimo, la posesión de zonas de mínimo oxígeno que modifican el comportamiento de sus seres vivos, la acumulación de sedimentos marinos que crea un registro histórico del ecosistema y el hecho de que su comportamiento está sujeto a fluctuaciones multidecadales o interanuales debido a su enlace con la Dinámica del Pacífico Ecuatorial [Chavez et al., 2008].

### 2.1.1. Clasificación Elemental del Ambiente Marino

Aunque el efecto mixto de la circulación oceánica asegura mayor diversidad, pero pequeña a través de grandes cantidades de agua, hay, sin embargo, algunas diferencias de mayor contraste entre las diferentes partes del mar. Una capa fría, oscura, lenta y hondo del océano profundo es obviamente muy diferente de una capa bien iluminada, sacudida por olas de la superficie marina, o fuertes corrientes y fluctuaciones de temperatura y salinidad que ocurren cerca a la costa. Por ello se tiene clasificaciones de subdivisiones del ambiente marino (Fig. 2.1), el cual toma en cuenta las diferentes condiciones de vida en diferentes partes de los océanos [Tait and Dipper, 1998].

En general, hay dos formas en las que los organismos viven en el mar; flotan o nadan en el agua, o habitan en el fondo marino. Según ello, hay dos grandes divisiones del ambiente marino, la Pelágica y la Béntica respectivamente; la División Pelágica corresponde al cuerpo entero de agua formando los mares y océanos mientras que la división Béntica, a todo el fondo marino (Fig. 2.1) [Tait

---

<sup>1</sup>Fenómeno provocado por vientos que desplazan las aguas de la superficie marina hacia zonas más profundas. Ello provoca la surgencia de aguas frías ricas en nutrientes.

and Dipper, 1998].

En aguas poco profundas hay usualmente mayor movimiento y variaciones de composición y temperatura que en zonas más profundas. Por lo tanto, se puede subdividir la División Pelágica en (a), la Zona Nerítica, las aguas poco profundas sobre la Plataforma Continental, y (b), la Zona Oceánica, aguas más profundas detrás del margen continental [Tait and Dipper, 1998].

Los oceanógrafos dividen las aguas oceánicas en capas separadas, llamadas zonas, desde la superficie marina hasta el fondo del mar. Se conoce más acerca de las capas más superficiales porque se nos facilita explorarlas y al mismo tiempo es en donde más vida puede encontrarse. A continuación mostramos la clasificación de las capas del Océano [Berkenkamp, 2009]:

(a) Epipelágico:

También llamada la zona eufótica, cubre desde la superficie hasta los 200 metros de profundidad. Es la capa donde llega la mayor iluminación del sol y es de lejos en donde más vida se encuentra. La vida en esta zona abarca desde los pequeños fitoplancton a el enorme tiburón ballena, el pez más grande en el océano. Esta capa no es considerada parte del fondo marino.

(b) Mesopelágico:

El fondo marino es considerado a partir de esta capa, en donde la corteza se quiebra. Es la zona más alta del fondo marino y comprende desde 200 hasta 1000 metros de profundidad. Esta zona es la capa oceánica con la mayor diferencia en temperatura. La comida es escasa debido a que no hay suficiente luz para que las plantas se desarrollen. Las criaturas que viven en esta zona están en continua competencia para su supervivencia por eso cuentan con distintos mecanismos de defensa como la bioluminiscencia.

(c) Batipelágico:

Esta zona es la parte donde más del 75 % del agua de los océanos recae. Abarca de los 1000 hasta los 4000 metros. Esta capa es denominada zona afótica por la absoluta falta de luz natural y Batipelágico por aguas profundas.

(d) Abisopelágico:

Comprende entre los 4000 y 6000 metros de profundidad. Su nombre significa sin fondo pero, en realidad, es donde la mayor parte del fondo marino se sienta. A menudo, es conocido como llano abisal porque la mayor parte del fondo marino es amplio y casi completamente plano.

(e) Zona Hadal:

Es absolutamente la parte más profunda del océano. Hadal significa *no visto* y es donde se encuentran las trincheras del océano profundo. Tiene un rango de profundidad de 6000 a 11000 metros.

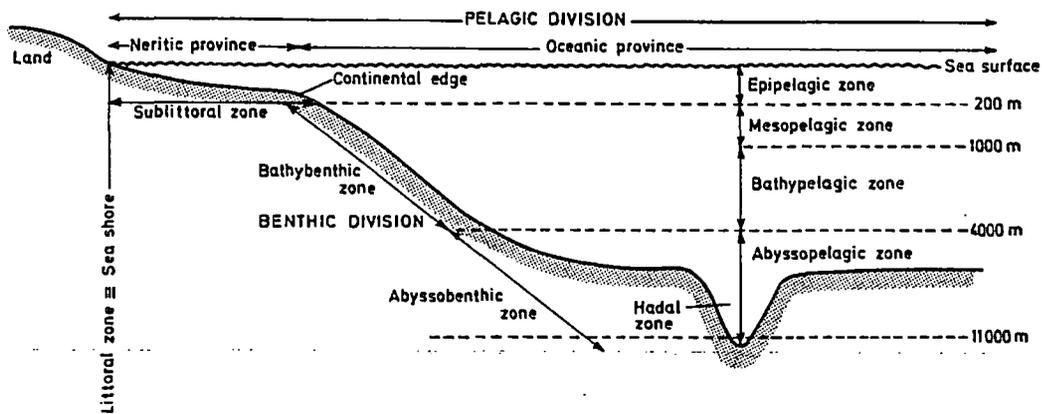


Figura 2.1: Principales divisiones del ambiente marino.  
Fuente: Elements of Marine Ecology [Tait and Dipper, 1998]

Según este criterio se puede categorizar a los peces de aguas profundas en grupos de mesopelágicos, batipelágicos y bentopelágicos. Los peces mesopelágicos viven por debajo de la zona fótica cerca de los 1000 metros de profundidad. Peces batipelágicos viven por debajo de los 1000 metros de profundidad y se adaptan a un ambiente muy pobre en alimento. Finalmente, las especies bentopelágicos viven muy cerca o en el fondo marino, esta zona es diferenciada por sus características biológicas [FAO, 2005].

### 2.1.2. Importancia del Ambiente Marino

El ecosistema marino nos brinda valiosos servicios ecológicos y económicos a la sociedad (Fig. 2.2). Entre estos servicios ecológicos tenemos la moderación del clima, la absorción de  $CO_2$ , ciclo y producción de nutrientes, tratamiento de residuos, reduce el impacto de las tormentas, habitats y áreas de crianza, recursos de genética y biodiversidad e información científica. Por otro lado, podemos encontrar servicios económicos como vital fuente de recursos alimenticios, comida para animales, fármacos, rutas de transporte, habitats costeros para los humanos, recreación, empleo, recursos naturales como gas y petróleo, minerales y materiales de construcción [Miller and Spoolman,

2011]. Por lo tanto es un ecosistema de vital importancia debido a sus dimensiones y al aporte vital que da a la sociedad. Sin embargo, como se mencionaba anteriormente, es uno de los ecosistemas menos explorado, en parte a su difícil acceso, y aún hay muchas incógnitas sobre este maravilloso ecosistema.

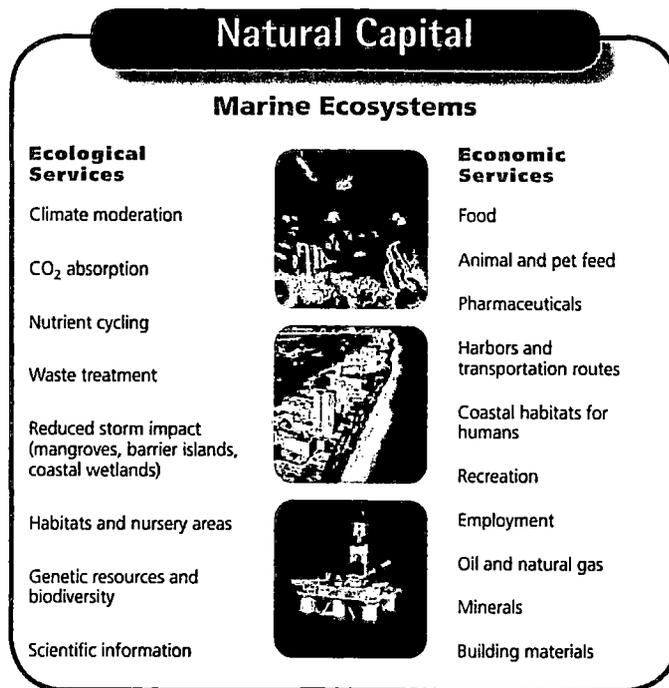


Figura 2.2: El Ecosistema Marino proporciona servicios ecológicos y económicas.  
Fuente: Living in the Environment [Miller and Spoolman, 2011]

## 2.2. El Fondo Marino y sus Sedimentos

La dinámica de la interacción entre los seres bióticos y los agentes externos de un ecosistema marino dejan rastros de lo pasado al erosionar fragmentos de los cuerpos bióticos o inertes. Estos fragmentos pueden ser, por mencionar, partículas de minerales erosionadas de rocas continentales, conchas, esqueletos de animales marinos y algas.

Las partículas son transportadas según las características de sus propio pesos, del lugar de la erosión, de las corrientes que afectan a los mismos, entre otros. Luego de este proceso, las partículas se hunden y se van depositando una sobre otra en el fondo marino, compactándose y generando los denominados sedimentos. Los cuales son una respuesta a la variedad ambiental, física y química

existente en el hábitat marino, por lo que existe diversidad de estos sedimentos caracterizado por el ambiente y los procesos involucrados desde el momento de la erosión de las partículas hasta la sedimentación de los mismos.

La morfología del fondo marino juega un rol importante porque este puede favorecer o no a que los fragmentos se acumulen en ciertas zonas del fondo marino. Por lo tanto, ya sea al inicio de la erosión, en el proceso de transporte o cuando la energía de los agentes externos se agote, estos fragmentos se acumulan en el fondo marino y luego se compactan por el efecto de la alta presión por millones de años formando sedimentos que almacenan la información histórica a diferentes escalas del ecosistema y clima pasado. Por tal motivo, los sedimentos son calificados como la memoria del océano.

Por lo tanto, Los sedimentos son importantes porque al almacenar información histórica del ecosistema pasado, nos permita analizar y entender la evolución del ecosistema, así como de su clima pasado. Por otro lado, el fondo marino, es un punto clave porque es el hábitat en donde se almacenan y preservan estos sedimentos, además que sus características también influyen en la forma de deposición de los sedimentos. Por esta razón en esta sección de muestra ciertas bases teóricas sobre el fondo marino y los sedimentos que almacena.

### 2.2.1. El Fondo Marino

Es uno de las ambientes menos entendidos en la Tierra debido a la dificultad encontrada en su exploración. Su entendimiento está entrelazado con los aspectos más atractivos de tecnologías para muestrear y penetrar a este ambiente [Tyler, 2003].

Por muchos años, se pensaba que el fondo marino era plano y sin características singulares [Blondel, 2012] o que existían grandes cuencas o depresiones de la corteza terrestre [Sverdrup et al., 2005], pero con el avance de técnicas modernas de ecosondas esta idea ha cambiado. Por ello, en los últimos años hay un gran número de estudios en relación a este tema y además se ha descubierto que existe una gran variedad de accidentes geográficos que subyacen en el fondo marino tales como los existentes en el relieve terrestre [Blondel, 2012]. Entre ellos podemos nombrar a los montes submarinos, los dorsales oceánicos, los valles, las cuencas marinas, los cañones marinos, entre otros.

Como puede notarse, el hábitat del fondo marino es muy variante. Del mismo modo, las diferentes variables que son de importancia para su análisis, tales como el nivel de oxígeno, la salinidad, temperatura, etc., varían dependiendo a la zona en que se encuentren y directamente de

los procesos biológicos, geológicos, químicos y físicos que ocurran. Una manera de entender este bioecosistema es a través de sus sedimentos que registran información útil de las características pasadas del ecosistema y de su cambio climático.

Finalmente, se puede decir que hay un gran interés de seguir explorando lo que ocurre en el fondo marino desde un punto de vista global, como un sistema complejo, que trae consigo un conjunto de procesos e individuos que se interrelacionan entre sí tanto directa como indirectamente. Parte de este entendimiento cubre los análisis de sedimentología con apoyo de información válida de la batimetría y morfología de este hábitat.

## 2.2.2. Los Sedimentos del Fondo Marino

Cualquier fragmento de material orgánico o inorgánico podría ser denominado sedimento. Los sedimentos nos revelan los procesos que forman las playas, la forma de las costas, los cañones submarinos, y controlan la abundancia de los organismos marinos, por nombrar algunos [Chamberlin and Dickey, 2007].

Estos fragmentos se establecen en el fondo marino por lo que, a través del océano la topografía del fondo marino es moderado por una manta de sedimentos que se forman a partir de partículas que descienden a las profundidades del océano día tras día, milenio tras milenio y era tras era [Day, 2007].

Su acumulación construye depósitos gruesos, en donde, sin perturbación, las capas más profundas constituyen los sedimentos más antiguos y en las más superficiales se encuentran los sedimentos más recientes debido al principio de superposición. De este modo, la acumulación también constituye un registro de la evolución del ecosistema y de su cambio climático [Chamberlin and Dickey, 2007].

Los depósitos más gruesos de sedimentos son generalmente encontrados cerca a los márgenes continentales, e donde el sedimento es depositado relativamente rápido; por lo contrario, en las zonas más profundas del océano, la acumulación es tan lenta que produce capas delgadas que varían con la edad de la corteza oceánica. Gran parte de estudiar a los sedimentos es analizar su tasa de acumulación, su distribución sobre el fondo marino, sus fuentes y abundancia, su química, y la historia que ellos registran capa a capa [Sverdrup et al., 2005].

### 1. Clasificación de los Sedimentos:

Para describir y catalogar a los sedimentos en un lenguaje común existen ciertas clasificaciones tales como:

(a) **Según el Tamaño de la Partícula:**

El sistema de clasificación más usada es la escala Udden-Wentworth. La cual se realiza según el tamaño promedio de las partículas del sedimento (granos). Según este, los sedimentos son clasificados de acuerdo a su diámetro (Fig. 2.3). En esta escala, la arcilla y limo constituyen los sedimentos de diámetros más pequeños mientras que las rocas y guijarros son los de mayor diámetro [Chamberlin and Dickey, 2007]. La escala considerada puede ser en milímetros (mm) o en la escala adimensional Phi ( $\Phi$ )<sup>2</sup>.

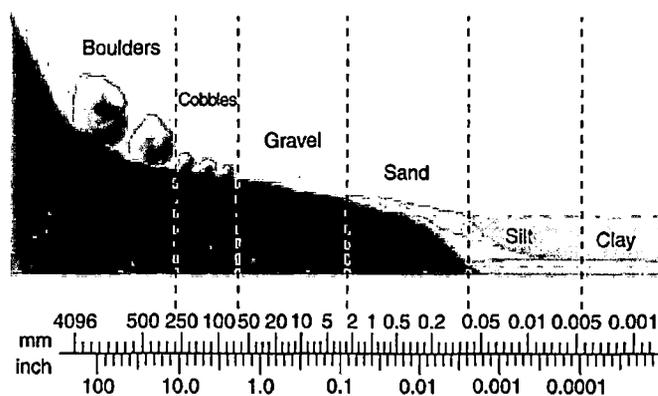


Figura 2.3: Clasificación del tamaño UddenWentworth de los sedimentos.  
Fuente: Exploring the World Ocean [Chamberlin and Dickey, 2007]

Los sedimentos gruesos son encontrados en lo alto de las costas porque al romper las ondas en la playa, las partículas más pesadas se sedimentan primero. Mientras que las partículas más finas permanecen en suspensión mayor tiempo y son llevadas hacia el mar por la resaca de las olas [Gray and Elliot, 2009].

Por lo general, sedimentos gruesos son respuesta a ambientes en donde hay mucha dinámica y por lo tanto están gobernados por un grado mayor de energía mientras que los ambientes menos dinámicos en el fondo marino, almacenan sedimentos más finos. Claramente, el tamaño del grano, medido como un diámetro de una partícula mediana es una propiedad clave [Gray and Elliot, 2009].

(b) **Según su Ubicación:**

<sup>2</sup>IMARPE emplea la escala adimensional Phi ( $\Phi$ ) como se verá en la sección de la metodología

Los sedimentos marinos son clasificados en neríticos(*de la costa*) y pelágicos(*del mar*) según la ubicación en donde se encuentran. Los sedimentos neríticos son encontrados cerca a los margenes continentales e islas y tienen un amplio rango de tamaños de partículas. La mayoría de estos son erosionados de rocas en la tierra y transportados a las costas por los ríos. Una vez que entran al océano, son propagados a lo largo de la Plataforma Continental y submergidos por las ondas y corrientes. Las partículas más grandes están cerca a las costas de las playas mientras que las más pequeñas son transportados muy remotamente de las costas [Sverdrup et al., 2005].

Los sedimentos pelágicos son de grano fino y se acumulan lentamente en el fondo marino. El grosor de estos es relacionado a la longitud de tiempo que ellos se han estado acumulando o la edad del fondo marino sobre el cual reposan. Por lo tanto su grosor tiende a incrementarse a distancias más grandes del dorsal oceánico [Sverdrup et al., 2005].

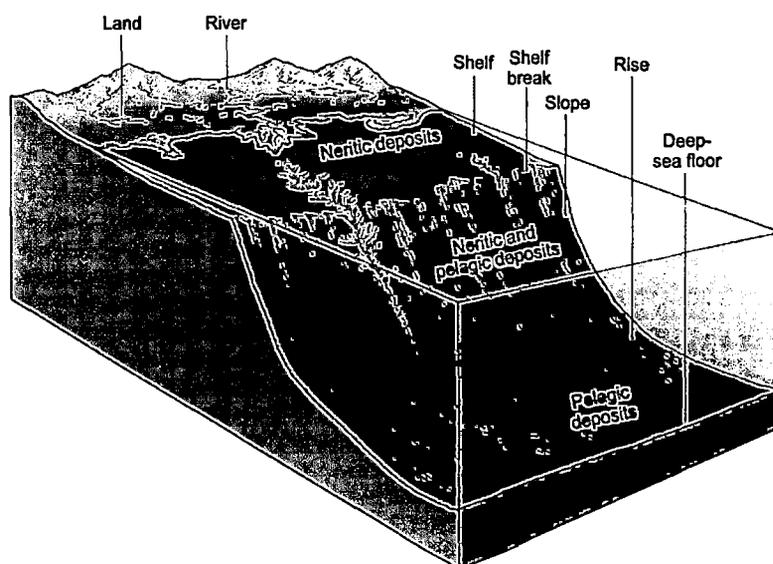


Figura 2.4: Clasificación de los sedimentos según su ubicación.  
Fuente: The Sea Floor and its Sediments [Sverdrup et al., 2005]

(c) **Según las Tasas de Depósito:**

La tasas de acumulación de sedimentos marinos tienen amplio rango debido a la variabilidad natural de los procesos que producen y transporta sedimentos. En estuarios de los ríos, esta tasa podría ser mayor a 8 metros por año. En bahías quietas, esta tasa podría ser 0.5 centímetros por año, y en la Plataforma y Talud Continental, valores de 10-40 centímetros

en 1000 años es típico [Sverdrup et al., 2005].

(d) **Según su Fuente:**

Los sedimentos marinos también pueden ser clasificados por la fuente de las partículas que conforman el sedimento que podrían ser de rocas preexistentes, organismos marinos, agua del mar y espacio. Aquellos que provienen de rocas preexistentes son clasificadas como sedimentos litógenos; comúnmente llamados sedimentos terrígenos cuando, técnicamente, este nombre incluye cualquier tipo de material viniendo del suelo terrestre.

Sedimentos derivados de organismos son clasificados como biógenos. Ellos podrían incluir fragmentos de corales y conchas al igual que partes de plantas y animales que viven en la superficie de marina. Aquellos provenientes del agua son clasificados como hidrógenos porque son producidos por el agua, debido a reacciones químicas. Finalmente, los sedimentos cosmógenos derivan de partículas que bombardean a la Tierra desde el espacio, la mayoría de esta partículas se quemas por pasar a través de la atmósfera pero aproximadamente el 10 % del material alcanza la superficie de la tierra [Sverdrup et al., 2005].

## 2. Ciclo del Sedimento:

Los procesos de erosión, transporte y deposición, en conjunto, es denominado proceso sedimentario, o simplemente sedimentación [Chamberlin and Dickey, 2007].

- **Erosión:** Desprendimiento de partículas arrastrados tales como rocas, organismos y especialmente microorganismos que forman esqueletos y conchas.
- **Transporte:** Circulación de las partículas mediante los agentes externos hasta que la energía se agote. Se da por agentes como los vientos, ríos, corrientes oceánicas, flujos de hielo, glaciares y procesos geológicos.
- **Deposición:** Cuando el proceso de transporte finaliza, las partículas empiezan a depositarse una sobre otra que con la presión logran compactarse y unirse formando el sedimento involucrando un número de procesos físicos, químicos y biológicos.

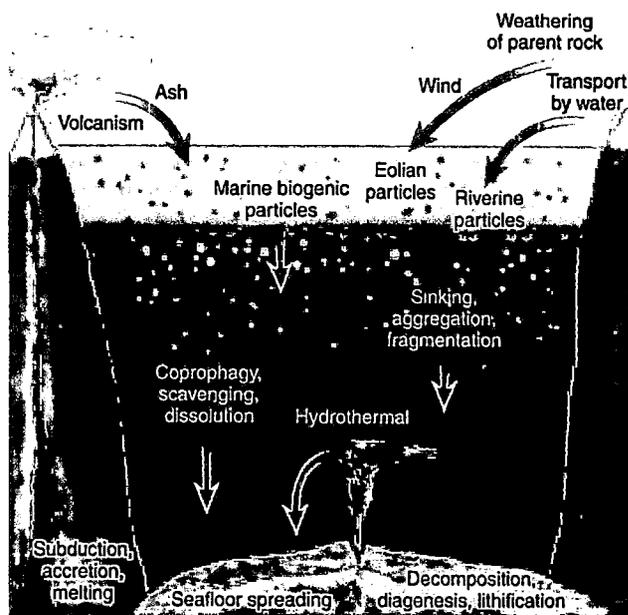


Figura 2.5: Ciclo del sedimento y procesos involucrados.  
Fuente: Ocean Sediments [Chamberlin and Dickey, 2007]

### 2.2.3. La Escala de Análisis en el Fondo Marino

El problema del patrón y escala es un punto clave de análisis en la ecología. El entendimiento de los procesos ecológicos involucra diversos fenómenos que ocurren en diferentes escalas de tiempo y espacio [Levin, 1992]. Dentro de un mismo fenómeno bajo estudio se encuentran diversos comportamientos a distintas escalas. Por mucho años no se ha considerado este factor tan importante porque se asumía que los patrones no son variables en la escala. A partir de los estudios de Levin se ha cambiado la forma de ver y abordar el problema del patrón y la escala.

Al hablar de escalas en la variación geográfica, Rahbek hace referencia a la unidad de muestreo y el espacio total cubierto. Según él, el primero es definido por el 'grano' y 'foco'. Siendo el grano el tamaño de la unidad de análisis y el foco, el área o espacio de inferencia representado por cada dato puntual. El espacio total cubierto es la 'extensión' y se refiere al espacio de inferencia al cual se aplica el conjunto entero de unidades muestrales [Rahbek, 2004].

Los procesos que se desarrollan en el hábitat del fondo del mar y su propia morfología varían según la escala de estudio. Por ello, la escala de observación influye la descripción de los patrones [Levin, 1992]. Si observamos un fenómeno a grandes escalas y poco a poco vamos dismi-

uyendo esa escala, es probable que encontremos un comportamiento distinto mientras pasamos de escala a escala. Por ello se ha incrementado la importancia de estudiar cuestiones ecológicas a través de las escalas [Chave, 2013]. Mientras más resolución tengamos en la observación de un fenómeno, es más probable encontrar más comportamientos o patrones porque permite analizar el fenómeno en una amplia gamma de escalas. Por ejemplo, en el caso de la predicción de densidades de especies, el uso de tamaños de grano demasiado grandes resultan en una excesiva pérdida de información y causa extrapolación falsa [Rahbek, 2004]. Además, existen muchos procesos de interés que se desenvuelven a escalas finas.

La sedimentología del fondo marino también tiene comportamientos a distintas escalas espaciales y temporales. Durante la última década, los avances en la tecnología del estudio marino y una mejor cobertura de los datos han mostrado cada vez más que los depósitos de Plataforma y Talud Continental son con frecuencia sometidos a procesos de pequeña escala [Reiche et al., 2011]. En relación al estudio presente, se sabe que el tamaño de grano del sedimento y la textura del sedimento varían a escalas finas espaciales. También se ha comprobado la asociación de la profundidad del fondo marino en la predicción de la distribución de las características sedimentarias mencionadas a escalas menores a la que se tiene en el mapa de batimetría del proyecto SRTM30\_PLUS (1000 metros aproximadamente) [Lark et al., 2012, Meilianda, 2012, Meilianda et al., 2011, Becker et al., 2009, Verfaillie et al., 2006]. En general, la batimetría ha sido empleada y recomendada como variables predictoras en estudios de distribución de las características de los sedimentos y variables relacionadas al ecosistema marino. En tales casos se ha necesitado información batimétrica a alta resolución para captar la variabilidad que existe en estos procesos.

Sin embargo, por muchos años no se conocía la forma del fondo marino y con el avance de la informática y modelos espaciales, se ha logrado obtener la cuantificación del fondo marino a escalas de 2km, 1km y en algunos casos (para ciertas zonas), a resoluciones menores a 1km. Hoy en día se intenta obtener grillas de batimetría a alta resolución para comprender la variabilidad a escalas finas y la interrelación que tiene está con procesos y patrones que se desarrollan a escalas finas. Ello debido a que se ha comprobado su asociación con diversos procesos del hábitat marino como se mencionó en el párrafo anterior.

Por lo tanto, en el presente estudio también se necesita obtener mapas batimétricos de alta resolución. De este modo, permitimos emplearla como covariable en estudios de procesos y fenómenos que se desenvuelven en el marco del Sistema de la Corriente de Humboldt del Norte. En nuestro caso

particular, permitimos su empleo en la predicción de ambientes de deposición continua de manera directa (batimetría) y de manera indirecta (morfometría).

## 2.3. Los Ambientes de Deposición Sedimentaria

Existen distintos tipos de ambientes de deposición según las características de los procesos físicos, químicos y biológicos ocurridos en el ecosistema marino, los cuales pueden verse explicados, en parte, por la dinámica de los agentes externos y los rasgos morfométricos de la topografía marina. En el presente estudio, estamos enfocados en los ambientes de deposición sedimentaria continua o depocentros sedimentarios porque en ellos se encuentran sedimentos finos de alta calidad para el análisis geológico, paleoceanográfico y estratigráfico caracterizados por poseer un tamaño granulométrico pequeño y por estar compuesto principalmente de arcilla y limo.

El depocentro es el punto de máxima acumulación sedimentaria en la cuenca. Dependiendo del flujo sedimentario, este depocentro podrían o no corresponder a la zona de máxima subsidencia. Este es raramente centrado en el fondo de una cuenca funcional sino más bien que uno de sus bordes donde las entradas con mayores [Biju-Duval, 2002].

Una cuenca sedimentaria puede ser definida como una zona particular de subsidencia que ha permitido la acumulación de sedimentos en un espesor de varios cientos de metros [Biju-Duval, 2002]. Estas son consideradas las grabadoras de la historia de la Tierra porque registran la variabilidad climática, historia tectónica y otros eventos dinámicos a grandes escalas temporales [Syvitski et al., 2003].

## 2.4. La Batimetría y Morfometría

### 2.4.1. Batimetría del Fondo Marino

La batimetría es la medida de la profundidad de agua en lagos, ríos y océanos. También es denominado la topografía del fondo marino (Fig. 2.6). El origen de la batimetría se remota alrededor del año 85 a.c., cuando un geógrafo griego le preguntó a un marinero acerca de la profundidad del océano. El marinero amarró una roca grande a una cuerda, se fue al medio del Mar Mediterráneo y deslizo la cuerda por casi 2 km hasta alcanzar el fondo marino. Este método fue aplicado, sin mayor modificación, para obtener un sondeo de las medidas de profundidad por casi los 2000 años

siguientes [Sverdrup et al., 2005].

No fue hasta 1920 que un equipo de sondeo acústico fue inventado, y la medida de la profundidad del fondo marino se convirtió en algo rutinario. El barco Germán *Meteor* hizo el primer uso a gran escala de la ecosonda en un crucero oceanográfico de investigación y detecto el dorsal oceánico del Atlántico [Sverdrup et al., 2005]. Al día de hoy las técnicas de ecosondas han tenido un gran desarrollo, por lo que se nos facilita, a comparación de años anteriores, mapear el fondo marino.

La importancia de la Batimetría se debe a que es base inicial de estudios morfológicos del fondo marino. Por este motivo diversas instituciones (NOAA<sup>3</sup>, Scripps<sup>4</sup>, IHO<sup>5</sup> y otros) hacen esfuerzos para obtener la visualización de mapas batimétricos conformando proyectos de investigación importantes como GEBCO<sup>6</sup> y SRTM30\_PLUS<sup>7</sup>.

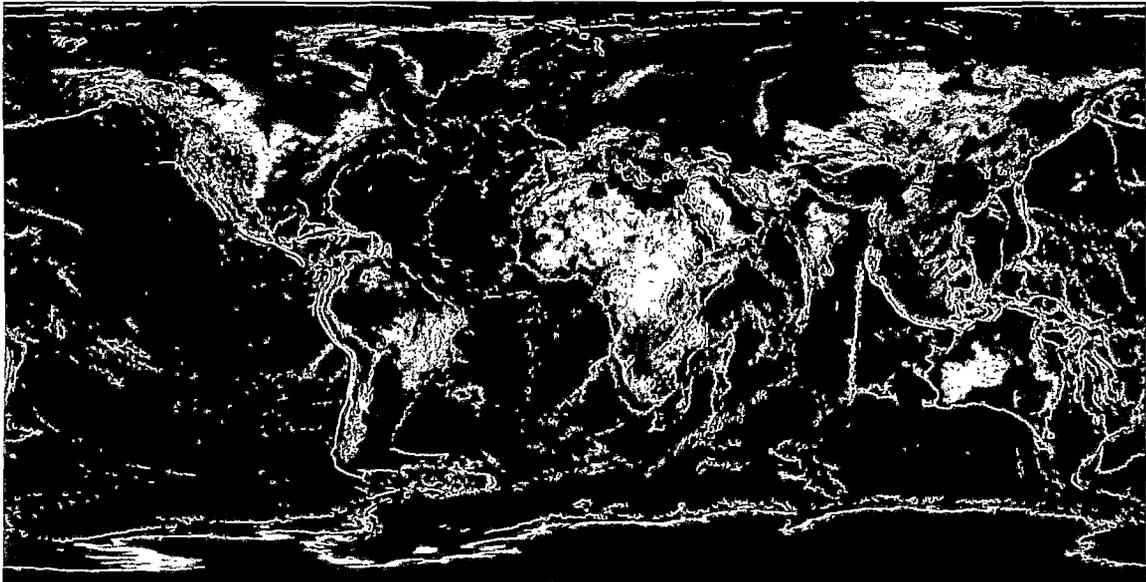


Figura 2.6: Batimetría y Topografía de la Tierra.  
Fuente: NOAA

---

<sup>3</sup>National Oceanic and Atmospheric Administration

<sup>4</sup>Scripps Institution of Oceanography

<sup>5</sup>International Hydrographic Organization

<sup>6</sup>General Bathymetric Chart of the Oceans

<sup>7</sup>Improving of Shuttle Radar Topography Mission Project

## 2.4.2. Morfometría del Fondo Marino

El estudio de las formas y estructura de las características del fondo marino. Por muchos años, se pensaba que el fondo marino era plano y sin características singulares. Pero con el avance de técnicas modernas de ecosondas esta idea ha cambiado. Por ello, en los últimos años hay un gran número de estudios en relación a este tema y además se ha descubierto que existe una gran variedad de accidentes geográficos que subyacen en el fondo marino tales como los existentes en el relieve terrestre [Blondel, 2012]. Entre ellos podemos nombrar a los montes submarinos, los dorsales oceánicos, los valles, las cuencas marinas, los cañones marinos, entre otros (Fig. 2.7).

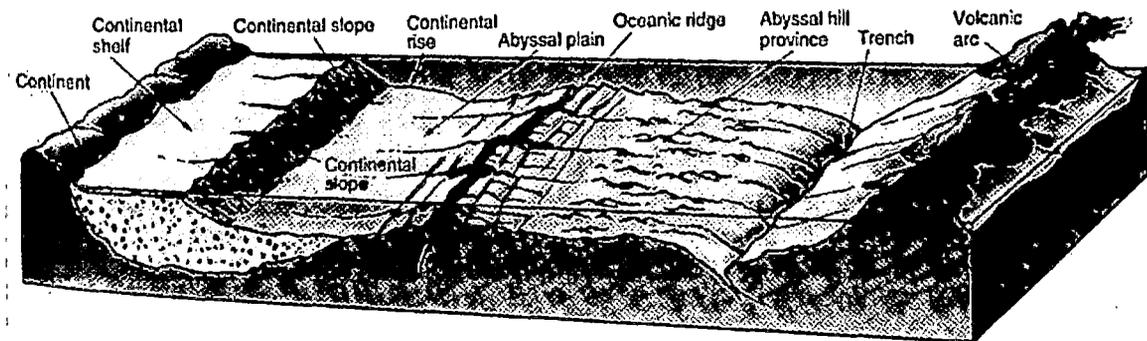


Figura 2.7: Morfología del Fondo Marino.  
Fuente: Universidad de Puerto Rico - Departamento de Geología

El estudio de la forma y estructura de la topografía marina tiene incidencia en la mayoría de las ciencias marinas y la topografía debe ser considerado como un factor del ambiente importante en todos los campos. Así muchos procesos podrían estar controlado, en parte, por la topografía así como la dinámica de surgencia y el transporte de sedimentos de un lado de una cuenca a otro [Laughton, 1962].

## Capítulo 3

# Metodología

En el presente capítulo se muestra los aspectos metodológicos considerados para la realización del estudio. Tales como la naturaleza de los datos, tipo de estudio, el pre-procesamiento de los datos, técnicas estadísticas a emplear, entre otras.

### 3.1. Diseño de Investigación

#### 3.1.1. Tipo de Estudio

Nuestro estudio se enfoca en predecir el valor de variables sedimentarias (zonas de depocentro) en función de ciertas características morfológicas. Por ello, el tipo de estudio es Predictivo-Correlacional. De esta manera evaluaremos la correlación existente entre las variables sedimentarias y morfológicas para luego proponer un modelo adecuado de predicción de depocentros de sedimentos.

#### 3.1.2. Área de Estudio

1. Para la predicción de la batimetría

El área de estudio cubre la región delimitada por los rangos de latitud  $[-20,7, -1,3]$  y de longitud  $[-85,6, -69,8]$  (Fig. 3.1). Note que esta área cubre la parte norte del Sistema de la Corriente de Humboldt.

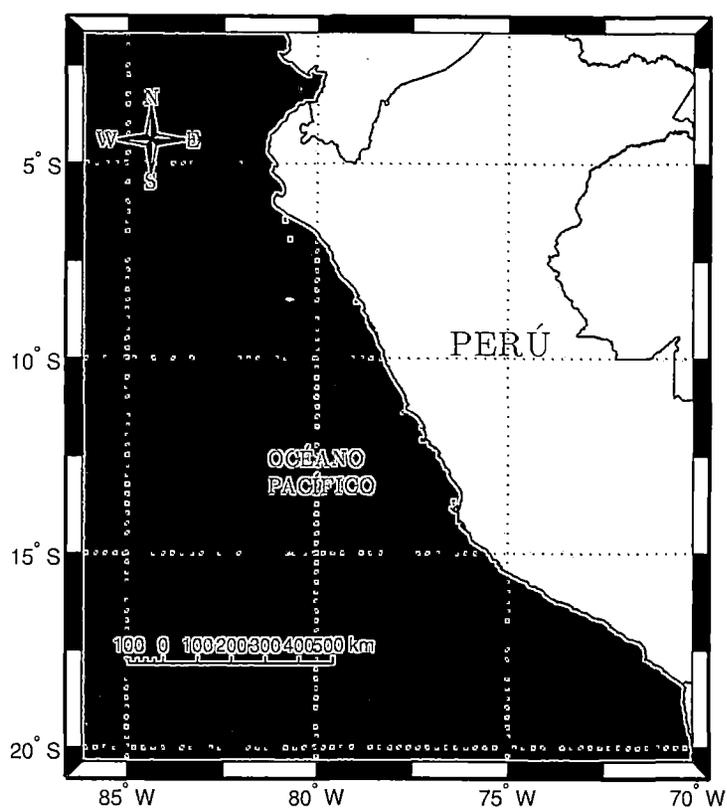


Figura 3.1: Área de Estudio para la Batimetría (Celeste)

2. Para la predicción de los depocentros

La ubicación de los depocentros sedimentarios (ambientes de deposición continua) que se plantea determinar en el presente estudio se encuentran en un amplio sector del margen continental, más precisamente en la Plataforma Continental y parte del Talud Continental superior (Fig. 3.2). Por ello, el área de estudio para la predicción de los ambientes de deposición continua esta delimitada a esta zona.

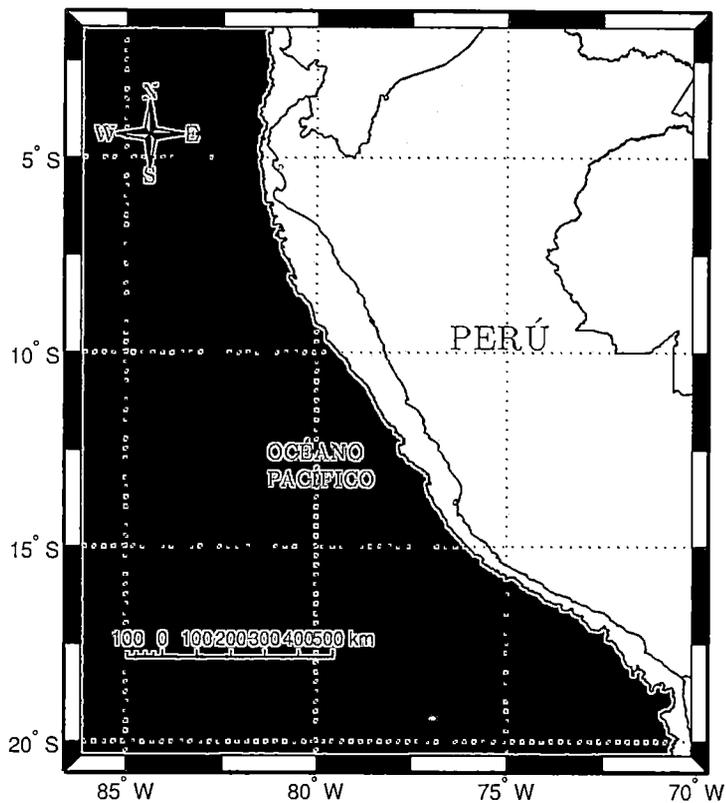


Figura 3.2: Área de Estudio para la Detección de Depocentros (Color beige)

### 3.1.3. Muestra

- En función a la Batimetría:

La muestra viene a ser los datos recolectados en las trayectorias de los cruceros científicos de IMARPE (Instituto del Mar del Perú), IRD (Institut de recherche pour le développement) y la NOAA (National Oceanic and Atmospheric Administration); además de grillas de algunas zonas de la DHN (Dirección de Hidrografía y Navegación), líneas costeras del Perú de GADM (Global Administrative Areas) y grillas espaciadas irregularmente a media milla náutica del proyecto SRTM30\_PLUS (Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution) que, en conjunto, vienen a ser alrededor de 200 millones de registros. En la figura 3.3 se puede ver la distribución espacial de las principales fuentes mencionadas anteriormente. Además, en la 3.4 se puede ver la distribución espacial de la muestra según la cantidad de fuente de datos que cubren el área.

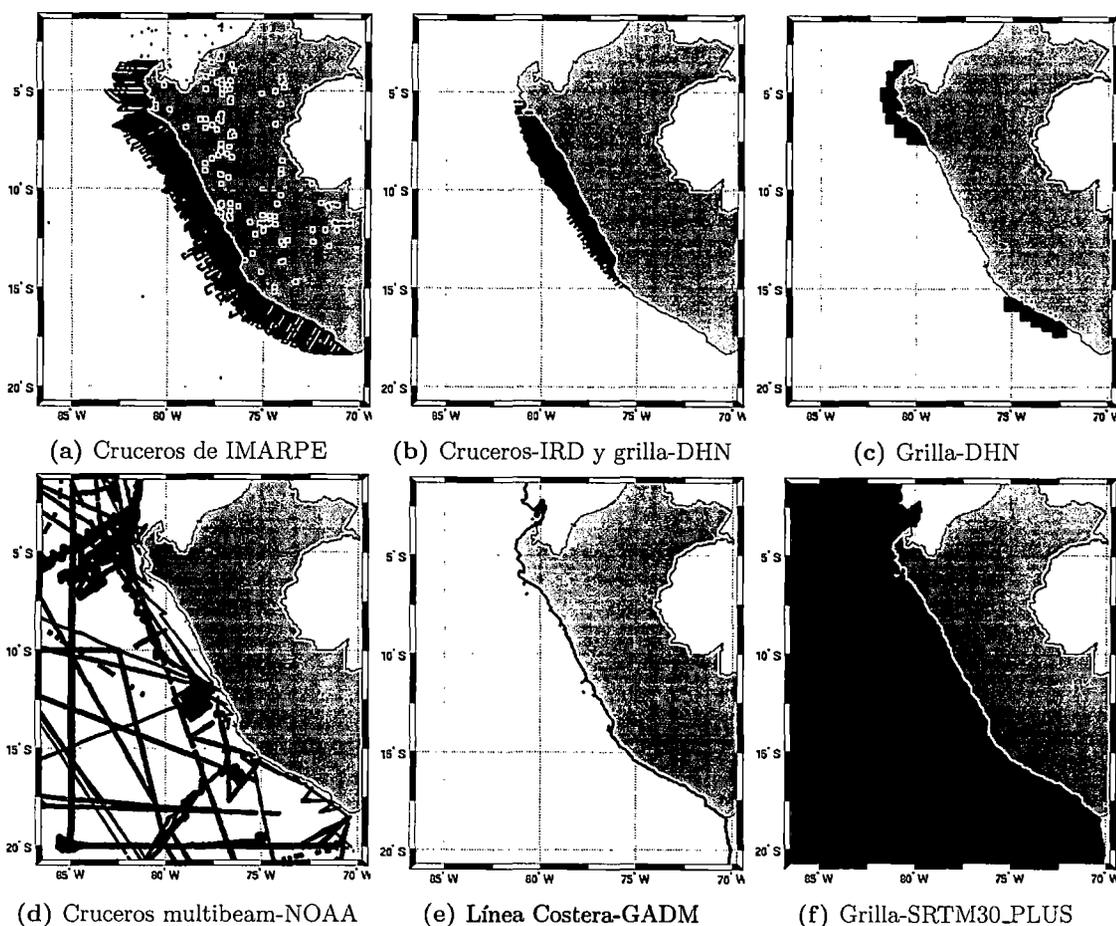


Figura 3.3: Distribución de las fuentes de datos batimétricos.

Se consideran los datos de estas fuentes porque son Institutos o Proyectos que parcialmente comparten los objetivos de este estudio.

**Instituto del Mar del Perú (IMARPE):** Organismo Técnico Especializado del Sector Producción, Subsector Pesquería, orientado a la investigación científica, así como al estudio y conocimiento del Mar Peruano y sus recursos, para asesorar al Estado en la toma de decisiones con respecto al uso racional de los recursos pesqueros y la conservación del ambiente marino, contribuyendo activamente con el desarrollo del país (<http://www.imarpe.pe/>).

**Institut de recherche pour le développement (IRD):** Organismo de investigación original y único en el panorama europeo de la investigación para el desarrollo, tiene como vocación realizar investigaciones en el Sur, para el Sur y con el Sur (<http://es.ird.fr/>). En el Perú tiene un convenio internacional con IMARPE a través del proyecto (LMI-

DISCOH)<sup>1</sup>, el cual tiene como objetivo el estudio de las dinámicas océano-atmósfera, bio-geoquímicas y ecológicas en el SCH con el fin de comprender y anticipar el efecto de las variabilidades intra-estacionarias, interanuales, decenales y de cambio climático sobre la dinámica del ecosistema costero (<http://es.discoh.ird.fr/>).

**Dirección de Hidrografía y Navegación (DHN):** Su misión es administrar, operar e investigar las actividades relacionadas con las ciencias del ambiente en el ámbito acuático, con el fin de contribuir al desarrollo nacional, brindar apoyo y seguridad en la navegación a las Unidades Navales y a los navegantes en general (<https://www.dhn.mil.pe/>).

**National Oceanic and Atmospheric Administration (NOAA):** Es una agencia que enriquece la vida a través de la ciencia. Su alcance va desde la superficie del sol hasta las profundidades del mar, mientras que trabajan para mantener a los ciudadanos informados de los cambios del entorno que les rodea (<http://www.noaa.gov/>).

**Global Administrative Areas (GADM):** Es una base de datos espacial de la localización de las áreas administrativas del mundo (o límites Administrativas) para su uso en el SIG y software similar. Áreas administrativas de esta base de datos son los países y subdivisiones de menor nivel, como provincias, departamentos, etc. GADM describe dónde están áreas administrativas (las características espaciales), y para cada área proporciona algunos atributos, como los nombres y variantes (<http://www.gadm.org/>).

**SRTM30\_PLUS (SRTM30\_PLUS):** Es un proyecto realizado con el fin de obtener datos globales de elevación y batimetría con una resolución de 30 segundos. Se ha desarrollado a partir de una amplia variedad de fuentes de datos. La tierra y la topografía del hielo proviene de la topografía SRTM30 y ICESat, respectivamente. La batimetría del Océano se basa en un nuevo modelo de satélite gravedad, donde la proporción de la gravedad de la topografía se calibra con 298 millones de sondeos editados [Becker et al., 2009].

---

<sup>1</sup>Dinámicas del sistema de la Corriente de Humboldt

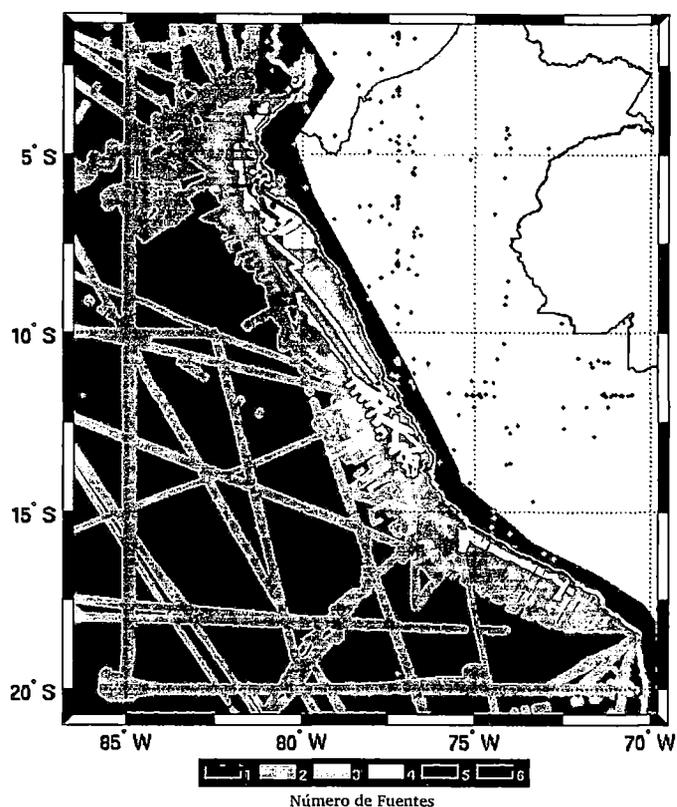


Figura 3.4: Distribución espacial de la muestra de batimetría según número de fuentes

- En función a la Sedimentología:

La muestra consta de los puntos de recolección de información sedimentaria en 556 puntos espaciados y distribuidos irregularmente en el margen continental peruano, provenientes de diferentes cruceros de investigación de IMARPE [Delgado et al., 1987, Delgado and Gomero, 1995, Velazco, 2012].

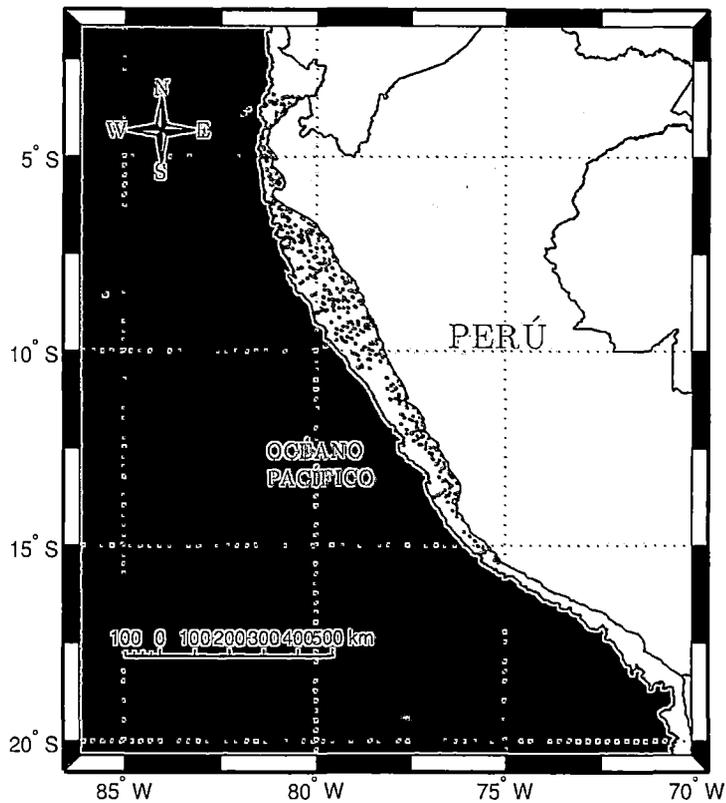


Figura 3.5: Distribución espacial de los puntos muestreados (color rojo) para obtener características sedimentarias (Composición de la textura y coeficiente de asimetría del tamaño de granos).

### 3.2. Operacionalización de las Variables

En los cuadros 3.1 y 3.2 se muestra la matriz de consistencia, que relaciona los problemas, los objetivos, las hipótesis y las variables a considerar en el estudio. Puede notar que el modelamiento de la presente tesis se basa en los factores batimétricos, morfológicos y sedimentarios para la predicción de depocentros de sedimentos. Sin embargo, las variables a emplear son muy generales, por tal motivo operacionalizamos dichas variables en esta sección hasta obtener indicadores medibles en cada uno de estos factores.

**Cuadro 3.1: Matriz de consistencia (i)**

Problema Principal	Objetivo Principal	Hipótesis Principal	Variables
Se desconoce la ubicación precisa de ambientes de deposición continua en el margen continental peruano.	Predecir la ubicación de probables ambientes de deposición continua en el margen continental peruano.	La ubicación de los ambientes de deposición continua puede ser identificada mediante el tamaño de grano y la composición de la textura del sedimento, y modelada a través de factores batimétricos y morfológicos.	<ol style="list-style-type: none"> <li>1. Variables de Batimetría</li> <li>2. Variables de la Morfología del fondo marino</li> <li>3. Variables de Sedimentología</li> </ol>

Cuadro 3.2: Matriz de consistencia

Problema	Objetivos	Hipótesis	Variables
1. No se tiene información a alta resolución que describa las características del fondo marino peruano. Lo cual es necesario para identificar características morfológicas relacionadas a los ambientes de sedimentación.	Modelar la distribución espacial de la profundidad del fondo marino a una alta resolución para obtener rasgos del fondo marino más detallados.	Existen datos de fuentes locales (Perú) que son más confiables en comparación a grillas batimétricas de baja resolución que se emplean a nivel internacional.	<ol style="list-style-type: none"> <li>Ubicación Espacial</li> <li>Profundidad</li> </ol>
2. Se necesita conocer las características sedimentarias de los ambientes de deposición continua.	Identificar el perfil sedimentario del ambiente de deposición continua mediante clusters.	Los tipos de ambientes de deposición son caracterizados según el tamaño de grano y la composición de la textura del sedimento.	<ol style="list-style-type: none"> <li>Fracciones granulométricas</li> <li>Parámetros estadísticos del tamaño de grano</li> </ol>
3. No se cuenta con índices morfológicos del fondo marino, así que es necesario definir indicadores de morfología y calcularlos de manera cuantitativa en base a un mapa de batimetría de alta resolución.	Determinar índices morfológicos (Pendiente, rugosidad, localización, dimensión, concavidad, entre otros) del fondo marino que sirvan como variables independientes para la detección de depocentros en el relieve marino.	Se puede obtener indicadores morfológicos a partir del mapa de batimetría.	<ol style="list-style-type: none"> <li>Grilla Batimétrica</li> </ol>
4. Teóricamente, se conoce que hay asociaciones entre los rasgos de morfología y los ambientes de deposición. Sin embargo, al momento, no se ha cuantificado y definido la significancia de estos rasgos morfológicos en el área del margen continental peruano.	Determinar un modelo espacial para la detección de ambientes de deposición continua en el relieve marino considerando las características morfológicas como covariables.	La presencia de depocentros es explicada por una tendencia en función a los índices de batimetría y morfometría y por un componente espacial.	<ol style="list-style-type: none"> <li>Cercanía a las costas</li> <li>Pendiente</li> <li>Curvatura</li> <li>Distancia a depresiones</li> <li>Pendiente con dirección</li> <li>Variabilidad</li> <li>Tipo de Ambiente Depositional obtenido en el objetivo 2</li> </ol>
5. Existen grandes áreas en las que no conocemos si hay presencia de ambientes de deposición continua, los cuales son importantes para obtener información de calidad de los procesos sedimentarios.	Detectar probables ambientes de deposición continua con el perfil de sedimentología identificado.	Se puede detectar probables ambientes de deposición en base al modelo espacial obtenido, incluso en zonas donde se tiene menor información de los perfiles sedimentarios.	<ol style="list-style-type: none"> <li>Tipo de Ambiente Depositional obtenido en el objetivo 2</li> <li>Variables morfológicas significativas obtenidas en el objetivo 4</li> </ol>

### 3.2.1. Factor Batimétrico

La batimetría es un factor sustancial para la determinación de cuencas y depocentros de sedimentación, por esta razón, consideramos un conjunto de variables con nos ayudan a caracterizar este factor:

- (a) Ubicación Espacial: La naturaleza espacial de las variables bajo estudio hacen que sea necesario cuantificar de manera estándar la locación de los datos, así que se considera la latitud y longitud en grados como dimensiones de esta característica. Ellos son ampliamente usados para determinar la ubicación de un punto en el globo terráqueo. La Latitud mide la distancia de un punto hacia el Ecuador mientras que la Longitud cuantifica la distancia de un punto a un meridiano considerado como punto inicial, generalmente al meridiano de Greenwich.
- (b) Profundidad: La profundidad puede ser considerado como la tercera dimensión para cuantificar la ubicación espacial, en nuestro caso representa la distancia del nivel del mar (0 metros) hasta tocar el piso marino. Para el estudio se considera la profundidad en una ubicación puntual y la profundidad promedio de su vecindad.
- (c) Error de Batimetría: No contamos con la medida de la profundidad a una distancia tan corta de 50 metros para un área tan grande como la considerada en este estudio por lo que la profundidad tiene que ser estimada. Por ello consideramos el error de esta predicción, ya que nos da información de la confiabilidad del dato de profundidad considerado.

**Cuadro 3.3:** Operacionalización de Variables de Batimetría

Variable	Dimensión	Indicador	Escala	Valor
Ubicación Espacial	Latitud	Distancia al Ecuador	Razón	$< -90, 90 >$ grados
	Longitud	Distancia al meridiano de Greenwich	Razón	$< -180, +180 >$ grados
Profundidad	Profundidad puntual	Distancia al fondo marino	Razón	$[0 - 8000 >$ metros
	Profundidad local	Distancia promedio al fondo marino	Razón	$[0 - 8000 >$ metros
Error de Batimetría	Error de Predicción	Raiz del Error Cuadrático Medio	Razón	$[0 - 300 >$ metros

### 3.2.2. Factor Morfológico

La morfología es uno de los aspectos importantes para la predicción de la ubicación de ambientes de deposición continua porque toma como base la batimetría, pero se analiza más allá para

obtener patrones estructurales del fondo marino.

- (a) Cercanía a las costas: La sedimentología tiene cierta variación debido a que la energía varía según la cercanía de las costas por efecto de las olas haciendo que partículas pesadas se sedimenten en las costas y partículas ligeras sean llevadas al mar por la resaca de las olas. En el estudio consideramos la distancia mínima de un punto a las costas. Un indicador no considera islas para este cálculo mientras que el otro si las considera como parte de las líneas costera.
- (b) Pendiente: La pendiente también es un factor importante para la caracterización de los sedimentos del fondo marino. Con este podemos evaluar si la zona es plana o empinada. Consideramos la pendiente puntual cuando se calcula la pendiente directamente de los datos de batimetría predichos mientras que se considera pendiente local cuando se calcula de un mapa suavizado en un radio de vecindad definido. Además se tiene en cuenta 3 tipos de pendientes. (i) Pendiente máxima de las ocho direcciones posibles de una vecindad regularmente espaciada. (ii) La pendiente en dirección de la latitud y (iii) la pendiente en dirección de la longitud.
- (c) Curvatura: Imagine una zona relativamente plana con pequeños montes de elevaciones y zonas profundas, sobre la cual se sueltan pequeñas esferas que se van acomodando según la morfología del terreno. Las esferas quedarán acomodadas en las depresiones en su mayoría, otras quedarán en zonas planas y ninguna en las elevaciones. Por ello la curvatura, que nos indica si es plana, elevada o profunda es un aspecto importante a considerar aunque este pequeño ejemplo no se compare a la dinámica y variabilidad gobernada en el ecosistema marino. La curvatura se calcula como la segunda derivada y se considera en la dirección de la mayor pendiente y perpendicular a la misma. También se evalúa el tipo de curvatura (elevación, depresión y plano) en ambas direcciones. Esta variable es cuantificada a nivel puntual y local igual que la pendiente.
- (d) Distancia a depresiones: La distancia a depresiones es calculada según la curvatura de la superficie, es decir, la distancia desde un punto cualquiera hasta el punto más cercano con curvatura en forma de lente de una depresión, para lo cual consideramos los dos tipos de curvaturas evaluadas. El nivel puntual y local también se halla en este caso.
- (e) Pendiente con dirección: En el caso anterior, la pendiente se considera solo como intensidad y no si es positiva o negativa. Por lo que en esta parte, se evalúa la intensidad y su dirección dependiendo de la curvatura. Este se calcula en el sentido de la pendiente máxima y perpendicular

a ella, y a nivel puntual y local.

(f) Variabilidad: La variabilidad es importante para la acumulación de sedimentos, se espera que en zonas más variables se acumulen más sedimentos porque esta variabilidad se resiste a la dinámica de las corrientes haciendo que sea protección de ambientes de deposición y también generando erosión. Se calcula mediante las desviaciones entre la grilla batimétrica estimada y una suavización de la misma de tal modo que la variabilidad no se calcule respecto a una forma lineal, cuadrática u otra, sino que se calcula según la tendencia local existente.

**Cuadro 3.4:** Operacionalización de Variables de Morfología

Variable	Dimensión	Indicador	Escala	Valor
Cercanía a las costas	Sin islas	Distancia mínima a las costas	Razón	< 0, 10 > grado
	Con islas	Distancia mínima a las costas	Razón	< 0, 10 > grado
Pendiente	Puntual	Máximo ángulo de inclinación	Razón	< 0, 90 > grado
		Ángulo de inclinación en el eje x	Razón	< 0, 90 > grado
		Ángulo de inclinación en el eje y	Razón	< 0, 90 > grado
	Local	Máximo ángulo de inclinación	Razón	< 0, 90 > grado
		Ángulo de inclinación en el eje x	Razón	< 0, 90 > grado
		Ángulo de inclinación en el eje y	Razón	< 0, 90 > grado
Curvatura	Puntual	Curvatura de la pendiente máxima (1)	Razón	< 0, 90 >
		Curvatura perpendicular a la pendiente máxima (2)	Razón	< 0, 90 >
		Tipo de curvatura de la pendiente máxima (1)	Nominal	< -1, 0, 1 >
		Tipo de curvatura perpendicular a la pendiente máxima (2)	Nominal	< -1, 0, 1 >
	Local	Curvatura de la pendiente máxima (1)	Razón	< 0, 90 >
		Curvatura perpendicular a la pendiente máxima (2)	Razón	< 0, 90 >
		Tipo de curvatura de la pendiente máxima (1)	Nominal	< -1, 0, 1 >
		Tipo de curvatura perpendicular a la pendiente máxima (2)	Nominal	< -1, 0, 1 >
Distancia a depresiones	Puntual	Distancia mínima con la curvatura 1	Razón	< 0, 10 > grado
		Distancia mínima con la curvatura 2	Razón	< 0, 10 > grado
	Local	Distancia mínima con la curvatura 1	Razón	< 0, 10 > grado
		Distancia mínima con la curvatura 2	Razón	< 0, 10 > grado
Pendiente con dirección	Puntual	Máximo ángulo de inclinación con dirección	Razón	< -90, 90 > gra
		Máximo ángulo de inclinación con dirección perpendicular	Razón	< -90, 90 > gra
	Local	Máximo ángulo de inclinación con dirección	Razón	< -90, 90 > gra
		Máximo ángulo de inclinación con dirección perpendicular	Razón	< -90, 90 > gra
Variabilidad	Local	Raíz promedio de la suma cuadrática de desviaciones	Razón	< 0, 300 >

### 3.2.3. Factor Sedimentario

Los sedimentos son el corazón de este estudio, de ellos consideramos tres variables importantes; las fracciones granulométricas, parámetros estadísticos del tamaño de grano (partícula) y el ambiente deposicional. A continuación describimos cada una de estas.

(a) Fracciones granulométricas de los sedimentos: Determinado en función a la combinación de partículas de determinado tamaño de grano. El tamaño de grano (mm) de las partículas (y su equivalencia en una escala adimensional Phi( $\Phi$ )) es definido en la escala de Wentworth (Fig. 3.6).

Millimeters	$\mu\text{m}$	Phi ( $\phi$ )	Wentworth size class	
4096		-20		
1024		-12		
256		-10	Boulder (-8 to -12 $\phi$ )	
64		-8		
16		-6	Pebble (-6 to -8 $\phi$ )	
4		-4		
		-2	Pebble (-2 to -6 $\phi$ )	
		-1.75		Gravel
3.36		-1.50	Gravel	
2.83		-1.25		
2.38		-1.00		
2.00		-0.75		
1.68		-0.50	Very coarse sand	
1.41		-0.25		
1.19		0.00		
1.00		0.25		
0.84		0.50	Coarse sand	
0.71		0.75		
0.59		1.00		
1/2	0.50	1.25		Sand
0.42	420	1.50	Medium sand	
0.35	350	1.75		
0.30	300	2.00		
1/4	0.25	2.25		
	0.210	2.50		
	0.177	2.75	Fine sand	
	0.149	3.00		
1/8	0.125	3.25		
	0.105	3.50		
	0.088	3.75	Very fine sand	
	0.074	4.00		
1/16	0.0625	4.25		
	0.0530	4.50		
	0.0440	4.75	Coarse silt	
	0.0370	5		
1/32	0.0310	6		
	0.0156	7	Medium silt	
1/64	0.0078	8	Fine silt	
1/128	0.0039	9	Very fine silt	
1/256	0.0020	10		Mud
	0.00098	11		
	0.00049	12		
	0.00024	13	Clay	
	0.00012	14		
	0.00006			

Figura 3.6: Escala de tamaño de grano de las partículas  
Fuente: A Scale of Grade and Class Terms for Clastic Sediments [Wentworth, 1922].

Las fracciones de grava y arena conforman también la denominada de modo más general “fracción gruesa”. En tanto que las partículas del tamaño del limo y arcilla conforman la denominada de modo más general “fracción fina”.

Las combinaciones de proporciones relativas de las fracciones granulométricas permiten definir

la textura de los sedimentos. La cantidad relativa de estas fracciones se ve reflejada por diversos procesos por los que han pasado las partículas que conforman el sedimento por eso son punto clave para los ambientes de deposición que buscamos. En el presente trabajo buscamos zonas en donde el limo y la arcilla representen el mayor porcentaje. Estos porcentajes provienen de los datos obtenidos en las publicaciones de IMARPE [Delgado and Gomero, 1995, Velazco, 2012].

- (b) **Parámetros estadísticos del tamaño de grano:** Constituyen las variables más importantes para el presente estudio. Se obtienen a partir del cálculo de valores del tamaño de grano de determinados percentiles de la distribución del tamaño de grano de los sedimentos en su forma gráfica (computación gráfica) ó del cálculo de los valores de toda la distribución del tamaño de grano de una muestra (medidas de momentos) [Folk, 1968]. Los parámetros estadísticos del tamaño de grano de los sedimentos son la mediana, el tamaño promedio, selección, asimetría y kurtosis. Para el presente estudio, se ha seleccionado a la asimetría como el parámetro que define ambientes de sedimentación reciente, según lo obtenido por [Delgado and Gomero, 1995, Velazco, 2012] e interpretan para el caso de la Plataforma Continental y el Talud Continental superior que cuando el coeficiente es negativo representa tamaños de partículas pequeñas.
- (c) **Ambiente Depositional:** Existen distintas clasificaciones y tipos de ambientes de sedimentación, nosotros empleamos este término para referirnos al ambiente reciente en que ocurren procesos de acumulación continua de sedimentos en un depocentro. Este indicador no es calculado directamente, sino con la ayuda de la fracción granulométrica fina y parámetros estadísticos del tamaño de grano determinamos clusters que definen este evento.

**Cuadro 3.5: Operacionalización de Variables de Sedimentología**

Variable	Dimensión	Indicador	Escala	Valor
Fracciones granulométricas	Limo	Porcentaje de Limo	Razón	[0 % – 100 %]
	Arcilla	Porcentaje de Arcilla	Razón	[0 % – 100 %]
	Arena	Porcentaje de Arena	Razón	[0 % – 100 %]
Parámetros estadísticos del tamaño de grano	Tamaño del Grano	Coefficiente de Asimetría	Razón	[-1, 1]
Ambiente Depositional	Tipo de ambiente	¿Es depocentro?	Nominal	[0, 1]

### 3.3. Recolección y Tratamiento de los datos

#### 3.3.1. Recolección de Datos

No se realizó ningún trabajo de campo, exclusivamente, para la realización de este trabajo, sino que se aprovecho las bases de datos disponibles a nivel nacional (IMARPE, IRD y DHN) y a nivel internacional (NOAA, GEBCO, SRTM30\_PLUS, GADM y GSHHS). Estos datos se recolectan con ayuda de cruceros científicos que explicaremos a continuación.

##### ▪ Datos de Batimetría:

Los datos de batimetría se pueden tomar mediante cruceros con ecosondas de haz simple o de haz múltiple, en otras ocasiones mediante gravas.

- Sonda náutica de brazo: En zonas poco profundas se emplea este tipo de técnicas en donde se suelta una cuerda con una parte pesada para que llegue al fondo marino, aparte de la profundidad se puede obtener muestras físicas del sustrato. Es muy limitado porque se realiza puntualmente.
- Ecosonda de haz simple: La recolección de datos a través de la ecosonda de haz simple consiste del crucero científico implementado con la ecosonda de haz simple de tal modo que mientras el crucero hace un recorrido, se envía una sonda que al hacer contacto con el fondo marino genera un eco. Mediante este eco y el perfil del mismo a través del agua, se puede cuantificar la profundidad. Sin embargo, nótese que este tipo de recolección es limitada por el área que se puede cubrir.
- Ecosonda de haz múltiple: Debido a la limitación de área barrida por las ecosonda de haz simple, se creo la de haz múltiple que envía varios haces desde la posición del crucero capturando una gran área, definitivamente es mucho mas ventajoso que emplear el haz simple.

- ##### ▪ Datos de Sedimentología:
- Para la obtención de las físicas de sedimentos de las cuales provienen los datos sedimentológicos (fracciones granulométricas, parámetros estadísticos), se emplearon técnicas de extracción, con dragas tipo Van Veen. Las dragas Van Veen son equipos de muestreo que permiten obtener una muestra del fondo marino desde la cubierta de un buque. Poseen dos mandíbulas metálicas articuladas en brazos metálicos que descienden hacia el fondo

sujetas en un extremo por un cable y en cubierta (el otro extremo) está el cable conectado a un winche que la izará del fondo. Al disminuir la tensión del cable se desactiva un mecanismo de seguro que solo funciona cuando el cable está en tensión. Al posarse en el fondo y disminuir la tensión las mandíbulas metálicas de la draga se cierran y capturan una muestra de sedimentos de la superficie del fondo. Los sedimentos son descritos a bordo y posteriormente trasladados en congelamiento al laboratorio donde son analizados por su tamaño de grano empleando tamices y pipetas [Ingram, 1974, Galehouse, 1974].

### 3.3.2. Pre-procesamiento de los Datos

Los datos provienen de distintas fuentes de datos tales como IMARPE-IRD, DHN, NOAA y Scripps Institution (Fig. 3.7), por lo cual se requiere dar un formato único a los ficheros pertenecientes a cada una de estas fuentes. Para ello, se emplea el software MATLAB obteniendo un fichero \*.mat con los siguientes atributos:

- (I) **Lat:** Latitud correspondiente a la ubicación de los datos, de -90 grados(SUR) a 90 grados(NORTE).
- (II) **Long:** Longitud correspondiente a la ubicación de los datos, de -180 grados(OESTE) a 90 grados(ESTE).
- (III) **Depth:** Profundidad correspondiente a la posición (*Lat, Long*) respecto al nivel del mar.
- (IV) **Source:** Descripción de la fuente proveniente: nombre del archivo final, nombre del archivo inicial, instituto, tipo de datos, proyecto para el cual se obtuvo los datos, contacto o responsable, cantidad de puntos y resolución.

**Ejemplo de Fuente:** BathyData8.mat, AMLR92T, NOAA, CRUISES, MULTIBEAM CRUISES, Daniel Price, 920415, 0.00002 millas náuticas.

### 3.3.3. Procesamiento de los Datos

En los aproximadamente 200 millones de puntos que se tienen de muestra, existen ciertos problemas que deben ser solucionados antes de realizar el modelamiento de los datos. Entre estos problemas encontramos, puntos con ubicaciones inconsistentes, con ubicaciones repetidas, con ubicaciones pseudo-repetidas y datos atípicos espaciales. Aunque los tres primeros casos son problemas

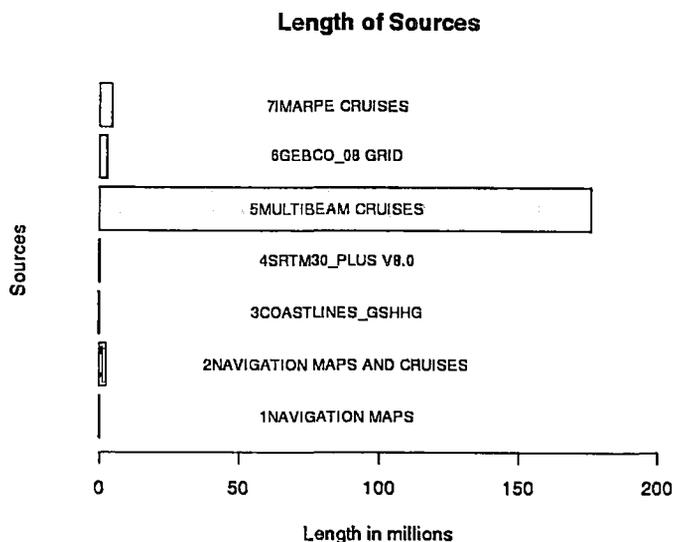


Figura 3.7: Cantidad de Datos según Fuente: 185,849,891 registros

muy sencillos, el trabajar con grandes cantidades de datos lo convierte en un tema más complejo por los recursos computacionales respecto a la memoria y tiempo, los cuales pueden dificultar el desarrollo de algoritmos como de identificar si un punto se encuentra dentro de un polígono, identificar los puntos más cercanos, identificar las distancias más cortas, entre otros.

#### 1. Detección de ubicaciones inconsistentes:

##### (a) Descripción:

Estos son datos que se identifican en ubicaciones ilógicas o contradiciendo la información de otras fuentes de datos. Entre los casos encontramos a: (i) aquellos puntos de profundidad ubicados en zonas de topografía, (ii) puntos en que se alejan demasiado del recorrido del crucero, (iii) ubicaciones que divergen con la información de la línea costera.

##### (b) Desventajas:

Estos puntos pueden generar predicciones inconsistentes si ellos están cerca de la línea costera peruana, creando conflicto entre la línea costera y los puntos de batimetría.

##### (c) Metodología de detección:

Por su definición, estos puntos son errores y por ello fueron eliminados del siguiente modo:

- (i) Se selecciona la línea costera más coherente con la mayoría de fuentes de datos y con ayuda de las imágenes satelitales de Google Earth, entre GSHHS Coastline (A

Global Self-consistent, Hierarchical, High-resolution Geography Database) y GADM (Global Administrative Areas).

- (II) Se definió el área de interés de la batimetría en base a la línea costera de la fuente de datos escogida.
- (III) Se detecta si el punto se encuentra fuera del polígono definido en (i).
- (IV) En caso de que los puntos resulten fuera del polígono definido, son eliminados.

## 2. Detección de datos repetidos:

### (a) Descripción:

Como se mencionó anteriormente, este trabajo aprovecha de la disponibilidad de datos de batimetría a nivel nacional e internacional, lo cual significa que muchos usuarios pueden haber usado los archivos y existe el riesgo de que algún usuario haya manipulado mal la base de datos, la haya abierto en algún software que modifica su precisión o que simplemente hubieran incoherencias desde la recolección de datos. Así que se generan datos repetidos, es decir, con la misma ubicación que dependiendo del caso podrían tener la misma profundidad, ligera diferencia en la profundidad o demasiada diferencia entre sus profundidades.

### (b) Desventajas:

En el mejor de los casos, cuando los puntos repetidos tienen la misma o ligera diferencia entre las profundidades, existen problemas técnicos en el modelo geoestadístico ya que este se basa en la matriz de correlaciones que depende de la distancia espacial entre puntos, lo que genera un perfil duplicado y hace que la matriz de correlación sea no invertible. Esto imposibilita la predicción espacial de la variable. En caso de que la diferencia sea mayor, genera inconsistencia en los datos. Por otro lado, estos puntos repetidos pueden ser útiles para evaluar el error del instrumento de medición.

### (c) Metodología de detección:

Este problema se analizó dentro de cada fuente de datos y se detectó una gran cantidad de datos repetidos que según el caso fueron eliminados o ponderados. La limpieza de datos en este proceso consta de:

- (I) Detectar los grupos de puntos repetidos.

- (II) Verificar la diferencia existente en cada uno de estos grupos.
- (III) Si la variabilidad dentro del grupo es demasiado grande, se elimina todo el grupo; caso contrario, promediamos el valor de profundidad en la ubicación repetida manteniendo solo un perfil de esta locación.

### 3. Detección de datos pseudo-repetidos:

#### (a) Descripción:

Existen datos que son muy diferentes, respecto a la profundidad, a distancias muy pequeñas (menores a la escala de la predicción). Computacionalmente, estos datos no permiten realizar la inversión de matrices y además generan ambigüedad, incertidumbre y sesgo en el modelo. Por ello, es necesario eliminarlos bajo el mismo criterio de la detección de datos repetidos. Aunque, técnicamente, el algoritmo cambio un poco.

#### (b) Desventajas:

Aunque estos no son puntos exactamente repetidos, es de conocimiento que los algoritmos computacionales trabajan con cierto nivel de exactitud. Lo cual hace que no se puedan realizar ciertas partes computacionales del modelamiento.

#### (c) Metodología de detección:

El proceso de detección es similar al de los puntos repetidos.

- (I) Detectar los grupos de puntos distanciado por una medida de  $k$  metros.
- (II) Verificar la variabilidad existente en cada uno de estos grupos.
- (III) Si la variabilidad dentro del grupo es demasiado grande, se elimina todo el grupo; caso contrario, promediamos las ubicaciones espaciales y el valor de profundidad en las ubicación pseudo-repetidas manteniendo solo un perfil del grupo.

### 4. Detección de datos atípicos:

#### (a) Descripción:

Un dato atípico espacial es aquel que se comporta de manera muy diferente a su vecindad. A través de la visualización de datos, se vio que existen datos atípicos espaciales. En el presente estudio, se utiliza este término para hacer referencia a errores que pudieron pasar en el trabajo de campo ya sea mediante dragas o cruceros científicos. Existen datos atípicos

que son obvios por la gran diferencia existente con su vecindad y por otro lado, hay otros en los que no es tan obvio, por ello necesitamos algoritmos adecuados como los que se mostrarán más adelante.

(b) Desventajas:

La presencia de datos atípicos u errores influye significativamente en la obtención de un mapa batimétrico porque generan artefactos como elevaciones o depresiones falsas muy pronunciadas, según sea el caso.

(c) Metodología de Detección de Datos Atípicos Espaciales: Enfoques conocidos y Propuesto

(I) Según los Vecinos más Cercanos

Sea un conjunto de datos geoestadístico  $y_1, y_2, y_3, \dots, y_n$  asociados a un espacio  $x_1, x_2, x_3, \dots, x_n$  respectivamente. Denotaremos  $NNk_i$  como la vecindad del dato  $y_i$  conformado por  $k$  veciones y  $g(NNk_i)$  como un estadístico de resumen de la vecindad  $NNk_i$ .

La detección de datos atípicos espaciales en base a su vecindad considera; (i) un estadístico que cuantifique la diferencia entre el dato y su vecindad  $f(g(NNk_i), y_i)$ , (ii) la distribución del estadístico  $f(\cdot)$  y (iii) un límite de tolerancia asociado a  $f(\cdot)$ .

(a) El estadístico:

La diferencia entre la media de los datos de la vecindad  $NNk_i$  y el valor de  $y_i$  es el estadístico de diferencia  $f(g(NNk_i), y_i)$  más básico. En tal caso  $g(NNk_i)$  vendría a ser  $\sum_{y_j \in NNk_i} \frac{y_j}{n}$ . El promedio aritmético es altamente afectado por los datos atípicos mientras que la mediana no lo es. Una mejora a la función  $f(\cdot)$  se da al considerar:

$$f(g(NNk_i), y_i) = med(NNk_i) - y_i \quad (3.1)$$

La función  $f(\cdot)$  en 3.1 genera resultados satisfactorios cuando la varianza del proceso estocástico es homogénea. Sin embargo, la varianza de la profundidad del fondo marino es heterogénea y es más alta en las zonas con mayor pendiente como en el talúd continental. Empleando la varianza de las vecindades como argumento de penalización, se puede obtener un mejor estadístico de diferencia

$f(\cdot)$  asumiendo que la relación entre la varianza y la variable es  $med(NNk) = made(NNk)^\lambda$ .

$$f(g(NNk_i), y_i) = \frac{med(NNk_i) - y_i}{made(NNk)^\lambda} \quad (3.2)$$

$\lambda$  es estimado mediante un ajuste de regresión.

(b) La distribución:

Obtener la distribución teórica de  $f(\cdot)$  en 3.2 no es una tarea sencilla, así que preferimos obtener la distribución empírica a través de un remuestreo.

(c) El límite de tolerancia:

Los límites de tolerancia son obtenidos mediante el remuestreo a un nivel de significancia  $\alpha$  a través de los percentiles. Si el terreno no es muy variable, se puede obtener resultados adecuados estandarizando la variable  $f(\cdot)$  y asumiendo que se comporta como una distribución normal estándar. Caso contrario es preferible considerar los percentiles del remuestreo y ciertas variaciones cuando las colas son muy pesadas debido a la gran cantidad de datos espaciales atípicos.

$$I.C. :< P_{\alpha/2} - k, P_{1-\alpha/2} + k > \quad (3.3)$$

Si el estadístico cae fuera del intervalo presentado en la ecuación 3.3, el dato es considerado atípico y es eliminado.

## (II) Por regresión robusta:

Cuando los datos son obtenidos de cruceros científicos, se puede aprovechar la continuidad de la toma de datos para detectar los datos atípicos según el recorrido del crucero. Se ajusta una curva de regresión robusta por zonas pequeñas y luego consideramos como datos atípicos a aquellos con los residuales más grandes; es decir, a aquellos que se alejan de la superficie ilustrada por la mayoría de puntos.

$$Y = F(X, \beta) + E \quad (3.4)$$

en donde  $Y$  es la profundidad del fondo marino,  $X$  representa las variables independientes que pueden ser una suma de funciones polinomiales de las coordenadas de

Latitud y Longitud para nuestro caso,  $E$  representa el término del error y  $\beta$  es el parámetro a ser estimado.

El método de regresión robusta consta de ajustar una curva de regresión a una superficie de tal modo que la regresión no se vea afectada por datos atípicos y grupos de datos atípicos. Por ello el estimador  $\beta_R$  de  $\beta$  no debe ser afectado por datos atípicos.

$$\beta_R = h(X_R) \quad (3.5)$$

Existen diversos modelos de regresión robusta, no hacemos énfasis en alguno porque, en general, la mayoría de los modelos cumplen con nuestro principal requerimiento (que no se vea afectado por un grupo de datos atípicos) y asumimos que el lector puede obtener  $\beta_R$ .

Con ayuda del estimador  $\beta_R$ , despejamos  $E$  de la Eq. 3.4.

$$\hat{E} = F(X, \beta_R) - Y \quad (3.6)$$

$\hat{E}$  puede ser estandarizado ( $E^* = \frac{\hat{E} - \text{median}(\hat{E})}{\text{mad}(\hat{E})}$ ) para facilitar la definición de los límites de tolerancia pudiendo ser 2,3,4,.. como límite superior dependiendo del nivel de confianza deseado.

(d) Aplicación:

Mediante las técnicas mostradas anteriormente, se detecta los datos atípicos espaciales; el método de vecindades se aplican a fuentes como DHN e IMARPE cuando los datos fueron obtenidos por dragas y empleamos la detección de datos atípicos mediante regresión robusta en caso de que los datos provengan de cruceros científicos con haz simple y múltiple.

### 3.4. Metodología del Modelamiento

Después de realizar el procesamiento de los datos y un análisis exploratorio de datos, continuamos con el modelamiento de la batimetría y de la predicción de los depocentros. Los cuales involucran técnicas estadísticas y computacionales que consideran el factor espacial, estos se muestran a continuación con el esquema general empleado para modelar la batimetría y la ubicación de

depocentros.

### 3.4.1. Modelamiento de la Batimetría

El primer análisis y predicción realizado es con respecto a la batimetría. Para plantear un esquema general del modelamiento ha de considerarse ciertos aspectos que deben ser contemplados por esta metodología.

(A) Consideraciones:

(a) Gran cantidad de datos:

Los modelos geoestadísticos que emplean el método de máxima verosimilitud son costosos computacionalmente al compararlos con algoritmos simples de interpolación, sin embargo el resultado es más favorable y nos da la posibilidad de cuantificar el error de predicción.

Ante este tema, nace la necesidad de encarar el modelamiento de un área tan grande con gran cantidad de datos.

(b) No estacionariedad

El relieve marino tiene un comportamiento no estacionario, gobernado por tendencias dependientes de la latitud, longitud y distancia a las costas.

(c) Anisotropía

Existe efectos de anisotropía porque la correlación espacial se mantiene más prolongadamente a través de la latitud y se pierde a través de la longitud.

(d) Líneas de la costa muy irregular

Las líneas costeras peruanas son muy irregulares haciendo que la tendencia estacionario y los efectos direccionales de dependencia espacial varíen constantemente en función de la irregularidad de la línea costera.

(e) Datos de distintas fuentes de datos

Otro factor importante es el hecho de que los datos provienen de diferentes tipos de datos y por lo tanto presentan distinta confiabilidad. A diferencia de los problemas anteriores, este no ha sido considerado en modelos geoestadísticos y es motivo por lo que presentamos una propuesta para estas situaciones en la sección de Técnicas de Análisis.

(B) Abarcando las consideraciones:

(a) Gran cantidad de datos:

Debido a la gran cantidad de datos, el modelamiento se realiza a través de ventanas móviles.

(b) No estacionariedad:

Se consideran covariables para ajustar la tendencia no estacionaria. Y trabajamos con los residuales para el modelo espacial.

(c) Anisotropía:

Se realiza la transformación de coordenadas para obtener correlaciones espaciales isotrópicas.

(d) Líneas de la costa muy irregular:

Se considera clusters espaciales, de tal modo que se evalúa la tendencia no estacionaria y el factor de anisotropía en cada una de ellas.

(e) Datos de distintas fuentes de datos:

La correlación espacial propuesta contempla la diferencia entre la fuentes de datos a través de la confiabilidad, mediante un factor de penalización.

(C) Esquema de modelamiento

El esquema del modelamiento de la batimetría contempla las consideraciones anteriores y se describe a continuación según la Fig. 3.8. En primer lugar estandarizamos el formato de los datos y limpiamos las bases de datos mediante el pre-procesamiento y procesamiento de los datos. Cuando obtenemos datos unificados, coherentes y limpios empezamos con el criterio del modelamiento. Segundo, como el área bajo estudio es muy amplia y muy variante, conformamos cluster espaciales cubriendo todo el área de estudio para empezar a analizarlos espacialmente uno por uno. Tercero, para cada cluster, realizamos un conjunto de procesos para obtener el modelo que caracterice a tal grupo (estos procesos se muestran de color verde en la Fig. 3.8). Después de definir el modelo, lo aplicamos mediante ventanas móviles cubriendo todo el área del cluster y almacenamos las grillas predichas en una general definida inicialmente. Este proceso se repite hasta obtener la predicción espacial del fondo marino en todos el alcance espacial de cada cluster.

Dentro del proceso que implica obtener el modelo para cada cluster (procesos de color verde de la Fig 3.8 ), iniciamos con la limpieza de datos entre fuentes de datos. Luego tomamos una muestra para obtener parámetros del ajuste al componente no estacionario. Se modela espacialmente los residuales obteniendo parámetros iniciales del modelo para luego ajustar un modelo geoestadístico y almacenarlo. Una vez obtenido el modelo, lo actualizamos y aplicamos para cada ventana móvil.

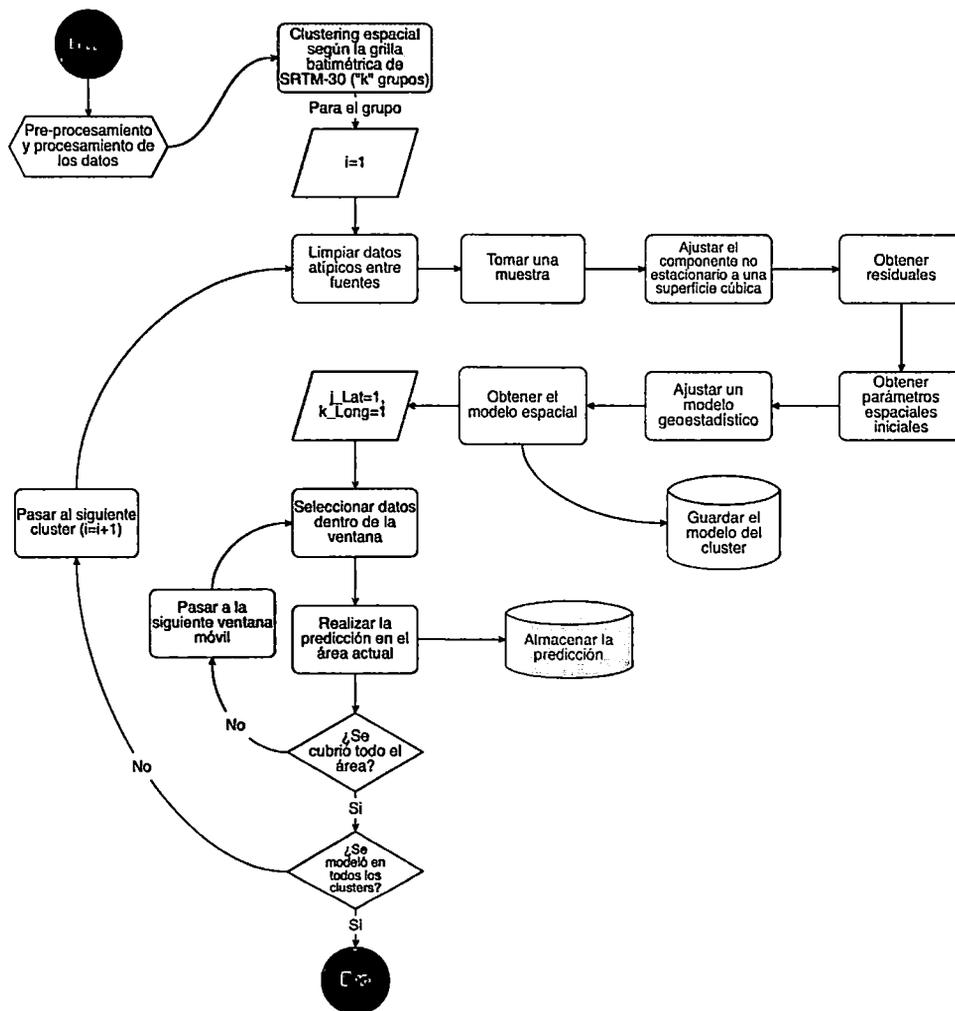


Figura 3.8: Esquema del modelamiento de la batimetría

### 3.4.2. Modelamiento de la Detección de Ambientes Deposicionales

Para el estudio, se cuenta con información de datos de la profundidad marina espaciados irregularmente y datos de sedimentología. Con ellos, no podemos modelar la presencia de datos directamente, así que es necesario hacer un poco de análisis y modelamiento para emplearlos con el fin de detectar posibles depocentros.

El esquema general se muestra en la Fig. 3.9. Empezando con los datos primarios de batimetría, empleamos técnicas de modelamiento geoestadístico con datos provenientes de distintas fuentes obteniendo una grilla batimétrica de alta resolución con su respectiva grilla de error. Al obtener la grilla completa, es posible caracterizar la morfología del relieve marino mediante el cálculo de índices morfométricos (principalmente), entre otros.

Luego, con los datos de sedimentación, aplicamos el método de clustering con el fin de encontrar posibles perfiles de sedimentación y entre uno de ellos el perfil de depocentro que deseamos (ambientes de deposición continua). Y generamos una variable respuesta en función a un proceso de bernoulli (es o no un depocentro)

Finalmente, se propone un modelo logístico espacial para caracterizar los depocentros sedimentarios según las características morfométricas para ser aplicado en la predicción del proceso bernoulli (presencia del depocentro) en el margen continental peruano.

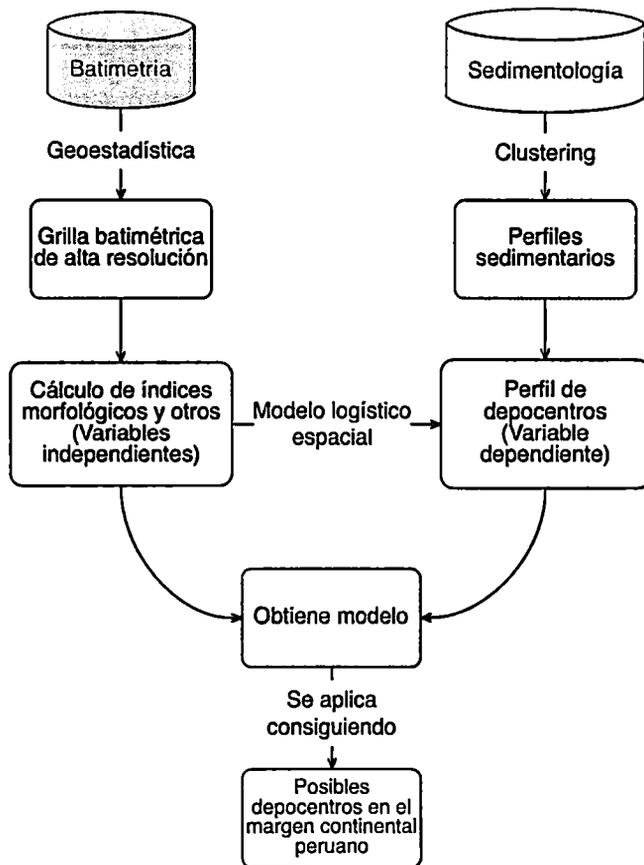


Figura 3.9: Esquema del modelamiento general del estudio

# Capítulo 4

## Técnicas de Análisis

### 4.1. Geoestadística

#### 4.1.1. Introducción Histórica

La Geoestadística es la sub-rama de la Estadística Espacial, en el cual los datos consisten de una muestra finita de medidas relacionadas a un fenómeno espacialmente continuo. Por ejemplo, la profundidad en un sondeo batimétrico. Los avances de la Geoestadística no se dieron exactamente en la teoría estadística sino en distintas áreas de aplicación. En agricultura, Fisher aportó con su pensamiento espacial de aleatoriedad y bloques. En minería, Matheron y sus colegas desarrollaron la mayoría de la teoría clásica de Geoestadística. Y en forestales, que fue el área de elección para la tesis de doctorado de Matheron (Considerado el padre de la Geoestadística) [Gelfand et al., 2010].

La Geoestadística<sup>1</sup> clásica tuvo sus inicios en la aplicación de la distribución log-normal en la minería del oro por H.S. Sichel. Esto fue seguido por la valiosa contribución de D.G. Krige en la aplicación del análisis de regresión. Luego consolidado por G. Matheron en su tesis doctoral de 1965. Más adelante hubieron aportes para la geoestadística no lineal, el cambio de soporte, la función de transferencia, selectividad de distribuciones y otros [Matheron and Kleingeld, 1987].

En la actualidad, todavía se están desarrollando mayores aportes a esta rama incorporando perspectivas bayesianas, modelos mixtos, anidados, entre otros y casos particulares de aplicación tal como en esta tesis se desea mostrar (una herramienta para aplicar geoestadística en datos prove-

---

<sup>1</sup>Denominada como tal por Matheron en 1962

nientes de diversas fuentes con distinta confiabilidad).

#### 4.1.2. Descripción General

Imagine que estamos interesados en conocer la forma de la topografía del fondo marino en un área cuadrada de 100 metros de latitud por 100 metros de longitud. Para ello, mediante trabajos *in situ*, se obtuvo los niveles de profundidad  $y_1, y_2, y_3, \dots, y_{10}$  en 10 ubicaciones puntuales  $x_1, x_2, x_3, \dots, x_{10}$ . La idea de conocer la topografía, indica que estamos interesados de obtener el valor de la profundidad en una ubicación arbitraria  $x_0$  en base a los datos disponibles de la muestra  $y_1, y_2, y_3, \dots, y_{10}$  (Fig. 4.1).

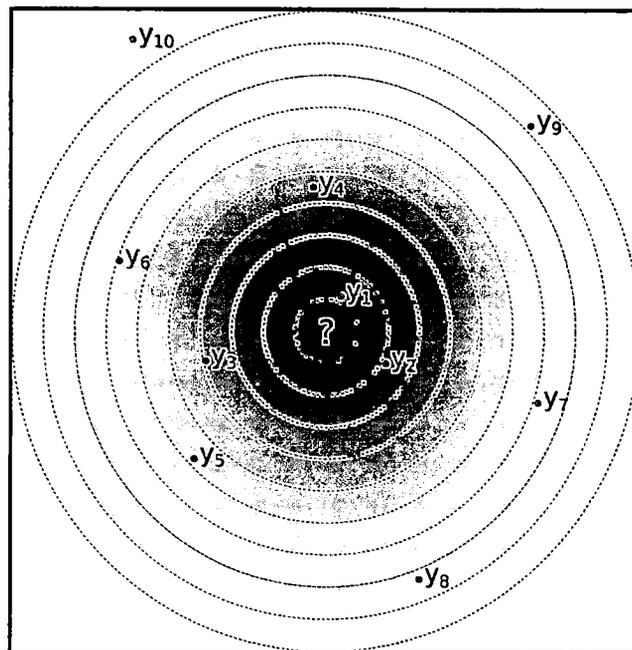


Figura 4.1: Ejemplo de Aplicación de Geoestadística. El signo de interrogación representa la ubicación en donde se desea estimar el valor de la profundidad ( $y_0$ ); los  $y_i$ , para  $i = 1, 2, 3, \dots, 10$ , representan los datos obtenidos del trabajo de campo. La degradación del color del fondo significa que a mayor distancia, existe menor similitud entre el dato  $y_i$  y el valor de la profundidad de interés ( $y_0$ ).

Antes de presentar el enfoque regularmente usado para este tipo de situaciones, consideremos las características que podría tener la topografía del fondo marino en cierta zona. Generalmente presenta una tendencia global que es decreciente según nos vamos alejando de las costas y localmente podemos encontrar elevaciones, depresiones, trincheras entre otros accidentes geográficos. Estas características no solo las posee la topografía; sino que muchas variables de la naturaleza también tienen este tipo de comportamiento, es decir que presentan un patrón global (constante, lineal, cuadrática,

entre otros), un patrón a nivel local (comportamiento espacial), además de un componente aleatorio que rige en la naturaleza.

En primer lugar, asumamos que el patrón global es constante (estacionario) y que el componente aleatorio es despreciable; en tal caso, nuestro problema se reduce a entender el patrón local existente que para nuestro ejemplo vienen a ser las elevaciones, depresiones, quebraduras y ciertos accidentes geográficos. Además, se puede decir que tiene un comportamiento continuo en el sentido que **las profundidades de dos ubicaciones cercanas deben ser semejantes** y presenta un comportamiento espacial (*Dos objetos cercanos son más similares que dos a mayor distancia*).

Según estas características, el problema puede ser resuelto estimando  $y_0$  a partir de los datos muestrales  $y_1, y_2, y_3, \dots, y_{10}$ . Es decir,  $\hat{y}_0 = g(y_1, y_2, y_3, \dots, y_{10})$ . Generalmente se considera un estimador lineal, para tal caso  $g(\cdot)$  puede ser representado como una combinación lineal de los datos muestrales,  $g(y_1, y_2, y_3, \dots, y_{10}) = \sum \lambda_i \cdot y_i$  para finalmente obtener:

$$\hat{y}(x_0) = \sum_{i=1}^n (\lambda_i y(x_i)) \quad (4.1)$$

A este momento nace la pregunta de como estimar los valores de  $\lambda_i$ , que aspectos debemos considerar o si hay diferencias al tratar esta situación con estadística clásica u otras. En primer lugar, los  $\lambda_i$  pueden ser considerados como el peso que aporta cada dato de la muestra a la estimación de  $y_0$  y segundo, ha de aclarar que los términos de comportamiento local y correlación espacial mencionados en las características de la topografía del fondo marino son conceptos importantes para determinar los valores de  $\lambda_i$ . De este modo, según nuestro ejemplo (Fig. 4.1), se puede esperar que  $(|\lambda_1| > |\lambda_2| > |\lambda_3| > \dots > |\lambda_{10}|)^2$  ya que el nivel de profundidad en  $x_1$  es más parecido al nivel de  $x_0$  que  $x_2$  a  $x_0$ , que el nivel de profundidad en  $x_2$  es más parecido al nivel de  $x_0$  que  $x_3$  a  $x_0$ , que el nivel de profundidad en  $x_3$  es más parecido al nivel de  $x_0$  que  $x_4$  a  $x_0$ , y así sucesivamente.

A este punto, nos preguntamos en como medir ese nivel de similaridad entre puntos más cercanos y puntos más lejanos. Existen diversas técnicas de interpolación (incluyendo geoestadística) para realizar esta fase y obtener los  $\lambda_i$  tales como la interpolación por medias móviles, medianas móviles, distancia inversa ponderada, spline, etc. Particularmente, la geoestadística considera el modelamiento de la estructura espacial por medio de funciones de varianza y correlación espacial permitiendo mayor flexibilidad para adaptar estas funciones a la situación bajo estudio.

---

<sup>2</sup>Es algo de esperar, pero en geoestadística, ello depende también de la redundancia de la información.

Por ahora solo presentaremos la función de correlación, esta tiene una caída según la distancia, entre pares de puntos, se va incrementando y ello hace que los puntos más cercanos al área de interés tengan mayor influencia sobre la estimación (mayor  $|\lambda_i|$ ).

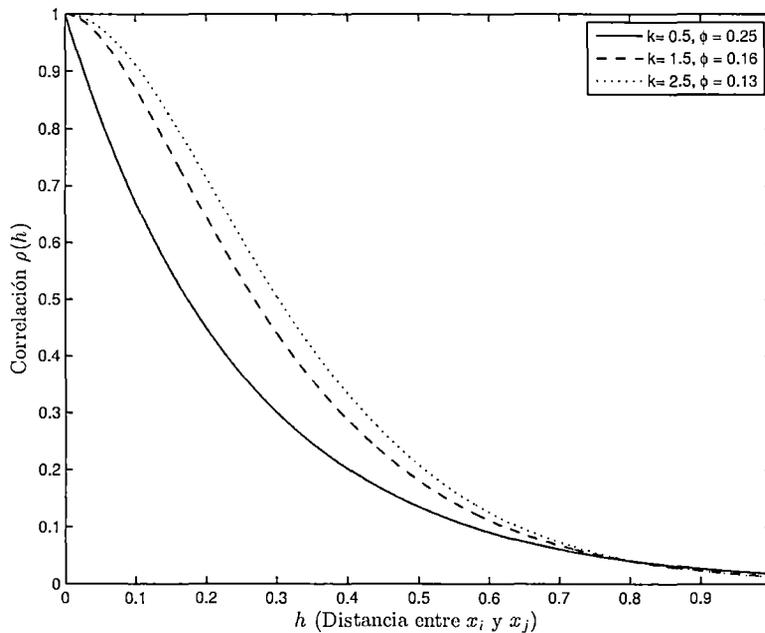


Figura 4.2: Correlación de Matern.

En resumen, la aplicación de geoestadística consta de dos fases; el primero, es un análisis de la estructura de correlación espacial mediante las funciones de varianza o correlación espacial y segundo, la interpolación en base a los parámetros estimados de de la estructura de correlación de tal modo que obtengamos los valores  $\lambda_i$  y por lo tanto la estimación de la variables bajo estudio en cualquier parte de la región de interés. En las próximas secciones, daremos conceptos más formales y mayores detalles del modelamiento geoestadístico.

### 4.1.3. Terminología y Notaciones

#### 1. Variable Aleatoria:

Una *variable aleatoria*  $Y$  con valores en el conjunto  $E$  es una función que asigna un valor  $Y(\omega)$  para cada resultado  $\omega$  en el espacio muestral  $\Omega$  [Cinlar, 2013].

#### 2. Proceso estocástico:

Un proceso estocástico con espacio de estado  $E$  es una colección  $Y_t; t \in T$  de variables aleatorias

$Y_t$  definidas en el mismo espacio de probabilidad y tomando valores en  $E$  [Cinlar, 2013].

### 3. Variable Regionalizada:

Variable medida en el espacio de forma que presenta una estructura de correlación. Es un proceso estocástico con dominio contenido en un espacio euclidiano  $d$ -dimensional  $R^d$ ,  $S(x) : x \in D \subset R^d$ .

Si  $d = 2$ ,  $Z(x)$  puede asociarse a una variable medida en un punto  $x$  del plano [Giraldo, 2002].

### 4. Estacionariedad:

Un proceso es estacionario si el esperado de  $S(x)$  es el mismo para todo  $x$ , la varianza de  $S(x)$  es el mismo para todo  $x$  y la correlación entre  $S(x)$  y  $S(x')$  depende solamente de  $u = \|x - x'\|$  [Diggle and Ribeiro, 2007].

$$E(S(x)) = u, V(S(x)) = \sigma^2, Corr(S(x), S(x+h)) = C(h) \quad (4.2)$$

Este mismo concepto se maneja en Series de tiempo, solo que la dirección es tal caso es única mientras que en Geoestadística existen múltiples direcciones. En Geoestadística, el variograma es una herramienta útil para verificar la estacionariedad mientras que en Series de Tiempo se suele emplear el correlograma.

### 5. Isotropía y Anisotropía:

Un proceso estocástico espacial es isotrópico si la correlación entre los datos no depende de la dirección en que se calcula y será anisotrópico si dicha correlación depende de la dirección hacia cual se realiza el cálculo. La forma más simple de efecto direccional en la estructura de correlación es denominada *anisotropía geométrica* [Diggle and Ribeiro, 2007], la cual se considera para este trabajo. Problemas respecto a la anisotropía zonal o combinada no se toma en cuenta, para ello puede verse el planteamiento de Isaaks y Srivastava [Isaaks, 1989].

Visualmente podemos detectar cuando un proceso estocástico  $S(x)$  es isotrópico o no. Se puede decir que cuando  $S(x)$  es isotrópico se ve que la dependencia espacial tiene un área circular de influencia mientras que en el caso contrario, se vé que el área de influencia es como un elipse, en el que existe una dirección principal (la dependencia decae de manera más lenta) y una dirección secundaria (la dependencia decae de manera rápida). En la figura 4.3 se muestra un

proceso isotrópico a la izquierda y uno anisotrópico a la derecha, fácilmente puede verse que el área de influencia del primero es circular mientras que en el otro es elipsoide.

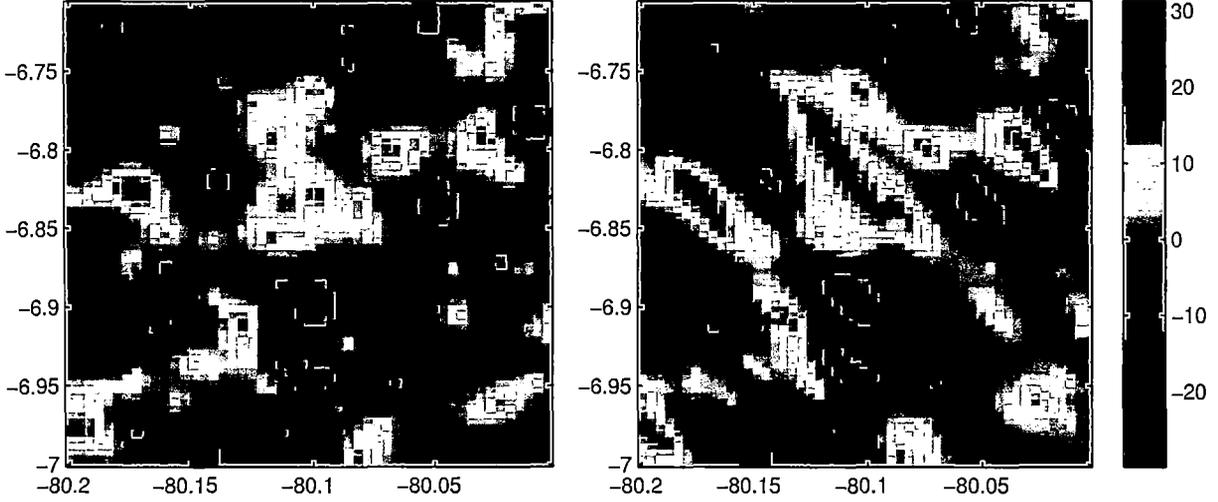


Figura 4.3: Proceso isotrópico (izquierda) y proceso anisotrópico (derecha).

La diferencia entre un proceso isotrópico y otro anisotrópico se ve de manera más clara en la correlación espacial. En la Fig. 4.4 se muestran funciones de correlación espacial isotrópica y anisotrópica correspondientes a los procesos mostrados en la Fig. 4.3. Se ve que el área de dependencia es circular cuando el proceso es estacionario y elipsoide en caso contrario. De esta manera, la correlación anisotrópica decrece de manera rápida en una dirección, mientras que en la dirección perpendicular decae de manera más lenta.

Ante la situación de anisotropía geométrica, Diggle y Ribeiro proponen realizar una transformación para convertir el área de influencia elíptica a un área circular [Diggle and Ribeiro, 2007]. Dos parámetros son suficientes para realizar tal transformación; el ángulo de anisotropía  $\psi_A$  que indica la dirección en donde se prolonga la dependencia espacial y la razón de anisotropía  $\psi_R$ , representando la razón entre el rango de dependencia en la dirección principal y secundaria (perpendicular a la dirección principal).

$$(\mathbf{x}'_1, \mathbf{x}'_2) = (\mathbf{x}_1, \mathbf{x}_2) \begin{pmatrix} \cos(\psi_A) & -\sin(\psi_A) \\ \sin(\psi_A) & \cos(\psi_A) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \psi_R^{-1} \end{pmatrix} \quad (4.3)$$

La Fig. 4.5 muestra el proceso anisotrópico de la Fig. 4.3 en la parte izquierda y a la derecha

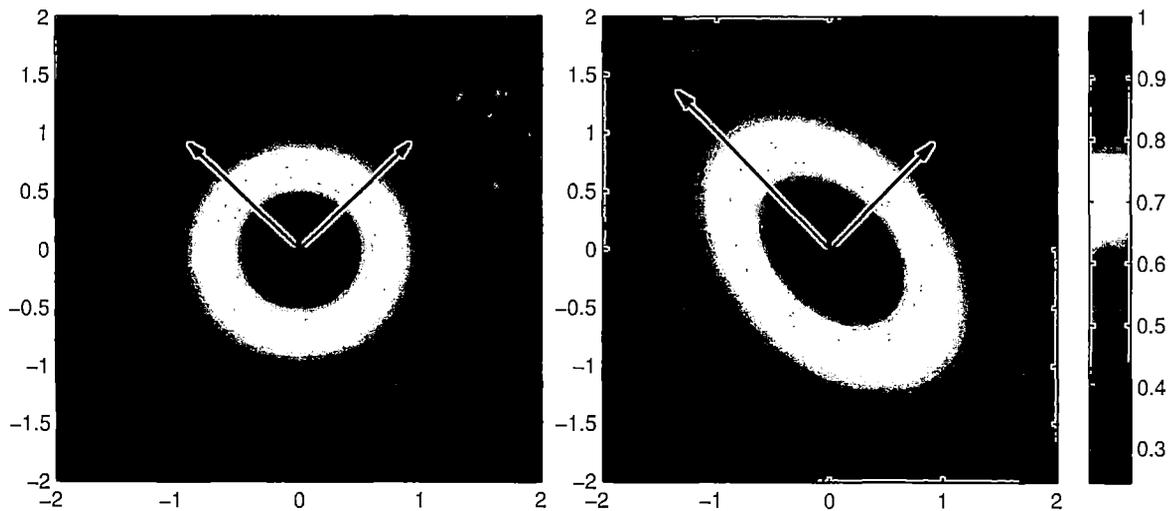


Figura 4.4: Correlación para un proceso isotrópico (izquierda) y un proceso anisotrópico con  $\theta = 135$  (derecha). Las flechas negras indican las direcciones principales.

se muestra la transformación planteada en 4.3. El proceso estocástico con coordenadas transformadas ahora muestra un comportamiento relativamente isotrópico.

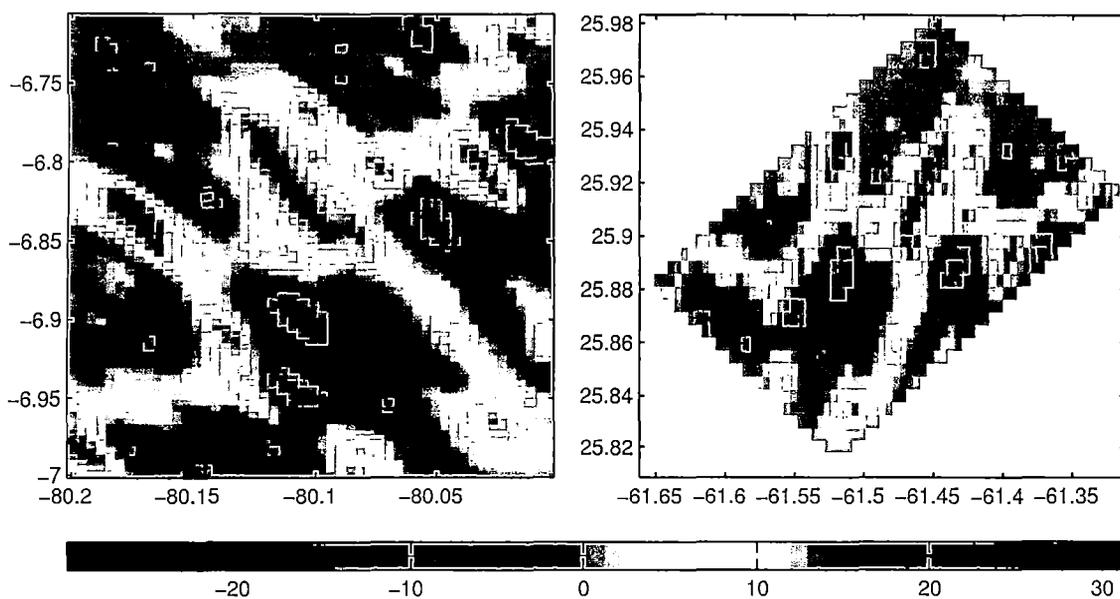


Figura 4.5: Proceso anisotrópico (izquierda) y Proceso con coordenadas transformadas - Isotrópico (derecha)

Esta manera de tratar la anisotropía geométrica es sencilla considerando los parámetros de ángulo  $\psi_A$  y razón  $\psi_R$  como conocidos, de manera a-priori; sin embargo, en la aplicación no

conocemos de dichos parámetros lo que conduce a la estimación de dos parámetros adicionales en caso de considerar la anisotropía. También ha de considerarse que si realizamos la transformación de coordenadas, debemos devolverlas a su escala original después del kriging.

#### 6. Distancia:

Una medida importante en la Geoestadística es la distancia, ya que el variograma y la correlación espacial de un proceso estacionario sólo depende de la distancia entre las ubicaciones. Generalmente, se emplea la distancia euclidiana, sin embargo dependiendo de la situación esta puede ser definida de una manera adecuado. Cuando se trabaja con datos de latitud y longitud a gran escala es recomendable emplear la distancia esférica ya que la Tierra tiene aproximadamente esa forma. Waller y Gotway definen varios tipos de distancias y realizan un mayor análisis de su importancia [Waller and Gotway, 2004].

#### 4.1.4. Correlación Espacial

El análisis de la correlación espacial, es la primera etapa de un análisis geoestadístico, algunos lo denominan estructura espacial (*Dos objetos cercanos son más similares que dos a mayor distancia*). Se realiza esta etapa porque se considera que los datos son autocorrelacionados espacialmente de acuerdo a la distancia entre las observaciones. Esta etapa enfatiza la diferencia con un análisis estadístico clásico, porque en este se considera que las observaciones son independientes. La correlación espacial puede ser evaluada por funciones como la semivarianza, covarianza y correlación, las cuales pueden ser vistas por medio de las herramientas variograma, covariograma y correlograma. A continuación presentaremos estos conceptos a mayor detalle.

##### 1. Funciones de Correlación o Estructura Espacial:

Como se mencionó anteriormente, estas funciones nos ayudan a entender la dependencia espacial existente en el fenómeno bajo estudio, e incluso nos pueden ayudar a verificar la existencia de la estructura espacial de manera visual y otras características de la estructura espacial como estacionariedad e isotropía.

##### a) Variograma:

##### 1) Variograma Teórico:

El variograma teórico es la varianza de los incrementos de una variable regionalizada la cual caracteriza la dependencia de segundo orden en un proceso estocástico espacial  $S(x)$  [Diggle and Ribeiro, 2007].

$$V(x, x') = V(S(x+h) - S(x)) \quad (4.4)$$

En donde,  $S(x)$  es el valor de la variable en una posición  $x$  y  $S(x+h)$  es el valor a una distancia  $h$  de  $x$ .

La función del variograma puede ser descompuesta del siguiente modo:

$$V(x, x') = \frac{1}{2}[Var\{S(x)\} + Var\{S(x+h)\} - 2.Cov\{S(x), S(x+h)\}]$$

En caso de que el proceso sea estacionario se cumple que:

$$V(h) = \sigma^2\{1 - \rho(h)\} \quad (4.5)$$

Por lo que, para un proceso estocástico espacial estacionario, el variograma debe ser creciente ya que  $\{1 - \rho(h)\}$  es creciente porque  $\rho(h)$  es decreciente mientras  $h$  se incrementa. Además el variograma debe alcanzar un límite según la ecuación 4.5. Este límite o la altura del variograma es igual a la varianza poblacional  $\sigma^2$  [Barnes, 1991].

## 2) Variograma Empírico:

Es la colección de pares de distancias y sus correspondientes ordenadas del variograma  $(u_{ij}, v_{ij}) : j > i$  para el conjunto de datos  $(x_i, Y_i) : i = 1, \dots, n$  [Diggle and Ribeiro, 2007]. En donde:

$$v_{ij} = \frac{1}{2}E[(Y_i - Y_j)^2] = \frac{1}{2n} \sum_1^n (y_i - y_j)^2 \quad (4.6)$$

El variograma empírico ha sido la herramienta más utilizada desde los inicios de la Geoestadística tanto como herramienta visual como para la estimación de los parámetros de la estructura espacial, al día de hoy, empiezan a considerarla como una herramienta del análisis exploratorio de los datos espaciales.

## b) Covarianza Espacial

La covarianza es otra función que puede ayudar a observar las características de la es-

estructura espacial. Es parte de la descomposición del variograma.

$$C(h) = COV(S(x), S(x+h)) = E[\{S(x+h) - E(S(x+h))\}\{S(x) - E(S(x))\}] \quad (4.7)$$

Si el proceso estocástico no tuviera un patrón espacial, la covarianza espacial sería igual a 0.

**c) Correlación Espacial:**

Para la presente tesis, la correlación espacial es la función más importante para la aplicación de los modelos geoestadísticos tanto en la estimación de la estructura espacial como en la predicción espacial por medio del Kriging.

Inicialmente, podemos definir a la correlación espacial como  $\rho(h) = corr(S(x), S(x+h))$ . Existe una relación interesante entre las tres funciones presentadas (variograma, covariograma y correlograma), en la ecuación 4.8 se muestra la correlación en función de la covarianza.

$$\rho(h) = \frac{C(h)}{C(0)} \quad (4.8)$$

En caso de que el proceso estocástico  $S(x)$  sea estacionario, la correlación puede deducirse de la ecuación 4.5.

$$\rho(h) = \frac{\sigma^2 - V(x, x+h)}{\sigma^2} \quad (4.9)$$

Cuando  $h \rightarrow 0$ ,  $V(x, x+h) \rightarrow 0$  lo que hace que  $\rho(h) \rightarrow 1$ . Por otro lado, mientras  $V(x, x+h) \rightarrow \sigma^2$  cuando  $h \rightarrow \infty$ , la correlación  $\rho(h) \rightarrow 0$ .

**2. Modelos Teóricos:**

Los modelos teóricos son empleados para ajustarse a los datos y poder hacer un cálculo de la matriz de correlación  $R$  en base a esta función ya que solo debe depender de  $h$  en los casos estacionario. Los modelos presentados a continuación solo son para procesos estocásticos estacionarios.

**a) Funciones de Correlación Espacial:**

La condición necesaria y suficiente de una función de correlación es que sea definida positiva. Sin embargo no es fácil comprobar este supuesto directamente, por ello es conveniente tener a la mano funciones de correlación espacial conocidas, que cumplan con esta condición y además sean flexibles para su aplicación en diferentes situaciones [Diggle and Ribeiro, 2007].

1) La familia Matern

El comportamiento más común de una estructura de covarianza estacionaria es que la correlación entre  $S(x)$  y  $S(x')$  decaiga mientras la distancia  $h = \|x - x'\|$  se incrementa. También se espera que según los parámetros de la función de correlación, la función exhiba diferentes grados de suavidad en el proceso estocástico adyacente [Diggle and Ribeiro, 2007].

La familia Matern cumple con estos requerimientos y consta de dos parámetros.

$$\rho(h) = \{2^{k-1}\Gamma(h)\}^{-1} \left(\frac{h}{\phi}\right)^k K_k\left(\frac{h}{\phi}\right) \quad (4.10)$$

en donde  $K_k(\cdot)$  representa una función Bessel modificada de orden  $k$ ,  $\phi > 0$  es un parámetro de escala con las dimensiones de distancia, y  $k > 0$ , llamado el orden, es un parámetro de forma el cual determina la suavidad de proceso estocástico  $S(x)$  [Diggle and Ribeiro, 2007].

Actualmente, la familia Matern es una de las más usadas debido a la flexibilidad que tienes para obtener formas suavizadas y a las distintas funciones que abarca como la función exponencial y gaussiana. Dependiendo de el parámetro  $k$  pueden obtenerse algunas funciones ya conocidas: cuando  $k = 0,5$ , se obtiene la función de correlación exponencial; cuando  $k = 1$ , se obtiene la función propuesta por Whittle; cuando  $k \rightarrow \infty$ , obtenemos la función Gaussiana.

2) La familia exponencial con parámetro de potencia

La función de correlación exponencial se ajusta muy bien a ciertas situaciones, cuando la caída de la función es monótona según la distancia se incrementa. La familia exponencial con un parámetro de potencia se define del siguiente modo:

$$\rho(h) = \exp\left\{-\left(\frac{h}{\phi}\right)^k\right\} \quad (4.11)$$

Esta función tiene un parámetro de escala  $\phi > 0$  y un parámetro de forma  $k$  limitado por  $0 < k \leq 2$ . En el caso extremo de  $k = 2$  se obtiene un equivalente a la función de Matérn cuando  $k \rightarrow \infty$ .

### 3) Otras familias

Además de las familias ya mencionadas, existen otras funciones que son de utilidad dependiendo del fenómeno bajo estudio. Sin embargo, estas no tienen tanta flexibilidad como las anteriores.

- La función cilíndrica:

Esta función decrece rápidamente cerca del origen pero decae más lento mientras la distancia se va haciendo mayor. Presenta un rango finito y su expresión matemática es del siguiente modo:

$$\rho(h) = \begin{cases} 1 - \frac{3}{2}\left(\frac{h}{\phi}\right) + \frac{1}{2}\left(\frac{h}{\phi}\right)^3 & : 0 \leq h \leq \phi \\ 0 & : h > \phi \end{cases} \quad (4.12)$$

El hecho de que presente un rango finito es de mucha ventaja cuando se trabaja con gran cantidad de datos y con matrices dispersas [Sun et al., 2012].

- La función cubica:

Decae de manera lenta en el origen y cuando se aproxima al rango.

$$\rho(h) = \begin{cases} 1 - 3\left(\frac{h}{\phi}\right)^2 + 2\left(\frac{h}{\phi}\right)^3 & : 0 \leq h \leq \phi \\ 0 & : h > \phi \end{cases} \quad (4.13)$$

- La función lineal:

La función lineal decae a una razón constante para cualquier valor de la distancia hasta que alcanza el rango y la correlación se convierte en cero, tiene un rango finito.

$$\rho(h) = \begin{cases} 1 - \left(\frac{h}{\phi}\right) & : 0 \leq h \leq \phi \\ 0 & : h > \phi \end{cases} \quad (4.14)$$

- La función spline:

Esta función decae rápidamente y su forma es muy semejante a la función lineal

solo que este si presenta puntos de inflexión a distancias cortas y cuando se aproxima a su rango finito.

$$\rho(h) = \begin{cases} 1 - 15\left(\frac{h}{\phi}\right)^2 + 30\left(\frac{h}{\phi}\right)^3 & : 0 \leq \left(\frac{h}{\phi}\right) \leq 0,2 \\ 1,25\left(1 - \left(\frac{h}{\phi}\right)\right)^3 & : 0,2 \leq \left(\frac{h}{\phi}\right) < 1 \\ 0 & : \left(\frac{h}{\phi}\right) \geq 1 \end{cases} \quad (4.15)$$

En general, cada una de estas funciones tiene sus ventajas en la flexibilidad de poderse ajustar a distintas situaciones y también presentan su complejidad al momento de realizar cálculos con la matriz de correlación  $R$ . Las familias de Matérn y exponencial se pueden ajustar a una amplia gama de situaciones pero estas generan matrices de correlaciones pesadas ya que no presentan un rango práctico finito. Por otro lado, la función esférica, lineal, spline y cúbica si producen matrices de correlación  $R$  dispersas que facilitan los cálculos con la matriz  $R$  permitiendo optimizar el tiempo y trabajar con grandes bases de datos [Sun et al., 2012].

#### 4.1.5. Estimación Gaussiana de Parámetros

El formato básico de una base de datos geoestadística es  $(x_i, y_i); i = 1, \dots, n$ ; en la que  $x_i$  representa la ubicación y  $y_i$  es un valor escalar asociado a la ubicación  $x_i$ . Además  $y_i$  es considerada como la variable medida o respuesta. Cada  $y_i$  es una realización de  $Y_i$  cuya distribución es independiente de los valores del proceso estocástico adyacente  $S(x)$ , en las ubicaciones  $x_i$ , el cual no es directamente observable. Es preferible mantener la diferencia entre  $Y_i$  y  $S(x_i)$  [Diggle and Ribeiro, 2007]. La forma básica de un modelo geoestadístico incorpora al menos dos elementos:

- Un proceso estocástico evaluado en los reales  $\{S(x) : x \in A\}$ .
- Una distribución multivariada para la variable aleatoria  $Y = (Y_1, Y_2, \dots, Y_n)$  condicional en  $S(\cdot)$ .  $Y_i$  puede ser considerado como una versión ruidosa de la señal  $S(x_i)$ .

La gran diferencia entre la estimación de parámetros para un modelo gaussiano y otros se debe al supuesto que  $S = (S_1, S_2, \dots, S_n)$  sigue una distribución multivariada normal.

Un modelo geoestadístico general puede ser presentado del siguiente modo:

$$Y_i = u(x_i) + S(x_i) + Z(x_i) \quad (4.16)$$

En donde  $Y_i$  representa la variables observada o medida,  $u(x_i)$  representa la tendencia existente en el proceso estocástico espacial bajo estudio  $S'(x_i) = u(x_i) + S(x_i)$ ,  $S(x_i)$  es el componente estacionario de  $S'(x_i)$  que mantiene una estructura espacial y  $Z(x_i)$  es el error aleatorio.

### 1. Proceso Estacionario:

Cuando el proceso  $S'(x_i)$  es estacionario,  $u(x_i)$  es constante  $u$  para cualquier valor de  $x_i$  de tal manera que el modelo de la ecuación 4.16 se reduce a  $Y_i = u + S(x_i) + Z(x_i)$  lo cual, sin pérdida de generalidad, se reduce a:

$$Y'_i = S(x_i) + Z(x_i) \quad (4.17)$$

Como  $S(x) = (S(x_1), \dots, S(x_n))$  de distribuye como una normal multivariada con media 0 y matriz de varianzas  $\Sigma$ ,  $Y'_i$  tiene distribución normal multivariada con media 0 y matriz de varianzas  $\sigma^2 R(\phi) + \tau^2 I$ .

La función de log-verosimilitud es:

$$L(\tau^2, \sigma^2, \phi) = -\frac{1}{2} \{n \log(2\phi) + \log\{|\sigma^2 R(\phi) + \tau^2 I|\}\} + y^T (\sigma^2 R(\phi) + \tau^2 I)^{-1} y \quad (4.18)$$

Considerando la matriz de Correlación de  $Y$  ( $R_Y = \frac{(\sigma^2 R(\phi) + \tau^2 I)}{\sigma^2 + \tau^2}$ ) conocida, la función log-verosimilitud de maximiza cuando  $\hat{\sigma}^2 = \frac{\{y(R_Y)\}^T R_Y^{-1} \{y\}}{n}$ .

Al sustituir este estimado, obtenemos una función de log-verosimilitud:

$$L(\phi) = -\frac{1}{2} \{n \log(2\phi) + n \log \hat{\sigma}^2(R_Y) + \log(|R_Y|) + n\} \quad (4.19)$$

La estimación de los parámetros de la estructura de correlación son obtenidos optimizando la función concentrada de verosimilitud (Ecu. 4.19).

### 2. Proceso no Estacionario:

En el caso de que el proceso  $S'(x_i)$  sea no estacionario,  $u(x_i)$  puede ser estimado a partir de variables externas, incluyendo las variables del espacio (Ejm. Latitud y longitud). De este modo  $u(x_i)$  es representado por  $F\beta$  obteniendo el equivalente de la ecuacion 4.16.

$$Y_i = u(x_i) + S(x_i) + Z(x_i) \quad (4.20)$$

Por consiguiente  $Y \sim N(F\beta, \sigma^2 R(\phi) + \tau^2 I)$ . La función log-verosimilitud depende de un parámetro multivariado adicional  $\beta$ .

$$L(\beta, \tau^2, \sigma^2, \phi) = -\frac{1}{2} \{n \log(2\phi) + \log\{|\sigma^2 R(\phi) + \tau^2 I|\} + (y - D\beta)^T (\sigma^2 R(\phi) + \tau^2 I)^{-1} (y - D\beta)\} \quad (4.21)$$

Considerando la matriz de Correlación de  $Y$  ( $R_Y = \frac{(\sigma^2 R(\phi) + \tau^2 I)}{\sigma^2 + \tau^2}$ ) conocida, la función log-verosimilitud de maximiza cuando  $\hat{\beta}(R_Y) = (F^T R_Y^{-1} F)^{-1} F^T R_Y^{-1} y$  y  $\hat{\sigma}^2 = \frac{\{y - F\hat{\beta}(R_Y)\}^T R_Y^{-1} \{y - F\hat{\beta}(R_Y)\}}{n}$ .

Note que  $\hat{\beta}(R_y)$  se reduce a un estimado de mínimos cuadrados generalizados si  $R_y$  es conocido.

Al sustituir estos dos estimados, obtenemos una función de log-verosimilitud:

$$L(\phi) = -\frac{1}{2} \{n \log(2\phi) + n \log \hat{\sigma}^2(R_Y) + \log(|R_Y|) + n\} \quad (4.22)$$

La estimación de los parámetros de la estructura de correlación son obtenidos optimizando la función concentrada de verosimilitud (Ecu. 4.22).

#### 4.1.6. Predicción Espacial

Después de analizar la correlación espacial en una muestra, nuestro objetivo será predecir los valores de la variable en una región definida. Para tal caso, el kriging es una de las técnicas más empleadas.

##### 1. Kriging Ordinario:

Suponga que tienen  $n$  observaciones de una variable  $Z(x_1), Z(x_2), Z(x_3), \dots, Z(x_n)$  y que desea predecir el valor de la variable  $Z(x_0)$ . En tal caso, la mayoría de los métodos y el kriging proponen un estimado  $Z(x_0)$  de como una función lineal de las  $n$  observaciones.

$$Z^*(x_0) = \sum_{i=1}^n (\lambda_i Z(x_i)) \quad (4.23)$$

En donde,  $\lambda_i$  representa las ponderaciones para cada observación de acuerdo a la correlación espacial. Para realizar la predicción se necesita estimadores de  $\lambda_i$  que son del siguiente modo:

$$\lambda = C^{-1}C_{i0} \quad (4.24)$$

En donde  $C$  y  $C_{i0}$  tienen la siguiente forma:

$$C = \begin{pmatrix} C_{11} & \cdots & C_{1n} & 1 \\ \vdots & \ddots & \vdots & 1 \\ C_{n1} & \cdots & C_{nn} & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}, C_{i0} = \begin{pmatrix} C_{10} \\ \vdots \\ C_{n0} \\ 1 \end{pmatrix} \quad (4.25)$$

$$C_{ij} = COV(Z(x_i), Z(x_j)) \quad (4.26)$$

De tal modo con las ponderaciones calculadas en la Ec. 4.24 podemos obtener el estimado  $Z^*(x_0)$  con la Ec. 4.23.

## 2. Kriging Universal:

Considere un predictor lineal:

$$\hat{y}(x) = \lambda^T Y \quad (4.27)$$

El error de estimación es :

$$\begin{aligned} y - \hat{y} &= \lambda^T Y - y(x) \\ &= \lambda^T (F\beta + Z) - (f(x)^T \beta + z) \\ &= \lambda^T Z - z + (F^T \lambda - f(x))^T \beta \end{aligned}$$

Para obtener un estimador insesgado se debe cumplir que:

$$F^T \lambda = f(x) \quad (4.28)$$

De este modo podemos obtener los estimados de  $\lambda$  en base al criterio del menor ECM (error cuadrático medio):

$$\begin{aligned}
 \varphi(x) &= E[(\hat{y}(x) - y(x))^2] \\
 &= E[(\lambda^t Z - z)^2] \\
 &= E[z^2 + \lambda^T Z Z^T \lambda - 2\lambda^T Z z] \\
 &= \sigma^2(1 + \lambda^T R \lambda - 2\lambda^T r)
 \end{aligned} \tag{4.29}$$

La optimización puede ser resuelta por medio de las restricciones de Lagrange:

$$L(\lambda, c) = \sigma^2(1 + \lambda^T R \lambda - 2\lambda^T r) - c^T(F^T \lambda - f) \tag{4.30}$$

A partir de la ecuación anterior obtenemos el sistema de ecuaciones:

$$\begin{bmatrix} R & F \\ F & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ \tilde{c} \end{bmatrix} = \begin{bmatrix} r \\ f \end{bmatrix} \tag{4.31}$$

en donde se definió:

$$\tilde{c} = \frac{c}{2\sigma^2} \tag{4.32}$$

La solución al sistema 4.31 está dada por:

$$\begin{aligned}
 \tilde{c} &= (F^T R^{-1} F)^{-1} (F^T R^{-1} r - f) \\
 \lambda &= R^{-1} (r - F \tilde{c})
 \end{aligned} \tag{4.33}$$

## 4.2. Geoestadística para un Procesos de Bernoulli: Modelo Logístico Espacial

Anteriormente hemos presentado el modelo geoestadístico cuando la variable respuesta es continua y se comporta como un proceso estocástico gaussiano, ahora prestamos atención al modelo

geoestadístico cuando la variable respuesta se comporta como un proceso de Bernoulli<sup>3</sup>. Los modelos de este tipo de respuesta se encuentran dentro de los denominados modelos de elección discreta por su relación con la elección de una respuesta y también se encuentran dentro de los modelos lineales generalizados por el uso de la función de enlace para modelar este tipo de eventos. En el presente estudio se extiende este concepto considerando el comportamiento espacial. Para ello, inicialmente, presentaremos el modelo logístico de regresión y luego extenderemos su aplicación considerando el comportamiento espacial, sin adentrar demasiado en la formalidad del modelo sino con el objetivo de encarar estas situaciones pero manteniendo la teoría como sustento.

#### 4.2.1. Modelo Logístico:

A continuación consideramos un modelo lineal generalizado cuando la variable respuesta es dicotómica, cuyas respuestas podrían ser “éxito” o “fracaso”, vivo o muerto, presente o ausente, es un depocentro o no lo es, etc. Así que, sea la función de probabilidad de una variable aleatoria binaria  $Z$ :

$$P(Y) = \begin{cases} \pi & , 1 : \text{el resultado es éxito} \\ 1 - \pi & , 0 : \text{el resultado es fracaso} \end{cases}$$

Se define un modelo de regresión logístico como:

$$\text{logit}(\pi_i) = X_i^T \beta + Z_i \quad (4.34)$$

En donde, la función la enlace es logit que tiene la siguiente forma  $\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right)$ . Además,  $X_i$  son las variables independientes explicativas,  $\beta$  es el vector de parámetros y  $Z_i$  son los residuales.

#### 4.2.2. Modelo Logístico Espacial:

Para plantear la solución al modelamiento espacial de una variable dicotómica, empleamos la perspectiva del método kriging con regresión, mediante el cual se ajusta un modelo de regresión a la variable y luego de modela espacialmente a las desviaciones. Para ello, retomemos la ecuación 4.34 y recuerdese la ecuación 4.20.

---

<sup>3</sup>Es aquel proceso con espacio muestral o posibles resultados del experimento {Si, No}

$$\text{logit}(\pi_i) = X_i^T \beta + Z_i$$

Ahora considérese que  $Z_i$  se comporta como un proceso estocástico espacial gaussiano. El modelamiento es como se muestra a continuación:

**ESTIMACIÓN:**

- (I) Ajustar un modelo de regresión logístico para obtener los estimadores de  $\beta$ .
- (II) Calcular los residuales  $Z_i = \text{logit}(\hat{\pi}_i) - X_i^T \hat{\beta}$
- (III) En caso de haber residuales muy atípicos, eliminarlos y rehacer el paso I y II hasta que no existan residuales atípicos.
- (IV) Modelar espacialmente los residuales obteniendo los parámetros del patrón espacial.

**PREDICCIÓN:**

- (I) Generar datos de entrada  $X'_i$ .
- (II) Obtener el estimador  $\text{logit}(\hat{\pi}_i) = (X'_i)^T \hat{\beta}$  para los  $X'_i$  con el modelo logístico obtenido en el paso I de la estimación.
- (III) Predecir espacialmente  $Z'_i$  para cada posición de  $X'_i$  en base al modelo espacial obtenido en el paso IV de la estimación.
- (IV) Sumar los resultados de II y III:  $(X'_i)^T \hat{\beta} + Z'_i$  que es igual a  $Y_0 = \text{logit}(\hat{\pi}_i)$ .
- (V) Obtenemos el estimador de  $\pi$  a través de la función inversa del logit:  $\frac{e^{Y_0}}{1+e^{Y_0}}$

### 4.3. Geoestadística para Datos de Fuentes con Distinta Confiabilidad: Una extensión de la función de correlación

En la sección anterior se presentó los conceptos de la Teoría Geoestadística, su modelación, estimación y predicción. Todo ese marco teórico es útil cuando los datos provienen de una misma fuente y por lo tanto presentan la misma confiabilidad. Sin embargo, para la realización de la presente

tesis se cuenta con información de distintas fuentes son distintos grados de confiabilidad. Por lo cual no es adecuado emplear el marco teórico de la sección anterior en relación a la dependencia espacial.

En la 4.6 presentamos la diferencia existente entre las fuentes. En la imagen (a) se muestra la distribución espacial de tres fuentes de datos que se superponen en una región de la Plataforma Continental peruana. Los puntos rojos son datos obtenidos por la DHN. Los puntos azules regularmente espaciados son los datos de la grilla de batimetría del proyecto SRTM30\_PLUS y los puntos verdes representan los datos obtenidos en los trayectos de los cruceros científicos de IMARPE. Los gráficos (b,c,d) visualizan los datos de acuerdo a su profundidad en distintos ángulos. En las imágenes (b y c) se va notando la discrepancia entre la fuente de SRTM30\_PLUS y las demás fuentes. En la imagen (d) se nota claramente la diferencia de la superficie de la fuente SRTM30\_PLUS alrededor de la longitud de  $-80,315$  (diferencia promedio alrededor de 20 metros) y con mayor dimensión alrededor de  $-80,23$  grados de longitud (diferencia promedio de 80 metros aproximadamente). La diferencia entre la fuente de DHN e IMARPE es mucho menor en comparación con la fuente de SRTM30\_PLUS (5 metros como máximo).

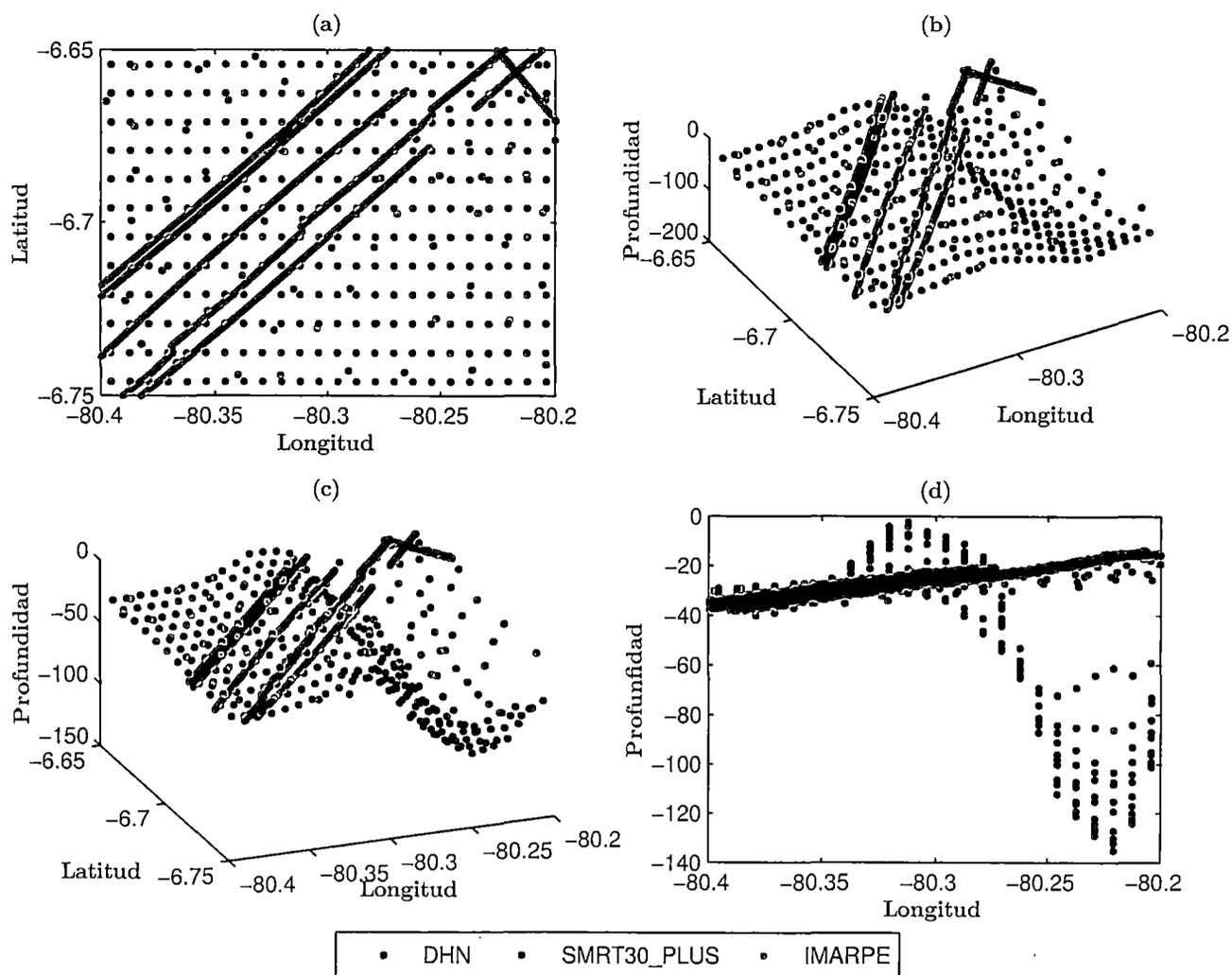


Figura 4.6: Comparación entre los datos espaciales de las fuentes DHN, SMTR30\_PLUS e IMARPE. (a) Distribución espacial de las fuentes. (b-c) Profundidad de las superficies según latitud y longitud por fuente. (d) Profundidad de las superficies según longitud por fuente

Las diferencias existentes entre las diversas fuentes de datos se debe, principalmente, a la metodología de obtención de datos. Los datos de los cruceros científicos de IMARPE son más confiables porque son datos capturados de manera continua. Su confiabilidad aumenta mientras los puntos se encuentran más cerca a las costas y disminuye mientras más se aleja de la Plataforma Continental porque no esta preparado para zonas profundas. La DHN e IMARPE tienen mediciones muy parecidas, lo cual aumenta su confiabilidad por respaldarse mutuamente. Finalmente, los datos de la grilla de SRTM30\_PLUS son datos interpolados por lo que presentan menor confiabilidad, sin embargo, es uno de los más usados debido a su alcance en todo el globo terráqueo.

Las diferencias encontradas deben ser proyectadas a un trasfondo geoestadístico de modo que se incorpore esta característica. Así que, en primer lugar evaluamos el impacto que tiene estas discrepancias en las herramientas de visualización y modelación espacial (variograma y correlograma) para luego proponer soluciones en un marco teórico geoestadístico.

### 4.3.1. Correlación Espacial

La correlación espacial es el aspecto más importante en el modelamiento geoestadístico. La podemos analizar indirectamente en el variograma y en el correlograma de manera directa. A continuación mostraremos los efectos de la discrepancia de fuentes en el variograma empírico y se va a proponer un modelo teórico para el variograma y el correlograma. Utilizamos como herramienta visual al variograma por su facilidad de cálculo a comparación del correlograma que necesita más parámetros.

#### 1. Variograma Empírico

La Fig. 4.7 es el variograma calculado por fuente y entre fuentes correspondientes a los datos distribuidos espacialmente en la Fig. 4.6. La figura se muestra para el caso de que se cuente con dos fuentes de datos ( (1) IMARPE y (2) SRTM\_30 para tal gráfico). Esta figura muestra que el comportamiento del variograma es distinto en tres casos: (i) cuando los pares de datos provienen de la fuente 1, (ii) cuando los pares de datos provienen de la fuente 2 y (iii) cuando los pares están conformados por un dato de la fuente 1 y uno de la fuente 2.

El efecto pepita es despreciable en el variograma de IMARPE, seguido por el de SRTM30\_PLUS (500 aprox) y es el más grande para el variograma cruzado (800 aprox). Esto indica que el error de medición en la fuente de IMARPE es despreciable y que capta la correlación espacial a pequeñas escalas. El error de medición es mayor en la fuente de SRTM30\_PLUS. En el variograma cruzado, el efecto pepita más grande (800 aprox) indica que **hay un desfase a distancias muy pequeñas ( $\cong 0$ ) entre las dos fuentes evaluadas.**

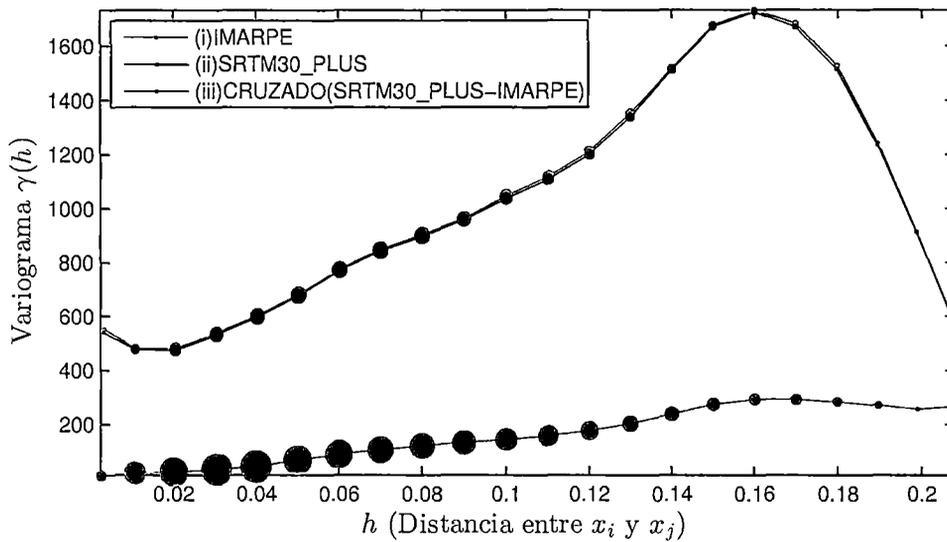


Figura 4.7: Variograma Empírico por fuente de IMARPE y SRTM30.PLUS, y entre las dos fuentes. El tamaño de los círculos es proporcional a la cantidad de pares encontrados a la distancia  $h$

A menor valor del variograma mayor similitud o correlación y a mayor valor del variograma mayor diferencia entre puntos a una distancia  $h$ . El variograma calculado para los datos de IMARPE tiene valores pequeños, lo cual indica mayor semejanza o correlación en esta fuente. Mientras que es mayor en los datos de la fuente de SRTM30.PLUS, indicando menor correlación espacial. Esta diferencia se debe a que los datos de IMARPE son menos variables que los de SRTM30.PLUS (Fig. 4.6-(d)). Por otro lado, el variograma cruzado es mayor a los dos anteriores a distancias cortas ( $< 0 - 0,06 >$ ) y tiene un comportamiento semejante al de SRTM30.PLUS porque este tienen mayor variabilidad. En general, se puede ver que la similitud es menor en el caso (iii) y que es mayor en el caso (i) ya que es la fuente con mayor confiabilidad. Más adelante se presenta los efectos del comportamiento de los tres variogramas en las funciones de correlación.

## 2. Correlación

Las características de los variogramas mostrados en la Fig. 4.7 implica cierto comportamiento en las funciones de correlación espacial tal como la Fig. 4.8-(a). En ella se puede ver que la correlación espacial de la fuente de IMARPE es mayor respecto a las tres funciones, su alcance es mayor y no existe desfase a una distancia equivalente a cero ( $\rho(0) = 1$ ). Para la fuente de datos de SRTM30.PLUS existe un desfase cuando los puntos se encuentran en la misma

ubicación ( $\rho(0) = 0,7$ ) indicando mayor error de medición en esta fuente de datos. Luego la correlación espacial cruzada indica que hay un desfase mayor a una distancia cero ( $\rho(0) = 0,6$ ) debido a un valor grande del efecto pepita en la Fig. 4.7. La interpretación de esta última función en la predicción del fondo marino es muy importante. Así que si se desea realizar una predicción de la estructura del fondo marino basado en la fuente de IMARPE, los datos de SRTM30\_PLUS no deben tener demasiada influencia en la predicción cuando hay discrepancia entre ellos. Ello debido a que su correlación con la fuente principal es mucho menor a 1 incluso cuando la distancia es equivalente a cero.

Por otro lado, la imagen (a) muestra que cada función tiene distinto alcance debido a que la varianza es distinta para cada fuente en la región de la Fig. 4.6. Esto hace que la estimación de parámetros se dificulte al agregar varias varianzas y varios efectos pepitas. Por lo que analizamos la batimetría en regiones más amplias de tal modo que la varianzas sean semejantes para cada fuente y podamos proponer una estructura de correlación espacial multi-fuente mas sencilla que considere el desfase que hay entre las fuentes a pequeñas distancias (Fig. 4.8-(b)). Además, a distancias grandes deben converger a 0 con un mismo alcance.

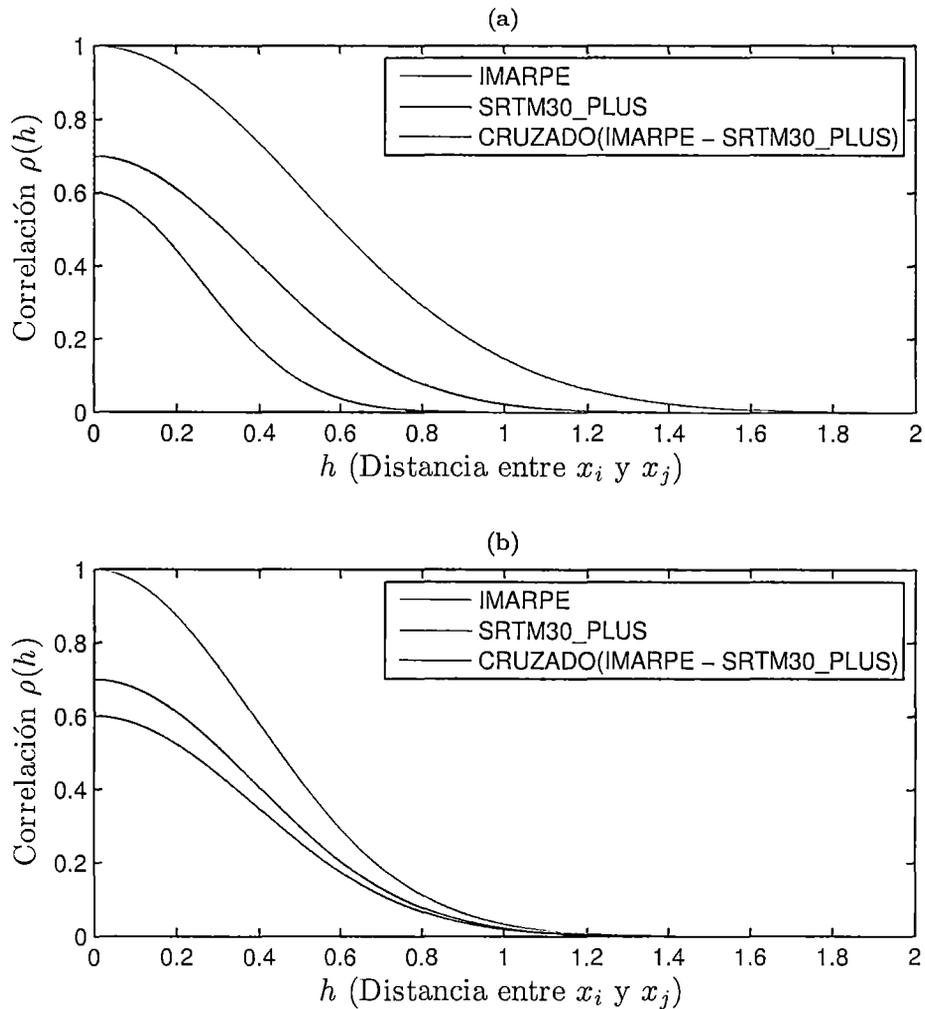


Figura 4.8: Correlación Teórica por fuente de IMARPE y SRTM30\_PLUS, y entre las dos fuentes. La escala de la distancia  $h$  es referencial. (a) Correlación teórica cuando las varianzas, alcance y los efectos pepita de cada función es diferente. (b) Correlación teórica cuando los efectos pepita son distintos pero con alcances y varianzas iguales.

Por lo tanto, la solución planteada en esta situación (Fig. 4.8-(b)) es extender el concepto de distancia y utilizar una sola función de correlación para capturar la estructura espacial multi-fuente, de tal modo que datos provenientes de distintas fuentes tengan mayor diferencia que datos que provienen de la misma fuente. De este modo, la correlación será mayor para datos que provienen de la misma fuente que en aquellos que provienen de distintas fuentes. Notamos que ello se puede obtener si aumentamos dos atributos a la función de correlación que indiquen las fuentes de la que provienen el par de datos para los que se calcula la distancia. Y consideramos un efecto pepita para cada caso ((i) Cuando el par de datos proviene de la

fuente, (ii) Proviene de la fuente 2 y (iii) Un dato es de la fuente 1 y el otro de la fuente 2).

Según la descripción anterior, la función de correlación espacial multi-fuente graficada en la Fig. 4.8-(b) queda expresada de la siguiente manera:

$$\rho_Y(h) = \begin{cases} 1 & : h = 0 \\ \frac{\sigma^2 \rho(h)}{\sigma^2 + \tau_1^2} & : x_s = x'_s = 1 \\ \frac{\sigma^2 \rho(h)}{\sigma^2 + \tau_2^2} & : x_s = x'_s = 2 \\ \frac{\sigma^2 \rho(h)}{\sigma^2 + \tau_3^2} & : x_s \neq x'_s \end{cases} \quad (4.35)$$

En donde  $h$  es la distancia espacial entre dos puntos cualquiera,  $\rho(y)$  es la función de correlación espacial del fenómeno bajo estudio sin considerar el error de medición.  $x_s$  representa la fuente del primer dato y  $x'_s$  es la fuente del segundo dato. La varianza general entre todos los datos es  $\sigma^2$  y  $\tau_1^2$ ,  $\tau_2^2$  y  $\tau_3^2$  son el desfase debido al error de medición en cada caso.

La extensión del concepto de distancia respecto a la distancia espacial y la distancia entre fuentes de datos modifica la estructura de la correlación espacial. Sin embargo, el enfoque que presentamos mantiene la forma de estimar los parámetros y realizar las predicciones cambiando solo la forma de la matriz de correlación  $R$ . Lo cual es una ventaja de esta propuesta, debido a su relativa fácil aplicación. La generalización de esta estructura de correlación para más de dos fuentes de datos puede darse de manera lógica.

### 4.3.2. Estimación de Parámetros

La estimación de parámetros se realiza del mismo modo que la sección anterior. Sin embargo, en el presente caso, se incrementa la cantidad de parámetros a estimar para la estructura de correlación. Así la estimación de los parámetros de la estructura de correlación son obtenidos optimizando la función concentrada de verosimilitud mostrada en la Ecu. 4.19:

$$L(\phi) = -\frac{1}{2} \{n \log(2\phi) + n \log \hat{\sigma}^2(R_Y) + \log(|R_Y|) + n\}$$

En donde  $R_Y$  está expresado en la Ecu. 4.35.

Solo trabajamos con el modelo para dos fuentes de datos porque se asegura la optimización del algoritmo. En caso de extender este modelo a más de dos fuentes, se debe analizar los métodos

de optimización para asegurar la optimización de los parámetros.

## Capítulo 5

# Resultados

### 5.1. Predicción de la Batimetría del Mar Peruano

El primer resultado que se obtuvo es el mapa de batimetría en alta resolución a través de un trabajo exhaustivo que considera el conocimiento de la topografía del relieve marino, modelamiento estadístico, programación y procesos para generar grillas batimétricas confiables<sup>1</sup>.

#### 5.1.1. Características Generales del Mapa de Batimetría Obtenido

La batimetría fue modelada mediante el kriging Universal y la estimación de parámetros por máxima verosimilitud, considerando su aplicación a través de ventanas móviles. La grilla es de alta resolución a una distancia de 50 metros, siendo el mapa de mayor resolución obtenido hasta el momento para la región perteneciente al mar peruano, y por lo tanto se puede obtener la grilla a resoluciones más grandes como 100, 150 u otro nivel mayor. Además, se obtuvo una grilla de errores correspondiente a la grilla de batimetría. Es decir, para cada punto de predicción de la profundidad del fondo marino, se calculó el error de predicción máximo, el cual depende del nivel de confianza deseado. Ello es esencial para ser usado en otros contextos y estudios como variable de entrada.

En la Fig. 5.1 se muestra la característica del mapa en una zona del margen y talud continental. El detalle del mapa varía dependiendo de la cantidad y calidad de datos alrededor de una zona. En zonas donde existen datos de batimetría capturados a través de los cruceros científicos de haz múltiple, se puede ver mayor detalle de la morfología del fondo marino (parte inferior del mapa)

---

<sup>1</sup>Como los presentados en el libro CookBook de GEBCO

mientras que las demás fuentes nos ayudan a capturar la tendencia de la superficie a una escala mayor.



Figura 5.1: Predicción de la batimetría a alta resolución (100 metros, exageración vertical x 20).

Enlazado al mapa de batimetría, tenemos el mapa de errores de predicción según el nivel de confianza deseado. En la Fig. 5.2 se tiene el mapa de errores de predicción al 95 % de confianza de la batimetría de la Fig. 5.1. Cerca a las costas el error de predicción es menor mientras que en el talud continental es mayor. El error también es mayor en zonas donde se tienen pocos datos de muestreo o cuando la confiabilidad de la fuente de los datos es baja.

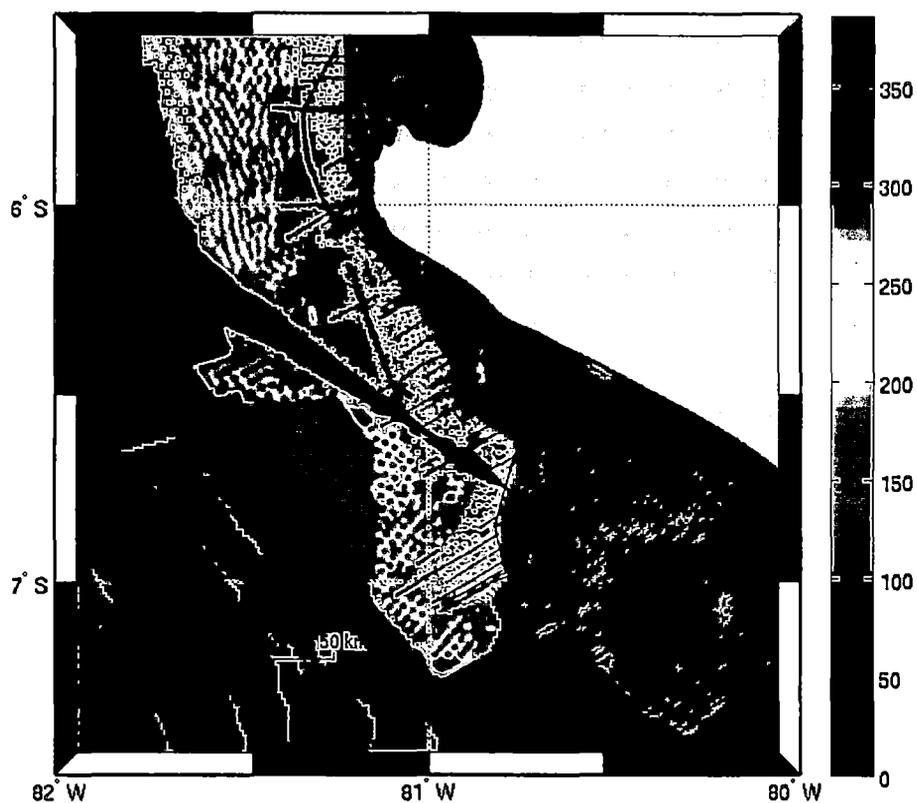


Figura 5.2: Error de predicción de la batimetría a alta resolución (100 metros, exageración vertical x 20).

En la Fig. 5.3, se muestra el mapa obtenido en una menor resolución (900 metros) para poder verlo de forma completa. Se puede observar claramente la Plataforma Continental, el Talud Continental y la zona abisal del fondo marino. Es posible observar los distintos accidentes que subyacen en el fondo marino, entre ellos los cañones cuando la Plataforma se quiebra y empieza el descenso del Talud Continental.



Figura 5.3: Predicción de la batimetría a baja resolución (900 metros, exageración vertical x 20).

El error de predicción de todo el área de estudio se muestra en la Fig. 5.4 al 95 % de confianza. En este se puede ver que el rango del error de predicción es de 0 a 450 metros aproximadamente. Los errores de predicción mayores se encuentran en la zona abisal (zonas profundas) entre los 80 grados de longitud al oeste y 17.5 grados de latitud al sur. Además el error varía entre cada cluster espacial según el ajuste del modelo que se hizo y varía dentro de cada cluster espacial según la cercanía a datos muestrales.

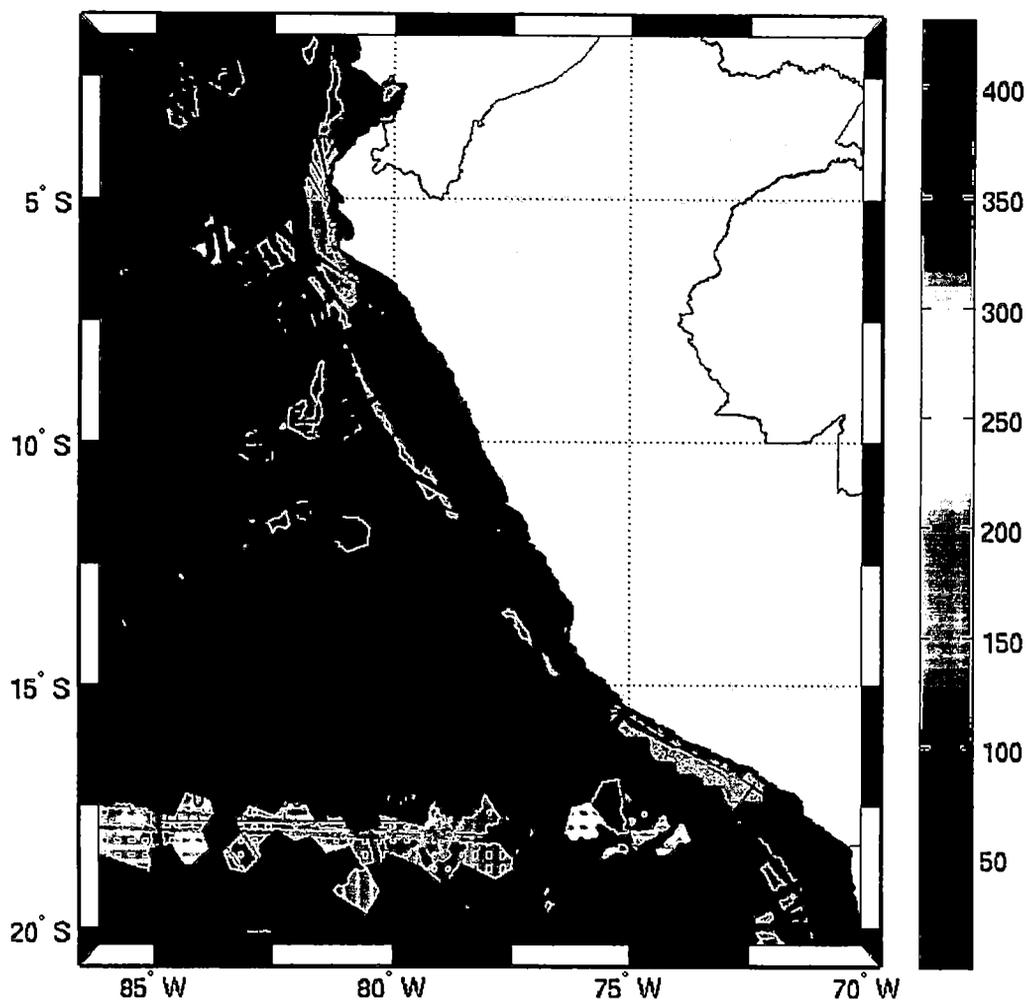
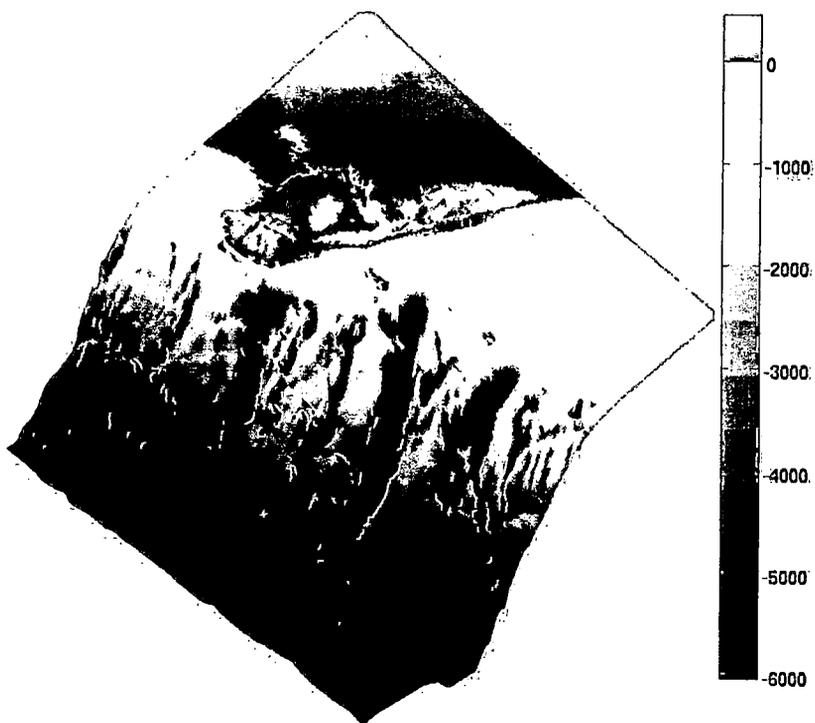


Figura 5.4: Error de predicción de la batimetría a baja resolución (900 metros, exageración vertical x 20).

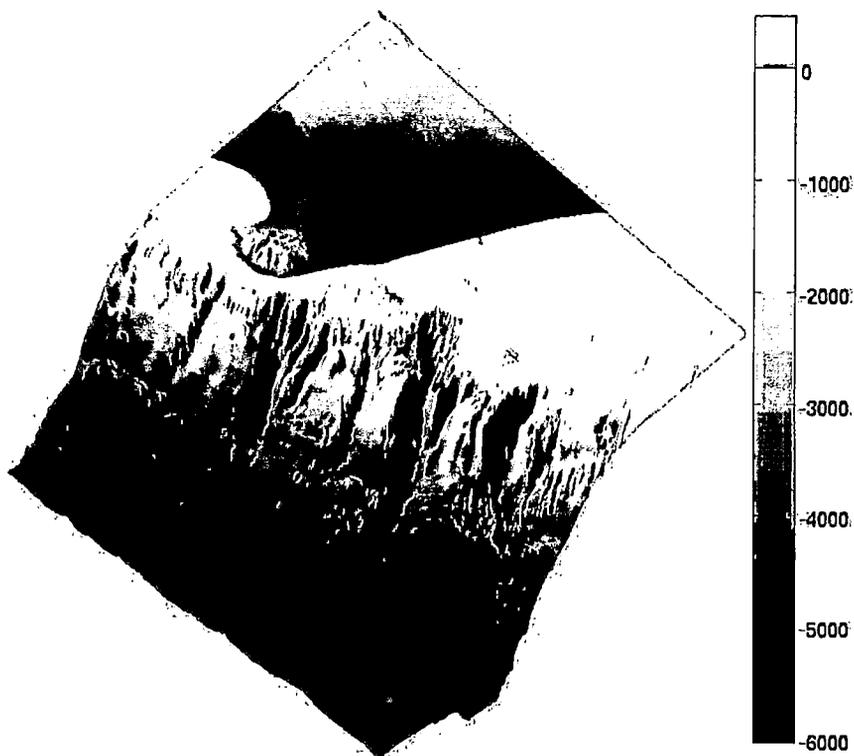
### 5.1.2. Comparación del Mapa de Batimetría Obtenido con la Batimetría del Proyecto SRTM30

Definitivamente, el hecho de que el modelamiento geoestadístico se realice en una alta resolución hace que no se pierda las características morfológicas a pequeñas escalas y se capture el comportamiento del relieve marino a mayor detalle. Por ello, mostramos la comparación entre el mapa de batimetría obtenido y el del proyecto SRTM30\_PLUS que se observan en la Fig. 5.5 en una región más pequeña. Es de notar que en la imagen (b) se obtiene una caracterización a menor escala del fondo marino observándose mayor detalle en las zonas de quiebre y en la parte inferior del Talud Continental (zona inferior de la imagen). Mientras que la imagen (a) muestra las tendencias

existentes a una escala mayor. Por lo tanto, se demuestra la mejora de los resultados obtenidos para el empleo de la batimetría de alta resolución en los estudios que comprendan sistemas que se desarrollan y caracterizan a pequeña escala.



(a) Batimetría a baja resolución (aprox. a 926 metros) según el proyecto SRTM30\_PLUS



(b) Batimetría a alta resolución (aprox. a 100 metros) según el resultado del presente estudio

Figura 5.5: Comparación a alta resolución entre el mapa de batimetría de SRTM30\_3PLUS y el mapa obtenido (exageración vertical x 20).

Una fuerte desventaja del mapa de SRTM30 es que no brinda los errores de batimetría y ello hace que no se tenga conocimiento de la confiabilidad a nivel local de esta grilla. Mientras que el resultado del presente estudio, si contempla brindar una grilla de errores de predicción haciendo que el usuario tenga mayor conocimiento sobre la validez del mismo. En la tabla 5.1 se muestra las principales características entre las dos grillas.

**Cuadro 5.1:** Comparación del Mapa de Batimetría Obtenido con la Batimetría del Proyecto SRTM30

Características	Mapa batimétrico del estudio	Mapa batimétrico de SRTM-30
Resolución	Alta (50 metros)	Baja (1000 metros aprox.)
Detalle	Mayor detalle en zonas con mayor información	Menor detalle
Error de predicción	Si tiene	No tiene
Detalle	Mayor detalle en zonas con mayor información	Menor detalle
Información local	Cuenta con datos de DHN, IRD e IMARPE	No tiene
Cruceros internacionales	Cuenta con datos de la NOAA	Cuenta con datos de la NOAA

Ya se mostró la mejora de la batimetría obtenida en pequeñas regiones respecto al mapa de batimetría del proyecto SRTM30\_PLUS. Sin embargo, los mapas deben ser muy semejantes a nivel global. A modo de validación de nuestro mapa de batimetría del fondo marino, se muestra la comparación entre ambos mapas a baja resolución en la Fig. 5.6. En ella encontramos que el mapa obtenido y la del SRTM30\_PLUS son muy parecidos a una resolución de 900 metros aproximadamente lo cual nos ayuda verificar la validez de nuestro mapa a nivel global.



(a) Batimetría a baja resolución (aprox. a 926 metros) según el proyecto SRTM30\_PLUS



(b) Batimetría a baja resolución (aprox. a 900 metros) según el resultado del presente estudio

Figura 5.6: Comparación a baja resolución entre el mapa de batimetría de SRTM30\_PLUS y el mapa obtenido (exageración vertical x 20).

### 5.1.3. Comparación entre un Modelo que no Considera la Diferencia entre Fuentes de Datos y el Modelo Empleado

Una gran fortaleza del presente estudio, es que su predicción considera datos de varias fuentes. Por lo que se logra abarcar un gran área de estudio; sin embargo, este aspecto también ha hecho complicado la predicción del fondo marino debido a la discrepancia entre las fuentes de datos en diversas zonas. Ante ello, no habían modelos geoestadísticos que consideren este aspecto y por ello consideramos un resultado a la extensión de la función de correlación considerada para la modelación de la correlación espacial, el cual fue presentado en el capítulo anterior (Sección 4.3: Geoestadística para Datos de Fuentes con Distinta Confiabilidad: Una extensión de la función de correlación).

La discrepancia entre fuentes de datos no es homogénea espacialmente o su presencia no siempre es significativa. Es decir, que en ciertas zonas existen diferencias significativas y en otras no. Además en caso de la existencia de discrepancia, la diferencia promedio tampoco es homogénea de tal modo que en algunas zonas puede ser 10 metros en otras 20 y así sucesivamente. Por lo que es necesario modelar esta discrepancia de modo que sea flexible según la zona.

En la Fig. 5.7, se presenta el variograma de tres fuentes (IMARPE, SMTR30\_PLUS y DHN) en una zona donde sí existe discrepancia entre las fuentes. Los datos son de una zona dentro de la Plataforma Continental, relativamente plana, para ver la distribución espacial de los datos véase la Fig. 4.6. La forma del variograma de SRTM30\_PLUS indica que la variabilidad es alta en esta fuente y que existe errores de medición por el desfase en  $h = 0$  o que existe una correlación espacial a menor escala que no es captada por estos datos. Según el variograma de la DHN, se puede decir que la variabilidad es baja y que no se capta una estructura de correlación espacial que probablemente se encuentre a menor escala de la medida en esta fuente. Los datos de IMARPE no muestran demasiada variabilidad a distancias pequeñas  $h \cong 0$  y además son capaces de captar la estructura espacial a pequeña escala. Los variogramas cruzados entre IMARPE y las demás fuentes son influenciados por la fuente que tiene mayor variabilidad y muy poca variabilidad

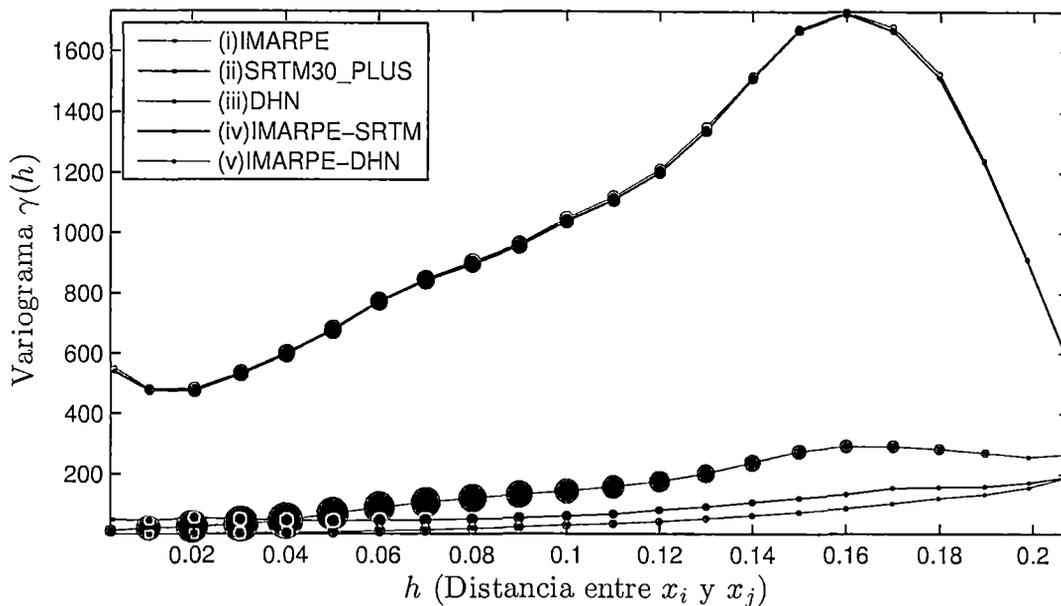


Figura 5.7: Variograma Empírico para la fuente de IMARPE, SRTM30\_PLUS y DHN, y entre IMARPE con cada una de las dos fuentes. El tamaño de los círculos es proporcional a la cantidad de pares encontrados a una distancia  $h$

Por otro lado, en la Fig. 5.8, se muestra una interpolación kriging que no considera la discrepancia entre las bases de datos, resaltando la gran influencia que tienen los datos pocos confiables en la predicción de la superficie mientras que la Fig. 5.9 ejemplifica el resultado bajo el modelo de correlación espacial que presentamos en el estudio (en la sección de las técnicas de análisis), en ella se puede verificar que los datos con poca confiabilidad no influyen del mismo modo en la superficie del relieve marino haciendo que nuestro modelo tenga la capacidad de poder discernir entre los datos más confiables y menos confiables, logrando traducir esa información en la función de correlación.

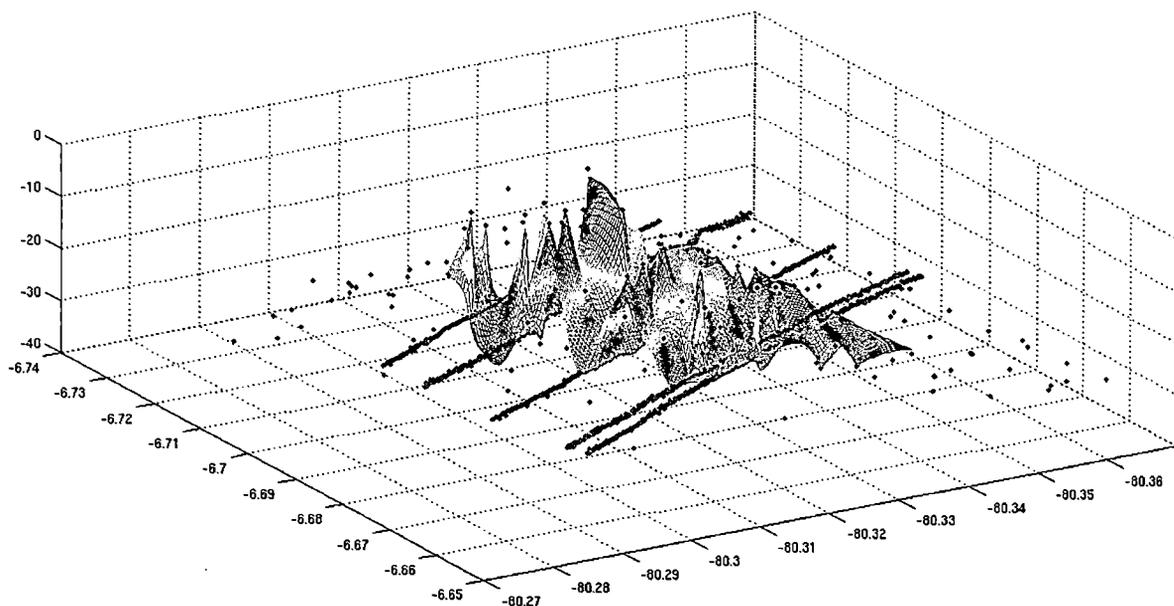


Figura 5.8: Kriging sin considerar nuestro modelo de correlación.

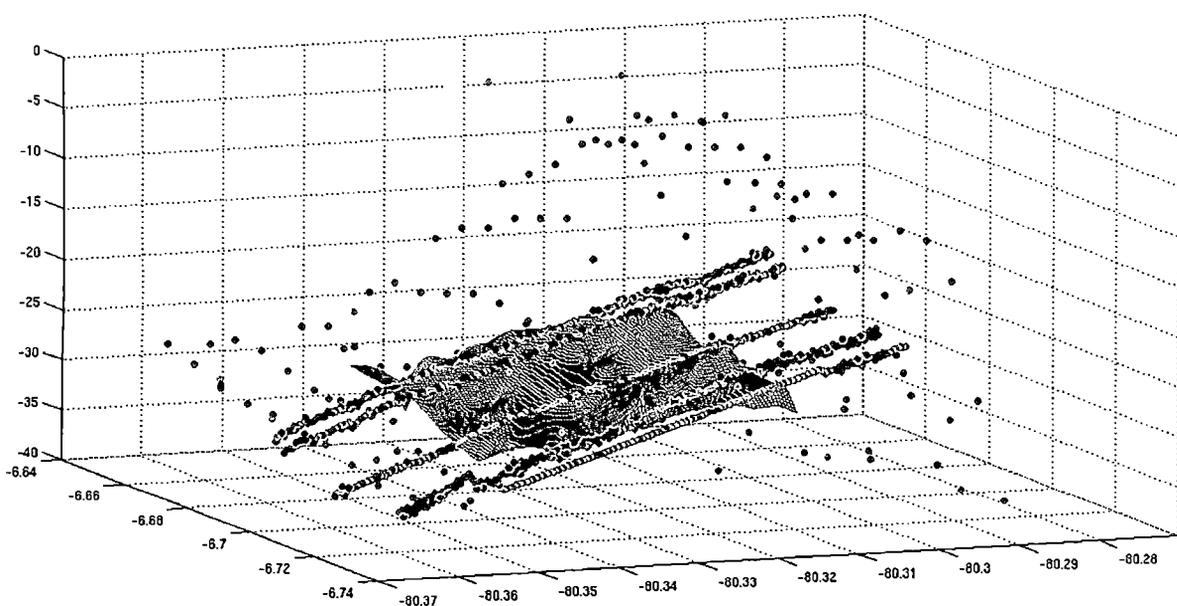


Figura 5.9: Kriging considerando nuestro modelo de correlación.

## 5.2. Identificación de Clusters Sedimentarios

Antes de proponer un modelo para estimar la probabilidad de que cierta zona sea o no un depocentro, se requiere obtener el perfil sedimentario de sedimentos finos. Para ello se evaluó los 500

perfiles sedimentarios que se tienen ubicados dentro del margen continental peruano tratando de encontrar grupos de perfiles sedimentarios y esperando que uno de los grupos corresponda al perfil de sedimentos finos deseado.

### 5.2.1. Descripción General de la Distribución de las Características Sedimentarias en la Plataforma Continental

#### 1. Según el tamaño del grano

El coeficiente de asimetría es mostrado en la Fig. 5.10, los colores van desde el más oscuro (negativo) hasta el más claro (positivo). Cuando el coeficiente de asimetría es negativo indica que el tamaño del grano es fino mientras que si es positivo indica que el tamaño del grano es grande. Por lo tanto, estamos interesados en aquellos perfiles de color oscuro. Por lo general, están distribuidos en la zona sur de la Plataforma Continental, en donde la energía de la hidrodinámica es menor que la zona norte (La Libertad, Lambayeque). También existen perfiles en donde el tamaño del grano es pequeño por la parte central (Lima, Ancash) pero con menor incidencia y encontramos algunos casos en la zona norte (Piura, Tumbes). Existe un patrón espacial que no es de mayor radio, indicando que esta variable debe estar explicada por otros factores (morfometría y batimetría).

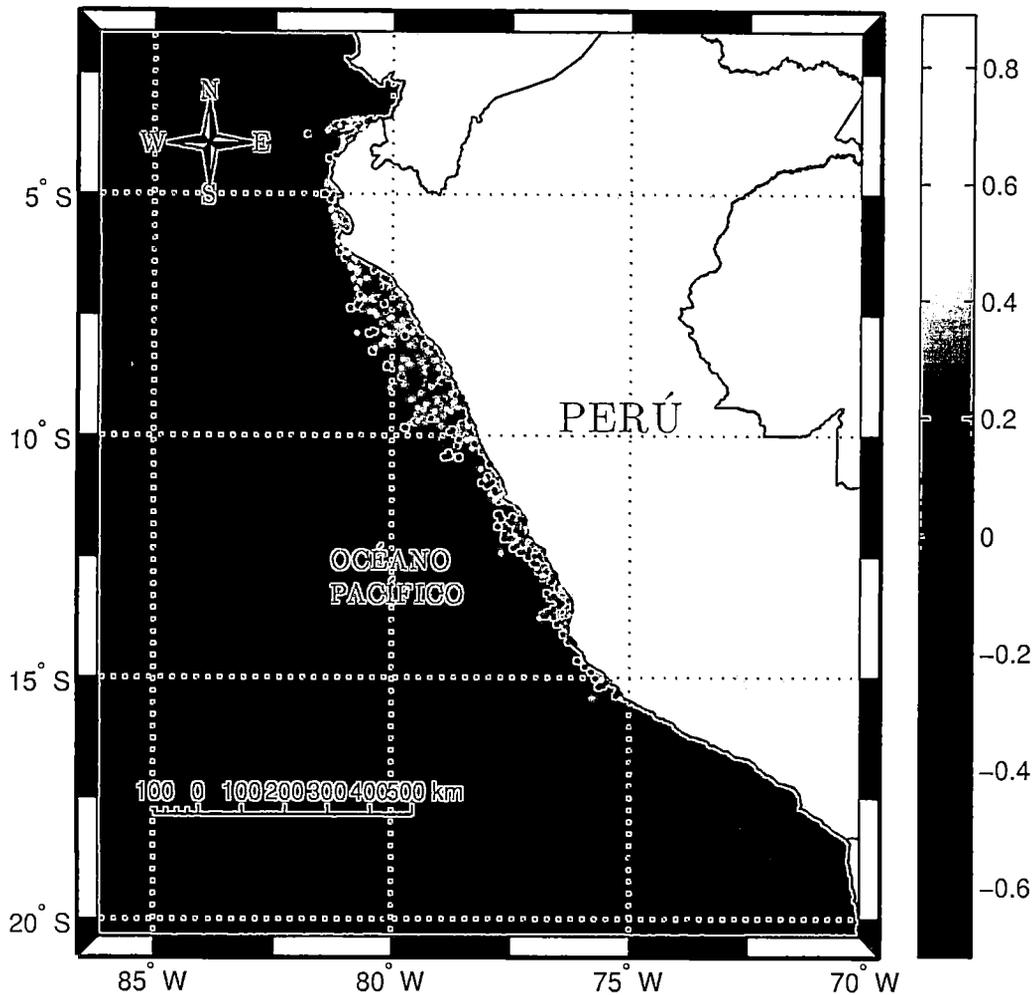


Figura 5.10: Distribución espacial del tamaño de grano según el coeficiente de asimetría. Valores negativos indican granos finos y valores positivos representan granos más gruesos.

## 2. Según la composición de la textura

Una composición general de la textura contempla los materiales de limo, arcilla y arena. Deseamos encontrar los perfiles que estén compuestos principalmente por limo y arcilla. En la Fig. 5.11 se presenta la distribución espacial del porcentaje de limo y arcilla en conjunto. En este se observa un patrón espacial más fuerte que en el coeficiente de asimetría del tamaño de los granos y una tendencia a encontrar sedimentos compuestos principalmente por limo y arcilla en la zona central sur. Se detectan perfiles deseados en la parte central norte pegada a las costas y en la parte superior del Talud Continental peruano. Y algunos casos en la zona norte.

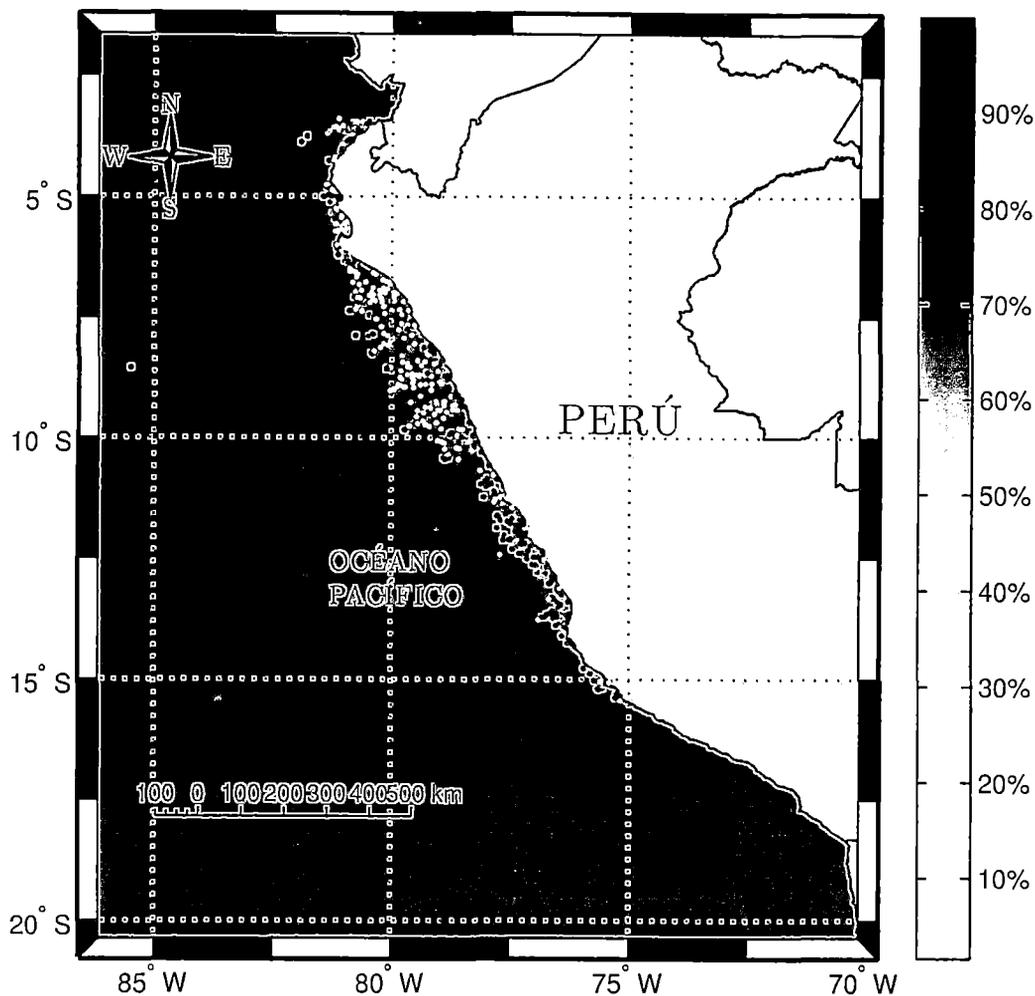


Figura 5.11: Cluster de los perfiles sedimentarios

### 5.2.2. Clusters Encontrados Según los Perfiles de Sedimentología

Se encontraron cuatro clusters de perfiles sedimentarios según la granulometría y la composición de textura de los sedimentos. La distribución espacial de estos clusters se muestra en la Fig. 5.12. El primer y tercer cluster (color rojo y azul respectivamente) se aglomeran en la zona norte y centro-norte. El segundo cluster, que es de nuestro interés, se concentra principalmente en la zona centro-sur y con menor intensidad en algunas regiones centro-norte muy cerca a las costas y en la parte superior del Talud Continental. Además, se encuentran algunas concentraciones en la parte norte. Por último, el cuarto cluster tiene concentraciones menos intensas y se encuentran en algunas regiones del sur y otras de la zona centro-sur cerca a las costas. Es importante notar que los clusters

no se distribuyen aleatoriamente, sino que por lo contrario presentan una estructura espacial que ha de considerarse en la predicción de depocentros.

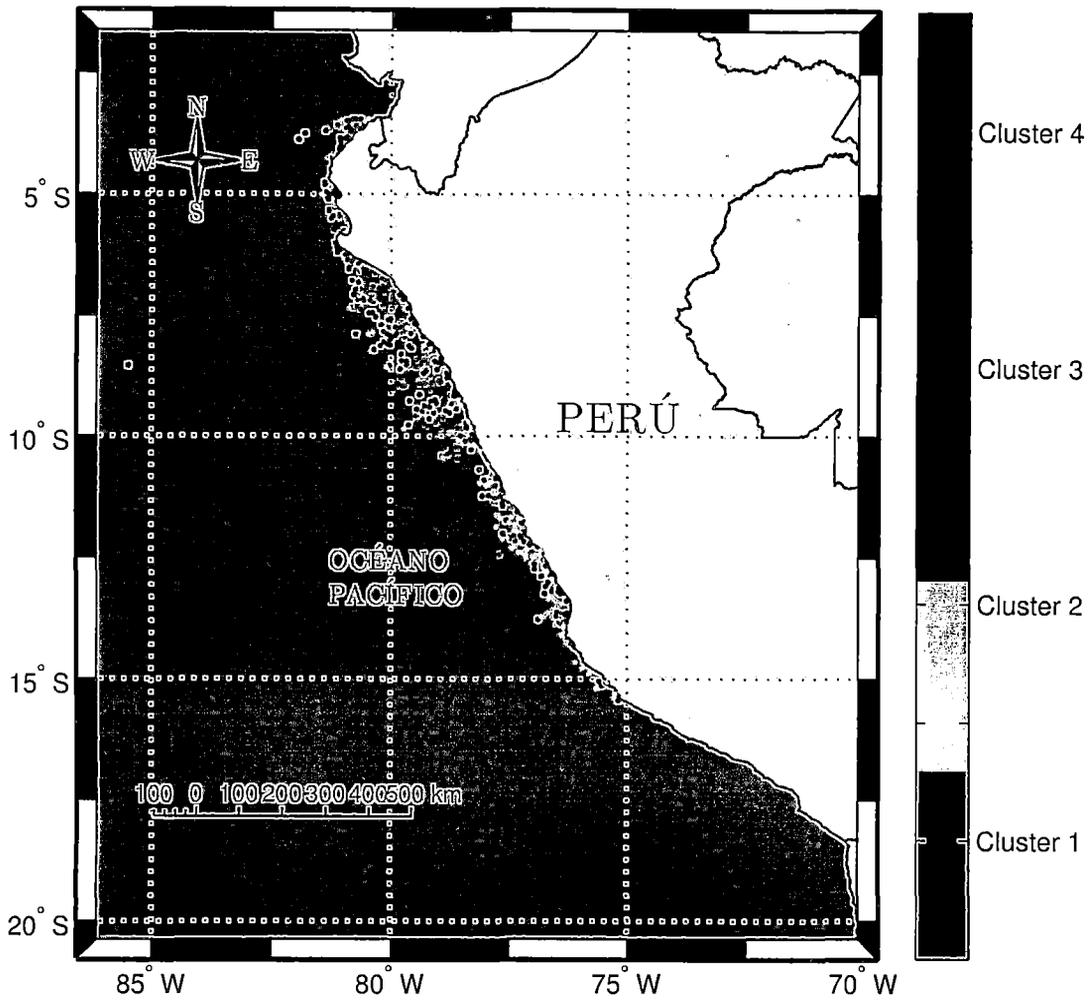


Figura 5.12: Distribución espacial de los clusters de perfiles sedimentarios

A continuación describimos las características de los cuatro clusters encontrados según el coeficiente de asimetría del tamaño del grano y el porcentaje total de limo y arcilla (Fig. 5.13) mediante la técnica de clustering K-means. El primer perfil tiene bajo porcentaje de limo y arcilla, alrededor del 16% y un alto coeficiente de asimetría, entre 0.25 y 0.5, que representa tamaño grande de grano. El segundo perfil se caracteriza por ser compuesto principalmente por limo y arcilla y poseer granos finos, típico de zonas de sedimentación continua como las que se busca en esta tesis. El tercer grupo posee alta variabilidad en el porcentaje de limo y arcilla entre 0.3 y 0.6, y es el

grupo con los granos de sedimentos más grandes (coeficiente de asimetría igual alrededor de 0.7). Finalmente, el cuarto grupo posee alta composición de limo y arcilla pero el tamaño del grano es moderadamente grande. Por lo tanto, el segundo grupo es de nuestro interés porque presenta las dos características deseadas (tamaño de grano pequeño y compuesto principalmente por limo y arcilla). A este grupo lo consideramos como el perfil de ambientes de deposición continua, zonas en donde la deposición es continua y fina.

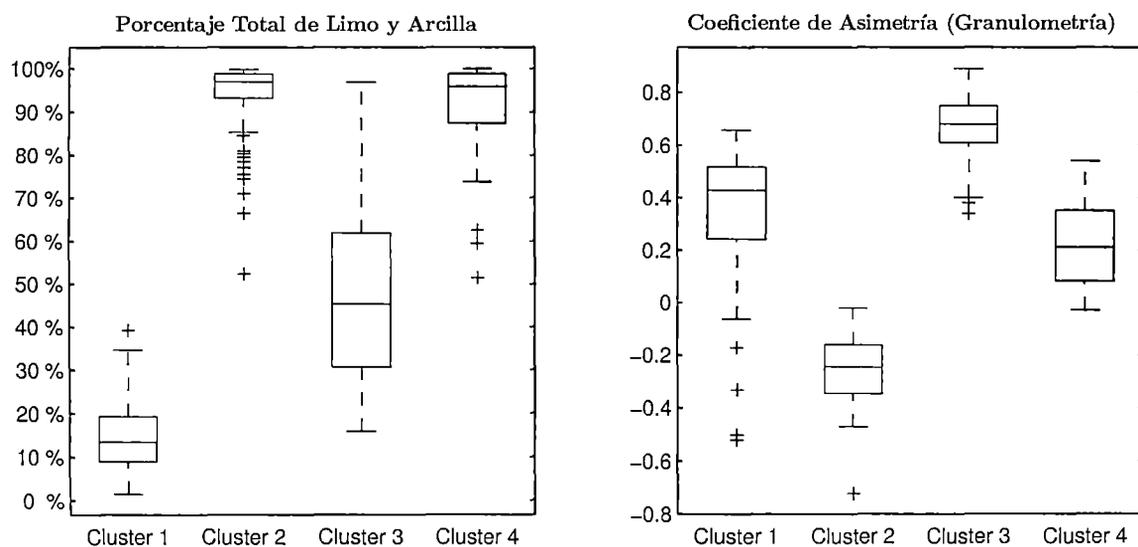


Figura 5.13: Características sedimentarias según cada cluster.

### 5.3. Modelamiento de Presencia de Depocentros Sedimentarios

El modelamiento de la presencia de depocentros sedimentarios tiene dos componentes importantes, el primero es la tendencia llevada por el factor morfológica y el segundo es la estructura espacial del mismo.

#### 5.3.1. Variables Consideradas y Factores Incorporados al Modelo

En primer lugar se considera un modelo lineal generalizado para capturar la tendencia explicada por la morfología y batimetría del fondo marino. Este modelo incorpora variables a nivel puntual (en la posición exacta o una vecindad de los 8 puntos a su alrededor) y a nivel local (en una

vecindad más grande). Las variables son presentadas y definidas en la sección 3.2.

Hemos considerado varios modelos para hacer una selección de aquel que nos brinde mayor bondad de ajuste y un buen nivel predictivo. Los dividimos en dos grupos. El primero no considera interacciones entre las covariables mientras que el segundo si lo hace. Dentro de cada grupo se ha considerado tres funciones de enlace (logit, probit y log log complementario). A continuación se presenta los modelos obtenidos, su bondad de ajuste y la comparación entre los modelos obtenidos para escoger el que será aplicado en la predicción de los ambientes de deposición continua en el margen continental peruano.

El número de datos utilizados en la muestra fueron  $n = 458$ , la selección de las variables se hace por el método paso a paso (stepwise) en donde se puede añadir y quitar variables del modelo. Este proceso considera el criterio de Akaike (AIC) como indicador de decisión.

### 1. Modelo Sin Interacciones

#### (a) Descripción del Modelo

$$\text{Modelo Logit: } \text{logit}(\hat{\pi}_i) = \log\left(\frac{\hat{\pi}_i}{1-\hat{\pi}_i}\right) = X_i^T \hat{\beta}$$

El modelo logit incorpora 5 variables en el proceso de selección de variables paso a paso. Dentro de ellas, la distancia mínima a la costa, el ángulo de inclinación y la curvatura perpendicular a la pendiente máxima son las más significativas. Los factores influyentes de manera positiva en la presencia de depocentros son el ángulo de inclinación en el eje y, curvatura perpendicular a la pendiente máxima y la profundidad. Por otro lado, la distancia mínima a la costa y la profundidad ajustada local tienen influencia negativa.

**Cuadro 5.2:** Variables incorporadas al modelo logit sin interacciones y su significancia

ID Variable	Nombre de la Variable	Coefficiente Estimado	Error Estándar	Estadístico T	Valor de p
Constante	Intercepto	-0.8037	0.1629	-4.933	0.0000
x3	Distancia mínima a la costa	-1.8798	0.4983	-3.772	0.0002
x7	Ángulo de inclinación en el eje y	0.2099	0.0863	2.431	0.0150
x9	Curvatura perpendicular a la pendiente máxima	2070.4055	1122.4213	1.845	0.0651
x16	Profundidad ajustada - local	-0.0119	0.0087	-1.367	0.1715
x1	Profundidad ajustada	0.0111	0.0087	1.269	0.2044

$$\text{Modelo Probit: } \Phi^{-1}(\hat{\pi}_i) = X_i^T \hat{\beta}$$

El modelo probit incorpora 6 variables en el proceso de selección de variables paso a paso. Dentro de ellas, la distancia mínima a la costa, la gradiente en el eje x local, la curvatura de la pendiente máxima local y el ángulo de inclinación en el eje x son las más significativas. Los factores influyentes de manera positiva en la presencia de depocentros son la gradiente en el eje x local, la curvatura de la pendiente máxima local y la curvatura perpendicular a la pendiente máxima. Por otro lado, la distancia mínima a la costa, el ángulo de inclinación en el eje x y la profundidad ajustada tienen influencia negativa.

**Cuadro 5.3:** Variables incorporadas al modelo probit sin interacciones y su significancia

ID Variable	Nombre de la Variable	Coficiente Estimado	Error Estándar	Estadístico T	Valor de p
Constante	Intercepto	-0.5258	0.0991	-5.306	0.0000
x3	Distancia mínima a la costa	-1.0998	0.2760	-3.985	0.0001
x18	Gradiente en el eje x - local	0.2422	0.0989	2.450	0.0143
x19	Curvatura de la pendiente máxima - local	0.1821	0.0767	2.373	0.0176
x6	Ángulo de inclinación en el eje x	-0.1532	0.0717	-2.135	0.0327
x1	Profundidad Ajustada	-0.0004	0.0002	-1.631	0.1029
x9	Curvatura perpendicular a la pendiente máxima	901.5249	608.6038	1.481	0.1385

$$\text{Modelo ComLogLog: } \log(-\log(1 - \hat{\pi}_i)) = X_i^T \hat{\beta}$$

El modelo comloglog incorpora 5 variables en el proceso de selección de variables paso a paso. Dentro de ellas, la distancia mínima a la costa, la curvatura perpendicular a la pendiente máxima, gradiente en el eje x - local y el ángulo de inclinación en el eje y son las más significativas. Los factores influyentes de manera positiva en la presencia de depocentros son la curvatura perpendicular a la pendiente máxima, gradiente en el eje x - local y el ángulo de inclinación en el eje y. Por otro lado, la distancia mínima a la costa y el ángulo de inclinación en el eje x tienen influencia negativa.

Cuadro 5.4: Variables incorporadas al modelo comloglog sin interacciones y su significancia

ID Variable	Nombre de la Variable	Coefficiente Estimado	Error Estándar	Estadístico T	Valor de p
Constante	Intercepto	-0.9919	0.1419	-6.990	0.0000
x3	Distancia mínima a la costa	-1.4487	0.4117	-3.519	0.0004
x9	Curvatura perpendicular a la pendiente máxima	1925.8582	744.1693	2.588	0.0097
x18	Gradiente en el eje x - local	0.3351	0.1397	2.399	0.0164
x7	Ángulo de inclinación en el eje y	0.1354	0.0649	2.087	0.0368
x6	Ángulo de inclinación en el eje x	-0.1869	0.1115	-1.676	0.0937

(b) Bondad de Ajuste

En el cuadro 5.5 se muestra algunos estadísticos para probar la bondad de ajuste de los modelos tales como el criterio de AIC, la devianza y una comparación del modelo con un modelo constante.

Según el criterio de Akaike y la devianza, el modelo probit es el mejor estadísticamente. Sin embargo los tres son significativos al compararlos con un modelo constante. En conclusión, el modelo probit es el más adecuado sin considerar las interacciones y además es significativo al compararlo contra un modelo constante.

Cuadro 5.5: Indicadores estadísticos de la bondad de ajuste

Prueba o Indicador	Logit	Probit	Valor Extremo
Criterio AIC	492.1594	486.9405	490.9814
Devianza del modelo			
Devianza ( $D_1$ )	480.1594	472.9405	478.9814
Grados de libertad ( $N - p$ )	452	451	452
Límite superior 95 % ( $X_{est}$ )	508.8931	501.5110	501.5110
Comparación con un modelo constante			
Devianza modelo constante ( $D_0$ )	509.53	509.53	509.53
Grados de libertad modelo constante ( $N - 1$ )	457	457	457
Estadístico de prueba ( $D_0 - D_1$ )	29.373	36.592	30.551
Significancia de la prueba ( $P$ )	<b>1.9587e-05</b>	<b>2.1143e-06</b>	<b>1.1485e-05</b>

A pesar de que los modelos sean significativos y que el modelo probit sea el más adecuado en términos de ajuste estadístico, los modelos no predicen bien las zonas de ambientes de deposición y además el modelo comloglog (valor extremo) es el que tiene mayor ca-

pacidad de predicción. Ello puede ser debido a que el modelo necesita más covariables o interacciones para que se pueda lograr un mejor ajuste.

**Cuadro 5.6:** Capacidad de predicción de los modelos

	Logit		Probit		Valor Extremo	
	Verdad	Falso	Verdad	Falso	Verdad	Falso
Verdad	343	3	341	5	343	3
Falso	106	6	104	8	105	7
Porcentaje de correcta predicción	76.2 %		76.2 %		76.42 %	

## 2. Modelo Con Interacciones

En la sección anterior presentamos los modelos logit, probit y comloglog para la predicción de depocentros. Aunque estos tienen un nivel predictivo en general, no se identifican bien los ambientes de deposición continua por lo que tratamos de mejorar aquellos modelos incorporando la interacción de las covariables.

### (a) Descripción del Modelo

En los cuadros 5.7, 5.8 y 5.9 se muestran las variables e interacciones incorporadas a los modelos logit, probit y comloglog respectivamente mediante el stepwise con el criterio de Akaike.

**Modelo Logit:**  $logit(\hat{\pi}_i) = \log\left(\frac{\hat{\pi}_i}{1-\hat{\pi}_i}\right) = X_i^T \hat{\beta}$

Cuadro 5.7: Variables incorporadas al modelo logit con interacciones y su significancia

ID Variable	Nombre de la Variable	Coefficiente Estimado	Error Estándar	Estadístico T	Valor de P
x16	Profundidad ajustada - local	-0.0336	0.0054	-6.23	0.00000
(Intercepto)	Intercepto	-2.1609	0.4349	-4.97	0.00000
x16:x17	interacción	0.0414	0.0095	4.36	0.00001
x17	Máximo ángulo de inclinación - local	6.1803	1.4326	4.31	0.00002
x3	Distancia mínima a la costa	-7.0080	1.6377	-4.28	0.00002
x16:x18	interacción	-0.0402	0.0107	-3.77	0.00016
x3:x16	interacción	0.0178	0.0049	3.62	0.00029
x3:x19	interacción	8.8602	2.5999	3.41	0.00065
x3:x18	interacción	-9.5088	2.8226	-3.37	0.00075
x9:x19	interacción	1.21E+004	3.66E+003	3.32	0.00091
x6	Ángulo de inclinación en el eje x	-2.8482	0.8626	-3.30	0.00096
x2	Error de la batimetría	-0.0034	0.0011	-3.23	0.00123
x26	Máximo ángulo de inclinación con dirección - local	-0.9666	0.3062	-3.16	0.00159
x2:x3	interacción	0.0073	0.0023	3.12	0.00180
x7:x16	interacción	-0.0097	0.0032	-3.05	0.00229
x7:x18	interacción	1.5098	0.5161	2.93	0.00344
x6:x16	interacción	-0.0086	0.0030	-2.89	0.00379
x18	Gradiente en el eje x - local	-4.2191	1.5983	-2.64	0.00830
x7:x17	interacción	-1.5490	0.6102	-2.54	0.01113
x13:x17	interacción	-0.4614	0.1987	-2.32	0.02022
x13:x16	interacción	-0.0034	0.0015	-2.24	0.02487
x3:x9	interacción	4.95E+004	2.22E+004	2.23	0.02560
x24	Distancia mínima hacia zonas de depresiones - local	-0.0025	0.0011	-2.20	0.02799
x3:x13	interacción	-1.4105	0.6742	-2.09	0.03644
x9:x16	interacción	27.0392	12.9623	2.09	0.03698
x2:x13	interacción	0.0002	0.0001	1.99	0.04709
x15	Máximo ángulo de inclinación con dirección perpendicular	-0.6381	0.3441	-1.85	0.06368
x3:x7	interacción	-2.1767	1.2006	-1.81	0.06983
x9	Curvatura perpendicular a la pendiente máxima	-1.43E+004	8.57E+003	-1.67	0.09545
x21	Curvatura perpendicular a la pendiente máxima - local	1.18E+005	7.12E+004	1.65	0.09816
x16:x24	interacción	0.0000	0.0000	-1.65	0.09824
x7	Ángulo de inclinación en el eje y	0.4556	0.6218	0.73	0.46373
x7:x19	interacción	0.1495	0.3439	0.43	0.66375
x19	Curvatura de la pendiente máxima - local	0.2018	0.7803	0.26	0.79591
x13	Distancia mínima hacia zonas de depresiones según 13	0.0113	0.1435	0.08	0.93720

$$\text{Modelo Probit: } \Phi^{-1}(\hat{\pi}_i) = X_i^T \hat{\beta}$$

**Cuadro 5.8:** Variables incorporadas al modelo probit con interacciones y su significancia

ID Variable	Nombre de la Variable	Coefficiente Estimado	Error Estándar	Estadístico T	Valor de p
x1	Profundidad ajustada	-0.0206	0.0027	-7.635	0.0000
Constante	Intercepto	-1.3563	0.2494	-5.439	0.0000
x3	Distancia mínima a la costa	-3.9567	0.9344	-4.235	0.0000
x17	Máximo ángulo de inclinación - local	3.1060	0.7498	4.142	0.0000
x1:x17	interacción	0.0251	0.0066	3.797	0.0001
x1:x3	interacción	0.0132	0.0035	3.782	0.0002
x1:x18	interacción	-0.0177	0.0051	-3.476	0.0005
x1:x19	interacción	-0.0077	0.0024	-3.179	0.0015
x18	Gradiente en el eje x - local	-3.1485	1.0583	-2.975	0.0029
x6:x17	interacción	0.6651	0.2246	2.962	0.0031
x1:x6	interacción	-0.0125	0.0045	-2.779	0.0055
x24	Distancia mínima hacia zonas de depresiones - local	-0.0022	0.0008	-2.701	0.0069
x5:x6	interacción	-0.8906	0.3452	-2.580	0.0099
x5:x22	interacción	-0.6408	0.2884	-2.222	0.0263
x17:x18	interacción	0.7847	0.3548	2.212	0.0270
x3:x6	interacción	-3.9515	1.7974	-2.199	0.0279
x1:x24	interacción	0.0000	0.0000	-2.190	0.0285
x3:x19	interacción	3.8529	1.7810	2.163	0.0305
x5:x17	interacción	-0.8072	0.3799	-2.124	0.0336
x17:x19	interacción	-0.6279	0.3367	-1.865	0.0622
x6	Ángulo de inclinación en el eje x	-1.0356	0.5613	-1.845	0.0650
x5:x19	interacción	1.0613	0.6164	1.722	0.0851
x22:x24	interacción	-0.9933	0.5784	-1.717	0.0859
x22	Tipo de curvatura de la pendiente máxima	0.3911	0.3099	1.262	0.2069
interacción	Máximo ángulo de inclinación	0.4557	0.3735	1.220	0.2224
x19	Curvatura de la pendiente máxima - local	-0.3212	0.5601	-0.573	0.5664

$$\text{Modelo ComLogLog: } \log(-\log(1 - \hat{\pi}_i)) = X_i^T \hat{\beta}$$

**Cuadro 5.9:** Variables incorporadas al modelo comloglog con interacciones y su significancia

ID Variable	Nombre de la Variable	Coefficiente Estimado	Error Estándar	Estadístico T	Valor de p
Constante	Intercepto	-0.9604	0.1797	-5.345	0.0000
x18	Gradiente en el eje x - local	1.3792	0.3308	4.169	0.0000
x3	Distancia mínima a la costa	-2.1590	0.5433	-3.974	0.0001
x7	Ángulo de inclinación en el eje y	0.7798	0.2320	3.361	0.0008
x3:x6	interacción	0.9999	0.3481	2.872	0.0041
x7:x18	interacción	-0.2485	0.0873	-2.846	0.0044
x6:x18	interacción	-0.1802	0.0640	-2.815	0.0049
x6:x7	interacción	0.1372	0.0546	2.511	0.0120
x24	Distancia mínima hacia zonas de depresiones - local	-0.0008	0.0004	-2.289	0.0221
x3:x24	interacción	0.0012	0.0006	2.198	0.0279
x6	Ángulo de inclinación en el eje x	-0.5265	0.2577	-2.043	0.0410
x5	Máximo ángulo de inclinación	-0.2539	0.1425	-1.782	0.0747
x3:x7	interacción	-0.4147	0.2421	-1.713	0.0868
x9	Curvatura perpendicular a la pendiente máxima	1010.9412	641.0009	1.577	0.1148

(b) Bondad de AJuste

Según el AIC y la devianza, los modelos probit y logit presentan mejor bondad de ajuste. Al comparar los modelos contra un modelo constante, los tres resultan significativos principalmente los modelos logit y probit. Además para probar la importancia de incorporar interacciones, se contrastó los modelos con interacciones vs. los modelos sin interacciones. Los tres modelos resultaron significativos al añadir las interacciones, es decir, se justifica las interacciones añadidas porque se mejoró la bondad de ajuste. Principalmente mejorar los modelos logit y probit con mayor significancia.

**Cuadro 5.10:** Indicadores estadísticos de la bondad de ajuste de los modelos con interacciones

Prueba o Indicador	Logit	Probit	Valor Extremo
Criterio AIC	367.6082	363.9161	474.5028
Devianza del modelo			
Devianza ( $D_2$ )	297.6082	311.9161	446.5028
Grados de libertad ( $N - p$ )	423	432	444
Límite superior 95 % ( $X_{est}$ )	471.9521	481.4587	494.1260
Comparación con un modelo constante			
Devianza modelo constante ( $D_0$ )	509.53	509.53	509.53
Grados de libertad modelo constante ( $N - 1$ )	457	457	457
Estadístico de prueba ( $D_0 - D_2$ )	211.92	197.62	63.03
Significancia de la prueba ( $P$ )	<b>1.359e-27</b>	<b>8.8122e-29</b>	<b>1.4993e-08</b>
Comparación con un modelo sin interacciones			
Devianza modelo sin interacción ( $D_1$ )	480.1594	472.9405	478.9814
Grados de libertad modelo sin interacción ( $N - p$ )	452	451	452
Estadístico de prueba ( $D_1 - D_2$ )	182.5512	161.0245	32.4786
Significancia de la prueba ( $P$ )	$\cong 0$	$\cong 0$	<b>7.6436e-05</b>

Para evaluar la mejora de los modelos según su capacidad predictiva, se muestra la clasificación obtenida por cada modelo en el cuadro 5.11. Tanto el modelo logit como el probit mejoran drásticamente en la predicción obteniendo 83.62% y 84.93% de predicción correcta respectivamente. Dejamos de lado el modelo de valor extremo porque no es capaz de predecir los ambientes de deposición como lo hace el modelo logit y probit. Decidimos trabajar con el modelo logit porque su fácil interpretación con los odds y porque se puede incorporar a un modelo logístico de manera lógica como se mostró en la sección 4.2.

**Cuadro 5.11:** Capacidad predictiva del modelo

	Logit		Probit		Valor Extremo	
	Verdad	Falso	Verdad	Falso	Verdad	Falso
Verdad	324	22	327	19	337	9
Falso	53	59	50	62	96	16
Porcentaje de correcta predicción	83.62 %		84.93 %		77.07 %	

### 3. Modelo seleccionado e interpretación:

Como se menciona anteriormente, el modelo seleccionado fue el modelo lineal generalizado

con función de enlace logit con interacciones 5.7. Las variables más importantes están relacionadas a características de la vecindad y a su lejanía con las costas. La profundidad de la superficie suavizada del relieve marino (P-Value  $\cong 0$ ), el máximo ángulo de inclinación de la superficie suavizada (P-Value  $\cong 0,00002$ ) y la distancia mínima a la costa (P-Value  $\cong 0,00002$ ) además de otras variables como el ángulo de inclinación en el eje x (Longitud), Error de la batimetría, máximo ángulo de inclinación con dirección a nivel local, gradiente en el eje x de la superficie suavizada, distancia mínima hacia zonas de depresiones de la superficie suavizada, máximo ángulo de inclinación con dirección perpendicular, curvatura perpendicular a la pendiente máxima, curvatura perpendicular a la pendiente máxima de la superficie ajustada, ángulo de inclinación e el eje y, curvatura de la pendiente máxima a nivel local y distancia mínima a depresiones que se muestran a mayor detalle en el cuadro 5.7.

Las interacciones también juegan un rol importante, principalmente, la interacción entre variables como la profundidad ajustada local, el máximo ángulo de inclinación local, la gradiente en el eje x a nivel local, curvatura de la pendiente máxima a nivel local, distancia mínima a la costa y la curvatura perpendicular a la pendiente máxima. Mayor detalle puede verse en la Tab. 5.7, en donde se muestra la importancia de las variables e interacciones en orden descendente.

De este modo el modelo para la predicción de depocentros sedimentarios sin considerar la estructura espacial queda representado del siguiente modo:

$$\text{logit}(\hat{\pi}_i) = \log\left(\frac{\hat{\pi}_i}{1 - \hat{\pi}_i}\right) = X_i^T \hat{\beta} \quad (5.1)$$

En donde  $X_i$  representa el perfil i-ésimo [ $X_i^{16}, X_i^0, X_i^{16*17}, X_i^{17}, \dots$ ] con las variables de la Tab. 5.7 y  $\beta$  son los coeficientes estimados para el modelo que se encuentran en la tercera columna. Con este modelo se pudo clasificar correctamente el 83.62% de los perfiles sedimentarios confirmando que es posible predecir las zonas de depocentro en base a las características de batimetría y morfometría.

### 5.3.2. Comportamiento Espacial

El comportamiento espacial se analiza con los residuales del modelo logit, de tal modo que se termina de explicar la variabilidad de los datos casi al 100%. En la Fig. 5.14 se observa el variograma para los residuales, con los círculos azules indicando la cantidad de pares con los que se

hizo el cálculo. Al inicio del variograma existe un desfase debido al efecto pepita ( $\tau^2$ ) alrededor de 4.5, lo cual indica que existe una estructura espacial a una escala menor de la estudiada. Lo anterior es lógico porque con el modelo logit explicamos alrededor del 85% de los datos en la primera fase del modelamiento, y solo empleamos el kriging para ajustar el resto de la variabilidad.

Junto al variograma de la Fig. 5.14, se muestra los modelos teóricos de correlación espacial. Decidimos quedarnos con el modelo exponencial con un efecto pepita no tan grande (de color negro) para asegurar que puntos que no sean bien estimados por el modelo logístico tengan alta influencia a distancias cortas. Lo cual no sucede si se considera un efecto pepita grande ya que se estaría concluyendo que la estructura espacial es aleatoria. De este modo el alcance del modelo exponencial con  $\tau = 2,4677$  es aproximadamente 0.25 grados (equivalente a 27.7 km).

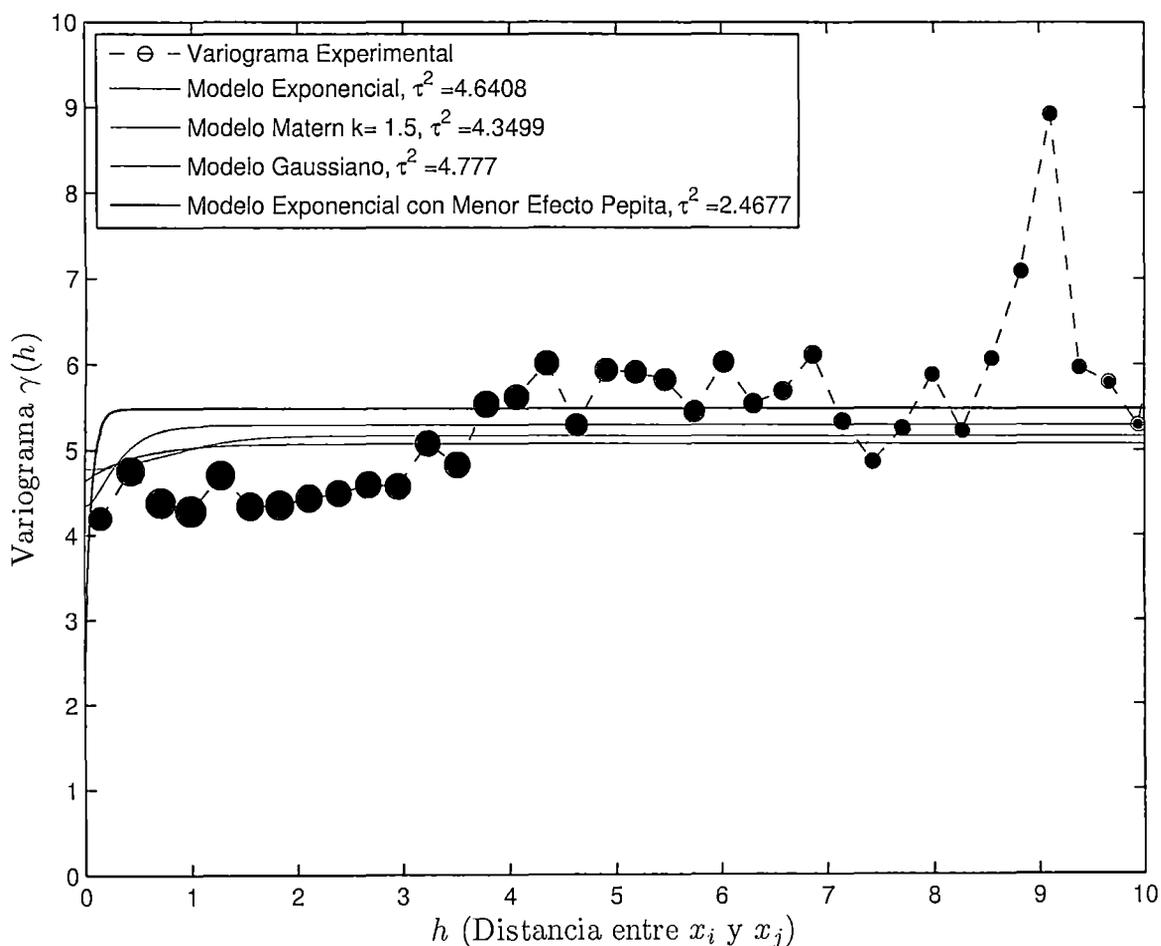


Figura 5.14: Variograma de los residuos del modelo logístico junto a los modelos ajustados.  $\tau^2$  representa el efecto pepita de cada modelo

## 5.4. Predicción de la Ubicación de Depocentros Sedimentarios

Considerando el modelo de regresión logístico y la estructura espacial mostrados en la sección anterior, se obtuvo un mapa de probabilidades en el margen continental peruano, es cual es presentado a continuación.

### 5.4.1. Características Generales del Mapa de Probabilidades sobre la Presencia de Depocentros

En primer lugar, este es un mapa de alta resolución, en el cual se observa la probabilidad de que una zona sea un ambiente de deposición continua fina. En la Fig. 5.15 se muestra el mapa a menor resolución para que pueda ser ilustrada. Su confiabilidad se reduce mientras nos alejamos de los puntos muestreados. La zona sur, centro-sur es en donde se encuentra zonas con mayor probabilidad a ser depocentros mientras que en la zona norte la ocurrencia es menor. Por otro lado, también se encuentran cerca al quiebre de la Plataforma peruana posiblemente por la dinámica y variabilidad que hay en estas zonas creando ambientes protegidos adecuados para la deposición sedimentaria.

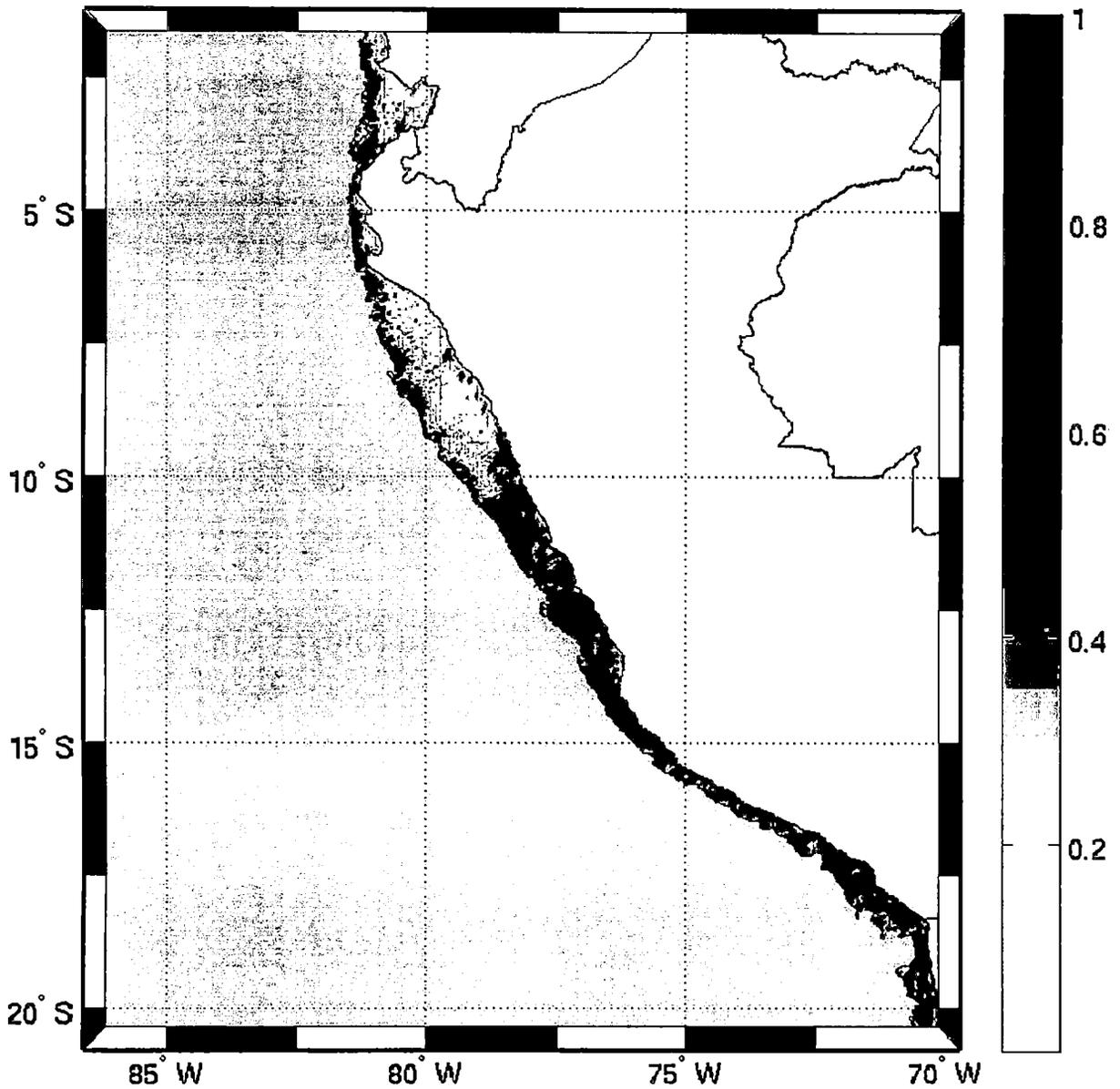


Figura 5.15: Predicción de probables zonas de deposición continua en la Plataforma Continental y la zona superior del Talud Continental.

La zona norte puede verse a mayor detalle en la Fig. 5.16. Existen pequeñas zonas cerca a las costas que son posibles ambientes de deposición. Sin embargo, la mayor cantidad de probables ambientes de deposición continua se encuentran en el quiebre de la Plataforma. Ello debe ser, principalmente por la alta variabilidad y rugosidad que existe en esas zonas.

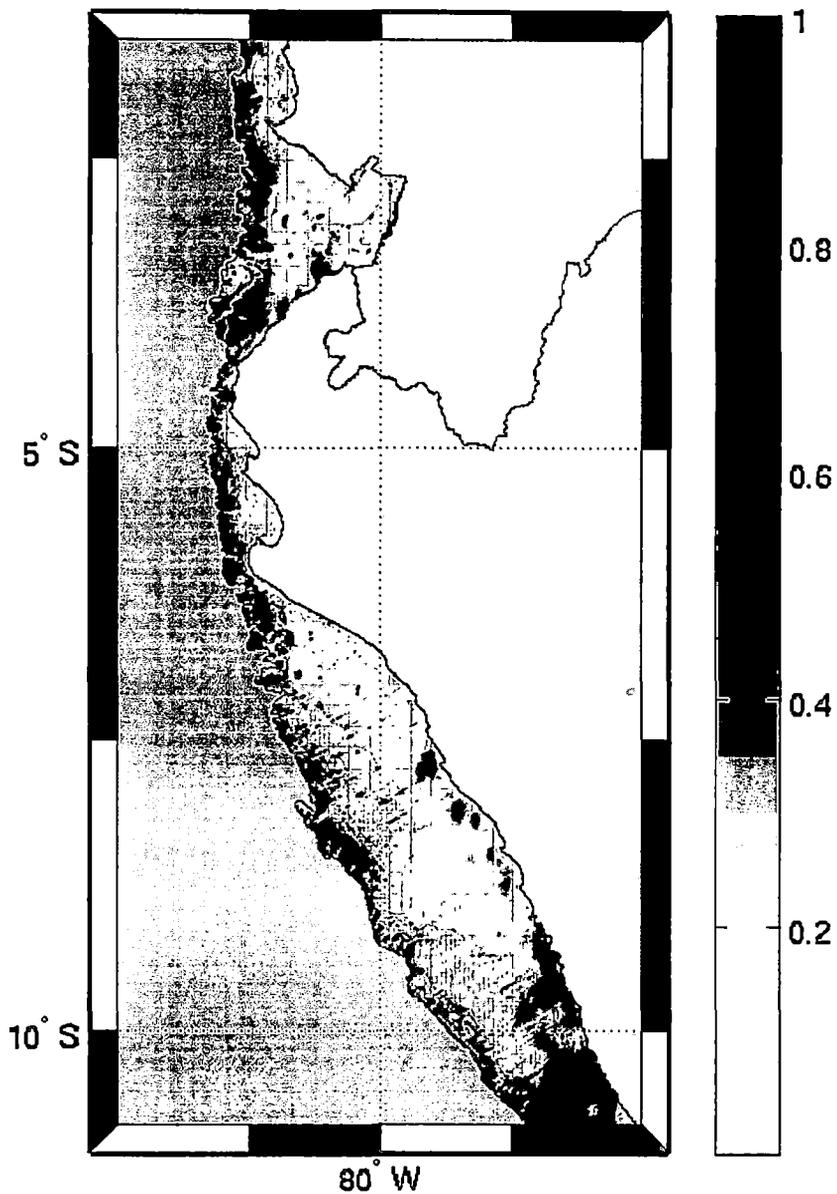


Figura 5.16: Predicción de probables zonas de deposición continua en la zona norte de la Plataforma Continental y la zona superior del Talud Continental.

La parte sur del área de estudio es una zona más tranquila respecto a la hidrodinámica. La Fig. 5.17 nos indica que en esta zona hay mayor probabilidad de encontrar más depocentros. Debido a la resolución de la imagen no se puede observar mayor detalle y pareciera que la predicción nos dice que casi toda el área es de deposición continua. Por lo que un pequeño zoom se muestra en la Fig. 5.18 mostrando a mayor detalle la creencia de presencia de depocentros en la zona sur según la

morfometría y la batimetría. Lo cual nos ayuda a notar los probables ambientes de deposición con mayor precisión.

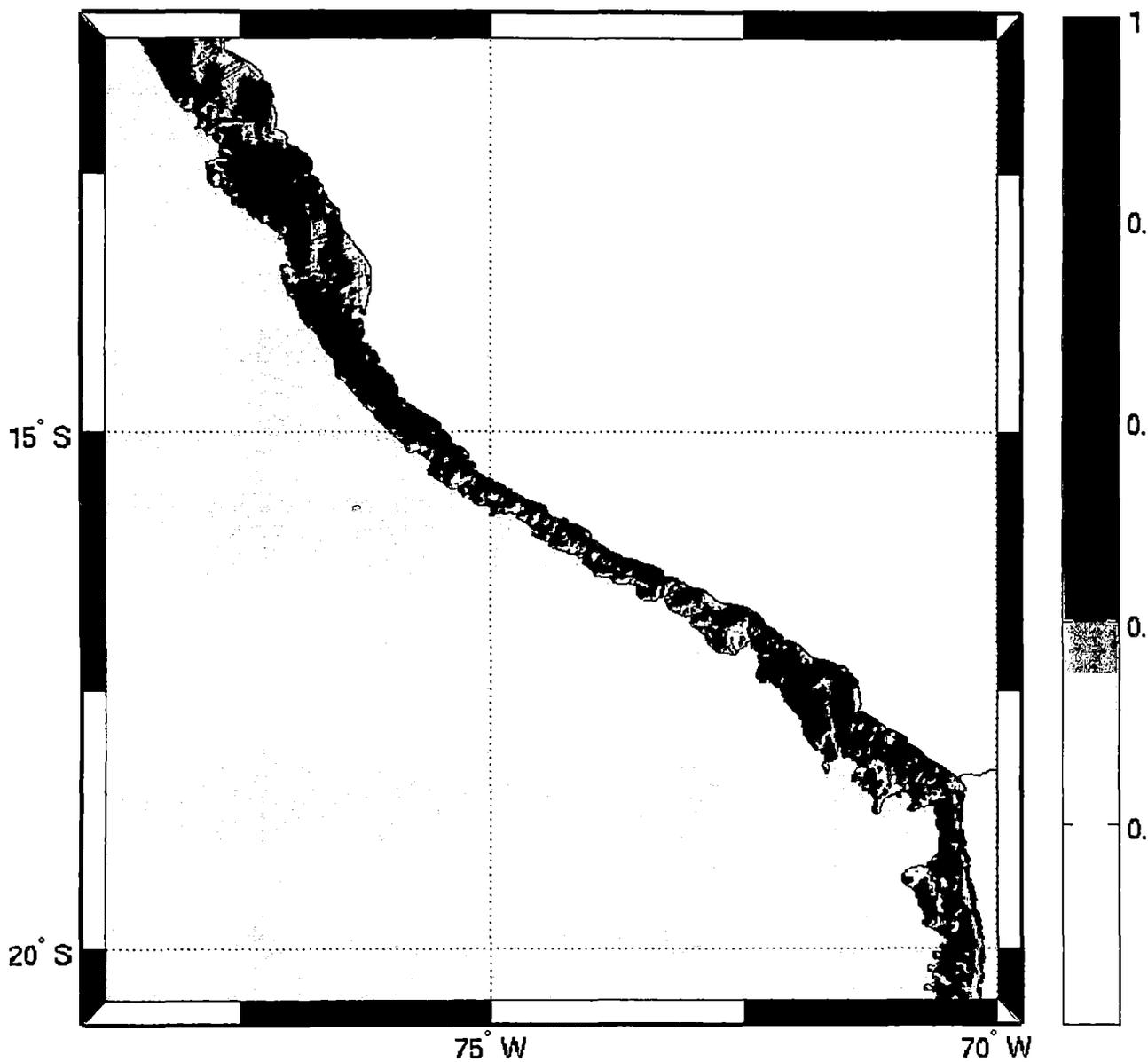


Figura 5.17: Predicción de probables zonas de deposición continua en la zona sur de la Plataforma Continental y la zona superior del Talud Continental.

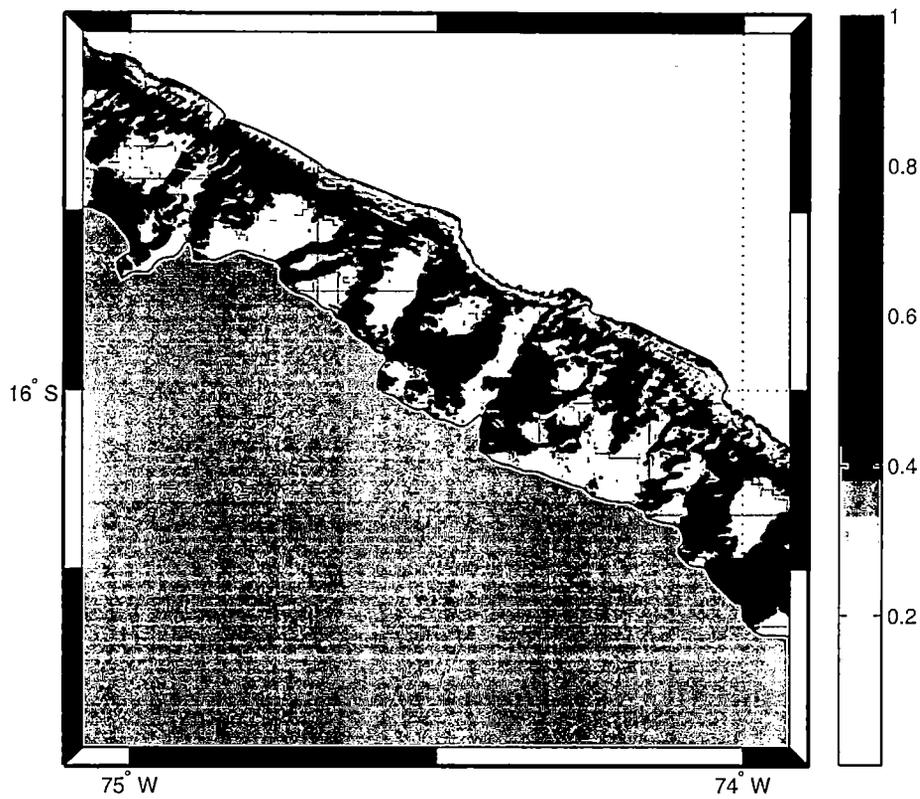


Figura 5.18: Predicción de probables zonas de deposición continua en la zona sur de la Plataforma Continental y la zona superior del Talud Continental a mayor resolución

# Capítulo 6

## Discusión

### 6.1. Acerca de la Batimetría

La grilla batimétrica obtenida en el estudio es de alta resolución y viene en conjunto con una grilla de errores de predicción, estas dos características hacen que sea de gran importancia para estudios de investigación en el ecosistema marino, estudios que involucran tanto el ecosistema bentónico como el pelágico. Además ha sido obtenido de un conjunto diverso de fuentes de datos, por lo que se planteó un enfoque de modelamiento para tal situación obteniendo resultados satisfactorios.

Por otro lado, el producto (los mapas obtenidos) no tienen los estándares cartográficos rigurosos, pero su información es de calidad y válida y es el insumo básico que posteriormente puede ser ploteada y empleada en la construcción de cartas y mapas con estos estándares. Además, el objetivo principal de esta grilla es ser útil para estudios de modelamiento en el marco del Sistema de la Corriente de Humboldt.

#### 6.1.1. ¿Qué se encontró?

El proyecto de elevación y batimetría SRTM30\_PLUS es uno de los más recientes y validados porque incorporó un conjunto de fuentes internacionales empleando modelos de gravimetría, también considero datos de cruceros multibeam de la NOAA y otras fuentes de institutos de investigación [Becker et al., 2009]. Sin embargo, de manera local, nosotros contamos con datos batimétricos de IMARPE, DHN e IRD distribuidas en la Plataforma Continental peruana y en algunas zonas

del Talud Continental superior. Ello hizo plantear que la grilla batimétrica del mar peruano podría mejorarse si consideramos nuestras fuentes locales. Mediante el modelamiento de la batimetría con las diversas fuentes pudimos comprobar que la grilla SRTM30\_PLUS tiene discrepancias con nuestras fuentes locales en ciertas zonas. Y que estas discrepancias van variando de zona en zona, sin que haya una diferencia estructural, constante o determinista. Esto nos hace pensar que la diferencia se debe a patrones falsos y aleatorios que se crearon al momento de modelar la batimetría mediante el proyecto SRTM30\_PLUS que puede ser por falta de información o porque los modelos gravimétricos son sensibles a ciertos materiales a pequeñas escalas. Sin ahondar en ello, logramos obtener una grilla batimétrica que este altamente influenciada por los datos de IMARPE, DHN e IRD en la Plataforma Continental y más allá, por los datos de la NOAA y en zonas donde no hay datos de campo, el dominio lo tiene los datos de SRTM30\_PLUS. Por lo tanto, el mapa obtenido muestra una mejora en función a la validez en la Plataforma Continental y en la mejora de resolución en zonas cercanas a datos multibeam.

### 6.1.2. Problemática del Trabajo de Modelamiento

Así como el mapa tiene sus fortalezas, también tiene sus debilidades y es que temas como la calidad de los datos y el trabajar con gran cantidad de datos limita el proceso del estudio y los resultados.

#### 1. Calidad de los datos

La calidad de los datos son un tema clave en la interpolación por medio del kriging por varios motivos. Primero, si la calidad no es buena, se podría estar sesgando los resultados en caso no se realice la limpieza de datos adecuada porque el kriging tratará de interpolar el punto sesgado creando patrones estructurales falsos (una elevación o profundidad falsa) . Segundo, existe mayor error de medición y ello involucra predicciones con grandes intervalos de confianza. Tercero, complica el proceso del estudio, realizando trabajos que se deberían haber hecho en la recolección de datos, entre otros.

#### 2. Modelos con enormes cantidad de datos

El método de estimación por máxima verosimilitud involucra inversión de matrices de correlación, lo cual dificulta la factibilidad de su aplicación en gran cantidad de datos por lo que

sería adecuado emplear técnicas de geoestadística para Big Data, como por ejemplo la estimación por matrices dispersas. Diversos estudios ya se han estado realizando gracias al empleo de la estimación por medio de matrices dispersas [Neumaier and Groeneveld, 1998, Barry and Pace, 1999, Reinhardt and Kudrass, 2003, Couckuyt et al., 2012, Sun et al., 2012]. En el estudio se considero su aplicación por clusters espaciales y ventanas móviles, sin embargo ello crea algunos desfases en los límites de estas áreas. Se consideró el empleo de matrices dispersas pero solo en las funciones de correlación spline, esférica y lineal [Lophaven et al., 2002]. En los demás casos (familia matern, exponencial) es difícil conseguir matrices dispersas de correlación que sean definidas positivas en un corte de distancia arbitrario.

Finalmente, se puede decir que el mapa trae consigo varias ventajas pero que se puede mejorar más adelante considerando tópicos de geoestadística robusta y su aplicación a gran cantidad de datos.

### **6.1.3. Aporte del Mapa Obtenido y Posibles Mejoras en Adelante**

Gracias al mapa obtenido, se puede conocer a mayor detalle las características del fondo marino en zonas donde pasaron los cruceros de haz múltiple y se pueden unir un conjunto de fuentes de datos que eran muy distintas en ciertas zonas.

Las posibles mejoras que se pueden realizar son el empleo de matrices dispersas para la estimación por máxima verosimilitud, incluir variables del ambiente como covariables para mejorar el modelo en zonas donde no se tiene información, actualizar el mapa con nuevos datos de cruceros científicos no considerados.

## **6.2. Acerca de los Posibles Ambientes de Deposición Continua**

El mapa de predicción de ambientes de deposición continua obtenido es de alta resolución, se realizó con un modelo kriging logístico. No se cuantificó el error de predicción debido a que lo hacemos en dos fases (1: Modelo de regresión logístico, 2: kriging ordinario). Sin embargo ha de considerarse que la creencia es más confiable en cuanto el posible ambiente de deposición no este aislado, mientras más cerca se encuentra de un dato muestral y en cuando se encuentre dentro del

polígono de datos muestrales.

### 6.2.1. ¿Qué se encontró?

Teóricamente se conoce que los ambientes de deposición continua tienen características de granos finos de sedimentos y composición dominada por arcilla y limo. Este perfil pudo ser encontrado de manera natural mediante clustering. Este cluster o grupo tiene las características de los ambientes de deposición continua (granos finos y composición dominada por arcilla y limo). Lo cual nos hace validar la información que tenemos y la segunda hipótesis específica planteada.

Por otra lado se encontró la asociación que hay entre ciertos rasgos morfológicos con los ambientes de deposición continua. La profundidad, el máximo ángulo de inclinación (pendiente), distancia mínima a la costa, ángulo de inclinación en el eje de longitud, error de batimetría y el máximo ángulo de inclinación con dirección (según si es elevación o depresión) son las variables con mayor relevancia.

Finalmente, se encontraron ciertos patrones generales en la predicción de los ambientes de deposición continua. Estos ambientes principalmente se ubican en la zona sur. Ello puede ser explicado por que en esta región existe menor hidrodinámica que en la zona norte. Además, las zonas de deposición continua se encuentran cerca al quiebre de la Plataforma debido a que en estas zonas existe mayor variabilidad (rugosidad).

### 6.2.2. Problemática del Trabajo de Modelamiento

#### 1. Limitación de los Datos

El estudio no consideró directamente variables relativamente importantes respecto a la hidrodinámica del océano. Se intentó incorporar la descarga de los ríos en el océano. Ello debido a la importancia que tiene los ríos al transportar los sedimentos al océano, concentrando el 95 % de sedimentos que entran al océano [Syvitski et al., 2003]. Además, la distribución de los muestras no es la más idónea para este tipo de estudios, porque están muy distanciados. Por lo que se estos factores limitan la confianza brindada en el resultado para zonas muy alejadas de las ubicaciones muestreadas, principalmente la zona sur y la zona norte (en sus extremos).

### **6.2.3. Aporte del Nuevo Mapa Obtenido**

Hasta la fecha, no se había realizado un estudio para la predicción de depocentros en el margen continental peruano, ni se había detectado los posibles perfiles sedimentarios que habían, tampoco se había detectado la relación entre la morfología y batimetría con la sedimentología de manera cuantitativa. Lo que si se hizo en el estudio y se mostró modelos con una buena capacidad predictiva basadas en las características batimétricas y morfológicas. Así que este nuevo mapa va a poder ser usado por expertos de geología para que sea un apoyo en la toma de decisiones a la hora de realizar trabajos de campo. Esperando que se confirme lo obtenido y sea un gran aporte para el conocimiento científico y para extender su aplicación a distintos estudios relacionados al ecosistema marino.

## Capítulo 7

# Conclusiones

El presente estudio tiene el objetivo principal de predecir ambientes de deposición continua en el margen continental peruano, sin embargo tenemos productos intermedios que son parte de los resultados y de los aportes finales del estudio. Así que a continuación abordamos las conclusiones principales del proyecto en función a la grilla de batimetría obtenido, al aporte acerca de la predicción de posibles ambientes de deposición continua y acerca de los modelos espaciales empleados.

### 7.1. De la Batimetría

**(Obj. 1)** La nueva grilla de batimetría obtenida para el área cubierta por los rangos de latitud  $[-20,7, -1,3]$  y de longitud  $[-85,6, -69,8]$  es una importante mejora de la grilla SRTM30\_PLUS en validez y resolución. Nuestra batimetría fue obtenida a partir de datos locales de cruceros científicos de IMARPE, IRD y DHN, y datos internacionales de la NOAA, GADM, SRTM30\_PLUS (para zonas sin información). Se logró identificar que la batimetría de SRTM30\_PLUS tiene diferencias con los datos locales en la Plataforma Continental peruana y fue mejorada al incorporar datos de las fuentes locales. Además, debido a disponibilidad de cruceros de alta resolución de la NOAA, IMARPE e IRD se mejoró la resolución del mapa capturando rasgos más detallados del fondo marino. Es importante mencionar que nuestra batimetría tiene concordancia con la línea costera peruana del proyecto GADM. Ello involucra una mejora porque GADM tiene mayor compatibilidad que GSHHS (una de las mejores líneas costeras a nivel mundial) con las demás fuentes de batimetría. Por lo tanto, la grilla batimétrica obtenida nos permitirá analizar procesos o fenómenos en el marco

del Sistema de la Corriente de Humboldt a escalas finas relacionadas con la batimetría.

(Obj. 3) Gracias a la batimetría obtenida y herramientas útiles de paquete TopoToolbox [Schwanghart and Kuhn, 2010] se pudo calcular indicadores morfológicos relevantes para los análisis de sedimentología, morfología y otros. Entre las principales variables se calcularon indicadores como la pendiente (en varias direcciones), la concavidad (elevación o depresión), rugosidad (variabilidad), distancia mínima a las costas, distancia mínima hacia una depresión, entre otros.

## 7.2. De los Posibles Ambientes de Deposición Sedimentaria

(Obj. 2) En primer lugar, se determinó el perfil sedimentario de los ambientes de deposición continua con una composición promedio de 94.22 % de limo y arcilla, y en un coeficiente de asimetría promedio de -0.2572.

(Obj. 4) Las principales variables para la predicción de los ambientes de deposición continua están relacionadas a características de la vecindad y a su lejanía con las costas. La profundidad de la superficie suavizada del relieve marino (P-Value  $\cong 0$ ), el máximo ángulo de inclinación de la superficie suavizada (P-Value  $\cong 0,00002$ ) y la distancia mínima a la costa (P-Value  $\cong 0,00002$ ) además de otras variables como el ángulo de inclinación en el eje x (Longitud), Error de la batimetría, máximo ángulo de inclinación con dirección a nivel local, gradiente en el eje x de la superficie suavizada, distancia mínima hacia zonas de depresiones de la superficie suavizada, máximo ángulo de inclinación con dirección perpendicular, curvatura perpendicular a la pendiente máxima, curvatura perpendicular a la pendiente máxima de la superficie ajustada, ángulo de inclinación e el eje y, curvatura de la pendiente máxima a nivel local y distancia mínima a depresiones que se muestran a mayor detalle en el cuadro 5.7.

(Obj. 5) El mapa de la distribución de probables ambientes de deposición sedimentaria continua representa una herramienta útil para los estudios de sedimentología '*in situ*' y analíticos. Fue obtenido a partir de características de morfología y batimetría de alta resolución. De manera global, los ambientes de deposición continua se encuentran principalmente en la zona sur de la Plataforma Continental. En la parte norte, se encuentran en el quiebre de la Plataforma debido a la variabilidad que hay en esas zonas. Gracias al mapa obtenido, se puede evaluar la presencia de posibles ambientes de deposición continua de manera más detallada. En conjunto con la habilidad y conocimiento del investigador geológico, servirá como una herramienta útil para mejorar la planificación de los trabajos

de campo.

### 7.3. De los Modelos Espaciales

La predicción de ambientes de deposición continua, se realizó a través de modelos lineales generalizados, modelo logístico para nuestro caso, y técnicas de modelamiento espacial (kriging). En conjunto mostraron resultados satisfactorios.

Además, la aplicación de técnicas espaciales para la limpieza, análisis y modelamiento de datos en un gran área como el realizado en el estudio, es de alta complejidad. Por lo que la metodología presentada para cada proceso será útil para este tipo de trabajos. Además se aconseja el empleo de técnicas de modelamiento espacial que sean aplicables a gran cantidad de datos como el empleo de matrices dispersas. Por otro lado, debido a la complejidad que existe en la morfología del mar peruano, se recomienda emplear técnicas que consideren que la tendencia polinómica y la anisotropía varía a lo largo de las costas [Gumiaux et al., 2003, Budrikaité, 2005, Boisvert et al., 2009, Lillah and Boisvert, 2012]. El análisis de la correlación espacial espectral (a través de varias escalas) también puede ayudar a mejora la comprensión y modelación espacial de la batimetría.

# Bibliografía

- [Barnes, 1991] Barnes, R. (1991). The variogram sill and the sample variance. *Mathematical Geology*, 23(4):673–678.
- [Barry and Pace, 1999] Barry, R. P. and Pace, R. K. (1999). Monte Carlo estimates of the log determinant of large sparse matrices. *Linear Algebra and its Applications*, 289:41–54.
- [Becker et al., 2009] Becker, J. J., Sandwell, D. T., Smith, W. H. F., Braud, J., Binder, B., Depner, J., Fabre, D., Factor, J., Ingalls, S., Kim, S.-H., Ladner, R., Marks, K., Nelson, S., Pharaoh, A., Trimmer, R., Von Rosenberg, J., Wallace, G., and Weatherall, P. (2009). Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution: SRTM30\_PLUS. *Marine Geodesy*, 32(4):355–371.
- [Bell et al., 2006] Bell, R., Studinger, M., and Karner, G. (2006). Identifying major sedimentary basins beneath the West Antarctic Ice Sheet from aeromagnetic data analysis. In Fütterer, D. K., Damaske, D., Kleinschmidt, G., Hubert, M., and Tessensohn, F., editors, *Antarctica*, chapter III, pages 117–121. Springer Berlin Heidelberg.
- [Berkenkamp, 2009] Berkenkamp, L. (2009). *Discover the Oceans: The World’s Largest Ecosystem (Discover Your World)*. Nomad Press.
- [Biju-Duval, 2002] Biju-Duval, B. (2002). *Sedimentary geology: sedimentary basins, depositional environments, petroleum formation*.
- [Blondel, 2012] Blondel, P. (2012). *Bathymetry and its Applications*. InTech.
- [Boisvert et al., 2009] Boisvert, J. B., Manchuk, J. G., and Deutsch, C. V. (2009). Kriging in the Presence of Locally Varying Anisotropy Using Non-Euclidean Distances. *Mathematical Geosciences*, 41(5):585–601.
- [Budrikaitė, 2005] Budrikaitė, L. (2005). Modeling of zonal anisotropic variograms. *Liet. matem. rink*, 45(2):339–342.
- [Chamberlin and Dickey, 2007] Chamberlin, W. S. and Dickey, T. D. (2007). Ocean Sediments. In *Exploring the World Ocean*, chapter V, page 416. McGraw-Hill Science/Engineering/Math.
- [Chave, 2013] Chave, J. (2013). The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecology letters*, 16 Suppl 1:4–16.
- [Chavez et al., 2008] Chavez, F. P., Bertrand, A., Guevara-Carrasco, R., Soler, P., and Csirke, J. (2008). The northern Humboldt Current System: Brief history, present status and a view towards the future. *Progress in Oceanography*, 79(2-4):95–105.
- [Cinlar, 2013] Cinlar, E. (2013). *Introduction to Stochastic Processes (Dover Books on Mathematics)*. Dover Publications.

- [Couckuyt et al., 2012] Couckuyt, I., Forrester, a., Gorissen, D., De Turck, F., and Dhaene, T. (2012). Blind Kriging: Implementation and performance analysis. *Advances in Engineering Software*, 49:1–13.
- [Cundy and Kershaw, 2013] Cundy, A. and Kershaw, S. (2013). *Oceanography: An Earth Science Perspective*, volume 38.
- [Cunningham and Cunningham, 2013] Cunningham, W. P. and Cunningham, M. A. (2013). *Principles of Environmental Science*. 7 edition.
- [Day, 2007] Day, T. (2007). *Oceans (Ecosystems)*. Facts on File.
- [Delgado and Gomero, 1995] Delgado, C. and Gomero, R. (1995). Atlas sedimentológico de la plataforma continental peruana. Technical report.
- [Delgado et al., 1987] Delgado, C., Gomero, R., Salzwedel, H., Flores, L., and Carbajal, G. (1987). Distribución de los sedimentos en el Margen Continental peruano mapa textural.
- [Diggle and Ribeiro, 2007] Diggle, P. J. and Ribeiro, P. J. (2007). *Model-based Geostatistics (Springer Series in Statistics)*, volume 1. Springer.
- [FAO, 2005] FAO (2005). *Review of the State of World Marine Fishery Resources, Issue 457*. Food & Agriculture Org.
- [Folk, 1968] Folk, R. (1968). Petrology of Sedimentary Rocks.
- [Gelfand et al., 2010] Gelfand, A., Diggle, P., Guttorp, P., and Fuentes, M. (2010). *Handbook of spatial statistics*. CRC Press.
- [Giraldo, 2002] Giraldo, R. (2002). *Introducción a la Geoestadística: Teoría y Aplicación*. Universidad Nacional de Colombia, Bogotá.
- [Gray and Elliot, 2009] Gray, J. S. and Elliot, M. (2009). *Ecology of Marine Sediments*. Oxford University Press, USA, third edition.
- [Gumiaux et al., 2003] Gumiaux, C., Gapais, D., and Brun, J. (2003). Geostatistics applied to best-fit interpolation of orientation data. *Tectonophysics*, 376(3-4):241–259.
- [Herbozo et al., 2006] Herbozo, G., Velazco, F., Gutiérrez, D., Soler, P., and Gutiérrez, M. (2006). Morfología y Cartografía de la Plataforma y del Talud Superior frente a la Costa del Perú Central (11° S y 15° S). *XIII Congreso Peruano de Geología*.
- [Isaaks, 1989] Isaaks, E. H. (1989). *Applied Geostatistics*, volume 2. Oxford University Press.
- [Lark et al., 2012] Lark, R., Dove, D., Green, S., a.E. Richardson, Stewart, H., and Stevenson, a. (2012). Spatial prediction of seabed sediment texture classes by cokriging from a legacy database of point observations. *Sedimentary Geology*, 281:35–49.
- [Laughton, 1962] Laughton, A. S. (1962). Sea Floor-Morphology. *Proceedings of the Royal Society of London.*, 265.
- [Levin, 1992] Levin, S. A. (1992). The Problem of Pattern and Scale in Ecology. *Ecology*, 73(6):1943–1967.
- [Lillah and Boisvert, 2012] Lillah, M. and Boisvert, J. (2012). Inference of 2D and 3D locally varying anisotropy fields for complex geological formations. In *Ninth International Geostatistics Congress*, pages 1–17.

- [Lophaven et al., 2002] Lophaven, S., Nielsen, H., and Søndergaard, J. (2002). Aspects of the DACE Matlab Toolbox. *IMM, Technical University of Denmark, Kongens Lyngby, Denmark*.
- [Matheron and Kleingeld, 1987] Matheron, G. and Kleingeld, W. (1987). The evolution of geostatistics. *Proceedings of the Twentieth International Symposium on the Application of Computers and Mathematics in the Mineral Industries*, 3(x):9–12.
- [Meilianda, 2012] Meilianda, E. (2012). Application of Multivariate Geostatistics to Investigate the Surface Sediment Distribution of the High-Energy and Shallow Sandy Spiekeroog Shelf at the German Bight, Southern North Sea. *Open Journal of Marine Science*, 02(04):103–118.
- [Meilianda et al., 2011] Meilianda, E., Alfian, D., and Huhn, K. (2011). Sediment grain-size distribution analysis at the shallow sandy shelf of the North Sea using multivariate geostatistics. *Procedia Environmental Sciences*, 7:317–322.
- [Miller and Spoolman, 2011] Miller, G. T. and Spoolman, S. E. (2011). *Living in the Environment*.
- [Mofjeld et al., 2004] Mofjeld, H., Symons, C., and Lonsdale, P. (2004). Tsunami scattering and earthquake faults in the deep Pacific Ocean. *Oceanography*, 17(1):38–46.
- [Neumaier and Groeneveld, 1998] Neumaier, A. and Groeneveld, E. (1998). Restricted maximum likelihood estimation of covariances in sparse linear models. *Genetics Selection Evolution*.
- [Pineda and Stefanoni, 2007] Pineda, J. B. and Stefanoni, J. L. H. (2007). Comparing the performance of two spatial interpolation methods for creating a digital bathymetric model of the Yucatan submerged platform. *Pan-American Journal of Aquatic Sciences*, 2:247–254.
- [Rahbek, 2004] Rahbek, C. (2004). The role of spatial scale and the perception of large-scale species-richness patterns. *Ecology Letters*, 8(2):224–239.
- [Reiche et al., 2011] Reiche, S., Hjelstuen, B. O., and Haffidason, H. (2011). High-resolution seismic stratigraphy, sedimentary processes and the origin of seabed cracks and pockmarks at Nyegga, mid-Norwegian margin. *Marine Geology*, 284(1-4):28–39.
- [Reinhardt and Kudrass, 2003] Reinhardt, L. and Kudrass, H. (2003). High-resolution sediment echosounding off Peru: Late Quaternary depositional sequences and sedimentary structures of a current-dominated shelf. *Marine Geophysical . . .*, 23:335–351.
- [Schwanghart and Kuhn, 2010] Schwanghart, W. and Kuhn, N. J. (2010). TopoToolbox: A set of Matlab functions for topographic analysis. *Environmental Modelling & Software*, 25(6):770–781.
- [Sun et al., 2012] Sun, Y., Li, B., and Genton, M. G. (2012). Geostatistics for Large Datasets. *Advances and Challenges in Space-time Modelling of Natural Events*, 207.
- [Sverdrup et al., 2005] Sverdrup, K., Duxbury, A., and Duxbury, A. (2005). The Sea Floor and Its Sediments. In *An introduction to the world's oceans*, chapter IV, pages 96–130.
- [Syvitski et al., 2003] Syvitski, J. P., Peckham, S. D., Hilberman, R., and Mulder, T. (2003). Predicting the terrestrial flux of sediment to the global ocean: a planetary perspective. *Sedimentary Geology*, 162(1-2):5–24.
- [Tait and Dipper, 1998] Tait, R. R. V. and Dipper, F. A. (1998). *Elements of marine ecology*.
- [Tatham, 1975] Tatham, R. H. (1975). Surface-wave dispersion applied to the detection of sedimentary basins. *Geophysics*, 40:40–55.

- [Tyler, 2003] Tyler, P. (2003). *Ecosystems of the Deep Ocean (Ecosystems of the World) Volume 28*. Elsevier Science.
- [Velazco, 2001] Velazco, F. (2001). Estudio sedimentológico de la bahía de Paracas.
- [Velazco, 2012] Velazco, F. (2012). Distribución de los sedimentos del margen continental peruano y su importancia en los estudios relacionados al sector pesquero. pages 1–5.
- [Verfaillie et al., 2006] Verfaillie, E., Van Lancker, V., and Van Meirvenne, M. (2006). Multivariate geostatistics for the predictive modelling of the surficial sand distribution in shelf seas. *Continental Shelf Research*, 26(19):2454–2468.
- [Waller and Gotway, 2004] Waller, L. A. and Gotway, C. A. (2004). *Applied Spatial Statistics for Public Health Data*.
- [Wenworth, 1922] Wenworth, K. (1922). A Scale of Grade and Class Terms for Clastic Sediments. *The Journal of Geology*, 30(5):377–392.