



Memoria del proyecto para optar al Título de
Ingeniero Civil Oceánico

ANÁLISIS MULTIVARIADO DE VALORES EXTREMOS DE OLEAJE EN CHILE

Claudio Alejandro Meza Valle

Junio 2016

ANÁLISIS MULTIVARIADO DE VALORES EXTREMOS DE OLEAJE EN CHILE

Claudio Alejandro Meza Valle

COMISIÓN REVISORA

NOTA

FIRMA

Sr. Hugo Acuña Sfrasani
Profesor guía

Sr. Roberto Salinas Naranjo
Integrante de Comisión

Sr. José Beyá Marshall
Integrante de Comisión

DECLARACIÓN

Este trabajo, o alguna de sus partes, no han sido presentados anteriormente en la Universidad de Valparaíso, institución universitaria chilena o extranjera u organismo de carácter estatal, para evaluación, comercialización u otros propósitos. Salvo las referencias citadas en el texto, confirmo que el contenido intelectual de este Proyecto de Título es resultado exclusivamente de mis esfuerzos personales.

La Universidad de Valparaíso reconoce expresamente la propiedad intelectual del autor sobre esta Memoria de Titulación. Sin embargo, en caso de ser sometida a evaluación para los propósitos de obtención del Título Profesional de Ingeniero Civil Oceánico, el autor renuncia a los derechos legales sobre la misma y los cede a la Universidad de Valparaíso, la que estará facultada para utilizarla con fines exclusivamente académicos.

Claudio Alejandro Meza Valle

AGRADECIMIENTOS

A ti Pamela Eileen, te agradezco el incondicional apoyo que me has dado en este largo trayecto, el amor y ánimo de cada día, y por sobre todas las cosas, por la ciega confianza que siempre has tenido en mí, aun cuando yo no he tenido las fuerzas suficientes. Te amo y sin ti nada de esto podría haber sido posible. Soy lo que soy gracias a tu persona. Infinitas gracias.

Kevin Alejandro, Isabella Agustina y Emilia Camille, mis niños y mi principal razón de ser.

A mis padres Claudio y Pamela, los amo, esto es para ustedes. Espero ser el reflejo de lo que siempre quisieron para mí. Esto representa toda la dedicación que a lo largo del tiempo pusieron en mí y en mi formación. La vida me eligió para ser el primero de muchos. Espero ser un ejemplo para los más pequeños de mi familia y puedan seguir el camino del estudio. Yo pondré todas mis ganas y energías para que así sea.

A los más locos, mis hermanos Esteban Andrés y Martín Alexis. Gracias por la vida de sonrisas y alegrías. Esto es para ustedes queridos míos.

A todos mis familiares que han estado haciendo el aguante durante toda mi vida, mis abuelos, mis tíos y tías, mis primos y primas. Nombrarlos uno por uno tomaría quizás más hojas de las que contiene este documento. Sin embargo, agradezco de manera especial a mi abuela Sonia Fuentes, a mi tío y amigo Mauricio Meza y a mi querida tía Guicela Valle, quienes han estado en mi caminar durante toda mi vida. Muchas gracias por su incondicional apoyo y cariño.

A mi amigo y profesor guía, Hugo Acuña Sfrasani, quien siempre tuvo una palabra de aliento y motivación, y quien depositó su confianza en mí para realizar esta memoria.

A mis amigos Matías Quezada y Benjamín Hernández, los cuales forman parte importante de mi carrera y en especial de este trabajo, ya que, con sus comentarios y sugerencias, cada vez fue mejorando.

A los profesores Álvaro Valdivia Olivares, Rubén Escanilla Camus y Pablo González Lever, motivadores que siempre me han apoyado hasta el día de hoy.

A mi gran amigo Felipe Concha, por los años de amistad y las innumerables jornadas de estudio.

A mis compañeros de estudio los brochas Ariel Gallardo, César Esparza, Rodrigo Leal, Janisse Ferrada por aquellas jornadas de discusiones a cerca del estudio y de la vida.

Les agradezco también a ustedes, a los que pusieron trabas para que esto saliera bien. Sin sus dificultades esto no podría haberse logrado.

A la compañía Baird & Associates S.A., por facilitarme el uso de sus datos para llevar a cabo este trabajo.

Agradecer no es una tarea fácil, ya que siempre queda algo por qué agradecer, o bien alguien a quién agradecer, por lo tanto, si usted lee esto, conoce mi historia personal hasta este punto, y sabe que de alguna forma contribuyó en este camino, agradezco que haya participado.

DEDICATORIA

*A mi mujer Pamela Eileen
y a mis hijos Kevin Alejandro,
Isabella Agustina
y Emilia Camille,
los amo, ustedes son el combustible perfecto de esta máquina que no se detiene...*

*...A mis amados Mario Meza y Santiago Valle, yo soy lo mejor de ustedes,
su partida fue temprana, sin embargo, están en cada acción, cada pensamiento, todos los días de mi
vida. Les dedico esto con el mayor amor que tengo.*

Claudio Alejandro Meza Valle

RESUMEN

En el presente trabajo se estimaron parámetros extremos de oleaje de manera simultánea a partir de análisis multivariados de valores extremos de oleaje en las zonas norte, centro y sur de Chile utilizando tres nodos de aguas profundas de la base de datos Olas Chile IV: nodo Iquique, nodo Valparaíso y nodo Chile 54S (54° latitud sur).

Previamente se realizaron agrupaciones de tormentas por medio del método de rastreo de *swells* (*Swell Tracking*, Hanson 2000) en aguas profundas, logrando identificar las características completas de los conjuntos de estados de mar que los constituyen en base a la física de sus parámetros componentes. A partir de las tormentas separadas se identificaron y extrajeron los mayores valores de altura de oleaje, además del periodo, dirección y duración asociada a la altura de oleaje máxima en cada tormenta.

Considerando los valores previamente identificados se realizó un análisis de valores extremos multivariados en las tres zonas a partir de un modelo multivariado de probabilidad conjunta (Heffernan & Tawn, 2004), extrapolando de manera simultánea pares de parámetros extremos de oleaje en aguas profundas. Con este método fue posible obtener curvas de distribución de probabilidad de periodo, dirección y duración para alturas de oleaje de diferentes periodos de retorno. Utilizando esta información fue posible definir los rangos que pueden adoptar estos parámetros con distintos valores de probabilidad.

ABSTRACT

In the present work extreme wave parameters were estimated in a simultaneous manner from a multivariate analysis of extreme wave values in the north, center and south of Chile using three deepwater wave nodes of the Olas Chile IV database: Iquique, Valparaiso and 54S (latitude 54° south) nodes.

Previously, storms grouping was performed using the swell tracking method proposed by Hanson (2000) in deep water; achieving identification of storm sets characteristics based on their physical components. From separated storms, the highest values of wave height and their associated parameters (periods, direction and duration) have been identified and extracted for each storm set.

Considering the previously identified values, a multivariate extreme value analysis in three points of Chile was performed using a multivariate joint probability model (Heffernan & Tawn, 2004), extrapolating simultaneously pairs of extreme wave parameters in deep water. With this method it was possible to obtain probability distribution curves of period, direction and duration associated to wave heights of different return periods. Using this information, it was possible to calculate the ranges that these parameters can take with different values of probability.

CONTENIDOS

1	INTRODUCCIÓN	1
2	OBJETIVOS	2
2.1	OBJETIVO GENERAL	2
2.2	OBJETIVOS ESPECÍFICOS	2
3	MARCO TEÓRICO	3
3.1	VALORES EXTREMOS UNIVARIADOS	4
3.2	VALORES EXTREMOS MULTIVARIADOS	6
3.3	APROXIMACIÓN CONDICIONAL MULTIVARIADA	7
3.3.1	ANTECEDENTES DEL MODELO	7
3.3.2	MODELO MARGINAL	8
3.3.3	MODELACIÓN ESTADÍSTICA	10
3.3.4	MODELO DE DEPENDENCIA CONDICIONAL	12
3.3.5	EXTRAPOLACIÓN DE DATOS	13
3.3.6	INFERENCIA ESTADÍSTICA	14
3.4	AVANCES SOBRE EL MODELO CONDICIONAL MULTIVARIADO	17
3.4.1	TRANSFORMACIÓN MARGINAL DE LAPLACE	17
3.5	RESUMEN DEL MÉTODO DE SIMULACIÓN MULTIVARIADA	18
4	METODOLOGÍA	19
4.1	MATERIALES	19
4.1.1	BASE DE DATOS DE OLEAJE	19
4.1.2	SOFTWARE	19
4.2	DEFINICIONES	20
4.3	METODOLOGÍA DE SIMULACIÓN	23
4.3.1	BASE DE DATOS	24
4.3.2	IDENTIFICACIÓN DE SISTEMAS DE OLEAJE	25
4.3.3	AGRUPACIÓN DE SWELLS	27
4.3.4	IDENTIFICACIÓN DE VALORES EXTREMOS	31
4.3.5	MODELACIÓN MULTIVARIADA DE PARÁMETROS EXTREMOS	35
5	RESULTADOS	39
5.1	AGRUPACIÓN DE SWELLS	39
5.2	LISTAS DE MAYORES TORMENTAS Y COMPARACIÓN CON POT	43

5.3	RESULTADOS DE LA MODELACIÓN CONJUNTA DE PARÁMETROS DE OLEAJE....	49
5.3.1	<i>DEFINICIÓN DE UMBRAL PARA SELECCIÓN DE DATOS.....</i>	49
5.3.2	<i>Cálculo de alturas de oleaje extremo</i>	51
5.3.3	<i>Cálculo de parámetros del modelo multivariado</i>	52
5.3.4	<i>Resultados de modelaciones condicionales multivariadas</i>	53
5.3.5	<i>Resultados de las modelaciones condicionales multivariadas: Zonas norte, centro y sur de Chile.....</i>	55
6	DISCUSIÓN SOBRE LA VALIDACIÓN DE LA METODOLOGÍA UTILIZADA	69
7	CONCLUSIONES Y RECOMENDACIONES.....	72
8	REFERENCIAS BIBLIOGRÁFICAS.....	74

LISTA DE FIGURAS

Figura 1.1: Rosa periodos de oleaje en aguas profundas	1
Figura 4.1: Esquema de fenómenos de oleaje considerados en el estudio	21
Figura 4.2: Rosa de puntos de periodo – Nodo Valparaíso.....	22
Figura 4.3: Diagrama de flujo. Modelación multivariada de parámetros extremos de oleaje	23
Figura 4.4: Comparación en aguas profundas nodo Valparaíso vs boya Watchkeeper	24
Figura 4.5: Sistemas de oleaje identificados – Zona norte, centro y sur	26
Figura 4.6: Ejemplo de evolución de sistemas de oleaje particionados. Estación 51028, 1 – 15 de Noviembre de 2000. a) Mediciones de boya NDBC. b) Hindcast WaveWatch III	28
Figura 4.7: Diagrama de flujo. Método de separación de tormentas	30
Figura 4.8: Esquema de método de valores sobre el umbral	31
Figura 4.9: Esquema de método de máximos anuales.....	32
Figura 4.10: Serie de tiempo de tormentas y sus parámetros	33
Figura 4.11: Identificación de parámetros – Tormenta típica.....	34
Figura 4.12: Dispersión de datos. Ejemplo de alturas máximas vs periodos <i>peak</i> de oleaje.....	35
Figura 4.13: Extrapolación condicional multivariada. a) Dispersión de datos extremos simulados. b) Densidad de probabilidad de periodos máximos en función de la altura de retorno	38
Figura 5.1: Separación se tormentas – <i>Swell</i> cercano SW, Zona central	40
Figura 5.2: Separación de tormentas – <i>Swell</i> remoto SW, Zona central.....	40
Figura 5.3: Separación de tormentas – <i>Swell</i> cercano NW E, Zona central.....	41
Figura 5.4: Separación de tormentas – <i>Swell</i> remoto NW, Zona central.....	41
Figura 5.5: Gráficos de estabilidad de parámetros – <i>Swell</i> remoto SW, Zona central	50
Figura 5.6: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> remoto SW – Zona central.....	54
Figura 5.7: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> remoto SW – Zona central	54
Figura 5.8: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> remoto SW – Zona central .	54
Figura 5.9: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> cercano SW – Zona norte.....	56
Figura 5.10: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> cercano SW – Zona norte	56
Figura 5.11: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> cercano SW – Zona norte .	56
Figura 5.12: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> remoto SW – Zona norte	57
Figura 5.13: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> remoto SW – Zona norte	57
Figura 5.14: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> remoto SW – Zona norte ...	57

Figura 5.15: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> cercano NE SE – Zona norte.....	58
Figura 5.16: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> cercano NE SE – Zona norte	58
Figura 5.17: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> cercano NE SE – Zona norte	58
Figura 5.18: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> remoto NW – Zona norte	59
Figura 5.19: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> remoto NW – Zona norte.....	59
Figura 5.20: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> remoto NW – Zona norte... ..	59
Figura 5.21: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> cercano SW – Zona central	60
Figura 5.22: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> cercano SW – Zona central.....	60
Figura 5.23: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> cercano SW – Zona central.....	60
Figura 5.24: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> remoto SW – Zona central.....	61
Figura 5.25: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> remoto SW – Zona central	61
Figura 5.26: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> remoto SW – Zona central ..	61
Figura 5.27: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> cercano NE E – Zona central	62
Figura 5.28: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> cercano NE E – Zona central	62
Figura 5.29: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> cercano NW E – Zona central	62
Figura 5.30: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> remoto NW – Zona central	63
Figura 5.31: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> remoto NW – Zona central	63
Figura 5.32: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> remoto NW – Zona central ..	63
Figura 5.33: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> cercano NE SW – Zona sur.....	64
Figura 5.34: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> cercano NE SW – Zona sur.....	64
Figura 5.35: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> cercano NE SW – Zona sur.....	64
Figura 5.36: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> SW NW – Zona sur.....	65
Figura 5.37: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> SW NW – Zona sur.....	65
Figura 5.38: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> SW NW – Zona sur	65
Figura 5.39: Resultados simulación multivariada Hm_0 vs Tp , <i>Swell</i> remoto NW – Zona sur	66
Figura 5.40: Resultados simulación multivariada Hm_0 vs Dpk , <i>Swell</i> remoto NW – Zona sur	66
Figura 5.41: Resultados simulación multivariada Hm_0 vs Duración, <i>Swell</i> remoto NW – Zona sur	66
Figura 42: Ajuste de Pareto vs Ajuste empírico sobre datos de oleaje. Nodo Valparaíso	70
Figura 43: Dispersión de datos reales y simulados - <i>Swell</i> remoto SW, Valparaíso	71

LISTA DE TABLAS

Tabla 4.1: Sistemas de oleaje identificados.....	25
Tabla 4.2: Valores del parámetro d utilizados para cada sistema de oleaje	29
Tabla 5.1: 30 mayores tormentas <i>swell</i> cercano SW - POT (izq.) y máximos por tormenta (der.)	44
Tabla 5.2: Estadística de alturas de ola identificadas - <i>swell</i> cercano SW	44
Tabla 5.3: 30 mayores tormentas <i>swell</i> remoto SW - POT (izq.) y máximos por tormenta (der.)	45
Tabla 5.4: Estadística de alturas de ola identificadas - <i>swell</i> remoto SW	45
Tabla 5.5: 30 mayores tormentas <i>swell</i> cercano NW E - POT (izq.) y máximos por tormenta (der.)...	46
Tabla 5.6: Estadística de alturas de ola identificadas - <i>swell</i> cercano NW E	46
Tabla 5.7: 30 mayores tormentas <i>swell</i> remoto NW - POT (izq.) y máximos por tormenta (der.)	47
Tabla 5.8: Estadística de alturas de ola identificadas - <i>swell</i> remoto NW	47
Tabla 5.9: Umbrales seleccionados para simulaciones.....	50
Tabla 5.10: Valores de retorno de altura de oleaje en distintas zonas	51
Tabla 5.11: Tabla de percentiles de periodos, <i>Swell</i> remoto SW	54
Tabla 5.12: Tabla de percentiles de direcciones, <i>Swell</i> remoto SW	54
Tabla 5.13: Tabla de percentiles de duraciones, <i>Swell</i> remoto SW	54
Tabla 5.14: Descripción de tablas y gráficos asociados a los resultados obtenidos	55
Tabla 5.15: Tabla de percentiles de periodos, <i>Swell</i> cercano SW	56
Tabla 5.16: Tabla de percentiles de direcciones, <i>Swell</i> cercano SW	56
Tabla 5.17: Tabla de percentiles de duraciones, <i>Swell</i> cercano SW	56
Tabla 5.18: Tabla de percentiles de periodos, <i>Swell</i> remoto SW	57
Tabla 5.19: Tabla de percentiles de direcciones, <i>Swell</i> remoto SW	57
Tabla 5.20: Tabla de percentiles de duraciones, <i>Swell</i> remoto SW	57
Tabla 5.21: Tabla de percentiles de periodos, <i>Swell</i> cercano NE E	58
Tabla 5.22: Tabla de percentiles de direcciones, <i>Swell</i> cercano NE E	58
Tabla 5.23: Tabla de percentiles de duraciones, <i>Swell</i> cercano NE E	58
Tabla 5.24: Tabla de percentiles de periodos, <i>Swell</i> remoto NW	59
Tabla 5.25: Tabla de percentiles de direcciones, <i>Swell</i> remoto NW	59
Tabla 5.26: Tabla de percentiles de direcciones, <i>Swell</i> remoto NW	59
Tabla 5.27: Tabla de percentiles de periodos, <i>Swell</i> cercano SW	60
Tabla 5.28: Tabla de percentiles de direcciones, <i>Swell</i> cercano SW	60

Tabla 5.29: Tabla de percentiles de duraciones, <i>Swell</i> cercano SW	60
Tabla 5.30: Tabla de percentiles de periodos, <i>Swell</i> remoto SW	61
Tabla 5.31: Tabla de percentiles de direcciones, <i>Swell</i> remoto SW	61
Tabla 5.32: Tabla de percentiles de duraciones, <i>Swell</i> remoto SW	61
Tabla 5.33: Tabla de percentiles de periodos, <i>Swell</i> cercano NE E	62
Tabla 5.34: Tabla de percentiles de direcciones, <i>Swell</i> cercano NE E	62
Tabla 5.35: Tabla de percentiles de duraciones, <i>Swell</i> cercano NE E	62
Tabla 5.36: Tabla de percentiles de periodos, <i>Swell</i> remoto NW	63
Tabla 5.37: Tabla de percentiles de direcciones, <i>Swell</i> remoto NW	63
Tabla 5.38: Tabla de percentiles de duraciones, <i>Swell</i> remoto NW	63
Tabla 5.39: Tabla de percentiles de periodos, <i>Swell</i> cercano NE SW	64
Tabla 5.40: Tabla de percentiles de direcciones, <i>Swell</i> cercano NE SW	64
Tabla 5.41: Tabla de percentiles de duraciones, <i>Swell</i> cercano NE SW	64
Tabla 5.42: Tabla de percentiles de periodos, <i>Swell</i> SW NW	65
Tabla 5.43: Tabla de percentiles de direcciones, <i>Swell</i> SW NW	65
Tabla 5.44: Tabla de percentiles de duraciones, <i>Swell</i> SW NW	65
Tabla 5.45: Tabla de percentiles de periodos, <i>Swell</i> remoto NW	66
Tabla 5.46: Tabla de percentiles de direcciones, <i>Swell</i> remoto NW	66
Tabla 5.47: Tabla de percentiles de duraciones, <i>Swell</i> remoto NW	66
Tabla 5.48: Resumen comparativo entre sistemas y nodos. Caso H_{tr} 100 años, ajuste superior	67

1 INTRODUCCIÓN

Para el diseño de obras en el borde costero, es importante conocer las principales características del oleaje generado durante los mayores eventos de tormentas. La mayor altura significativa de oleaje en una tormenta en conjunto con su periodo y dirección asociados, además de la duración de la tormenta, son considerados usualmente en el diseño de estructuras costeras.

La identificación de estos parámetros físicos se basa sobre la comprensión y análisis de dos puntos importantes relativos al clima marítimo de aguas profundas:

- a) La descripción del oleaje generado a partir de eventos históricos de tormenta de manera independiente de acuerdo a su origen.
- b) El análisis de eventos extremos de oleaje y el cálculo de la probabilidad de que éstos ocurran con una periodicidad previamente definida.

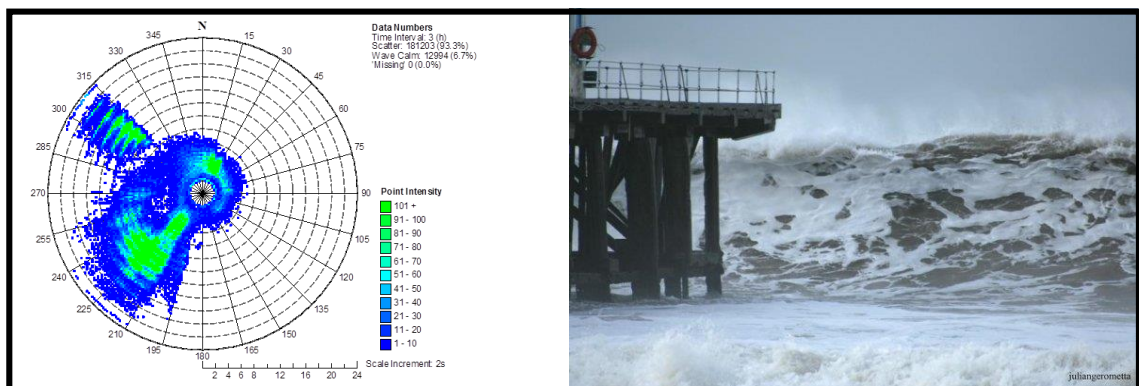


Figura 1.1: Rosa periodos de oleaje en aguas profundas
Fuente: Elaboración propia

Respecto de lo anterior, existe una uniformidad en la metodología para la definición de la altura de ola de diseño, la cual se basa en el análisis de valores extremos (EVA, por sus siglas en inglés) de una serie de datos históricos de las mayores alturas de oleaje. No obstante, no se cuenta con la misma uniformidad de criterios para la estimación del periodo y dirección asociados a cada altura de diseño.

En este trabajo se propone aplicar una metodología para la selección de tormentas en aguas profundas basada en el comportamiento físico de los parámetros que las componen, en conjunto con un modelo de probabilidad para valores extremos multivariados de oleaje en tres puntos distintos de nuestro país. Con este método es posible determinar cuantitativamente otras variables de resumen de oleaje asociadas a la altura para diferentes valores de periodo de retorno.

2 OBJETIVOS

2.1 OBJETIVO GENERAL

Aplicar un modelo de probabilidad condicional multivariada para la estimación de valores de periodo, dirección y duración asociados a alturas de oleaje de diseño en tres puntos de Chile.

2.2 OBJETIVOS ESPECÍFICOS

- Analizar un modelo de probabilidad conjunta de valores extremos de altura, periodo, dirección y duración de oleaje.
- Aplicar el modelo en tres puntos distintos de nuestro país (Norte, centro y sur).
- A partir de los resultados, en caso de que corresponda, generar recomendaciones de aplicación del método propuesto.

3 MARCO TEÓRICO

El presente trabajo parte de la base que el lector tiene conocimientos del oleaje, por lo cual no se darán detalles de los conceptos básicos de esta materia. Para aquellos lectores interesados en nociones más generales respecto al oleaje se recomienda consultar las siguientes referencias: Goda, Y., 2000, Liu, Z. y Frigaard, P., 2001, indicadas en la sección 7 de este documento (Referencias Bibliográficas).

El principal enfoque de la estadística clásica ha sido analizar el comportamiento promedio de los procesos estocásticos, desarrollando así inferencias y estimaciones que permitan conocer cómo se comportan las variables analizadas obteniendo conclusiones específicas acerca de ellas. Sin embargo, con este enfoque promedio no es posible analizar las propiedades ni el comportamiento de los valores más altos o más bajos de una variable en estudio, que en muchos casos son de mayor interés que los valores promedio. Cuando éste es el caso, se recurre a la Teoría de Valores Extremos, que en los últimos años se ha desarrollado significativamente, tanto desde el punto de vista metodológico como de las aplicaciones (Fougères, A., 2004)

La teoría de valores extremos fue desarrollada para analizar la ocurrencia de eventos poco frecuentes o atípicos dentro de un conjunto de datos generalmente asociados a valores en ambos extremos o colas de la distribución de una variable.

La evaluación del impacto y el riesgo asociado a diferentes fenómenos físicos ambientales atípicos, o bien procesos y/o situaciones relacionadas a áreas tales como la química, biología, economía, entre otros, a menudo requieren de desarrollar un análisis de valores extremos, el cual en numerosas ocasiones es inherentemente multivariado. Así es el caso de algunos procesos oceánicos, tales como el oleaje, corrientes y mareas, entre otros, los cuales poseen características físicas distintas, pero que pueden suceder de manera simultánea. El oleaje, que es el objeto de estudio de esta memoria, está definido por un conjunto de parámetros físicos cuyos valores extremos pueden ser analizados individualmente o de manera conjunta.

A continuación se presenta una descripción de la teoría de valores extremos univariados y multivariados según Fougères (2004).

3.1 VALORES EXTREMOS UNIVARIADOS

La teoría de valores extremos se centra en el análisis de datos atípicos dentro de una población, es decir, que tienen una baja probabilidad de ocurrencia dentro de un conjunto de datos, donde el problema a resolver corresponde a la modelación de las colas de las distribuciones univariadas.

La suposición fundamental sobre la cual se basan todos los métodos de análisis es la existencia de un dominio de atracción para los máximos, esto es: si X_1, X_2, \dots, X_n son observaciones independientes e idénticamente distribuidas de un fenómeno aleatorio con función de distribución F , entonces existen dos secuencias $(a_n)_n$ y $(b_n)_n$, donde $a_n > 0$ y $b_n \in \mathbb{R}$ y también una función de distribución no degenerada G , tal que:

$$\lim_{n \rightarrow \infty} P \left\{ \frac{\max X_i - b_n}{a_n} \leq x \right\} = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x) \quad (1)$$

El conjunto de funciones de distribuciones G , tal que la Ecuación (1) se mantenga están referidas como la familia de Valores Extremos Generalizada (GEV, en adelante y por su sigla en inglés). Esta Ecuación fue introducida por von Mises en el año 1954 y Jenkinson en 1955. La función de distribución G tiene la siguiente forma paramétrica:

$$G(x) = e^{\left\{ -\left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi}} \right\}} \quad (2)$$

Donde la notación mantiene el valor a si $a > 0$ y 0 si es de otra manera, y donde $\sigma > 0$ y $\mu, \xi \in \mathbb{R}$. Originalmente, esta función fue llamada distribución de Fréchet, Weibull o Gumbel, dependiendo de si su parámetro de forma $\xi > 0, \xi < 0$ o $\xi = 0$ (como caso límite). Los científicos Fisher y Tippett exhibieron estos tres tipos de distribuciones en el año 1928, y también el autor Gnedenko en 1943.

Bajo la hipótesis fundamental de la existencia de un dominio de atracción para los máximos, surgieron dos formas principales de modelar los eventos extremos:

1. En primera instancia, los modelos para bloque de máximos (*block máxima*), basados en la Ecuación 1, en los cuales la distribución asintótica de los máximos renormalizados es considerada como una aproximación para la distribución de los máximos sobre un número fijo (suficientemente grande) de observaciones.
2. Métodos de umbral basados sobre la forma asintótica de la distribución de los excesos sobre un valor dado. Específicamente, dos ramas han sido estudiadas: la

primera corresponde al método de valores sobre el umbral (*Peak Over Threshold* o POT, por sus siglas en inglés), el cual está basado sobre aproximaciones de Pareto generalizadas. Esto significa que si la función de distribución de una variable aleatoria X está en el dominio de atracción de una distribución GEV con parámetros (μ, σ, ξ) , entonces la función de distribución condicional de las excedencias tiene la siguiente propiedad:

$$P(X > u + x | X > u) \sim \left(1 + \xi \frac{x}{\tilde{\sigma}}\right)^{-\frac{1}{\xi}}, x \rightarrow \infty \quad (3)$$

Donde $\tilde{\sigma} = \sigma + \xi(u - \mu)$. Esta aproximación se debe a Balkema & Haan (1974) y Pickands (1975) y ha sido estudiada por otros autores hasta la actualidad.

La segunda forma de usar los modelos sobre un umbral corresponde a aproximar el proceso puntual asociado a observaciones mayores que u por un proceso de Poisson no homogéneo con medida de intensidad (x, ∞) dado por

$$\left(1 + \xi \frac{x - \mu}{\tilde{\sigma}}\right)^{-\frac{1}{\xi}} \quad (4)$$

Esto ha sido estudiado por los autores Pickands (1971) y Smith (1989). También han sido desarrollados métodos adaptados para cuando el interés está sobre secuencias de extremos dependientes, lo cual en la actualidad es de uso común cuando se consideran sobre todo datos medioambientales.

3.2 VALORES EXTREMOS MULTIVARIADOS

La teoría de valores extremos multivariada se enfoca en la caracterización, estimación y extrapolación de las colas conjuntas de distribuciones multidimensionales. La evaluación precisa de las probabilidades de eventos extremos son el foco de numerosas aplicaciones, desde evaluaciones de impactos ambientales (Coles & Tawn, 1994; Joe, 1994; de Haan & de Ronde, 1998, Schlather & Tawn, 2003) hasta administración de riesgo financiero (Embrechts et al., 1997; Longin, 2000; Stărică, 2000; Poon et al., 2004) y modelación de tráfico de internet (Maulik et al., 2002; Resnick & Rootzén, 2000).

Las primeras contribuciones a la teoría de valores extremos multivariados estuvieron centradas sobre el comportamiento límite de los máximos componente a componente de vectores aleatorios independientes e idénticamente distribuidos.

Toda la teoría desarrollada está basada, al igual que en el caso univariado, sobre la existencia de un dominio de atracción. De acuerdo a esto y de forma simplificada, se enuncia lo siguiente:

Sean los elementos $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$. Si los datos $X_i = (X_{i,1}, X_{i,2}, \dots, X_{i,d})$, con $i = 1, 2, \dots, n$, son vectores aleatorios independientes e idénticamente distribuidos de dimensión d con función de distribución F , se asume que existen \mathbb{R}^d secuencias (a_n) y (b_n) , donde $a_n > 0$ y $b_n \in \mathbb{R}$ para todo $j = 1, 2, \dots, d$, y una función de distribución G con márgenes no degenerados, tal que:

$$P \left\{ \frac{\max_{i=1, \dots, n} X_i - b_n}{a_n} \leq x \right\} = F^n(a_n x + b_n) \rightarrow G(x), n \rightarrow \infty \quad (5)$$

Así, la función de distribución G es denominada *Función de Distribución de Valores Extremos Multivariados*, y se dice que F está en el dominio de atracción (multivariado) de G para los máximos. Se debe notar que los márgenes univariados de G son distribuciones de valores extremos.

3.3 APROXIMACIÓN CONDICIONAL MULTIVARIADA

En la actualidad enormes esfuerzos han logrado el desarrollo de modelos estadísticos que permiten la estimación de más de una variable de manera conjunta. En este trabajo, se ha utilizado el modelo propuesto por Heffernan & Tawn (2004). Este modelo fue desarrollado y aplicado en primera instancia para conocer el comportamiento de eventos extremos de gases contaminantes en la ciudad de Leeds, Reino Unido.

De manera posterior, este modelo ha sido extendido a diversas áreas de estudio, y de manera específica al análisis de variables extremas de oleaje. Jonathan et al (2010) aplicaron esta metodología a eventos oceánicos de diferentes localidades.

La base teórica del modelo según los autores Heffernan & Tawn (2004) se presenta a continuación.

3.3.1 ANTECEDENTES DEL MODELO

Considere un vector continuo variable $X = (X_1, \dots, X_d)$ con función de distribución $F(x)$ desconocida. A partir de una muestra de n observaciones independientes e idénticamente distribuida de F se desea estimar los funcionales de la distribución de X cuando X es extrema en al menos una componente. Este método permite que cualquier funcional sea considerado y se enfoca principalmente en estimar $Pr(X \in C)$ donde C es un conjunto extremo tal que para todos los $x \in C$ al menos una componente de x es extrema. Comúnmente, ninguna observación habrá ocurrido en C . La estructura de C motiva la siguiente partición natural de C en d subconjuntos $C = \bigcup_{i=1}^d C_i$. Aquí, C_i es aquella parte de C para el cual X_i es la componente más grande de X , medida por los cuantiles de la distribución marginal de X_i ; luego

$$C_i = C \cap \left\{ x \in \mathbb{R}^d : F_{X_i}(X_i) > F_{X_j}(X_j); j = 1, \dots, d; j \neq i \right\}, \text{ Para } i = 1, \dots, d. \quad (6)$$

Se asume que los subconjuntos de C de la forma $C \cap \left\{ x \in \mathbb{R}^d : F_{X_i}(X_i) = F_{X_j}(X_j) \text{ para } i \neq j \right\}$ puede ser ignorados; estos son conjuntos nulos en cuyos subconjuntos no existen componentes singulares en la estructura de dependencia de X .

Con la partición de C definida de esta forma, C es un conjunto extremo si todos los valores x_j en un subconjunto no vacío C_i caen en la cola superior de F_{X_i} , esto es, si $v_{X_i} = \inf_{x \in C_i} (x_i)$, entonces $F_{X_i}(v_{X_i})$ es cercano a 1 para $i = 1, \dots, d$. De esta forma

$$Pr(X \in C) = \sum_{i=1}^d Pr(X \in C) = \sum_{i=1}^d Pr(X \in C_i | X_i > v_{X_i}) Pr(X_i > v_{X_i}) \quad (7)$$

Se debe considerar la estimación de $Pr(X \in C)$ usando la descomposición (7). Se necesita estimar la probabilidad $Pr(X_i > v_{X_i})$ y $Pr(X \in C_i | X_i > v_{X_i})$. Para la primera estimación se requiere de un modelo marginal de valores extremos y la última necesita de un modelo de valores extremos para la estructura de dependencia.

3.3.2 MODELO MARGINAL

Actualmente, los métodos para estimar extremos marginales son relativamente estándar. La teoría de valores extremos univariados provee una justificación asintótica para que la distribución de Pareto generalizada sea un modelo apropiado para estimar la distribución de excesos sobre un nivel alto de umbral (Pickands, 1975). Así, la cola marginal de X_i con $i = 1, \dots, d$, se puede modelar por:

$$Pr(X > x + u_{X_i} | X_i > u_{X_i}) = (1 + \xi_i x / \beta_i)_+^{-1/\xi_i}, \quad \text{donde } x > 0 \quad (8)$$

En esta expresión u_{X_i} es un nivel de umbral alto para la variable X_i , β_i y ξ_i son los parámetros de escala y de forma respectivamente, con $\beta_i > 0$ y $s_+ = \max(s, 0)$ para cualquier $s \in \mathbb{R}$. Se requiere de un modelo para la distribución marginal completa F_{X_i} de X_i para cada $i = 1, \dots, d$, ya que para estimar $Pr(X \in C_i | X_i > v_{X_i})$ se necesita describir todos los valores X_j que puedan ocurrir con cualquier X_i grande. Se adopta el modelo semiparamétrico \hat{F}_{X_i} para F_{X_i} , (Coles & Tawn, 1991, 1994), esto es:

$$\hat{F}_{X_i}(x) = \begin{cases} 1 - \{1 - \hat{F}_{X_i}(u_{X_i})\} \{(1 + \xi_i(x - u_{X_i})/\beta_i)_+^{-1/\xi_i}\}, & x > u_{X_i} \\ \hat{F}_{X_i}(x), & x \leq u_{X_i} \end{cases} \quad (9)$$

Donde \hat{F}_{X_i} es la distribución empírica de los valores X_i . El punto final superior de la distribución, esta denotado por x^{F_i} , el cual es ∞ si $\xi_i \geq 0$ y $u_{X_i} - \beta_i/\xi_i$ si $\xi_i < 0$. El modelo representado por la Ecuación (9), provee la base para estimar el término $Pr(X_i > v_{X_i})$ de la Ecuación (7).

Tanto la estructura marginal como la estructura de dependencia de X son necesarias para determinar $Pr(X \in C_i | X_i > v_{X_i})$. En este punto del estudio, estas dos contribuciones están separadas y se realiza un enfoque sobre la modelación de dependencia trabajando con los marginales asumiéndolos como conocidos. Todas las distribuciones marginales univariadas

deben ser transformadas a distribuciones de la forma Gumbel estándar por medio de la transformada integral de probabilidad, la cual para el modelo marginal dado por la Ecuación 9 tiene la siguiente forma:

$$Y_i = -\log[-\log\{\hat{F}_{X_i}(X_i)\}] \quad \text{para } i = 1, \dots, d$$

$$Y_i = t_i(X_i; \psi_i, \hat{F}_{X_i}), \tag{10}$$

$$Y_i = t_i(X_i)$$

Donde $\psi_i = (\beta_i, \xi_i)$ son los parámetros marginales. Esta transformación entrega que $Pr(Y_i \leq y) = \exp\{-\exp(-y)\}$ para cada i , de manera que $Pr(Y_i > y) \sim \exp(-y)$ a medida que $y \rightarrow \infty$, e Y_i tiene una cola superior exponencial. Para clarificar cual de las variables marginales se está utilizando, se define X para denotar la variable con su distribución marginal original e Y para la variable con marginales Gumbel.

3.3.2.1 Modelo de dependencia extrema

En esta sección se describe el enfoque realizado sobre la modelación de dependencia extrema de variables con distribuciones marginales Gumbel. Modelar la dependencia para valores extremos es más complejo que la modelación univariada a pesar de que existen diversos métodos propuestos que siguen evolucionando. Cuando el interés se encuentra sobre los extremos superiores de cada componente de Y , la estructura de dependencia cae en dos categorías: asintóticamente dependiente o asintóticamente independiente de la variable Y_i cuando el siguiente límite:

$$\lim_{y \rightarrow \infty} \{Pr(Y_{-i} > y | Y_i > y)\}$$

no es cero y es cero respectivamente. Aquí Y_{-i} denota el vector Y excluyendo la componente Y_i e y un vector de valores de la variable y . Todos los métodos existentes para valores extremos multivariados son apropiados para estimar $Pr(X \in C)$ bajo dependencia asintótica de los Y asociados, o por variables asintóticamente independientes tal que todos los $x \in C$ sean grandes en todas sus componentes.

Aquí se presenta una aproximación para valores extremos multivariados que constituye un cambio de dirección respecto de otros métodos de valores extremos. La estrategia de modelación está basada en una suposición sobre la forma asintótica de la distribución condicional de la variable dado que tiene una componente extrema, esto es, la distribución de $Y_{-i} | Y_i = y_i$ a medida que y_i va adoptando valores mayores. Esta aproximación condicional resulta en una extensión natural del modelo de distribución condicional univariada

generalizada de Pareto representado por la Ecuación (8) al caso multivariado $Pr(X \in C_i | X_i > v_{X_i})$ expresado como:

$$Pr(X \in C_i | X_i > v_{X_i}) = \int_{v_{X_i}}^{x^{F_i}} Pr(X \in C_i | X_i = x) d\hat{F}_{X_i}(x) / \{1 - \hat{F}_{X_i}(v_{X_i})\} \quad (11)$$

Donde el integrando es evaluado usando la distribución de $Y_{-i} | Y_i = y_i$ luego de la transformación marginal. Cuando $v_{X_i} > u_{X_i}$ la derivada de $\hat{F}_{X_i}(x) / \{1 - \hat{F}_{X_i}(v_{X_i})\}$ es la función de densidad generalizada de Pareto con parámetros de escala y forma iguales a $\beta_i + \xi_i(v_{X_i} - u_{X_i})$ y ξ_i respectivamente.

El modelo condicional aplica si las variables son asintóticamente dependientes o asintóticamente independientes; este puede ser usado para estimar $Pr(X \in C)$ para cualquier conjunto extremo C , y es aplicable a cualquier número de dimensiones. El modelo utilizado para la distribución condicional está motivado por una suposición distribucional asintótica y está respaldado por diversos ejemplos teóricos. Además, este es de tipo semiparamétrico y es aplicada una regresión paramétrica para estimar los parámetros de escala y de forma de los marginales de la distribución condicional conjunta, mientras que para estimar la estructura residual multivariada son empleados métodos no paramétricos. A medida que la aproximación utilizada carece de una caracterización asintótica completa de la estructura probabilística, tales como los que sustentan los métodos de valor extremo existentes, la modelación expone que esta metodología presenta ventajas matemáticas duras y prácticas en comparación con otros métodos de valores extremos multivariados existentes.

3.3.3 MODELACIÓN ESTADÍSTICA

A continuación se presentan los resultados teóricos que motivan la selección del método estadístico propuesto, las suposiciones relacionadas a la forma asintótica de la distribución condicional y sus consecuencias. Aquí también se identifican las condiciones que deben ser satisfechas por las funciones normalizantes fundamentando esta suposición por la representación límite que debe mantenerse. Se presentan también algunos ejemplos teóricos que sugieren que la suposición asintótica es apropiada para un amplio rango de distribuciones, y que la clase de funciones normalizantes es reducida.

3.3.3.1 Representación límite y sus propiedades

Se debe considerar la estructura asintótica de las distribuciones condicionales de una variable aleatoria $Y = (Y_1, \dots, Y_d)$ de dimensiones d con distribuciones marginales Gumbel. Para cada $i = 1, \dots, d$, se examina la distribución condicional $Pr(Y_{-i} \leq y_{-i} | Y_i = y_i)$, donde aquí y en adelante, el algebra vectorial se aplica componente a componente. Para examinar el

comportamiento límite de estas distribuciones a medida que $y_i \rightarrow \infty$ se requiere que la distribución límite sea no degenerada en ninguno de sus márgenes, de manera que se debe controlar el crecimiento de y_{-i} de acuerdo a la dependencia de Y_{-i} sobre Y_i .

De manera específica, se asume que para un valor i dado, existe un vector de funciones normalizantes $a_{|i}(y_i)$ y $b_{|i}(y_i)$, ambos con valores en el dominio $\mathbb{R} \rightarrow \mathbb{R}^{(d-1)}$, las cuales pueden ser seleccionadas de manera tal que, para todos los $z_{|i}$ y para cualquier secuencia de valores y_i tal que $y_i \rightarrow \infty$ se mantenga que:

$$\lim_{y_i \rightarrow \infty} P\{Y_{-i} \leq a_{|i}(y_i) + b_{|i}(y_i)Z_{-i} | Y_i = y_i\} = G_{|i}(Z_{-i}) \quad (12)$$

donde todos los marginales de la distribución límite $G_{|i}$ son no degenerados. Una expresión alternativa de esta suposición, la cual tiene una interpretación estadística más simple, es que las variables estandarizadas

$$Z_{-i} = \frac{Y_{-i} - a_{|i}(y_i)}{b_{|i}(y_i)} \quad (13)$$

tienen la propiedad que:

$$\lim_{y_i \rightarrow \infty} \{Pr(Z_{|i} \leq z_{|i} | Y_i = y_i)\} = G_{|i}(z_{|i}) \quad (14)$$

Donde la distribución límite $G_{|i}$ tiene distribuciones marginales no degeneradas.

Bajo la Ecuación 12 (ó equivalentemente la Ecuación 14, se tiene que, de manera condicional sobre $Y_i > u_i$, a medida que $u_i \rightarrow \infty$, las variables $Y_i - u_i$ y $Z_{|i}$ son independientes en el límite con las distribuciones marginales exponenciales límite y $G_{|i}(z_{|i})$ respectivamente. Para ver que este resultado se mantiene, sea $y_i = u_i + y$ con $y > 0$, entonces:

$$Pr(Z_{|i} \leq z_{|i}, Y_i - u_i = y | Y_i > u_i) \rightarrow G_{|i}(z_{|i})exp(-y) \quad (15)$$

La convergencia final de esta derivación está implicada por la cola exponencial de las variables Gumbel y la propiedad que el límite condicional (Ecuación 13) se mantiene independientemente de cómo los valores $y_i \rightarrow \infty$.

3.3.4 MODELO DE DEPENDENCIA CONDICIONAL

Utilizando la misma aproximación que en otros métodos de valores extremos univariados y multivariados, se realiza una suposición asintótica que se mantiene bajo condiciones débiles con el fin de proveer que la variable limitante es suficientemente extrema. Para ello se usa la formulación de la distribución condicional límite (Ecuación 12) y su propiedad de independencia limitante (Ecuación 15), para capturar el comportamiento de la variable Y_{-i} que ocurre con grandes Y_i . Se asume que para cada $i = 1, \dots, d$ existe un umbral alto u_{Y_i} para el cual se modela:

$$\begin{aligned} Pr\{Y_{-i} < a_{|i}(y_i) + b_{|i}(y_i)z_{|i} | Y_i = y_i\} &= Pr(Z_{|i} < z_{|i} | Y_i = y_i) \\ &= G_{|i}(z_{|i}), \text{ para todos los } y_i > u_{Y_i} \end{aligned} \quad (16)$$

donde $Z_{|i}$ es el residual estandarizado definido por la Ecuación 13, con función de distribución $G_{|i}$, y $Z_{|i}$ es independiente de Y_i para $Y_i > u_{Y_i}$. El comportamiento de dependencia extrema es entonces caracterizado por las funciones de posición y escala $a_{|i}(y_i)$ y $b_{|i}(y_i)$ y la función de distribución $G_{|i}$.

Primero se debe considerar la especificación de los modelos condicionales individuales, esto es $a_{|i}(y_i)$, $b_{|i}(y_i)$ y $G_{|i}$ para un i dado. Se adopta el siguiente modelo paramétrico:

$$\begin{aligned} a_{|i}(y) &= a_{|i} y + I_{\{a_{|i}=0, b_{|i}<0\}}\{c_{|i} - d_{|i} \log(y)\} \\ b_{|i}(y) &= y^{b_{|i}} \end{aligned} \quad (17)$$

Ecuación en cuyo lado derecho $a_{|i}$, $b_{|i}$, $c_{|i}$, $d_{|i}$ son vectores constantes e I es una función indicadora. Los vectores de constantes tienen componentes tales que $0 \leq a_{j|i} \leq 1$, $-\infty < b_{j|i} < 1$, $-\infty < c_{j|i} < \infty$, $0 \leq d_{j|i} \leq 1$ para todos los $j \neq i$. La familia paramétrica (Ecuación 17) tiene formulaciones estructurales distintas para $a_{|i}(y)$ para pares asociados tanto positiva como negativamente, obviando la asimetría de la distribución marginal Gumbel, para la cual la cola superior es más pesada estadísticamente que su cola inferior.

Se adopta el modelo paramétrico (Ecuación 17) ya que representa una familia paramétrica única de funciones normalizantes apropiadas para una amplio rango de ejemplos teóricos. Los parámetros $a_{|i}(y)$ y $b_{|i}(y)$ se denotan por $\theta_{|i} = (a_{|i}, b_{|i}, c_{|i}, d_{|i})$ y se adopta la convención que $c_{j|i} = d_{j|i} = 0$ a menos que $a_{j|i} = 0$ y $b_{j|i} < 0$.

En resumen, para $i = 1, \dots, d$ el modelo de dependencia es un modelo de regresión semiparamétrico multivariado de la forma:

$$Y_{-i} = a_{|i}(y_i) + b_{|i}(y_i)Z_{|i} \quad (18)$$

para $Y_i = y_i > u_{Y_i}$, donde $a_{|i}(y_i)$ y $b_{|i}(y_i)$ están dados por el modelo paramétrico (Ecuación 17), y la distribución de los residuales estandarizados es modelada de manera no paramétrica. Los parámetros del modelo global (θ) son $\theta = (\theta_{|1}, \dots, \theta_{|d})$. Cada modelo de regresión se aplica solo sobre el umbral u_{Y_i} para el cual la estructura de dependencia se aprecia bien descrita por el modelo (Ecuación 18).

La estructura de dependencia implicada por el modelo (Ecuación 18) se puede categorizar en cuatro clases que identifican el comportamiento de los cuantiles de la distribución de $Y_i|Y_j = y_j$ a medida que $y_i \rightarrow \infty$. Si los cuantiles de la distribución condicional crecen a la misma tasa que y_i , esto es, $a_{j|i} = 1$ y $b_{j|i} = 0$, entonces las variables (Y_i, Y_j) son asintóticamente dependientes, de otra forma ellas son asintóticamente independientes.

3.3.5 EXTRAPOLACIÓN DE DATOS

Se generan muestras aleatorias de distribuciones condicionales de $X|X_i > v_{X_i}$ para cada i , usando los modelos condicionales estimados. Estas muestras son usadas para obtener aproximaciones de Monte Carlo de los funcionales de las colas conjuntas de la distribución de X . Dado que se utiliza el modelo estimado, los parámetros son reemplazados por los valores de θ estimados ($\hat{\theta}$ y $\hat{\psi}$) los cuales son obtenidos al usar los métodos descritos en los puntos anteriores. Se emplea el siguiente algoritmo de muestreo:

Paso 1: Simular Y_i desde una distribución Gumbel condicional sobre su excedente $\tau_i(v_{X_i})$.

Paso 2: Muestrear $Z_{|i}$ desde $\hat{G}_{|i}$ independiente de Y_i .

Paso 3: Obtener $Y_{-i} = \hat{a}_{|i}(y_i) + \hat{b}_{|i}(y_i)Z_{-i}$

Paso 4: Transformar $Y = (Y_{-i}, Y_i)$ a la escala original usando la transformación inversa de probabilidad (Ecuación 10)

Paso 5: El vector transformado resultante X constituye un valor de la distribución condicional de $X|X_i > v_{X_i}$.

3.3.6 INFERENCIA ESTADÍSTICA

El modelo comprende un modelo de distribución marginal (Ecuación 9) y otro modelo de dependencia (Ecuación 18) ambos semiparamétricos, los cuales consisten de componentes que son especificados paramétricamente y componentes para los cuales ningún modelo paramétrico es adecuado. La estrategia de inferencia esta conducida por tres puntos importantes:

1. Carencia de distribuciones conjuntas especificadas de manera paramétrica para cada distribución condicional.
2. Ausencia de de restricciones prácticas para imponer auto-consistencia entre las diferentes distribuciones condicionales.
3. Necesidad de simplicidad.

Esto conduce a usar un algoritmo para estimación puntual, el cual pueda hacer suposiciones simplificadas y un algoritmo de *bootstrap* semiparamétrico para evaluar la incertidumbre no basada en estos supuestos.

La inferencia para las estructuras marginales y de dependencia es llevada a cabo por etapas estimando primero los parámetros marginales (ψ) y luego estimando los parámetros de dependencia (θ) asumiendo conocidos los parámetros marginales.

3.3.6.1 Estimación marginal

Se estiman las distribuciones marginales univariadas d de manera conjunta, ignorando la dependencia entre componentes. Específicamente, se asume independencia entre componentes de la variable construyendo la función log-verosimilitud (*log-likelihood*).

$$\log\{L(\psi)\} = \sum_{i=1}^d \sum_{k=1}^{n_{u_{X_i}}} \log\{\hat{f}_{X_i}(x_{i|i,k})\} \quad (19)$$

donde \hat{f}_{X_i} es la densidad asociada con la Ecuación 9, $n_{u_{X_i}}$ es el número de observaciones con i –ésima componente que excede el umbral marginal u_{X_i} y la j –ésima componente de la k –ésima observación está denotada por $x_{i|i,k}: j = 1, \dots, d; k = 1, \dots, n_{u_{X_i}}$. Si no existen enlaces entre los parámetros de las diversas componentes, entonces maximizar la Ecuación 20 es equivalente a ajustar la distribución de Pareto generalizada a los valores que exceden el umbral marginal de manera separada para cada margen. Cuando hay restricciones entre

los parámetros marginales, maximizar de manera conjunta la función log-verosimilitud (Ecuación 19) permite obtener eficiencia inferencial.

3.3.6.2 Condicional simple

Para cada i , se desea estimar $\theta_{|i}$ bajo suposiciones mínimas acerca de $G_{|i}$. Si se asume que $Z_{|i}$ tiene dos momentos marginales finitos, entonces $\theta_{|i}$ determina los promedios y varianzas marginales de la variable condicional $Y_{-i}|Y_i = y_i$ cuando $y_i > u_{y_i}$. Específicamente, si los $Z_{|i}$ tienen promedios y desviaciones estándar marginales denotados por vectores $\mu_{|i}$ y $\sigma_{|i}$ respectivamente, entonces las variables aleatorias $Y_{-i}|Y_i = y_i$ para $y_i > u_{y_i}$ tienen vectores promedio y desviación estándar dados por las expresiones:

$$\mu_{|i}(y) = a_{|i}(y) + \mu_{|i} b_{|i}(y)$$

$$\sigma_{|i}(y) = \sigma_{|i} b_{|i}(y)$$

las cuales son funciones de $y, \theta_{|i}$ y de las constantes $\lambda_{|i} = (\mu_{|i}, \sigma_{|i})$. Así $(\theta_{|i}, \lambda_{|i})$ son los parámetros de un modelo de regresión multivariado con varianza no constante y error de distribución no especificado. Se aprovecha la consistencia de las estimaciones de máxima verosimilitud de $\theta_{|i}$ obtenida al usar un modelo paramétrico para $G_{|i}$, la cual es susceptible de ser mal interpretada. De manera específica, se maximiza la función objetivo asociada sobre el espacio paramétrico para producir un estimador puntual consistente y válido para $\theta_{|i}$.

Así, la función objetivo utilizada para la estimación puntual de $\theta_{|i}$ y $\lambda_{|i}$ está dada por la expresión:

$$Q_{|i}(\theta_{|i}) = - \sum_{j \neq i} \sum_{k=1}^{n_{uy_i}} \left[\log\{\sigma_{j|i}(y_{i|i,k})\} + \frac{1}{2} \left\{ \frac{y_{j|i,k} - \mu_{j|i}(y_{i|i,k})}{\sigma_{j|i}(y_{i|i,k})} \right\}^2 \right] \quad (20)$$

A.3.6.3 Condicional total

Se considera ahora una estimación conjunta de los parámetros θ del modelo condicional. Al igual que en el punto anterior, se asume falsamente que existe independencia entre las diferentes distribuciones condicionales para dar la función objetivo:

$$Q(\theta, \lambda) = \sum_{i=1}^d Q_{|i}(\theta_{|i}, \lambda_{|i}) \quad (21)$$

donde $Q_{|i}(\theta_{|i}, \lambda_{|i})$ es como en la Ecuación 20 y $\lambda = (\lambda_{|1}, \dots, \lambda_{|d})$. Para errores de distribución Gaussianos se puede ver que la función objetivo (Ecuación 21) es una aproximación a la pseudoverosimilitud, introducida por Besag (1975) como aproximación a la función conjunta de verosimilitud.

3.4 AVANCES SOBRE EL MODELO CONDICIONAL MULTIVARIADO

En primera instancia, el modelo original de Heffernan & Tawn (2004) considera realizar una transformación de los marginales a una escala Gumbel utilizando la Transformada Integral de Probabilidad, con el fin de que la distribución de los marginales tenga una cola superior con comportamiento exponencial. No obstante, avances específicos sobre este modelo han sido realizados (Keef, Papastathopoulos & Tawn, 2013) con el fin de sobrellevar complicaciones que han sido experimentadas en cuanto a la modelación de variables negativamente asociadas se refiere.

A pesar de las experiencias positivas del modelo, existen complicaciones con el modelamiento de las variables con algunas componentes asociadas positiva y negativamente. Esto es debido a que el modelo de regresión semiparamétrico toma formas funcionales diferentes en estos dos casos mencionados. Se ha identificado que este problema ocurre debido a la elección de la distribución Gumbel para la distribución marginal.

Los avances realizados proponen transformar las variables marginales a distribuciones de Laplace como alternativa a la distribución de Gumbel. La distribución de Laplace tiene ambas colas exponenciales y posee simetría, lo cual permite capturar la cola exponencial superior (requerida por Gumbel) para modelar la dependencia positiva y también lo permite para las variables asociadas negativamente a ser incorporadas en el modelo, dada su simetría. A continuación se presenta la transformación a marginales de Laplace.

3.4.1 TRANSFORMACIÓN MARGINAL DE LAPLACE

En lugar de realizar la transformación (Ecuación 10) para producir marginales Gumbel, se realiza lo siguiente:

$$Y_{|i} = \begin{cases} \log\{2F_i(X_i)\}, & \text{para } X_i < F_i^{-1}(0.5) \\ -\log\{2[1 - F_i(X_i)]\}, & \text{para } X_i \geq F_i^{-1}(0.5) \end{cases} \quad (22)$$

para $i = 1, \dots, d$. Entonces $Y = (Y_1, \dots, Y_d)$ tiene distribuciones marginales de Laplace con la siguiente propiedad:

$$Pr(Y_i < y) = \begin{cases} \exp(y)/2, & \text{si } y < 0 \\ 1 - \frac{\exp(-y)}{2}, & \text{si } y \geq 0 \end{cases} \quad (23)$$

Para todos los $i = 1, \dots, d$. Así, las colas inferior y superior de Y_i son exponencialmente distribuidas de manera exacta y entonces para cualquier $u > 0$ las distribuciones de $Y_i - u | Y_i > u$ y $-(Y_i + u) | Y_i < -u$ son exponenciales con promedio igual a 1.

3.5 RESUMEN DEL MÉTODO DE SIMULACIÓN MULTIVARIADA

En esta subsección se presenta un resumen de la metodología de simulación multivariada de valores extremos. Los pasos enunciados de manera simplificada se exponen a continuación.

1. Generación de series bi-variadas de datos a partir de un método de selección de datos.
2. Realizar un ajuste marginal de probabilidad sobre cada variable considerando la distribución de Pareto propuesta por los autores, para valores sobre un umbral.
3. Transformar los marginales de Pareto a marginales tipo Laplace para cada variable utilizando la transformación expuesta en la Ecuación 23.
4. Ajustar un umbral condicional según la variable condicionante.
5. Considerando los datos del punto 4, utilizarlos en el modelo de regresión multivariada (Ecuación 18).
6. Una vez obtenidos los valores simulados, regresarlos al espacio de probabilidad de Pareto y a partir de esto último a su espacio real.
7. Estos datos corresponden a los datos extrapolados de manera simultánea.

4 METODOLOGÍA

En el presente capítulo se describen los materiales considerados y los procesos empleados para llevar a cabo el estudio propuesto.

4.1 MATERIALES

4.1.1 BASE DE DATOS DE OLEAJE

El análisis fue realizado en tres puntos distintos de Chile. De norte a sur, se utilizaron los nodos de espectros bidimensionales de aguas profundas de Iquique, Valparaíso y Chile 54S de la base de datos Olas Chile IV desarrollada por Baird. La elección de estos nodos tiene como finalidad obtener resultados representativos a lo largo de la costa de nuestro país. En la sección 4.2.1 se describe con más detalle la base de datos utilizada.

4.1.2 SOFTWARE

- MATLAB (MATrix LABoratory): Herramienta de *software* matemático que ofrece un entorno de desarrollo integrado (IDE por sus siglas en inglés) con un lenguaje de programación propio. Entre sus prestaciones básicas se destacan: la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos *hardware*.
- R: Lenguaje de programación y ambiente de *software* para cálculo estadístico y gráficos, soportados por la Fundación R para Cálculo Estadístico. Este lenguaje es ampliamente utilizado por estadísticos para el desarrollo de *software* y análisis de datos.
- XWAVE: *Software* desarrollado por Baird & Associates que tiene la capacidad de realizar diversos tipos de análisis estadísticos sobre las bases de datos de oleaje utilizadas.

4.2 DEFINICIONES

A continuación se definen algunos conceptos relativos a oleaje de manera de contextualizar el lenguaje utilizado en este capítulo y en la presentación de los resultados.

- **Estado de mar:** Representación de un registro de oleaje de 3 horas de extensión en un punto específico del océano mediante un espectro bidimensional de frecuencia y dirección.
- **Sistema de oleaje:** Corresponde al grupo de olas generado en una zona geográfica particular, en los cuales se presentan características similares en sus parámetros componentes.
- **Partición:** División o componente de un espectro que representa a un sistema de oleaje de características específicas.
- **Tormenta:** Grupos de particiones sucesivas en el tiempo, que presentan características físicas similares en sus parámetros componentes, con un inicio y final bien definidos.

Así mismo, en la Figura 4.1 se presenta un esquema de las definiciones mencionadas, en la cual se hace distinción (de manera gráfica, diferenciada por colores) entre cada componente de oleaje considerada en el estudio.

Por ejemplo, el estado de mar abarca varios sistemas de oleaje en un instante de tiempo. Cada uno de ellos es una partición. También se muestra un sistema de oleaje en que todas sus componentes provienen del E - SE y en un rango de periodos entre 4 y 7 segundos. También se expone una tormenta donde se aprecia una clara secuencia de particiones con características similares que evolucionan en el tiempo.

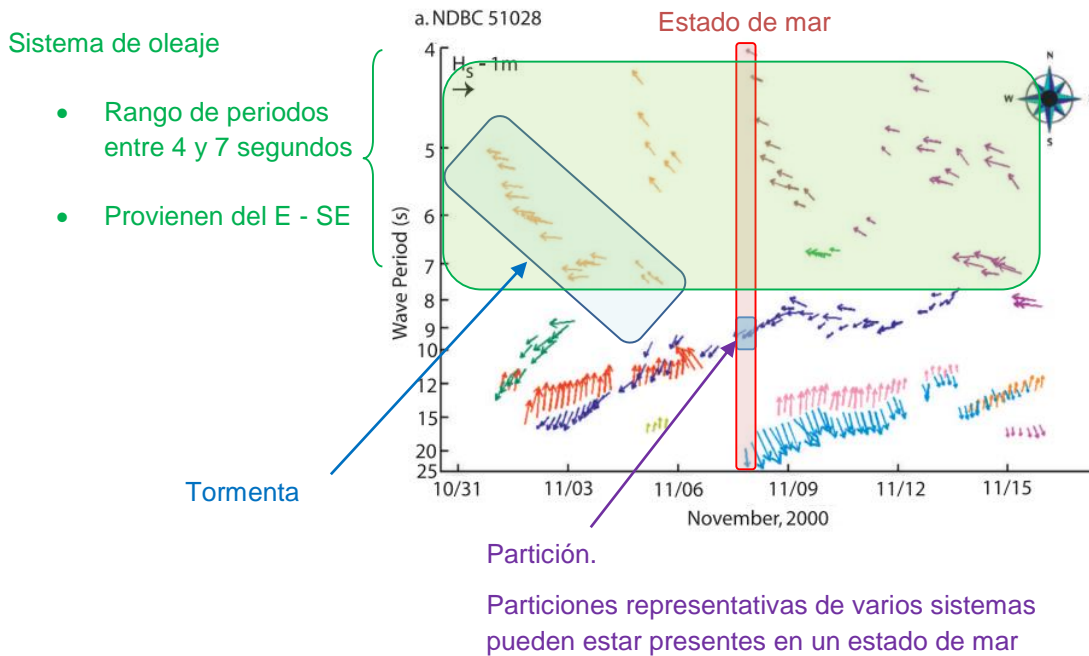


Figura 4.1: Esquema de fenómenos de oleaje considerados en el estudio

Fuente: Elaboración propia

Los vectores representan las direcciones incidentes de cada estado de mar en escala náutica. El periodo del oleaje está representado por el inicio de cada vector. Por su parte, la altura de oleaje se representa por la longitud de cada vector, cuya escala de referencia se encuentra en la esquina superior izquierda de la figura. Los resultados exponen una separación natural de los *swells* presentes diferenciados por los colores de los vectores, en los cuales es posible notar dónde ocurre el cambio de una tormenta a otra.

Respecto de las definiciones previamente enunciadas se muestra que el recuadro verde encierra todas las componentes que representan a un sistema de oleaje. Éstas presentan direcciones y periodos de similares características, pero que no son continuas en el tiempo. El recuadro celeste por su parte expone un evento de tormenta, en el cual los parámetros de altura, periodo y dirección son de similar naturaleza y son continuos en un intervalo de tiempo. Así mismo, el recuadro vertical rojo presenta varios sistemas presentes en un estado de mar de 3 horas. Finalmente, el cuadro azul muestra de manera gráfica el concepto de partición.

Otra manera de graficar los conceptos aquí definidos es con la rosa de puntos de periodo expuesta en la Figura 4.2. En ella, cada punto representa cada una de las flechas expuestas en el gráfico de la Figura 4.1, mostrando el periodo y la dirección *peak* asociados a cada partición o estado de mar de la base de datos espectral. Los grupos que se forman en la rosa representan los distintos sistemas de oleaje presentes en el lugar. Esta es otra forma de visualizar los conceptos previamente descritos sin considerar la evolución en el tiempo, pero sí todos los estados de mar o particiones de la base de datos.

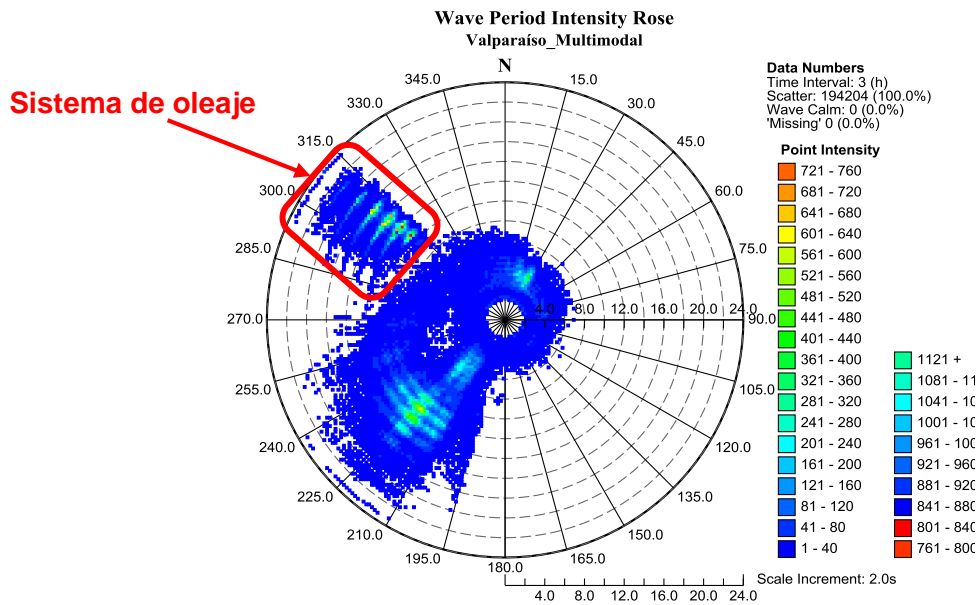


Figura 4.2: Rosa de puntos de periodo – Nodo Valparaíso

Fuente: Partición espectral multimodal aplicada a la determinación de olas de diseño aplicación en Valparaíso, Chile. Rojas, A., 2015

4.3 METODOLOGÍA DE SIMULACIÓN

En la Figura 4.3 se expone un diagrama de flujo indicando los datos, metodologías y procesos existentes para llevar a cabo un análisis de valores extremo de oleaje según la literatura. El trabajo propuesto aborda el proceso de simulación multivariada de valores extremos de oleaje de acuerdo a una serie de etapas, las cuales se destacan en el diagrama de flujo y se describen en las siguientes secciones del documento. Los cuadros rojos en el diagrama de flujo muestran las etapas del proceso, mientras que los cuadros en color rosado se nombran los distintos métodos que se pueden usar en cada etapa, marcándose con un *tick* aquellos métodos que finalmente fueron utilizados.

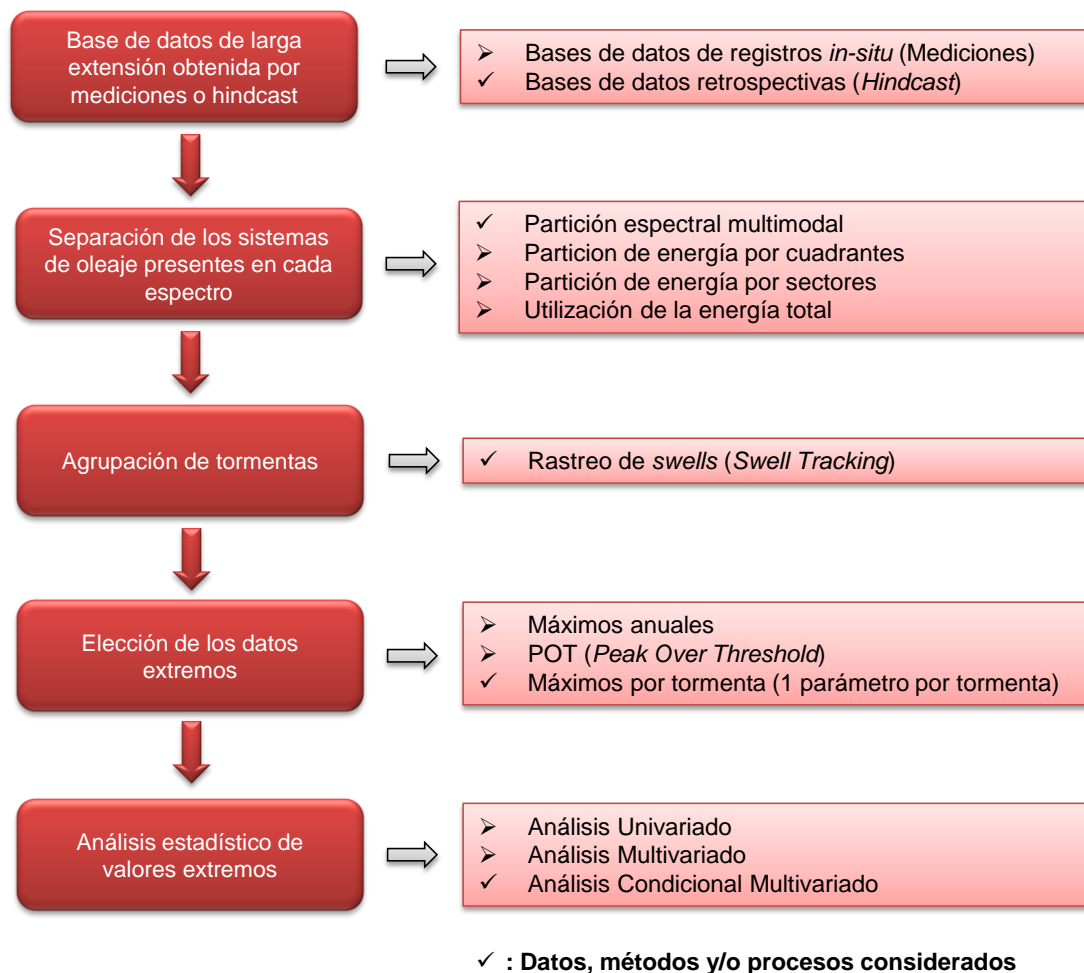


Figura 4.3: Diagrama de flujo. Modelación multivariada de parámetros extremos de oleaje

Fuente: Elaboración propia

4.3.1 BASE DE DATOS

Para el desarrollo de la memoria se ha utilizado la base de datos de aguas profundas Olas Chile IV, desarrollada por Baird. Los nodos utilizados contienen 30 años de oleaje espectral bidimensional en aguas profundas situados en 3 puntos de Chile con el fin de poder caracterizar la metodología propuesta a lo largo del país. Los nodos corresponden al nodo Iquique, nodo Valparaíso y nodo Chile 54S, de manera de utilizar oleaje representativo del norte, centro y sur de Chile, respectivamente.

El uso de este *hindcast* de larga data favorece el cálculo, ya que cuenta con una información histórica y robusta de las características del oleaje que incide sobre las costas de Chile.

La base de datos Olas Chile IV ha sido validada tanto en aguas profundas como en aguas someras con información de boyas y equipos de medición ADCP. Los resultados obtenidos a partir de estas validaciones han sido satisfactorios, ya que se establece que el *hindcast* representa de manera adecuada el oleaje registrado en ambos casos. La Figura 4.4 presenta una comparación de altura, periodo y dirección del nodo Valparaíso (línea verde) contra datos medidos de la boya de aguas profundas *Watchkeeper* (línea negra) para los años 2006 a 2007 en la Bahía de Valparaíso.

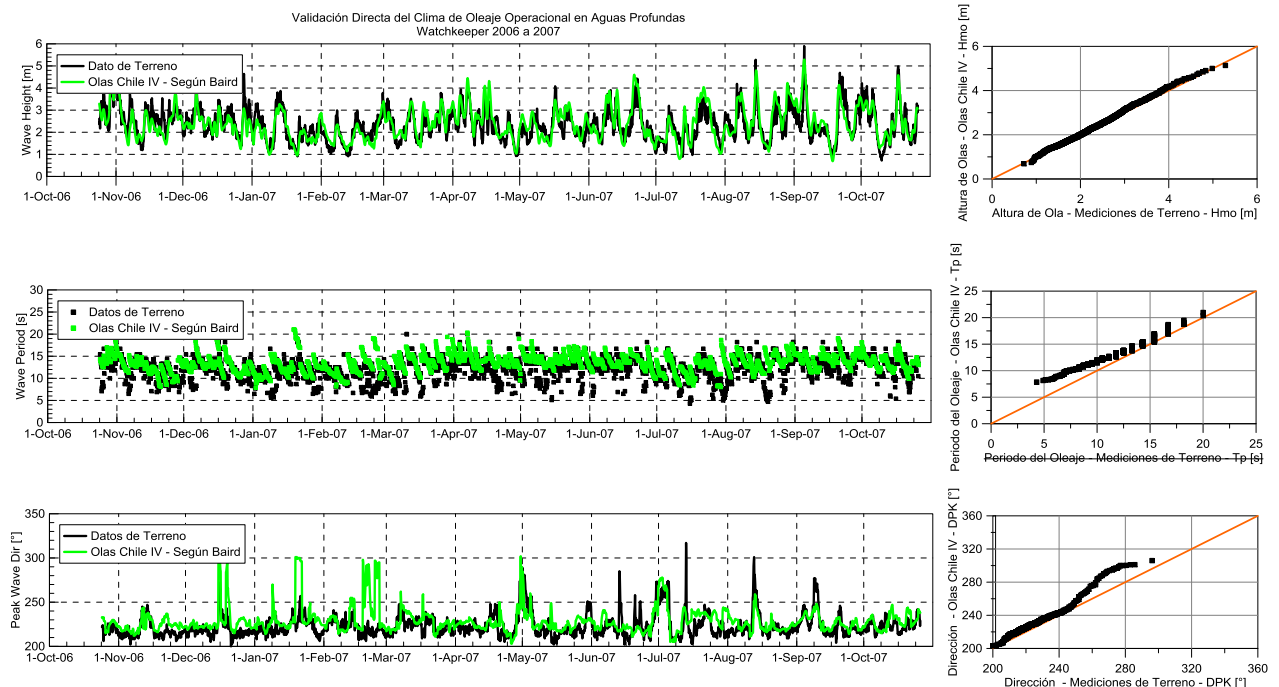


Figura 4.4: Comparación en aguas profundas nodo Valparaíso vs boya Watchkeeper

Fuente: Baird

4.3.2 IDENTIFICACIÓN DE SISTEMAS DE OLAJE

La definición de los sistemas de oleaje en aguas profundas se ha llevado a cabo a partir de un método de partición espectral multimodal sobre los espectros bidimensionales de aguas profundas de cada nodo utilizado, Iquique, Valparaíso y Chile54S por medio de la aplicación de algoritmos avanzados y automatizados. Este método permite aislar la energía de cada espectro bidimensional de la base de datos, resultando en una adecuada definición de cada sistema presente. A continuación se presenta un resumen del método de acuerdo a Rojas (2015):

1. Aislamiento del *peak* de energía espectral.
2. Identificación y combinación de los *peak* espectrales comunes.
3. Remoción de particiones con energías bajas.
4. Cálculo de estadísticas para cada partición.
5. Generación de series de tiempo asociadas a cada sistema de oleaje identificado (partición).

La Figura 4.5 expone un ejemplo de la partición espectral multimodal y de los sistemas de oleaje identificados para los nodos de Iquique, Valparaíso y Chile54S. Así mismo, la Tabla 4.1 presenta una lista con los nombres de estos sistemas. Todo el análisis efectuado en este trabajo se realizó para cada uno de los sistemas de oleaje en cada uno de los nodos de aguas profundas.

Tabla 4.1: Sistemas de oleaje identificados

Zona	Nodo	Sistemas de oleaje identificados
Norte	Iquique	<i>Swell</i> cercano SW
		<i>Swell</i> remoto SW
		<i>Swell</i> cercano NE SE
		<i>Swell</i> remoto NW
Centro	Valparaíso	<i>Swell</i> cercano SW
		<i>Swell</i> remoto SW
		<i>Swell</i> cercano NW E
		<i>Swell</i> remoto NW
Sur	Chile 54S	<i>Swell</i> cercano NE SW
		<i>Swell</i> SW NW
		<i>Swell</i> remoto NW

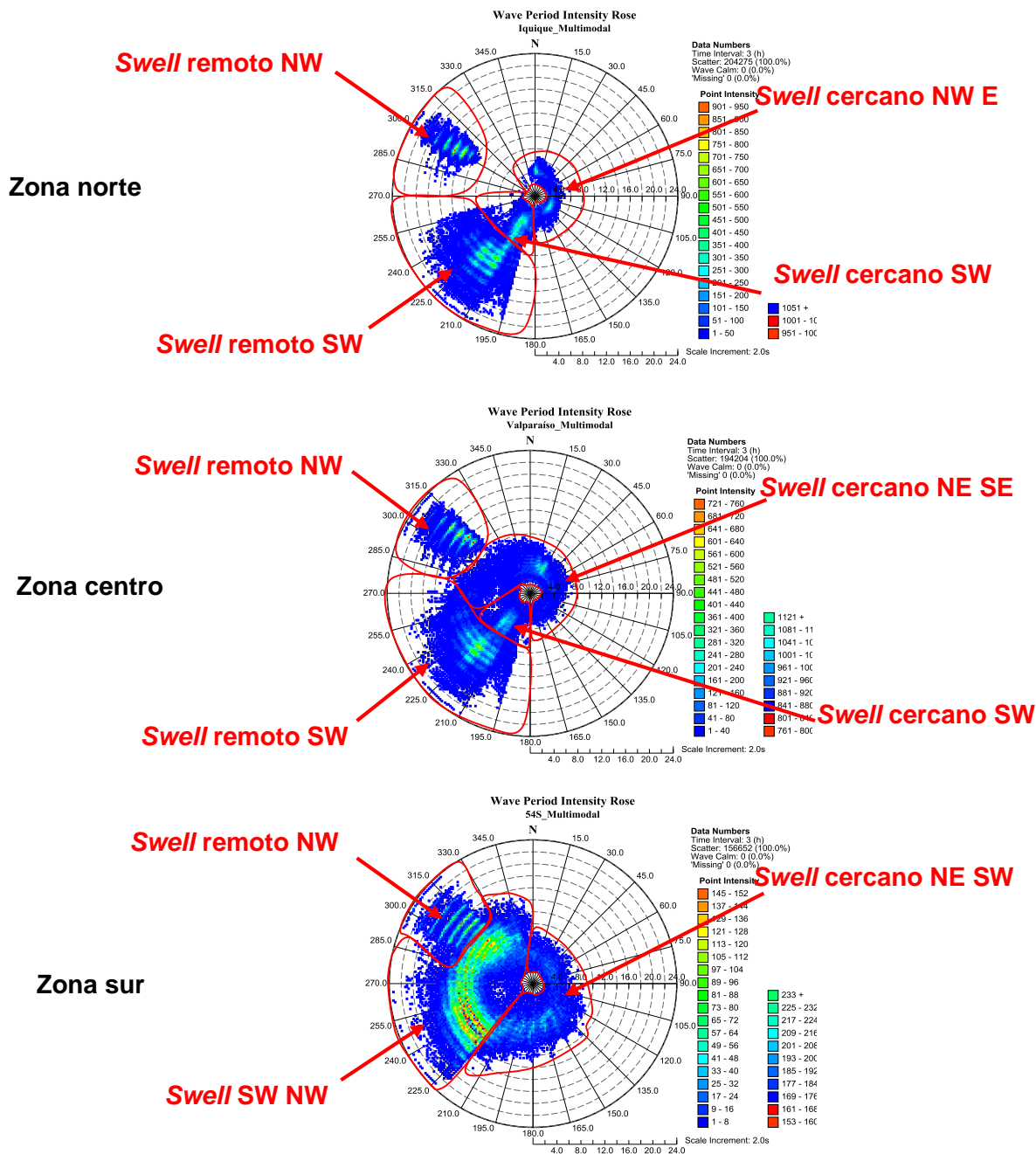


Figura 4.5: Sistemas de oleaje identificados – Zona norte, centro y sur

Fuente: Elaboración propia

4.3.3 AGRUPACIÓN DE SWELLS

La comprensión del comportamiento de las tormentas de aguas profundas ha resultado en avances que mejoran la estimación de las características y propiedades de los eventos, así también como la definición de su dirección de procedencia y distancia hasta su punto de generación.

De acuerdo a lo anterior, un método denominado Rastreo de *Swells* (*Swell Tracking* en inglés) ha sido desarrollado y propuesto por Hanson (2001), el cual permite determinar de manera adecuada las características físicas de las tormentas en un punto, permitiendo identificar además todos los parámetros que las componen, su duración y también la forma de cada una de ellas.

Para determinar si una nueva partición de oleaje pertenece a un grupo de particiones previamente identificadas, la posición de ésta en un espacio ponderado de altura-periodo-dirección es comparado con la ubicación de un grupo ya estimado calculado a partir de las propiedades promedio de las cinco particiones previas de cada grupo existente. Si el grupo contiene menos de cinco particiones, entonces todas las particiones son utilizadas para determinar la ubicación del grupo. La distancia ponderada entre una nueva observación y el valor del grupo ya calculado está determinada por:

$$d = \sqrt{\left(\frac{a_{rms,peak} - a_{rms,grupo}}{3a_{rms,grupo}}\right)^2 + \left(\frac{\theta_m - \theta_{grupo}}{\pi}\right)^2 + \left(\frac{\log f_m - \log f_{grupo}}{\frac{1}{3}}\right)^2} \quad (24)$$

Donde a_{rms} es la amplitud de oleaje raíz media cuadrática, θ_m corresponde a la dirección media del oleaje medida en radianes y $\log f_m$ equivale al logaritmo natural de la frecuencia media. Los coeficientes (3; π ; 1/3) han sido introducidos para enfatizar la importancia relativa de cada término. La nueva observación es asignada al grupo con el cual comparte el menor valor de d . Los valores de d mayores que un umbral asignado son considerados demasiado altos para una relación válida. Cuando una partición no puede ser asignada a un grupo existente, se crea un nuevo grupo para ella.

Los factores de ponderación de la Ecuación (24) han sido seleccionados por el autor a partir de un proceso iterativo empírico de ensayo-error, el cual ha entregado resultados confiables de acuerdo a los datos estudiados. Un ejemplo gráfico de los resultados del método se expone en la Figura 4.6. En ella se presentan todos los sistemas de oleaje (diferenciados por colores) presentes en la boya NDBC 51028 en el océano Pacífico, tanto para las mediciones (A) como para los resultados del modelo Wavewatch III (B). En ambas figuras las alturas están representadas por la longitud de los vectores de cada componente en el tiempo.

Respecto de los periodos, se observa que estos tienen una forma descendente para algunos sistemas de oleaje a medida que el tiempo transcurre, indicando que los periodos mayores se presentan primero por efecto de la dispersión del oleaje. Finalmente, la orientación de los vectores indican las direcciones de procedencia de cada componente dentro de cada uno de los sistemas. De esta forma es posible identificar su forma evolutiva, siendo posible también estimar su duración.

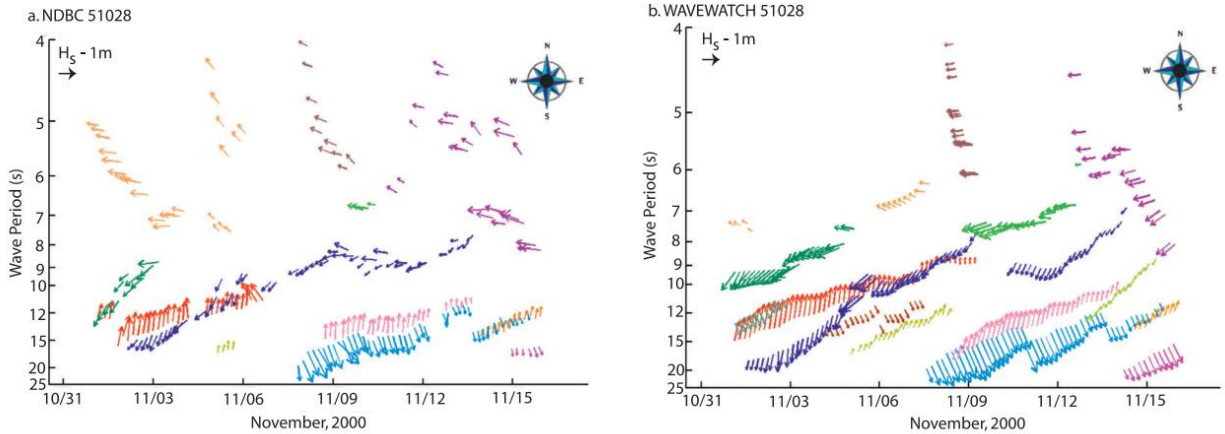


Figura 4.6: Ejemplo de evolución de sistemas de oleaje particionados. Estación 51028, 1 – 15 de Noviembre de 2000. a) Mediciones de boya NDBC. b) Hindcast WaveWatch III

Fuente: Pacific Hindcast Performance of Three Numerical Wave Models. Hanson, J., Tracy, B., Tolman, H., Scott, D. 2009

La ventaja de este método es que permite realizar una separación de tormentas a partir del comportamiento físico de sus parámetros componentes, específicamente alturas, periodos, direcciones y duración. De esta manera es posible aislar una tormenta y describirla desde el inicio hasta el término de su desarrollo en un sitio específico. Tomando en cuenta que los periodos y direcciones *peak* representan de mejor manera las características del oleaje, se ha efectuado una modificación sobre la Ecuación (24) propuesta por Hanson (2001). Esta considera la utilización de los parámetros *peak* de oleaje de cada partición en lugar de los valores medios propuestos por el autor, resultando en la Ecuación (25):

$$d = \sqrt{\left(\frac{a_{rms,peak} - a_{rms,grupo}}{3a_{rms,grupo}}\right)^2 + \left(\frac{\theta_{peak} - \theta_{grupo}}{\pi}\right)^2 + \left(\frac{\log f_{peak} - \log f_{grupo}}{\frac{1}{3}}\right)^2} \quad (25)$$

Considerando lo anterior, la metodología consta de los siguientes pasos:

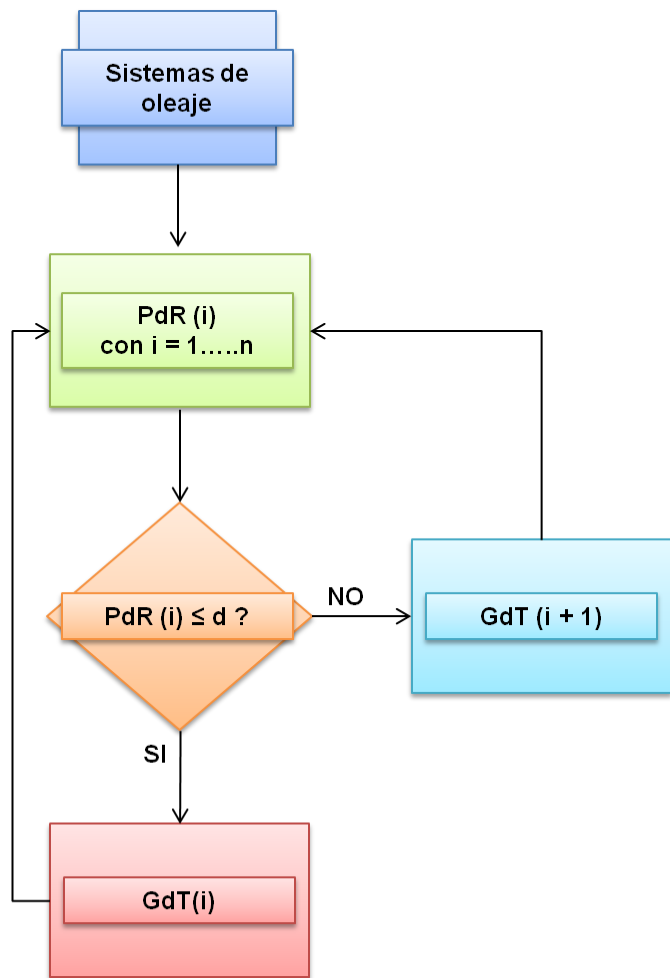
1. Utilizando cada componente de las series de tiempo obtenidas desde la separación espectral multimodal, se determina el valor de distancia espacio-temporal (d) entre componentes utilizando la Ecuación (25). La Tabla 4.2 presenta el resumen de los valores del parámetro d utilizados para cada uno de los sistemas de oleaje de los nodos analizados cuyos desempeños fueron adecuados.
2. Si el valor de $d \leq 0.6$ entonces las componentes de oleaje pertenecen al mismo grupo, es decir, forman parte de la misma tormenta, de lo contrario se comienza a crear una tormenta nueva.

Es importante destacar que con cada separación de tormenta realizada se crea grupo nuevo que almacena las propiedades de cada tormenta

La Figura 4.7 muestra un diagrama de flujo simplificado del método de separación de tormentas.

Tabla 4.2: Valores del parámetro d utilizados para cada sistema de oleaje

Nodo	Caso	Valor de parámetro d
Zona norte	<i>Swell cercano SW</i>	1.2
	<i>Swell remoto SW</i>	0.6
	<i>Swell cercano NE SE</i>	3.0
	<i>Swell remoto NW</i>	0.8
Zona centro	<i>Swell cercano SW</i>	1.2
	<i>Swell remoto SW</i>	0.8
	<i>Swell cercano NW E</i>	3.0
	<i>Swell remoto NW</i>	0.8
Zona sur	<i>Swell cercano NE SW</i>	3.0
	<i>Swell SW NW</i>	0.6
	<i>Swell remoto NW</i>	0.8



PdR: Parámetros de resumen

GdT: Grupo de tormenta

Figura 4.7: Diagrama de flujo. Método de separación de tormentas

Fuente: Elaboración propia

4.3.4 IDENTIFICACIÓN DE VALORES EXTREMOS

Los criterios más utilizados para identificar los valores extremos en una serie de oleaje son el *Peak Over Threshold* (POT) y los máximos anuales. Dado que en este trabajo se lograron aislar las tormentas con la partición espectral multimodal y la agrupación de *swells*, el POT puede ser reemplazado por la elección de los máximos por tormentas. Los tres métodos son explicados a continuación:

4.3.4.1 MÉTODO DE VALORES SOBRE EL UMBRAL

El método de valores sobre el umbral o *Peak Over Threshold* (en adelante POT por sus siglas en inglés) es ampliamente utilizado en la selección de datos de tormentas para realizar análisis de valores extremos (EVA). Éste consiste en determinar grupos de datos de una variable específica cuyas magnitudes sean mayores a un valor umbral fijo previamente seleccionado, eligiendo a continuación el valor máximo (*peak*) de cada grupo. La Figura 4.8, muestra un esquema de aplicación del método.

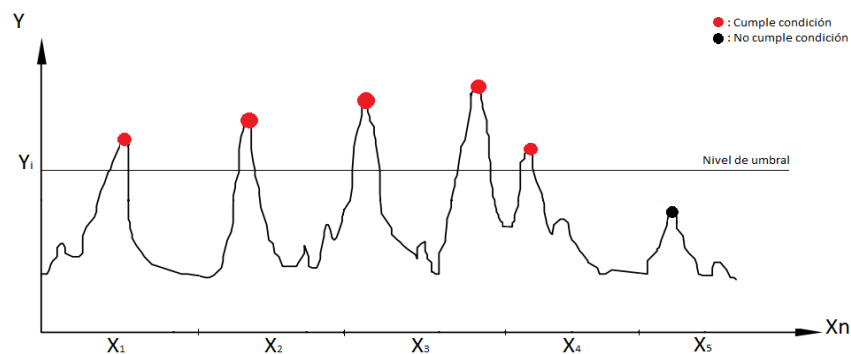


Figura 4.8: Esquema de método de valores sobre el umbral

Fuente: Elaboración propia

La selección de umbral, una duración máxima por tormenta y una distancia mínima entre tormentas en el método POT representan de manera simplificada la separación de tormentas aplicada en este trabajo.

4.3.4.2 MÉTODO DE MÁXIMOS ANUALES

Otro de los métodos para realizar análisis de valores extremos es el de máximos anuales. Este método consiste en seleccionar el valor máximo de la variable por año dentro de la base de datos. A continuación, la Figura 4.9 expone un esquema del método de máximos anuales.

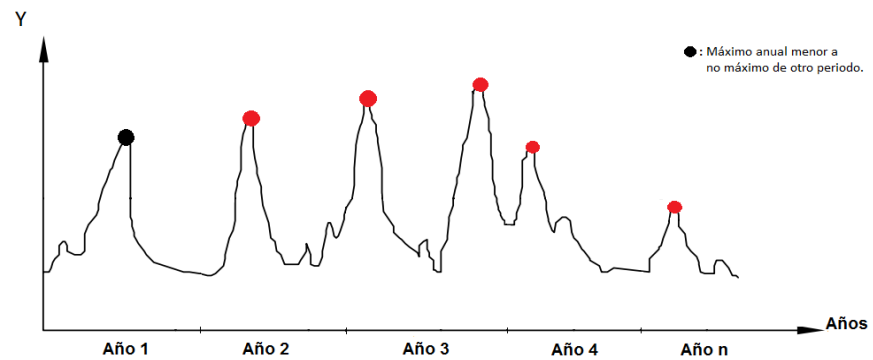


Figura 4.9: Esquema de método de máximos anuales

Fuente: Elaboración propia

4.3.4.3 MÉTODO DE LOS MÁXIMOS POR TORMENTA

La Figura 4.10 expone una esquematización de la identificación de tormentas dentro de una serie de tiempo de parámetros de resumen diferenciadas por colores luego de aplicar el método de agrupación de *swells*. En ellas se representan las evoluciones de alturas, periodos y direcciones a través del tiempo.

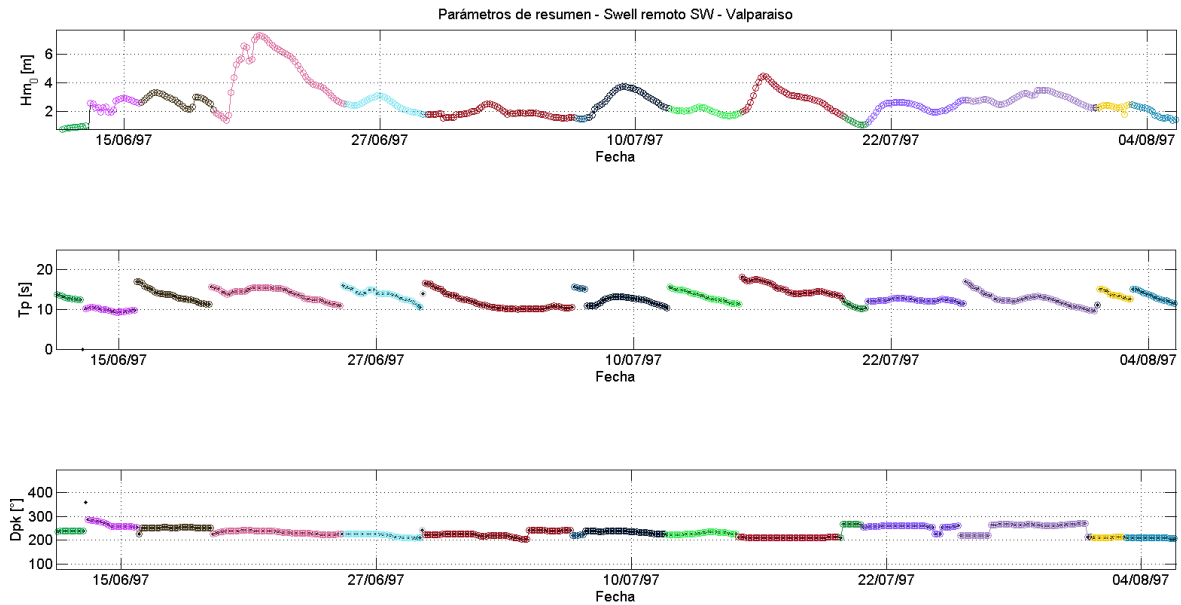


Figura 4.10: Serie de tiempo de tormentas y sus parámetros

Fuente: Elaboración propia

Se han extraído las alturas máximas, los periodos y las direcciones asociadas a éstas, además de las duraciones para cada conjunto de particiones identificado. Por lo tanto, cada tormenta queda representada por una altura de oleaje (H_{m0}), un periodo *peak* (T_p), una dirección *peak* (D_{pk}) y su duración.

Un esquema de cómo se han considerado los datos extraídos se presenta en la Figura 4.11. La altura máxima (H_{m0max}) se establece como la mayor altura de ola dentro de un evento de tormenta. A su vez, el periodo T_{pmax} y la duración D_{pkmax} corresponden a los valores de estas variables cuando en el instante de tiempo en que ocurre la mayor altura de ola en la tormenta H_{m0max} .

Esta identificación se ha llevado a cabo para todas las tormentas analizadas y almacenadas dentro de cada nodo de aguas profundas utilizado.

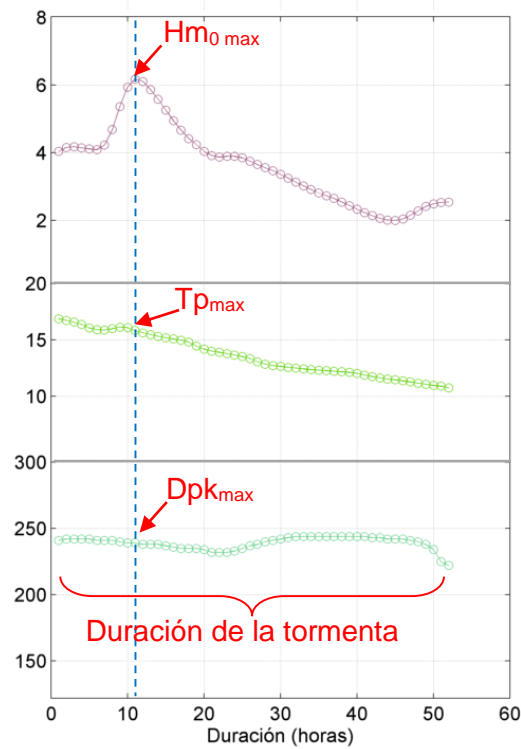


Figura 4.11: Identificación de parámetros – Tormenta típica

Fuente: Elaboración propia

Una vez creados los grupos de tormentas, se procede a extraer y seleccionar los parámetros a utilizar en las simulaciones de valores extremos multivariados.

En resumen, de cada tormenta se extrajeron los parámetros componentes para ser utilizados en las simulaciones de valores extremos multivariados, totalizando una lista de 4 variables las cuales corresponden a alturas máximas, periodos y direcciones asociados a las alturas, y también las duraciones.

4.3.5 MODELACIÓN MULTIVARIADA DE PARÁMETROS EXTREMOS

A partir de los datos generados por el método de los máximos por tormenta se generan gráficos de dispersión de los valores de altura máxima versus periodo, dirección y duración similares al expuesto en la Figura 4.12.

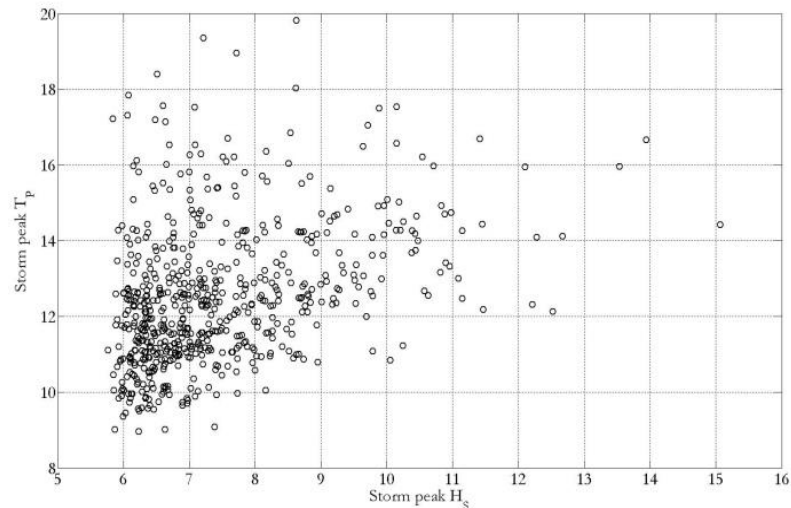


Figura 4.12: Dispersión de datos. Ejemplo de alturas máximas vs periodos *peak* de oleaje

Fuente: “Joint Modeling of Waves Spectral Parameters for Extreme Sea States”. Ocean Engineering. 2010. Jonathan, P., Flynn, J. Ewans, K., 2010

A continuación, la lista de parámetros es ordenada de manera descendente en función de la altura máxima de oleaje. Luego, se aplica la metodología de probabilidad conjunta multivariada utilizada para la estimación de variables extremas de oleaje, la cual para este estudio considera la estimación de dos variables de manera simultánea. Si bien esta puede ser extensible a n -dimensiones, la complejidad de cálculo y aplicación aumenta (Heffernan & Tawn, 2004).

El método consta de dos etapas:

- a) Simulación marginal univariada de las variables seleccionadas.
- b) Modelación conjunta de los parámetros de oleaje seleccionados.

4.3.5.1 MODELO MARGINAL

La modelación marginal de valores extremos permite estimar el comportamiento estadístico de las colas superiores, inferiores, o bien ambas de un conjunto de datos estudiado, en este caso, los parámetros identificados en la sección previa son tratados cada uno de manera independiente.

El procedimiento consiste en ajustar una distribución de probabilidad univariada a cada conjunto de parámetros seleccionados, ordenados de manera descendente en función de la altura máxima.

La distribución marginal utilizada corresponde a la de Pareto, la cual es aplicable a valores sobre un umbral previamente establecido. Con base en lo anterior, esta distribución es aplicada a los datos seleccionados para las simulaciones luego de escoger un umbral marginal, en este caso, seleccionado a partir del método de estabilidad de parámetros.

Luego de realizados los ajustes, se efectúa un cambio de dominio de los valores ajustados por medio de la transformada integral de probabilidad para convertir los marginales obtenidos desde una escala de Pareto a una escala de Laplace, con el fin de obtener valores ajustados que se comporten de manera exponencial en sus colas superior e inferior.

A partir de lo anterior, se seleccionan dos variables para ser aplicadas en el modelo condicional, una correspondiente a la variable condicionante y otra a la variable condicionada. En este estudio se consideró la altura de oleaje máxima como la variable condicionante, mientras que las variables condicionadas corresponden a las demás variables dentro de la lista de datos. De manera posterior, se procede a realizar la modelación condicional de las variables (periodos, direcciones y duraciones).

La Figura 4.13 (a) es un ejemplo de simulación multivariada. En ella se expone en color gris los datos ajustados y transformados a espacio de Laplace, donde la variable condicionante se grafica sobre el eje X , denotada por Y_i y la variable condicionada sobre el eje Y , denotada por Y_j .

4.3.5.2 MODELO CONDICIONAL

El modelo condicional permite conocer el comportamiento conjunto de las variables condicionante y condicionada de manera simultánea, utilizando al igual que en la simulación marginal, un nivel de umbral condicional.

Este comportamiento conjunto queda determinado por los coeficientes de dependencia extrema (a , b) estimados de manera iterativa a partir de un modelo de regresión multivariada. Este modelo considera un nivel de umbral condicional correspondiente a los datos mayores o iguales al percentil seleccionado por el modelador sobre las muestras marginales de

Laplace. En este estudio se ha considerado un nivel de umbral igual al utilizado en la simulación marginal, no obstante, este valor puede ser distinto, y es determinado según el criterio considerado para las simulaciones.

De manera posterior, los coeficientes de dependencia son utilizados en el modelo semiparamétrico multivariado (Ecuación 18) para la extrapolación conjunta de las variables mediante simulación.

4.3.5.3 NIVEL DE RETORNO Y ESTIMACIÓN DE PARÁMETROS EXTREMOS

El procedimiento utilizado para la estimación de parámetros extremos asociados a distintos niveles de retorno se expone en este apartado.

Para llevar a cabo la estimación de parámetros extremos se realiza el cálculo del valor de retorno requerido de la variable condicionante y su probabilidad asociada a la aparición del evento. Esta probabilidad es utilizada como nivel a partir del cual se desea realizar las simulaciones de Montecarlo de las variables extremas con el modelo condicional. Esto permite la generación de valores extremos simulados representativos del comportamiento de las variables condicionante y condicionada. A modo de ejemplo, en la Figura 4.13 (a) se expone un caso en el cual las simulaciones representadas por los puntos de color azul, indican el comportamiento conjunto de las colas extremas de los ajustes marginales de las variables simuladas sobre un umbral condicional.

Una vez obtenidos los valores extremos simulados, los cuales se encuentran también en un espacio de Laplace, son transformados de regreso a su distribución de origen (Pareto) y luego a su espacio real, para así obtener los valores extremos extrapolados de las variables de estudio en un espacio real.

Posteriormente, se ajusta una distribución *kernel* bidimensional a la dispersión de los valores extremos en espacio real. Esto permite estimar la densidad de probabilidad de los pares simulados asociados a las variables condicionante y condicionada. A partir de lo anterior y utilizando el valor de retorno de la variable condicionante, se efectúa un *slice* sobre las distribución *kernel* con el fin de encontrar la densidad de probabilidad de la variable condicionada asociada a la variable condicionante de retorno. De esta forma es posible conocer los valores probables de la variable condicionada, asociada al mismo nivel de retorno de la variable condicionante.

Como caso de ejemplo, la estimación de un periodo *peak* con una probabilidad del 50% de ocurrencia, asociado a una altura de oleaje de 100 años de periodo de retorno. La Figura 4.13 (b) muestra un esquema del ejemplo mencionado.

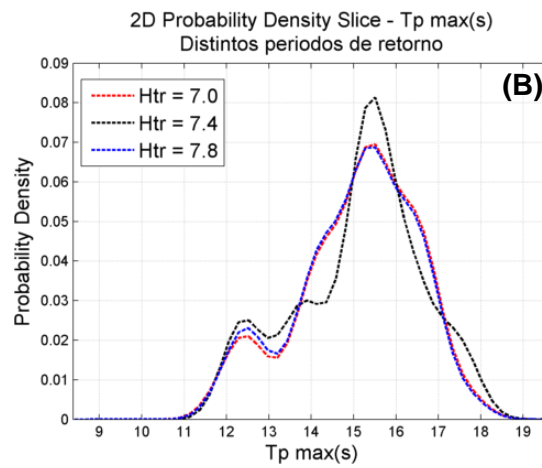
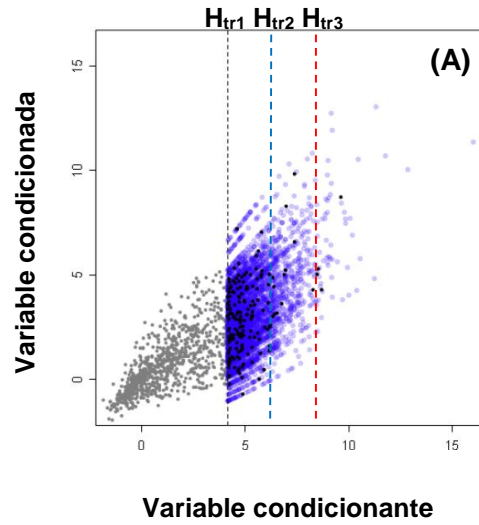


Figura 4.13: Extrapolación condicional multivariada. a) Dispersión de datos extremos simulados. b) Densidad de probabilidad de periodos máximos en función de la altura de retorno

Fuente: Elaboración propia

5 RESULTADOS

Los resultados del análisis multivariado de valores extremos han sido obtenidos para periodos de retorno correspondientes a 25, 50 y 100 años para los nodos de Iquique, Valparaíso y Chile 54S (latitud sur), respectivamente, los cuales se presentan en la sección 5.3.

En las secciones 5.1 y 5.2 de este capítulo se muestran como ejemplo resultados de las etapas previas del análisis multivariado de valores extremos en la zona central.

5.1 AGRUPACIÓN DE SWELLS

Utilizando el método de Hanson (2001) se han agrupado las tormentas para cada uno de los sistemas de oleaje identificados.

En las Figuras 5.2 a 5.5 se muestran las tormentas diferenciadas entre sí por colores para cada sistema de oleaje durante un intervalo de tiempo de un mes. Los vectores representan los parámetros del oleaje para cada componente de una tormenta.

Cada vector representa al sistema de oleaje en un estado de mar. El periodo del oleaje está representado por el punto de inicio (en el eje de las ordenadas) de los vectores componentes. Así mismo, la altura de oleaje se representa por la longitud de cada vector, cuya escala está en función del mayor valor de altura de ola encontrado en cada serie de tiempo desde 1980 a 2009. La dirección de incidencia de cada parámetro, por su parte, queda representada por la dirección hacia donde apunta cada uno de los vectores, en escala náutica.

Los resultados muestran la agrupación natural de los *swells* presentes dentro de las series de tiempo de oleaje, en los cuales es posible notar dónde ocurre el cambio de una tormenta a otra y cómo el algoritmo implementado es capaz de detectarlas.

El oleaje generado en una misma zona geográfica presenta características similares en sus parámetros, lo cual implica que un sistema de oleaje contiene parámetros de la misma naturaleza y origen. De esta manera, las tormentas diferenciadas por colores en las figuras antes mencionadas, fueron generadas en la misma zona y por eventos distintos.

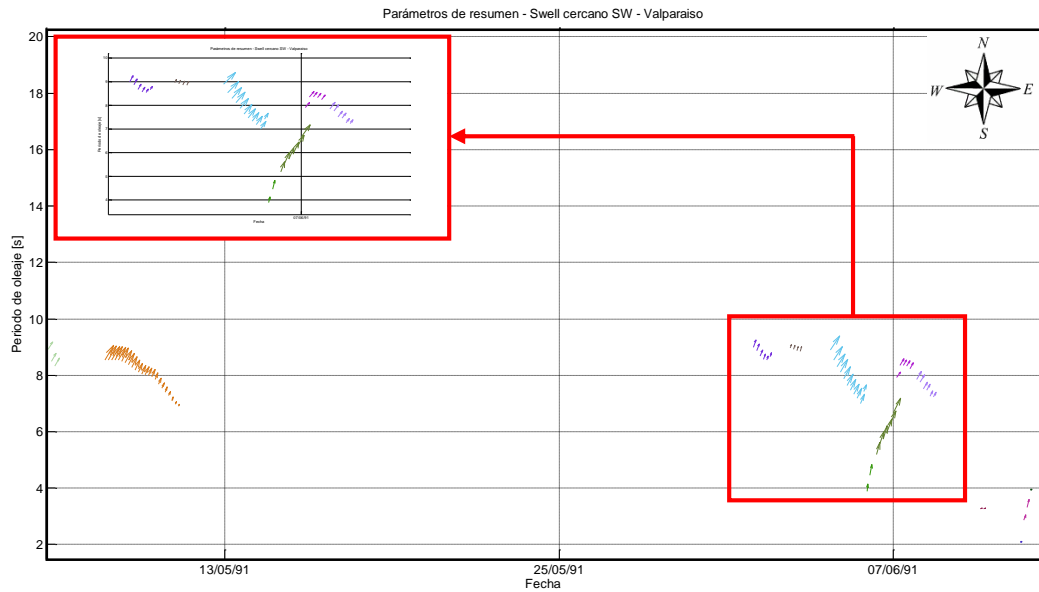


Figura 5.1: Separación se tormentas – Swell cercano SW, Zona central
Fuente: Elaboración propia

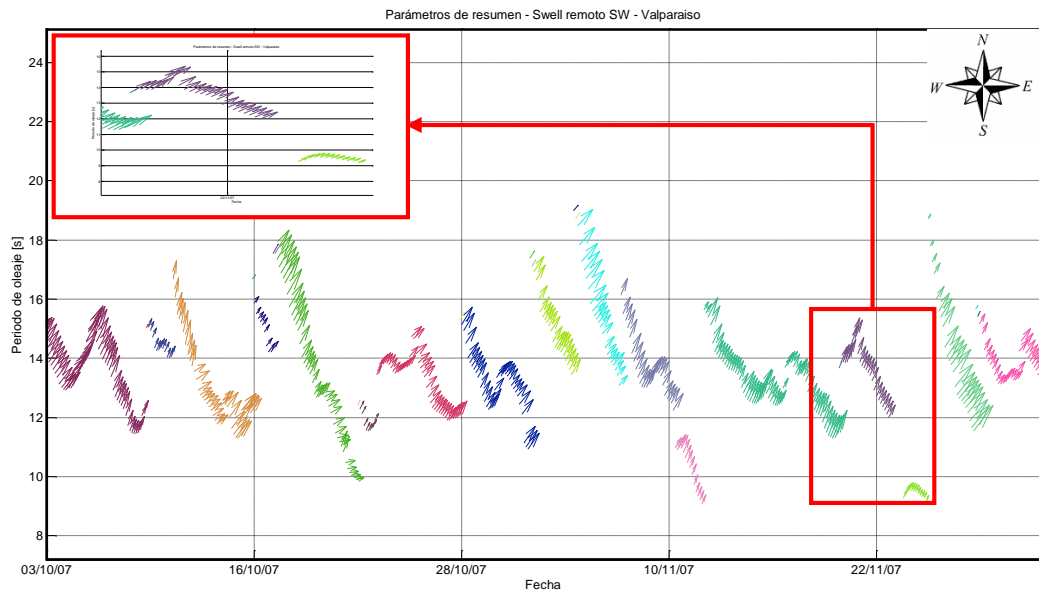


Figura 5.2: Separación de tormentas – Swell remoto SW, Zona central
Fuente: Elaboración propia

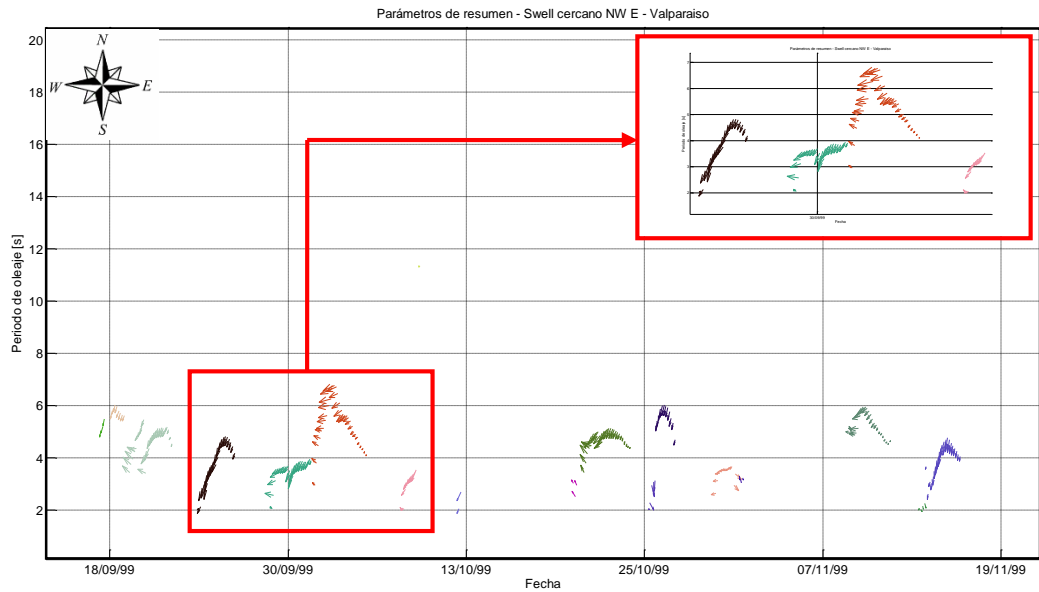


Figura 5.3: Separación de tormentas – Swell cercano NW E, Zona central
Fuente: Elaboración propia

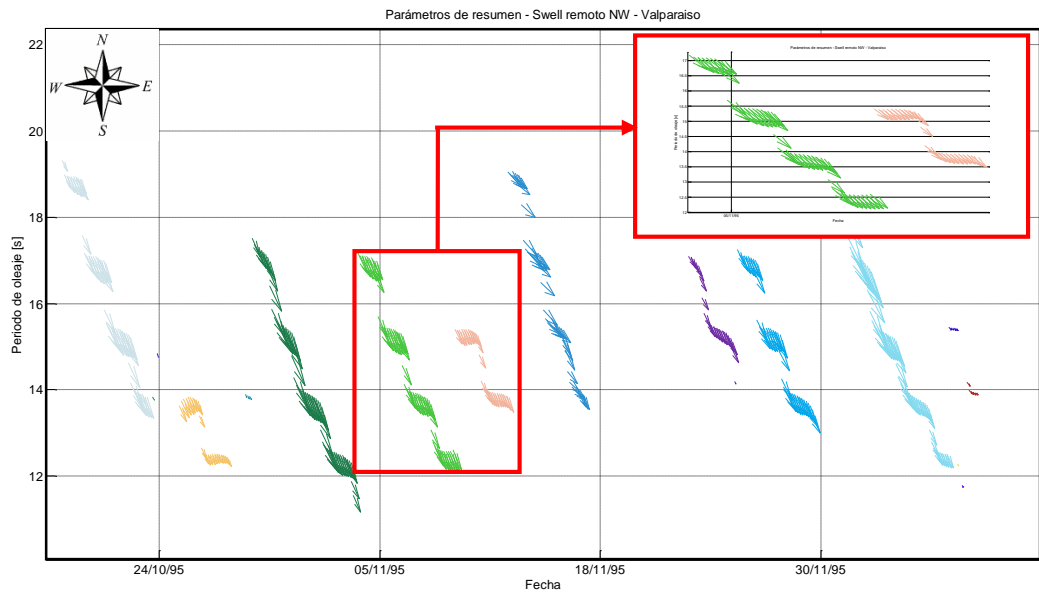


Figura 5.4: Separación de tormentas – Swell remoto NW, Zona central
Fuente: Elaboración propia

De las figuras se puede notar que para todos los sistemas de oleaje identificados, principalmente en los *swells* remotos, existe una separación adecuada de las tormentas dentro de las series de tiempo. En ellas se evidencia el efecto de la dispersión de frecuencias, a través de la cual se produce el efecto que los eventos con periodos más altos se presentan primero debido a que poseen una velocidad de propagación mayor, reflejándose en un ordenamiento descendente de los valores (si estos son graficados en función de los periodos, como es el caso expuesto).

Para los *swells* remotos se logra establecer una forma característica en sus tormentas debido a que estos sistemas de oleaje fueron generados por una fuente lejana, lo cual permite un desarrollo completo del oleaje a medida que viaja en el espacio oceánico. Por su parte, las tormentas del tipo *swell* cercanos presentan mayor dispersión en sus apariciones, ya que estos sistemas fueron generados por fuentes más próximas. Cabe destacar que la formulación utilizada para la separación de *swells* es sensible a estas variaciones, resultando ser más preciso al ser aplicado a los *swells* remotos, que tienen una menor variabilidad.

No obstante lo anterior, en todas los casos el método utiliza de manera adecuada las propiedades físicas del oleaje, ésto lo convierte en un proceso física y matemáticamente robusto.

5.2 LISTAS DE MAYORES TORMENTAS Y COMPARACIÓN CON POT

El objetivo de esta sección es comparar los resultados del método de agrupación de *swells* y selección de máximos por tormenta con los resultados que se obtienen con el método POT

En las Tablas 5.1, 5.3, 5.5 y 5.7 se muestran listas con las 30 mayores tormentas en cada uno de los sistemas de oleaje en la zona central (Valparaíso) obtenidas mediante agrupación de *swells* y la selección de los máximos por tormentas.

Éstas se comparan con las 30 mayores tormentas obtenidas mediante el método POT aplicado sobre las series de estados de mar, considerando umbrales de 3.3 m para el sistema *swell* cercano SW, 5.1 m para el *swell* remoto SW, 3.8 m para el sistema *swell* cercano NW E y finalmente un umbral de 1.32 m para el *swell* remoto NW. La distancia mínima entre tormentas es de 48 horas para todos los sistemas considerados.

La comparación muestra algunas diferencias entre las listas con ambos métodos, considerando que el método de separación de *swells* incluye las variaciones físicas de los parámetros de resumen del oleaje, mientras que el POT aplicado sobre las series de estados de mar es una aproximación matemática para separar las tormentas sobre cierto umbral en cada sistema de oleaje. Las barras de color rojizo indican los casos en que ocurren estas diferencias, mientras que las barras verdes muestran los porcentajes de las diferencias identificadas entre los valores de alturas de oleaje.

Cabe mencionar que las listas han sido ordenadas por orden decreciente de magnitudes de alturas de ola. Las diferencias entre otros parámetros se resumen en las Tablas 5.2, 5.4, 5.6 y 5.8 con las estadísticas bajo cada lista de tormentas.

Tabla 5.1: 30 mayores tormentas swell cercano SW - POT (izq.) y máximos por tormenta (der.)

Altura de Ola	Período Peak	Dirección Peak	Dur (horas)	Altura de Ola	Período Peak	Dirección Peak	Diferencia de magnitudes de alturas [%]	Dur (horas)
Hm0 [m]	Tp [s]	DPK [°]		Hm0 [m]	Tp [s]	DPK [°]		
4.0	9.0	221	24	4.1	8.7	222	1%	12
4.0	8.8	221	42	4.0	8.8	223	0%	9
4.0	8.8	223	99	3.8	8.6	219	5%	9
3.7	8.7	224	36	3.7	8.4	223	1%	36
3.7	8.2	221	51	3.7	8.7	224	1%	15
3.6	8.5	220	72	3.7	8.2	221	0%	30
3.6	8.6	223	33	3.6	8.5	220	1%	30
3.6	8.2	225	51	3.6	8.6	223	0%	15
3.6	8.5	219	36	3.6	8.2	225	1%	42
3.6	8.5	214	57	3.6	8.5	219	0%	75
3.5	7.9	221	33	3.6	8.5	214	1%	15
3.5	8.0	222	33	3.6	8.6	221	1%	9
3.5	8.2	224	24	3.5	8.6	223	1%	15
3.5	8.5	223	30	3.5	7.9	221	1%	45
3.5	8.5	218	33	3.5	8.0	222	1%	9
3.5	8.5	215	48	3.5	8.2	224	1%	42
3.5	8.3	213	27	3.5	8.5	223	1%	33
3.5	8.4	222	30	3.5	8.5	218	1%	93
3.5	8.4	210	27	3.5	8.5	215	1%	18
3.4	8.9	213	30	3.5	8.3	222	1%	15
3.4	8.4	215	57	3.5	8.3	213	1%	183
3.4	8.4	219	30	3.5	8.3	213	1%	18
3.4	8.2	223	36	3.5	8.4	222	1%	15
3.4	8.0	229	27	3.5	8.4	210	1%	186
3.4	8.8	220	51	3.4	8.9	213	1%	24
3.4	8.5	218	30	3.4	8.5	213	2%	18
3.4	8.8	221	24	3.4	8.6	212	3%	9
3.3	8.5	221	51	3.4	8.3	219	4%	42
3.3	8.5	216	45	3.4	8.4	219	4%	39
3.3	8.1	222	30	3.4	8.4	215	3%	186

Tabla 5.2: Estadística de alturas de ola identificadas - swell cercano SW

Swell cercano SW						
Variable	Método	Mín. (m)	Máx. (m)	Promedio	Sesgo	RMSE
Hm ₀ (m)	POT	3.3	4.0	3.5	0.1	0.07
	Separación de Swells	3.4	4.1	3.6		
Tp (s)	POT	7.9	9.0	8.4	0.0	0.32
	Separación de Swells	7.9	8.9	8.4		
Dpk (°)	POT	210.0	229.0	219.9	0.8	6.05
	Separación de Swells	210.0	225.0	219.0		
Duración (horas)	POT	24.0	99.0	39.9	3.0	54.00
	Separación de Swells	9.0	186.0	42.9		

Tabla 5.3: 30 mayores tormentas swell remoto SW - POT (izq.) y máximos por tormenta (der.)

Altura de Ola	Período Peak	Dirección Peak	Dur (horas)	Altura de Ola	Período Peak	Dirección Peak	Diferencia de magnitudes de alturas [%]	Dur (horas)
Hm0 [m]	Tp [s]	DPK [°]		Hm0 [m]	Tp [s]	DPK [°]		
7.4	15.5	240	78	7.4	15.5	240	0%	147
7.2	15.5	257	48	7.2	15.5	257	0%	138
7.0	15.5	253	39	7.0	15.5	253	0%	183
6.4	15.5	222	45	6.4	15.5	222	0%	165
6.4	16.8	240	45	6.4	16.8	240	0%	93
6.4	12.7	269	30	6.4	12.7	269	0%	15
6.2	15.8	239	27	6.2	15.8	239	0%	186
5.9	15.6	225	33	5.9	15.6	225	0%	186
5.9	13.6	258	39	5.9	13.6	258	0%	81
5.7	15.3	237	36	5.7	15.3	237	0%	132
5.7	12.8	253	21	5.7	12.8	253	0%	63
5.7	17.1	221	30	5.7	17.1	221	0%	240
5.6	14.6	231	24	5.6	13.5	244	0%	105
5.6	17.2	222	27	5.6	14.6	231	0%	315
5.5	16.9	224	33	5.6	17.2	222	1%	72
5.5	13.8	257	30	5.5	16.9	224	0%	99
5.5	15.4	237	30	5.5	13.8	257	0%	132
5.4	10.9	266	84	5.5	15.4	237	2%	195
5.4	10.8	282	42	5.4	11.6	276	1%	18
5.4	14.5	229	36	5.4	10.9	266	0%	138
5.4	14.2	236	39	5.4	10.8	282	1%	255
5.3	16.0	236	27	5.4	15.3	208	1%	57
5.3	14.8	223	39	5.4	14.5	229	1%	303
5.3	14.8	242	27	5.4	14.2	236	2%	66
5.2	15.3	241	21	5.3	16.0	236	2%	228
5.2	16.5	220	30	5.3	14.8	223	1%	153
5.2	15.7	222	21	5.3	14.8	242	1%	75
5.2	14.6	238	21	5.2	15.3	241	1%	96
5.2	17.7	225	75	5.2	16.5	220	1%	150
5.1	14.8	223	21	5.2	15.7	222	2%	108

Tabla 5.4: Estadística de alturas de ola identificadas - swell remoto SW

Swell remoto SW						
Variable	Método	Mín. (m)	Máx. (m)	Promedio	Sesgo	RMSE
Hm ₀ (m)	POT	5.1	7.4	5.7	0.1	0.04
	Separación de Swells	5.2	7.4	5.8		
Tp (s)	POT	10.8	17.7	15.0	0.2	1.57
	Separación de Swells	10.8	17.2	14.8		
Dpk (°)	POT	220.0	282.0	238.9	1.4	15.74
	Separación de Swells	208.0	282.0	240.3		
Duración (horas)	POT	21.0	84.0	36.6	103.2	127.02
	Separación de Swells	15.0	315.0	139.8		

Tabla 5.5: 30 mayores tormentas swell cercano NW E - POT (izq.) y máximos por tormenta (der.)

Altura de Ola	Período Peak	Dirección Peak	Dur (horas)	Altura de Ola	Período Peak	Dirección Peak	Diferencia de magnitudes de alturas [%]	Dur (horas)
Hm0 [m]	Tp [s]	DPK [°]		Hm0 [m]	Tp [s]	DPK [°]		
6.0	10.3	319	27	6.0	10.3	319	0%	90
5.6	10.0	329	54	5.6	10.0	329	0%	114
5.5	9.4	335	84	5.5	9.4	335	0%	105
5.3	9.4	337	39	5.3	9.4	337	0%	216
5.0	9.2	329	24	5.0	9.2	329	0%	159
5.0	8.7	331	27	5.0	8.7	331	0%	69
4.9	9.2	330	165	4.9	9.1	311	1%	30
4.9	8.6	345	30	4.9	9.2	330	0%	126
4.8	8.8	306	39	4.9	8.6	345	0%	150
4.8	9.5	328	42	4.8	8.8	306	1%	120
4.8	10.4	286	33	4.8	9.5	328	0%	165
4.8	9.5	312	36	4.8	10.4	286	0%	63
4.7	9.0	326	27	4.8	9.5	312	1%	84
4.7	8.6	330	21	4.7	9.0	326	0%	48
4.7	8.5	343	27	4.7	8.6	330	0%	72
4.7	9.6	314	27	4.7	8.5	343	1%	156
4.6	8.4	311	36	4.7	9.6	314	2%	108
4.6	8.4	363	90	4.7	9.9	288	2%	39
4.5	10.9	298	21	4.6	8.4	311	1%	36
4.5	9.3	307	78	4.6	8.4	363	1%	333
4.5	8.8	347	36	4.5	8.3	324	2%	51
4.4	8.7	348	33	4.5	10.2	289	3%	57
4.2	8.5	325	132	4.5	10.9	298	7%	177
4.1	9.2	308	21	4.5	9.3	307	8%	189
4.1	8.3	344	27	4.5	8.8	347	9%	72
4.0	8.3	352	21	4.4	8.7	348	9%	138
3.9	8.2	330	30	4.3	8.3	337	9%	30
3.9	8.1	304	30	4.3	8.3	333	9%	105
3.9	8.3	343	69	4.3	9.3	296	10%	39
3.8	7.9	329	21	4.2	9.2	299	10%	99

Tabla 5.6: Estadística de alturas de ola identificadas - swell cercano NW E

<i>Swell cercano NW E</i>						
Variable	Método	Mín. (m)	Máx. (m)	Promedio	Sesgo	RMSE
Hm ₀ (m)	POT	3.8	6.0	4.6	0.2	0.20
	Separación de Swells	4.2	6.0	4.8		
Tp (s)	POT	7.9	10.9	9.0	0.2	0.94
	Separación de Swells	8.3	10.9	9.2		
Dpk (°)	POT	286.0	363.0	327.0	5.3	28.26
	Separación de Swells	286.0	363.0	321.7		
Duración (horas)	POT	21.0	165.0	44.9	63.1	96.33
	Separación de Swells	30.0	333.0	108.0		

Tabla 5.7: 30 mayores tormentas swell remoto NW - POT (izq.) y máximos por tormenta (der.)

Altura de Ola	Período Peak	Dirección Peak	Dur (horas)	Altura de Ola	Período Peak	Dirección Peak	Diferencia de magnitudes de alturas [%]	Dur (horas)
Hm0 [m]	Tp [s]	DPK [°]		Hm0 [m]	Tp [s]	DPK [°]		
2.6	15.6	310	129	3.0	11.6	297	12%	21
2.5	15.4	309	81	2.6	15.6	310	5%	105
2.3	16.5	311	201	2.5	15.4	309	5%	60
2.3	18.8	300	141	2.4	18.3	309	2%	21
2.3	15.6	308	246	2.3	16.5	311	3%	81
2.3	17.6	307	33	2.3	18.8	300	2%	60
2.2	17.0	307	309	2.3	15.6	308	3%	63
2.1	17.1	300	45	2.3	17.6	307	6%	252
2.1	18.7	301	150	2.2	21.1	307	7%	18
2.0	15.7	307	294	2.2	17.0	307	8%	234
2.0	20.9	308	78	2.1	15.2	308	8%	108
1.8	19.2	310	87	2.1	17.1	300	14%	18
1.8	13.8	302	72	2.1	15.4	308	15%	21
1.8	15.1	308	75	2.1	18.7	301	15%	30
1.6	18.7	309	129	2.0	15.7	307	23%	66
1.6	16.9	311	21	2.0	16.9	301	22%	138
1.5	17.2	309	78	2.0	15.3	312	24%	213
1.5	19.0	309	72	2.0	19.1	309	24%	162
1.5	18.6	306	66	1.9	18.8	308	23%	57
1.5	19.0	304	240	1.9	16.8	307	24%	72
1.5	19.0	309	42	1.9	14.0	310	24%	39
1.5	18.9	308	57	1.9	15.3	310	24%	156
1.4	14.8	301	90	1.9	16.9	309	25%	66
1.4	18.6	311	45	1.9	13.8	308	25%	27
1.4	13.5	310	45	1.9	17.1	309	26%	24
1.4	17.0	308	114	1.9	15.9	308	27%	27
1.4	19.0	310	96	1.9	15.1	304	26%	105
1.4	18.8	302	42	1.9	18.7	308	26%	219
1.4	18.9	308	48	1.8	19.2	310	26%	147
1.4	18.8	303	42	1.8	15.1	306	26%	27

Tabla 5.8: Estadística de alturas de ola identificadas - swell remoto NW

Swell remoto NW						
Variable	Método	Mín. (m)	Máx. (m)	Promedio	Sesgo	RMSE
Hm ₀ (m)	POT	1.4	2.6	1.8	0.3	0.37
	Separación de Swells	1.8	3.0	2.1		
Tp (s)	POT	13.5	20.9	17.5	0.9	2.63
	Separación de Swells	11.6	21.1	16.6		
Dpk (°)	POT	300.0	311.0	306.9	0.0	5.40
	Separación de Swells	297.0	312.0	306.9		
Duración (horas)	POT	21.0	309.0	105.6	17.7	106.71
	Separación de Swells	18.0	252.0	87.9		

A partir del análisis de las tablas previamente expuestas, se puede apreciar que para el caso *swell* cercano SW se han detectado diferencias menores en las alturas *peak*. Además, para los eventos en los cuales existe equivalencia de magnitud, existen variaciones no significativas en periodos y direcciones. Así mismo, de la estadística de los resultados se desprende que los datos presentan un sesgo (*bias*) de +0.1 m hacia los valores detectados por el método de agrupación de *swells* con error cuadrático medio de 0.07 entre ambas listas.

En el caso *swell* remoto SW, se aprecia que para los mayores valores de altura, los parámetros de resumen asociados coinciden en su totalidad con la utilización de ambos métodos, no obstante, dentro de los valores restantes se puede encontrar una diferencia máxima de 2% en las alturas. Entre los datos existe un sesgo (*bias*) de +0.1 m hacia los valores identificados con la agrupación de *swells*, presentando un error cuadrático medio de 0.04.

Para el sistema de oleaje *swell* cercano NW E, se puede observar que existen diferencias en un 80% de los datos de alturas de olas con valores que varían en hasta un 10%. El rango de incidencia de las direcciones se encuentra entre los 286° y 3°, con periodos característicos que van desde los 8.3 s hasta los 10.9 s. Entre estos datos de alturas de ola existe un sesgo de +0.2 m hacia los valores detectados por el método de agrupación de *swells* y un error medio cuadrático de 0.2. En este caso se aprecia una diferencia de 0.4 m para los valores mínimos de ambas listas.

Finalmente, en las tormentas del caso *swell* remoto NW, se observan variaciones de un 2% a un 27% en los valores de alturas de ola identificados por el método de separación de *swells*, encontrando diferencias en el 100% de los datos identificados respecto del método POT. En este caso se presenta un sesgo (*bias*) de +0.3 m hacia los datos detectados por agrupación de *swells*. El error cuadrático medio para los valores en este sistema de oleaje alcanza un valor de 0.37. Las diferencias para los valores mínimos y máximos de ambas listas es de 0.4 m.

En general, las diferencias entre las alturas, periodos y direcciones no son significativas entre ambos métodos. En el caso de las alturas mayores de las listas, las diferencias tienden a ser nulas. En el caso de las alturas menores, las diferencias aumentan con un sesgo hacia los datos detectados por agrupación de *swells*.

Las diferencias mayores se producen en las duraciones, ya que ambos métodos la consideran de manera muy diferente. En general, el rango de duraciones obtenidas con la agrupación de *swells* es mayor al obtenido por medio del POT.

5.3 RESULTADOS DE LA MODELACIÓN CONJUNTA DE PARÁMETROS DE OLEAJE

A continuación se presentan los resultados de la modelación conjunta de parámetros de oleaje. Las simulaciones han sido realizadas considerando la variable altura máxima como variable condicionante, mientras que las demás fueron consideradas como variables condicionadas.

5.3.1 DEFINICIÓN DE UMBRAL PARA SELECCIÓN DE DATOS

Dentro del proceso de simulación de valores extremos se ha considerado el cálculo de estabilidad de los parámetros de escala y forma para distintos porcentajes del total de datos disponibles (umbral), con el fin de caracterizar todas las variables condicionadas en función de un umbral característico. Este método de estabilidad de umbral es un criterio gráfico a partir del cual se escogen valores donde los parámetros se mantienen aproximadamente constantes para la mayoría de las variables a analizar.

Esta técnica de selección de umbral es ampliamente utilizada para este tipo de análisis (Jonathan & Ewans, 2010). Cabe mencionar que la selección de un umbral apropiado es un proceso complejo cuando se realizan simulaciones de valores extremos, debido a que la selección debe tener en cuenta que debe ser suficientemente alto para que la aproximación asintótica de la cola extrema sea buena, pero suficientemente bajo para que exista la información suficiente para la estimación de los parámetros de escala, forma y posición, es decir, suficientes observaciones sobre el umbral.

Un ejemplo de gráficos de estabilidad de umbral se muestra en Figura 5.5. En ellos se puede observar que para todos los sistemas de oleaje las curvas se mantienen relativamente constantes para los distintos niveles de umbral, o bien se encuentran variaciones menores de los parámetros. No obstante, se aprecia que a medida que los niveles se acercan al máximo, los valores de escala y forma comienzan a ser inestables de manera que no se recomienda una elección demasiado elevada para las simulaciones. En el caso de las duraciones, los parámetros son un poco más inestables debido al amplio rango de valores que se encuentran.

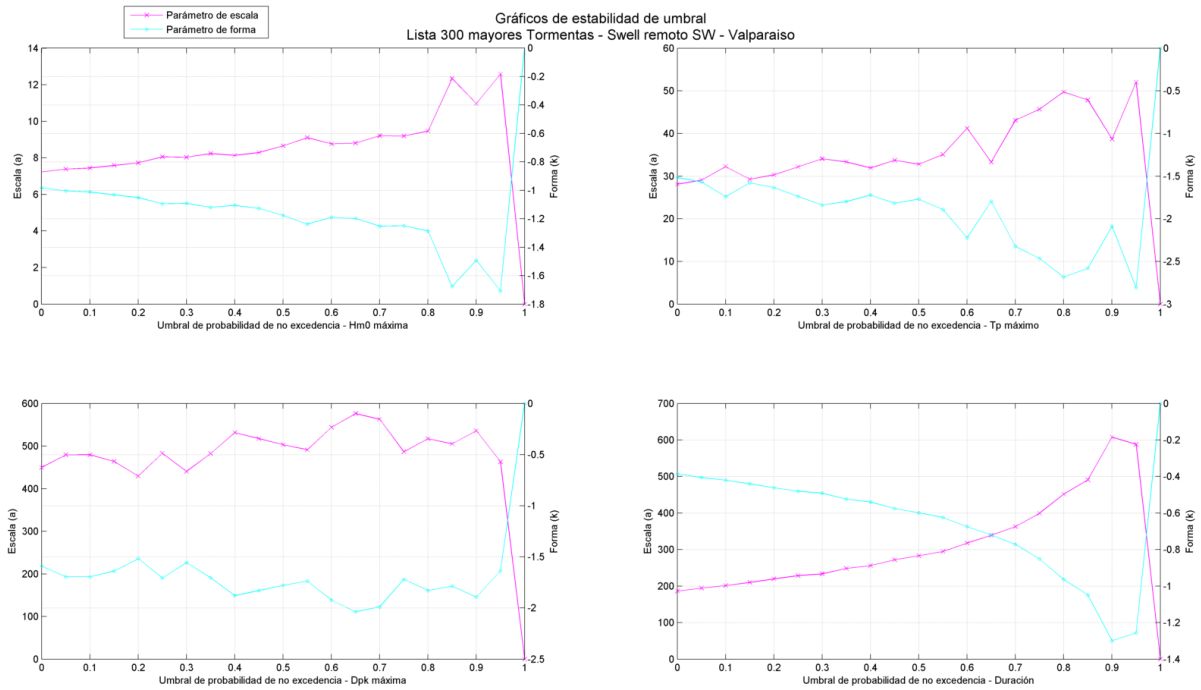


Figura 5.5: Gráficos de estabilidad de parámetros – Swell remoto SW, Zona central

Fuente: Elaboración propia

Los umbrales seleccionados para las variables por cada sistema de oleaje se presentan en las Tabla 5.9. Cabe mencionar que esta cota es considerada para todas las variables a modelar por sistema de oleaje, utilizándolos en la estimación de los niveles de retorno de las variables condicionantes y en las modelaciones multivariadas de las variables condicionadas.

Tabla 5.9: Umbrales seleccionados para simulaciones

Nodo	Caso	Nivel de umbral (%)	Cantidad de datos
Zona norte	<i>Swell</i> cercano SW	0.60	118
	<i>Swell</i> remoto SW	0.60	117
	<i>Swell</i> cercano NE SE	0.50	147
	<i>Swell</i> remoto NW	0.60	118
Zona centro	<i>Swell</i> cercano SW	0.70	90
	<i>Swell</i> remoto SW	0.70	89
	<i>Swell</i> cercano NW E	0.60	119
	<i>Swell</i> remoto NW	0.60	120
Zona sur	<i>Swell</i> cercano NE SW	0.50	150
	<i>Swell</i> SW NW	0.50	150
	<i>Swell</i> remoto NW	0.60	117

5.3.2 CÁLCULO DE ALTURAS DE OLEAJE EXTREMO

Utilizando las cantidades de datos ya expuestos, se han determinado las alturas de oleaje extremo para periodos de retorno de 25, 50 y 100 años por cada sistema de oleaje relativos a cada zona de estudio. La Tabla 5.10 presenta estos valores.

El ajuste probabilístico para la determinación de altura de olas extremas ha sido efectuado con el método de máxima verosimilitud (MLE, *Maximum Likelihood Estimation*) utilizando la distribución de valores extremos GPD (*Generalized Pareto Distribution*, por sus siglas en inglés). Este método estima de manera automática los parámetros para cada distribución candidata, en este caso GPD, la cual tiene la característica de ser triparamétrica, permitiendo un mejor ajuste de los datos en sus colas extremas.

Tabla 5.10: Valores de retorno de altura de oleaje en distintas zonas

Nodo	Caso	Periodo de retorno (años)	H _{tr} (m) - Ajuste promedio	H _{tr} (m) - Ajuste superior
Zona norte	Swell cercano SW	25	2.7	3.0
		50	2.8	3.3
		100	3.0	3.5
	Swell remoto SW	25	4.3	4.5
		50	4.4	4.7
		100	4.5	4.8
	Swell cercano NE SE	25	1.7	2.0
		50	7.8	2.3
		100	2.0	2.6
	Swell remoto NW	25	1.5	1.8
		50	1.7	2.0
		100	1.8	2.3
Zona centro	Swell cercano SW	25	3.9	4.1
		50	4.0	4.2
		100	4.0	4.3
	Swell remoto SW	25	7.1	7.9
		50	7.5	8.7
		100	7.9	9.5
	Swell cercano NW E	25	5.7	6.0
		50	5.8	6.2
		100	6.0	6.4
	Swell remoto NW	25	2.7	3.0
		50	2.8	3.2
		100	2.9	3.4
Zona sur	Swell cercano NE SW	25	10.1	11.6
		50	10.9	12.9
		100	11.7	14.2
	Swell SW NW	25	13.9	15.1
		50	14.5	16.0
		100	15.1	17.0
	Swell remoto NW	25	7.7	9.8
		50	8.8	11.7
		100	9.9	13.9

5.3.3 CÁLCULO DE PARÁMETROS DEL MODELO MULTIVARIADO

A continuación se calcularon los parámetros del modelo multivariado de acuerdo a la metodología descrita en la sección 4.3.5.

Las simulaciones conjuntas multivariadas se han llevado a cabo considerando el percentil 80% de los datos reales como valor mínimo de extrapolación para todas las series de sistemas de oleaje consideradas en el estudio.

Las simulaciones de Montecarlo con el modelo condicional permiten generar suficientes pares ordenados de las variables analizadas para ajustar una función de densidad de probabilidad.

Para conocer la densidad de probabilidad para las variables simuladas, se ha ajustado una distribución *kernel* bidimensional a todos los pares de variable condicionante vs variables condicionadas simuladas. Este proceso permite estimar dónde se encuentran las mayores concentraciones de pares simulados y cuantificar su densidad de probabilidad dentro del espacio de simulación.

Finalmente, utilizando los valores de alturas de olas de retorno de 25, 50 y 100 años, éstas fueron ingresadas sobre el espacio de simulación *kernel* y se extrajeron secciones o *slices* de la función bidimensional de probabilidad a lo largo de las variables condicionadas. Una vez obtenidas estas densidades unidimensionales, los valores se transformaron a probabilidades acumuladas, logrando obtener curvas de no excedencias para cada variable simulada en función del valor de retorno de la variable condicionante. A partir de esto último es que se estimaron valores extremos bivariados con distinta probabilidad de ocurrencias de manera simultánea.

5.3.4 RESULTADOS DE MODELACIONES CONDICIONALES MULTIVARIADAS

A continuación se presentan los resultados de las modelaciones condicionales multivariadas correspondientes al sistema de oleaje *swell* remoto SW para el nodo de Valparaíso. Las Figuras 5.6 (A) a 5.8 (A) muestran los gráficos de dispersión de parámetros identificados y valores simulados. De acuerdo a los resultados graficados, se aprecia que existe concordancia en las simulaciones, ya que estas presentan un comportamiento similar al de los valores reales en su dispersión, exhibiendo incluso la misma forma pero extrapolada, representando también donde se encuentran las mayores concentraciones de datos.

La interpretación de estos gráficos antes mencionados se basa sobre la disminución o aumento de la dispersión de las variables dependientes (T_p , D_{pk} y duración) al aumentar el valor de la variable independiente (altura de ola) para distintos niveles de retorno. De esta manera, se aprecia que los periodos simulados tienen una tendencia a la disminución de su dispersión a medida que la variable condicionante aumenta (forma cónica). Por su parte, las direcciones y duraciones muestran una tendencia más neutra, aunque igualmente hay una reducción en el rango de valores probables a medida que aumenta el periodo de retorno.

Así mismo, las Figuras 5.12 (B) a 5.14 (B) presentan los resultados de las distribuciones *kernel* bidimensionales sobre los valores simulados. En estas figuras también se grafican las alturas asociadas a los periodos de retorno de 25, 50 y 100 años de manera de indicar las secciones sobre las cuales se efectuaron los *slíces* sobre las densidades bidimensionales de cada variable simulada. Por su parte, las Figuras 5.12 (C) a 5.14 (C) muestran las curvas de densidad de probabilidad en función de los datos simulados para cada altura de oleaje asociada a los periodos de retorno ya indicados. De ellas se puede destacar que los valores simulados presentan valores similares para los distintos niveles de retorno de la variable condicionante.

Finalmente, el resumen de los resultados asociados a las alturas de ola para 25, 50 y 100 años de periodo de retorno para cada variable simulada son presentados en las Tabla 5.11 a 5.13, en las cuales es posible apreciar que la ola de 100 años de periodo de retorno con ajuste superior es de 9.5 m, con un periodo que podría estar entre 16.6 s y 18.1 s, una dirección que podría estar entre 254° y 267° y una duración que puede ser entre las 48 horas y 169 horas con un 90% de probabilidad (entre los percentiles 5 y 95). Al reducir la probabilidad, se podrían obtener rangos más acotados para el periodo, dirección y duración a partir de las Figuras 5.6 a 5.8. Los valores a utilizar en el diseño van a estar condicionados por las características particulares en cada sitio, pero con los resultados generados con las herramientas aquí presentadas es posible conocer las probabilidades de excedencia o no excedencia de los parámetros y así poder estimar el riesgo de adoptar un determinado valor o rango de valores.

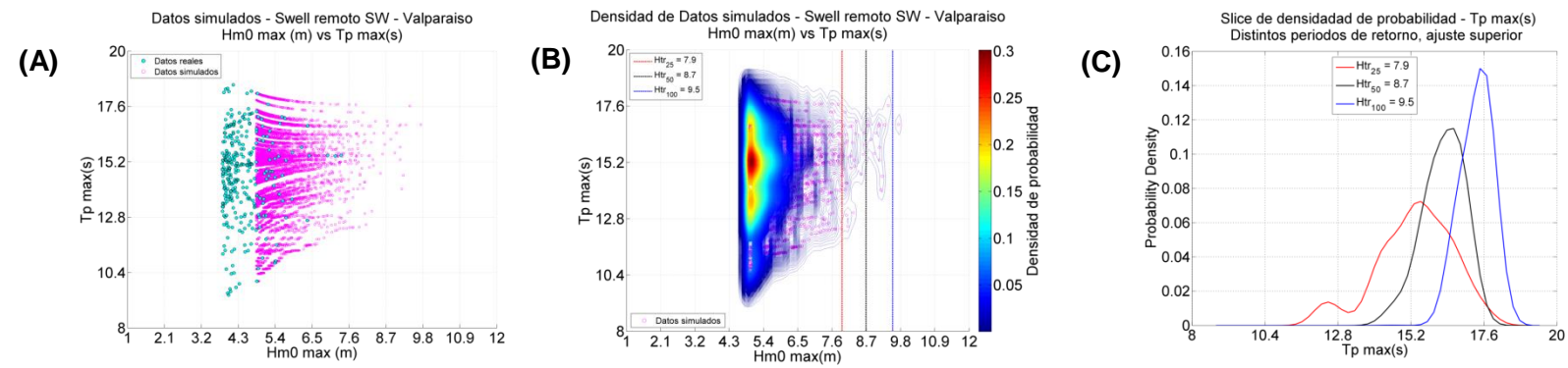


Figura 5.6: Resultados simulación multivariada Hm_0 vs Tp , *Swell* remoto SW – Zona central

Fuente: Elaboración propia

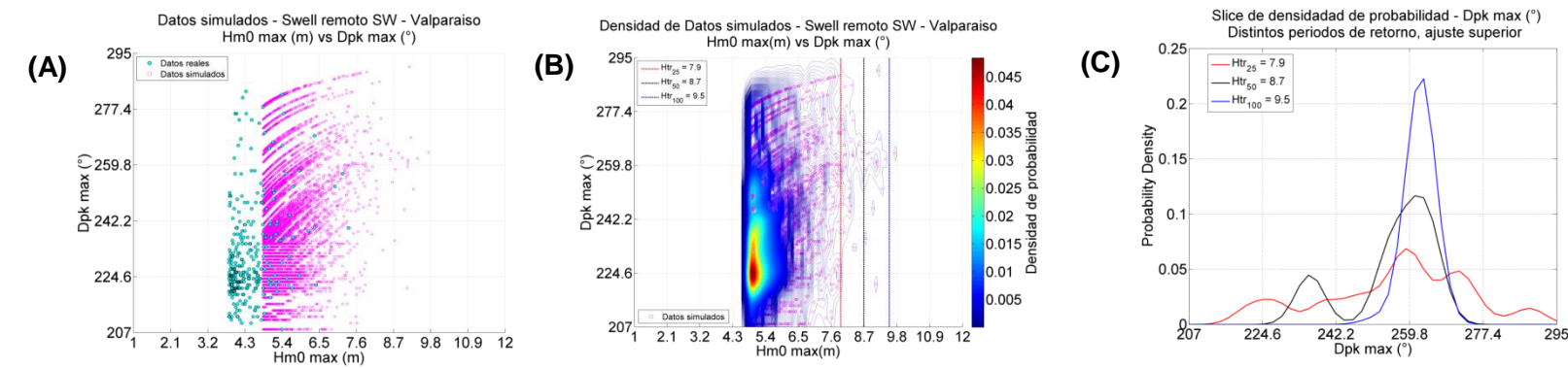


Figura 5.7: Resultados simulación multivariada Hm_0 vs Dpk , *Swell* remoto SW – Zona central

Fuente: Elaboración propia

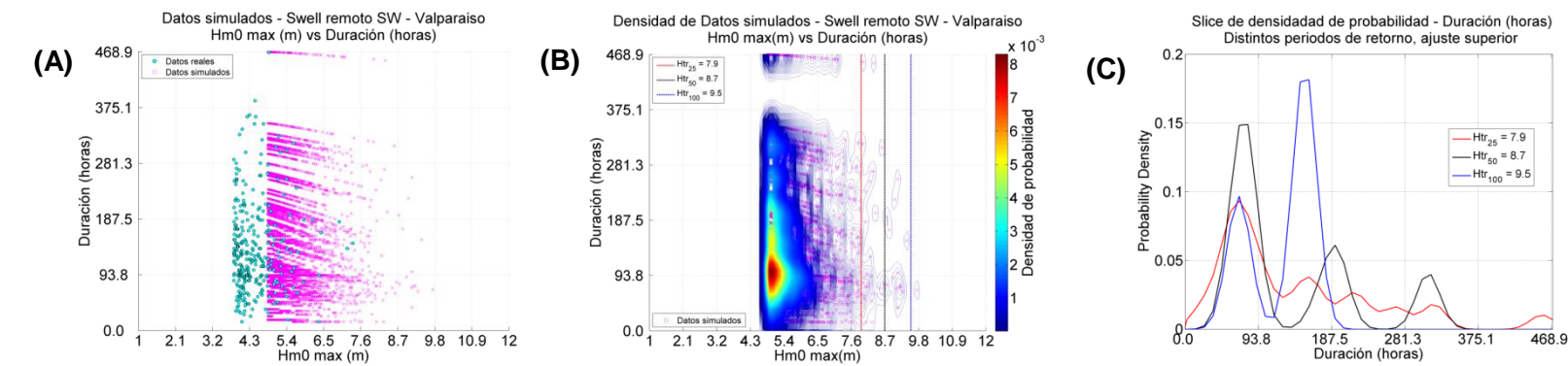


Figura 5.8: Resultados simulación multivariada Hm_0 vs Duración, *Swell* remoto SW – Zona central

Fuente: Elaboración propia

Tabla 5.11: Tabla de percentiles de periodos, *Swell* remoto SW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	7.9	5	12.6
		50	15.3
		95	17.2
50	8.7	5	14.8
		50	16.2
		95	17.2
100	9.5	5	16.6
		50	17.3
		95	18.1

Tabla 5.12: Tabla de percentiles de direcciones, *Swell* remoto SW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	7.9	5	221.9
		50	257.5
		95	284.5
50	8.7	5	232.6
		50	257.6
		95	267.2
100	9.5	5	254.6
		50	261.3
		95	267.3

Tabla 5.13: Tabla de percentiles de duraciones, *Swell* remoto SW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	7.9	5	25.0
		50	97.4
		95	324.3
50	8.7	5	44.2
		50	85.3
		95	311.0
100	9.5	5	47.5
		50	137.1
		95	169.2

5.3.5 RESULTADOS DE LAS MODELACIONES CONDICIONALES MULTIVARIADAS: ZONAS NORTE, CENTRO Y SUR DE CHILE

En esta sección se presentan los resultados de todos los nodos utilizados en el análisis multivariado de valores extremos. Aquí se exponen los gráficos correspondientes a las dispersiones de los datos reales y simulados, las distribuciones *kernel* bidimensionales sobre las simulaciones y también las curvas de densidad de probabilidad de las variables condicionadas asociadas a las alturas de oleaje con periodos de retorno de 25, 50 y 100 años. Además, se presentan las tablas con los valores de los periodos, direcciones y duraciones para distintos niveles de probabilidad (percentiles 5, 50 y 95) para las alturas de oleaje extremo asociadas a los niveles de retorno ya mencionadas. La Tabla 5.14 contiene la descripción de los gráficos y tablas con los resultados obtenidos.

Tabla 5.14: Descripción de tablas y gráficos asociados a los resultados obtenidos

REPRESENTACIÓN	Variables	Zona norte				Zona centro				Zona sur		
		<i>Swell cercano SW</i>	<i>Swell remoto SW</i>	<i>Swell cercano NE SE</i>	<i>Swell remoto NW</i>	<i>Swell cercano SW</i>	<i>Swell remoto SW</i>	<i>Swell cercano NE E</i>	<i>Swell remoto NW</i>	<i>Swell cercano NE SW</i>	<i>Swell SW NW</i>	<i>Swell remoto NW</i>
Dispersión de datos reales y simulados	Tp (s)	Figura 5.9 (A)	Figura 5.12 (A)	Figura 5.15 (A)	Figura 5.18 (A)	Figura 5.21 (A)	Figura 5.24 (A)	Figura 5.27 (A)	Figura 5.30 (A)	Figura 5.33 (A)	Figura 5.36 (A)	Figura 5.39 (A)
	DPK (°)	Figura 5.10 (A)	Figura 5.13 (A)	Figura 5.16 (A)	Figura 5.19 (A)	Figura 5.22 (A)	Figura 5.25 (A)	Figura 5.28 (A)	Figura 5.31 (A)	Figura 5.34 (A)	Figura 5.37 (A)	Figura 5.40 (A)
	Duración (horas)	Figura 5.11 (A)	Figura 5.14 (A)	Figura 5.17 (A)	Figura 5.20 (A)	Figura 5.23 (A)	Figura 5.26 (A)	Figura 5.29 (A)	Figura 5.32 (A)	Figura 5.35 (A)	Figura 5.38 (A)	Figura 5.41 (A)
Densidad de probabilidad bidimensional	Tp (s)	Figura 5.9 (B)	Figura 5.12 (B)	Figura 5.15 (B)	Figura 5.18 (B)	Figura 5.21 (B)	Figura 5.24 (B)	Figura 5.27 (B)	Figura 5.30 (B)	Figura 5.33 (B)	Figura 5.36 (B)	Figura 5.39 (B)
	DPK (°)	Figura 5.10 (B)	Figura 5.13 (B)	Figura 5.16 (B)	Figura 5.19 (B)	Figura 5.22 (B)	Figura 5.25 (B)	Figura 5.28 (B)	Figura 5.31 (B)	Figura 5.34 (B)	Figura 5.37 (B)	Figura 5.40 (B)
	Duración (horas)	Figura 5.11 (B)	Figura 5.14 (B)	Figura 5.17 (B)	Figura 5.20 (B)	Figura 5.23 (B)	Figura 5.26 (B)	Figura 5.29 (B)	Figura 5.32 (B)	Figura 5.35 (B)	Figura 5.38 (B)	Figura 5.41 (B)
Slice de densidad de probabilidad	Tp (s)	Figura 5.9 (C)	Figura 5.12 (C)	Figura 5.15 (C)	Figura 5.18 (C)	Figura 5.21 (C)	Figura 5.24 (C)	Figura 5.27 (C)	Figura 5.30 (C)	Figura 5.33 (C)	Figura 5.36 (C)	Figura 5.39 (C)
	DPK (°)	Figura 5.10 (C)	Figura 5.13 (C)	Figura 5.16 (C)	Figura 5.19 (C)	Figura 5.22 (C)	Figura 5.25 (C)	Figura 5.28 (C)	Figura 5.31 (C)	Figura 5.34 (C)	Figura 5.37 (C)	Figura 5.40 (C)
	Duración (horas)	Figura 5.11 (C)	Figura 5.14 (C)	Figura 5.17 (C)	Figura 5.20 (C)	Figura 5.23 (C)	Figura 5.26 (C)	Figura 5.29 (C)	Figura 5.32 (C)	Figura 5.35 (C)	Figura 5.38 (C)	Figura 5.41 (C)
Tabla de variable extrema simulada – Distintos niveles de retorno	Tp (s)	Tabla 5.15	Tabla 5.18	Tabla 5.21	Tabla 5.24	Tabla 5.27	Tabla 5.30	Tabla 5.33	Tabla 5.36	Tabla 5.39	Tabla 5.42	Tabla 5.45
	DPK (°)	Tabla 5.16	Tabla 5.19	Tabla 5.22	Tabla 5.25	Tabla 5.28	Tabla 5.31	Tabla 5.34	Tabla 5.37	Tabla 5.40	Tabla 5.43	Tabla 5.46
	Duración (horas)	Tabla 5.17	Tabla 5.20	Tabla 5.23	Tabla 5.26	Tabla 5.29	Tabla 5.32	Tabla 5.35	Tabla 5.38	Tabla 5.41	Tabla 5.44	Tabla 5.47

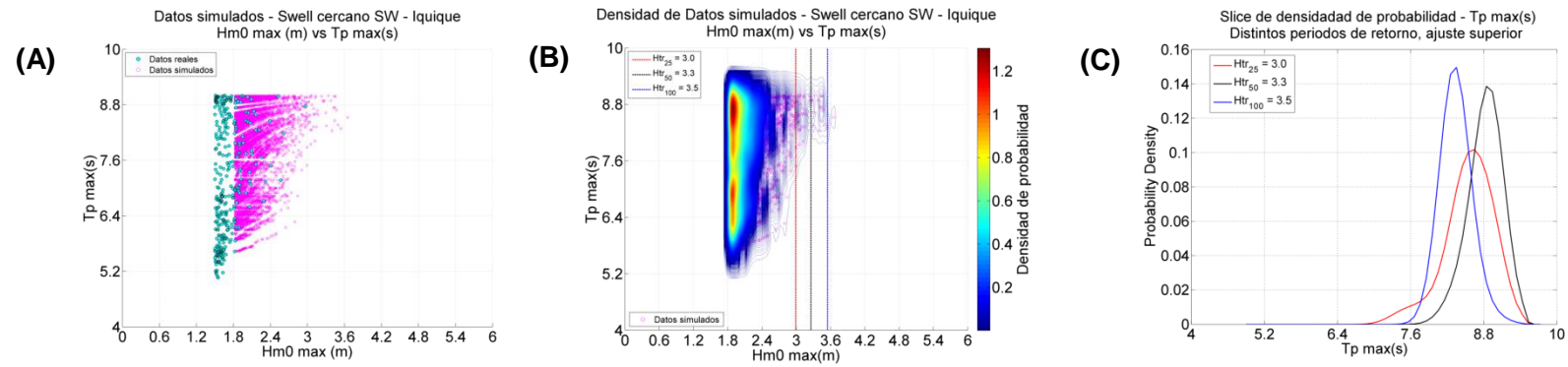


Figura 5.9: Resultados simulación multivariada H_{m0} vs T_p , *Swell* cercano SW – Zona norte

Fuente: Elaboración propia

Tabla 5.15: Tabla de percentiles de periodos, *Swell* cercano SW Zona norte. Ajuste superior

Periodos de retorno (años)	H_{m0} (m)	Percentiles (%)	Periodo (s)
25	3.0	5	7.7
		50	8.5
		95	9.1
50	3.3	5	8.2
		50	8.8
		95	9.2
100	3.5	5	7.9
		50	8.3
		95	8.8

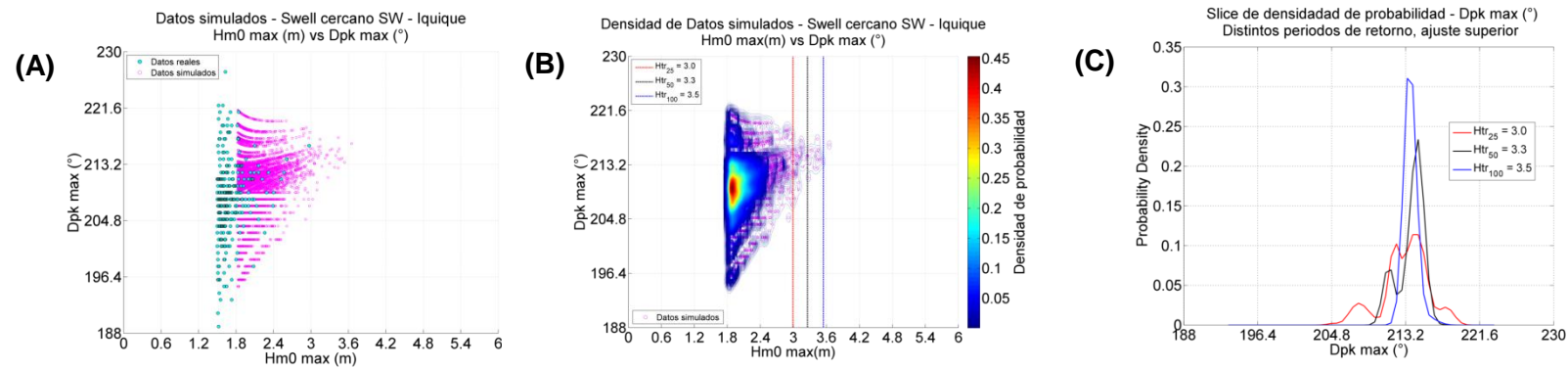


Figura 5.10: Resultados simulación multivariada H_{m0} vs D_{pk} , *Swell* cercano SW – Zona norte

Fuente: Elaboración propia

Tabla 5.16: Tabla de percentiles de direcciones, *Swell* cercano SW Zona norte. Ajuste superior

Periodos de retorno (años)	H_{m0} (m)	Percentiles (%)	Dirección (°)
25	3.0	5	207.3
		50	213.1
		95	217.2
50	3.3	5	210.5
		50	213.9
		95	215.5
100	3.5	5	212.2
		50	213.4
		95	214.8

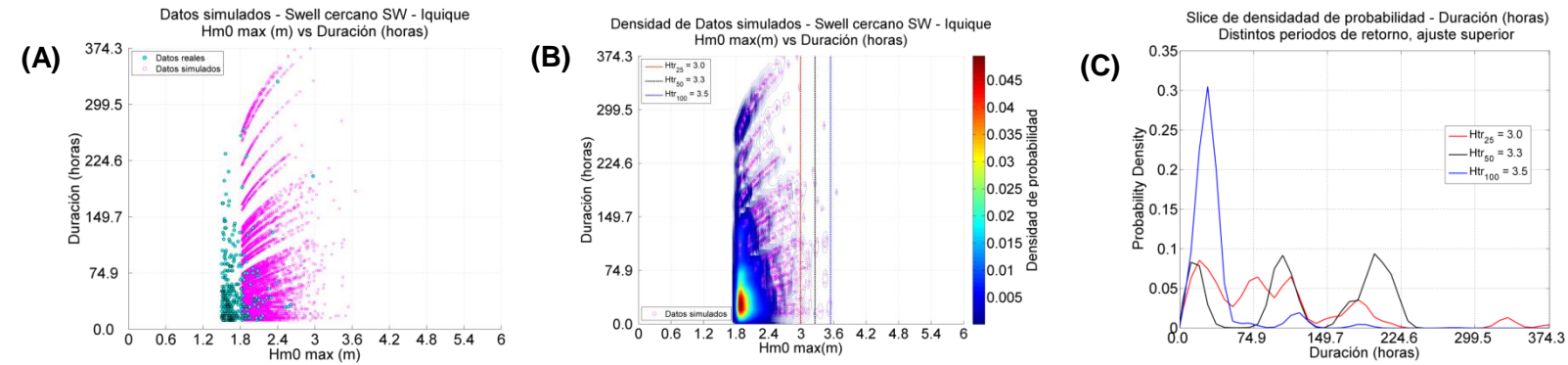


Figura 5.11: Resultados simulación multivariada H_{m0} vs Duración, *Swell* cercano SW – Zona norte

Fuente: Elaboración propia

Tabla 5.17: Tabla de percentiles de duraciones, *Swell* cercano SW Zona norte. Ajuste superior

Periodos de retorno (años)	H_{m0} (m)	Percentiles (%)	Duración (hr)
25	3.0	5	6.3
		50	74.8
		95	208.9
50	3.3	5	4.2
		50	112.2
		95	214.0
100	3.5	5	5.2
		50	24.2
		95	110.1

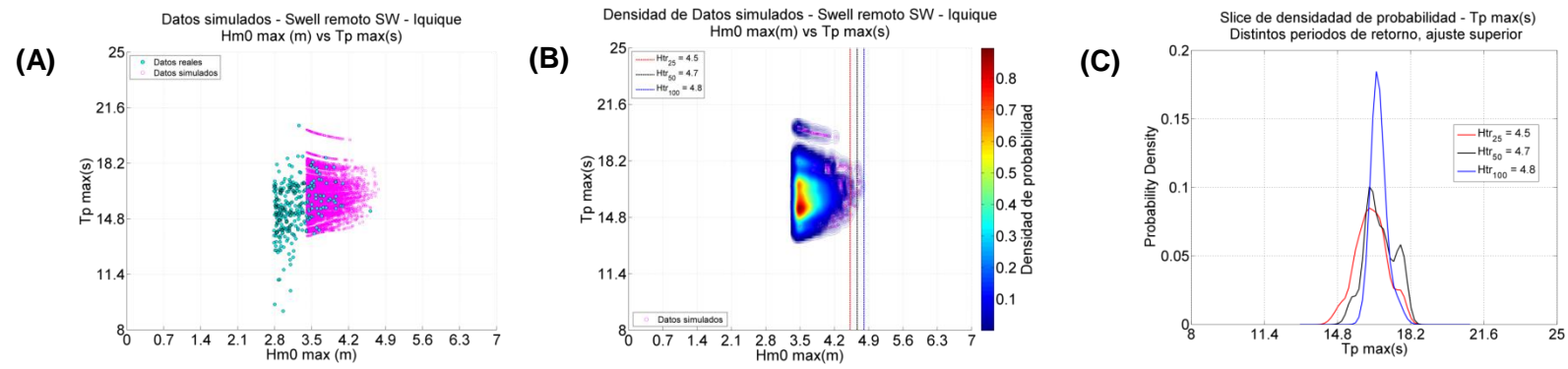


Figura 5.12: Resultados simulación multivariada Hm_0 vs Tp , Swell remoto SW – Zona norte

Fuente: Elaboración propia

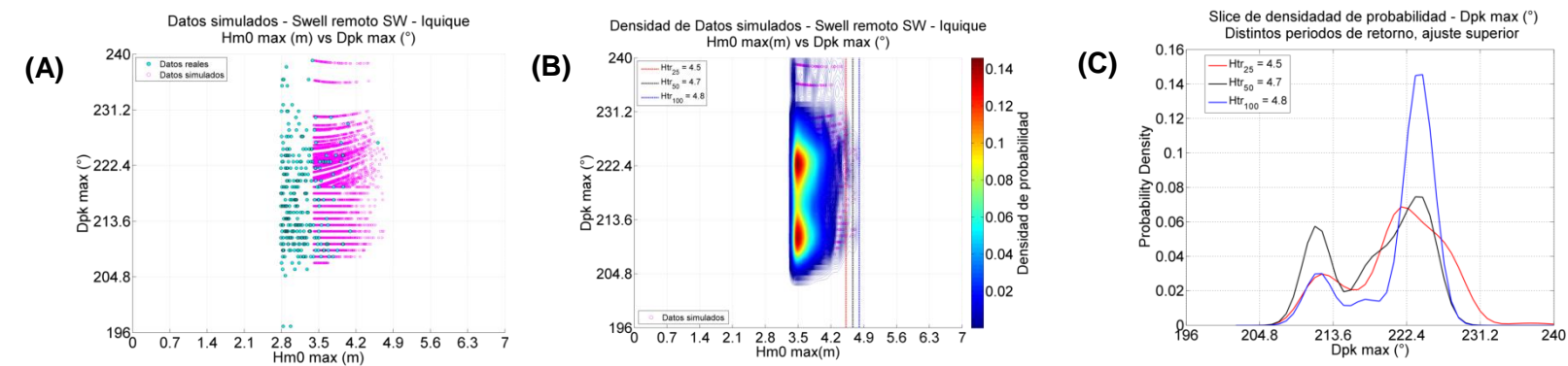


Figura 5.13: Resultados simulación multivariada Hm_0 vs Dpk , Swell remoto SW – Zona norte

Fuente: Elaboración propia

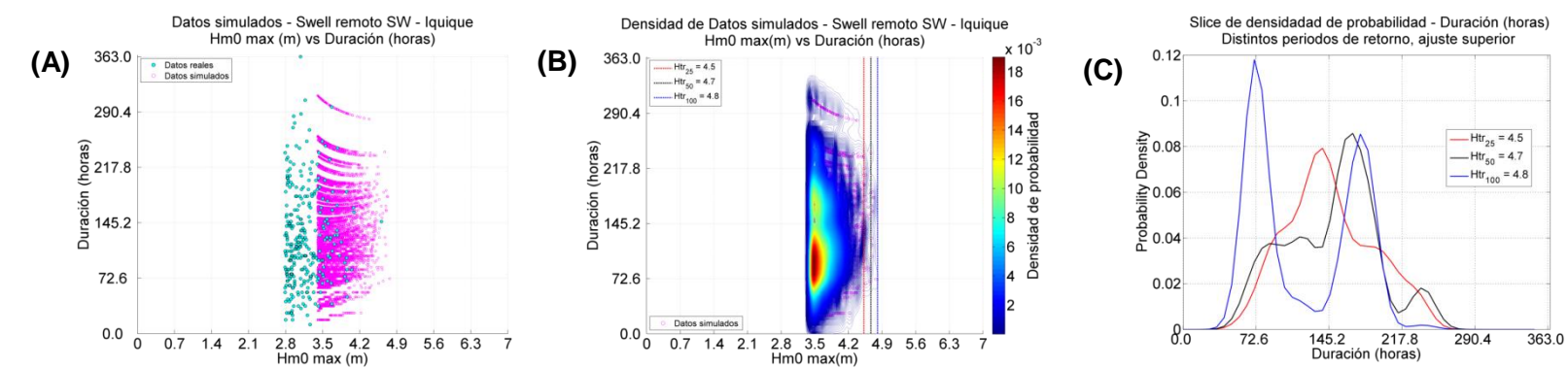


Figura 5.14: Resultados simulación multivariada Hm_0 vs Duración, Swell remoto SW – Zona norte

Fuente: Elaboración propia

Tabla 5.18: Tabla de percentiles de periodos, Swell remoto SW Zona norte. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	4.5	5	15.0
		50	16.3
		95	17.7
50	4.7	5	15.5
		50	16.6
		95	17.9
100	4.8	5	16.0
		50	16.6
		95	17.4

Tabla 5.19: Tabla de percentiles de direcciones, Swell remoto SW Zona norte. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	4.5	5	210.4
		50	221.5
		95	229.0
50	4.7	5	209.5
		50	219.7
		95	225.9
100	4.8	5	210.6
		50	222.8
		95	226.1

Tabla 5.20: Tabla de percentiles de duraciones, Swell remoto SW Zona norte. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	4.5	5	74.5
		50	137.7
		95	221.7
50	4.7	5	66.4
		50	151.4
		95	229.2
100	4.8	5	51.9
		50	96.4
		95	191.0

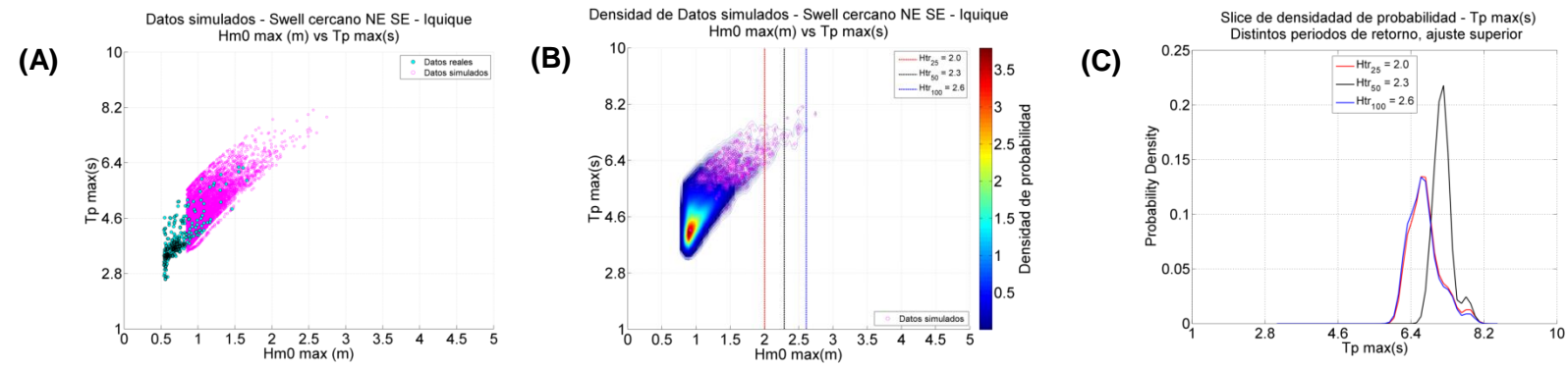


Figura 5.15: Resultados simulación multivariada Hm_0 vs Tp , Swell cercano NE SE – Zona norte

Fuente: Elaboración propia

Tabla 5.21: Tabla de percentiles de periodos, Swell cercano NE E Zona norte. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	2.0	5	6.1
		50	6.6
		95	7.5
50	2.3	5	6.6
		50	7.1
		95	7.7
100	2.6	5	6.1
		50	6.6
		95	7.4

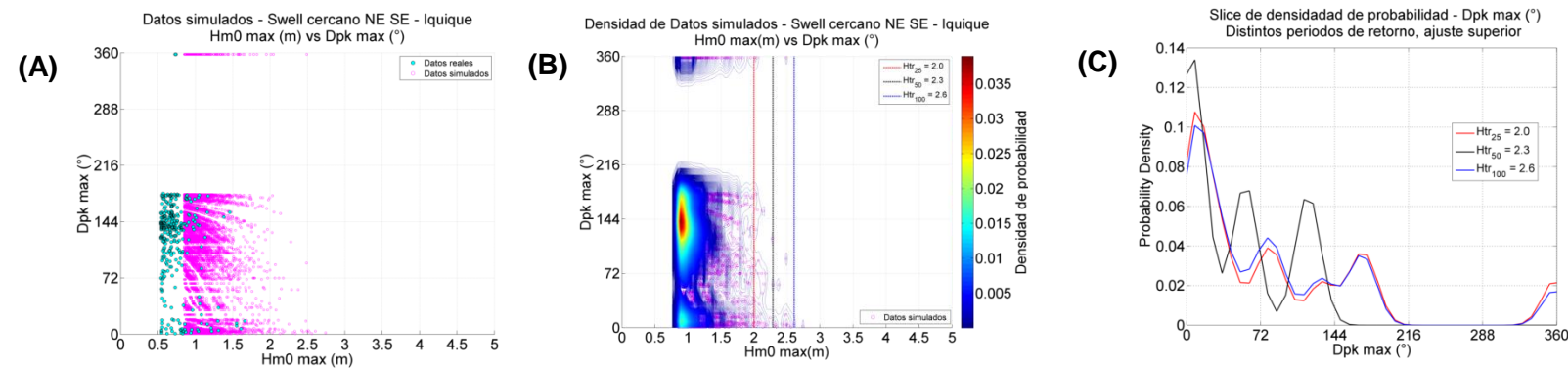


Figura 5.16: Resultados simulación multivariada Hm_0 vs Dpk , Swell cercano NE SE – Zona norte

Fuente: Elaboración propia

Tabla 5.22: Tabla de percentiles de direcciones, Swell cercano NE E Zona norte. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	2.0	5	360.0
		50	40.7
		95	348.8
50	2.3	5	360.0
		50	25.2
		95	123.3
100	2.6	5	360.0
		50	45.4
		95	342.3

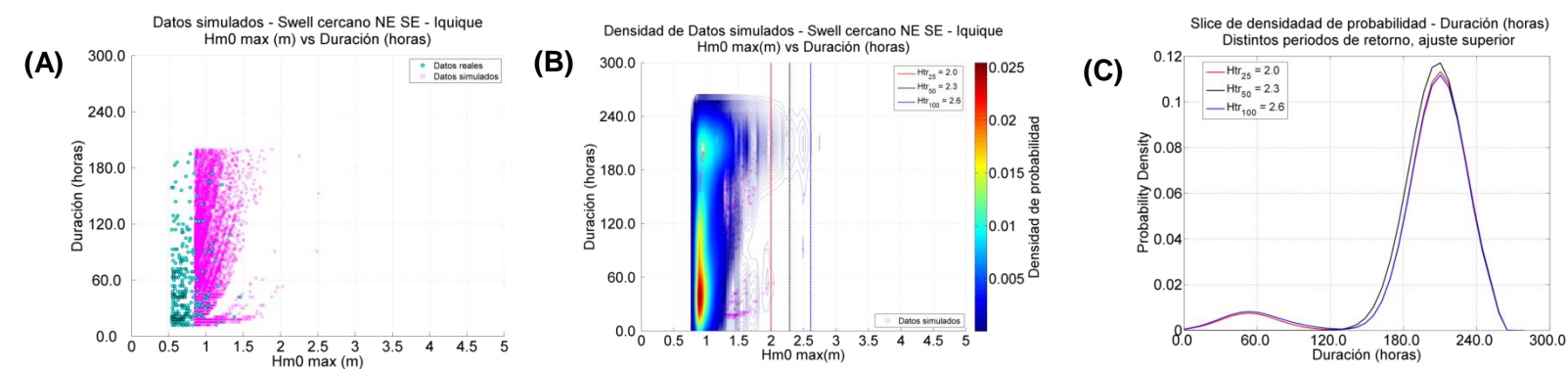


Figura 5.17: Resultados simulación multivariada Hm_0 vs Duración, Swell cercano NE SE – Zona norte

Fuente: Elaboración propia

Tabla 5.23: Tabla de percentiles de duraciones, Swell cercano NE E Zona norte. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	2.0	5	67.4
		50	204.1
		95	240.8
50	2.3	5	164.8
		50	204.0
		95	240.4
100	2.6	5	62.0
		50	203.7
		95	240.6

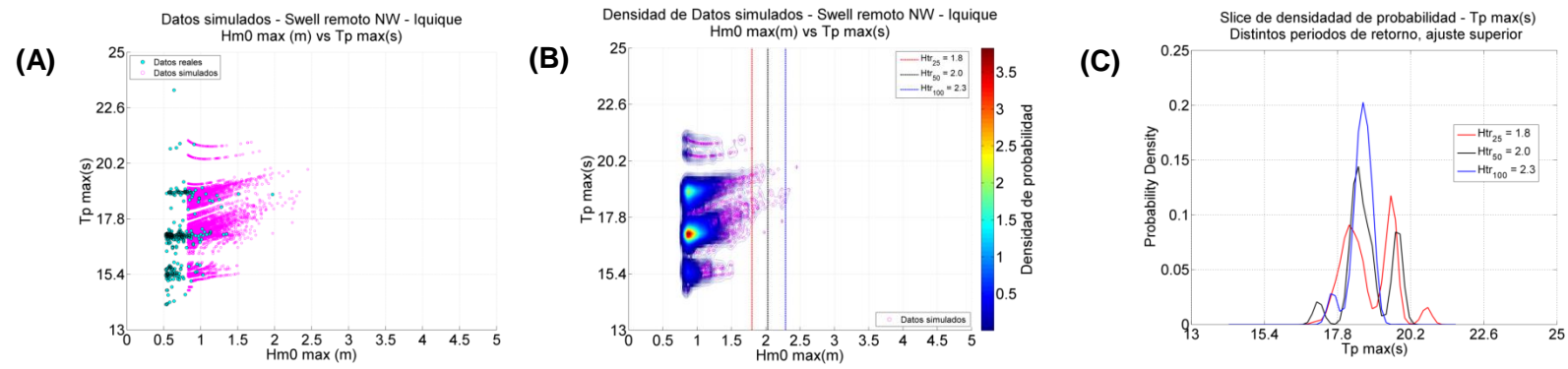


Figura 5.18: Resultados simulación multivariada Hm_0 vs Tp , *Swell* remoto NW – Zona norte

Fuente: Elaboración propia

Tabla 5.24: Tabla de percentiles de periodos, *Swell* remoto NW Zona norte. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	1.8	5	17.6
		50	18.6
		95	19.8
50	2.0	5	17.3
		50	18.6
		95	19.8
100	2.3	5	17.6
		50	18.5
		95	19

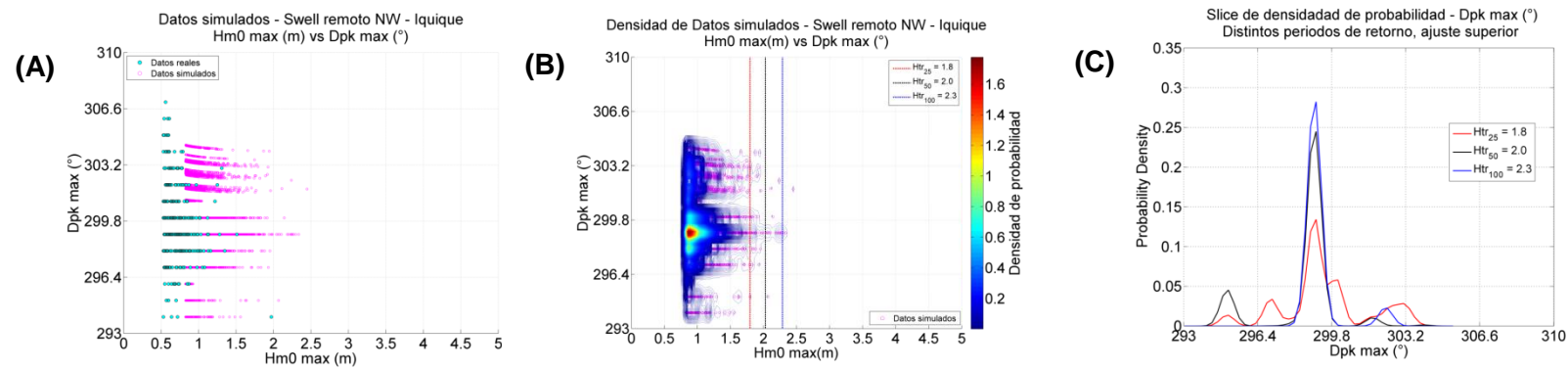


Figura 5.19: Resultados simulación multivariada Hm_0 vs Dpk , *Swell* remoto NW – Zona norte

Fuente: Elaboración propia

Tabla 5.25: Tabla de percentiles de direcciones, *Swell* remoto NW Zona norte. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	1.8	5	296.4
		50	299.1
		95	303.0
50	2.0	5	294.7
		50	298.8
		95	299.6
100	2.3	5	298.4
		50	298.9
		95	302.0

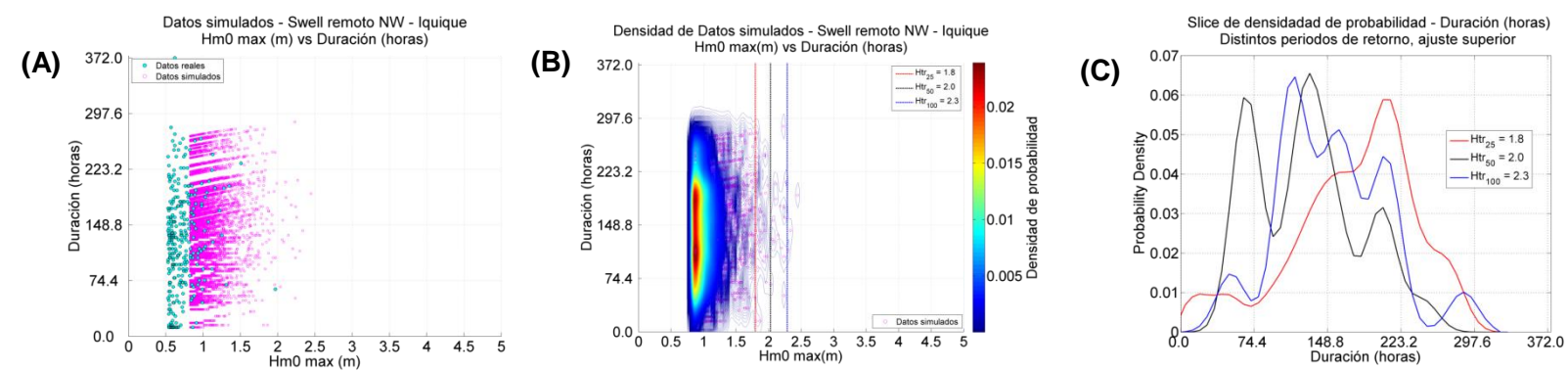


Figura 5.20: Resultados simulación multivariada Hm_0 vs Duración, *Swell* remoto NW – Zona norte

Fuente: Elaboración propia

Tabla 5.26: Tabla de percentiles de direcciones, *Swell* remoto NW Zona norte. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	1.8	5	33.2
		50	181.0
		95	265.2
50	2.0	5	47.9
		50	145.4
		95	220.6
100	2.3	5	53.9
		50	145.4
		95	249.7

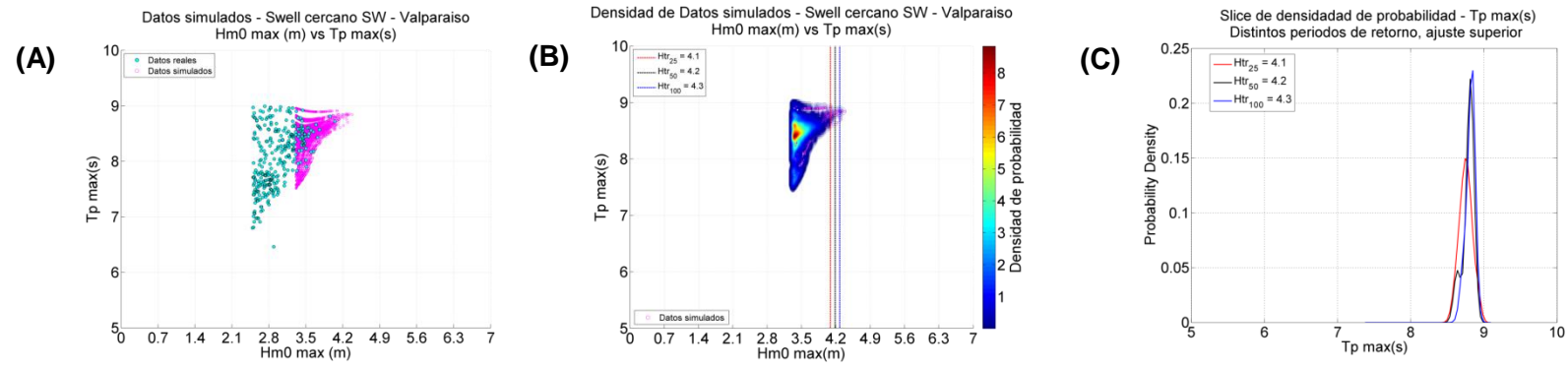


Figura 5.21: Resultados simulación multivariada Hm_0 vs Tp , *Swell* cercano SW – Zona central

Fuente: Elaboración propia

Tabla 5.27: Tabla de percentiles de periodos, *Swell* cercano SW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	4.1	5	8.6
		50	8.7
		95	8.9
50	4.2	5	8.6
		50	8.8
		95	8.9
100	4.3	5	8.7
		50	8.8
		95	8.9

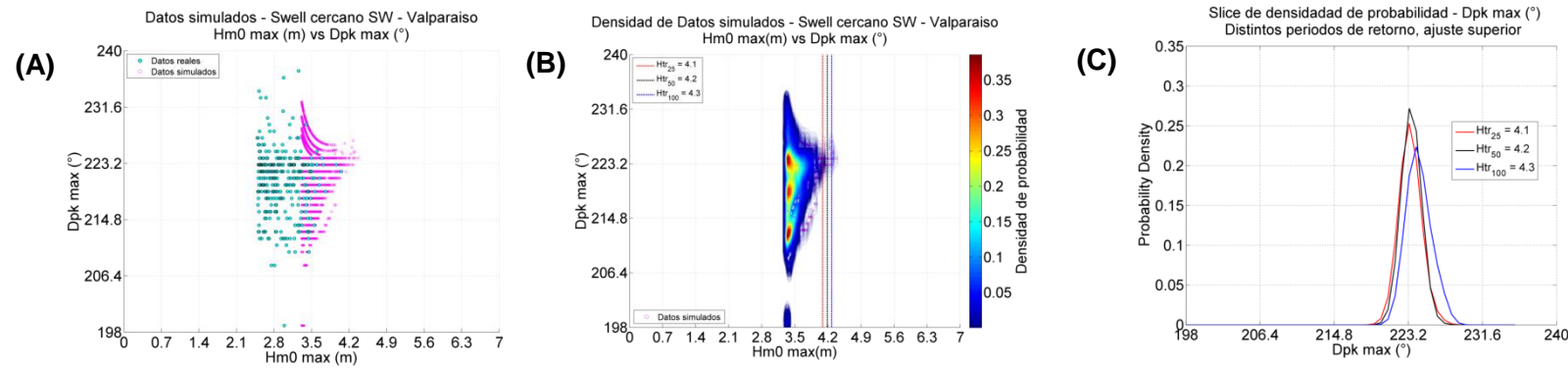


Figura 5.22: Resultados simulación multivariada Hm_0 vs Dpk , *Swell* cercano SW – Zona central

Fuente: Elaboración propia

Tabla 5.28: Tabla de percentiles de direcciones, *Swell* cercano SW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	4.1	5	220.9
		50	223.0
		95	225.3
50	4.2	5	221.2
		50	223.1
		95	225.1
100	4.3	5	221.7
		50	223.9
		95	226.6

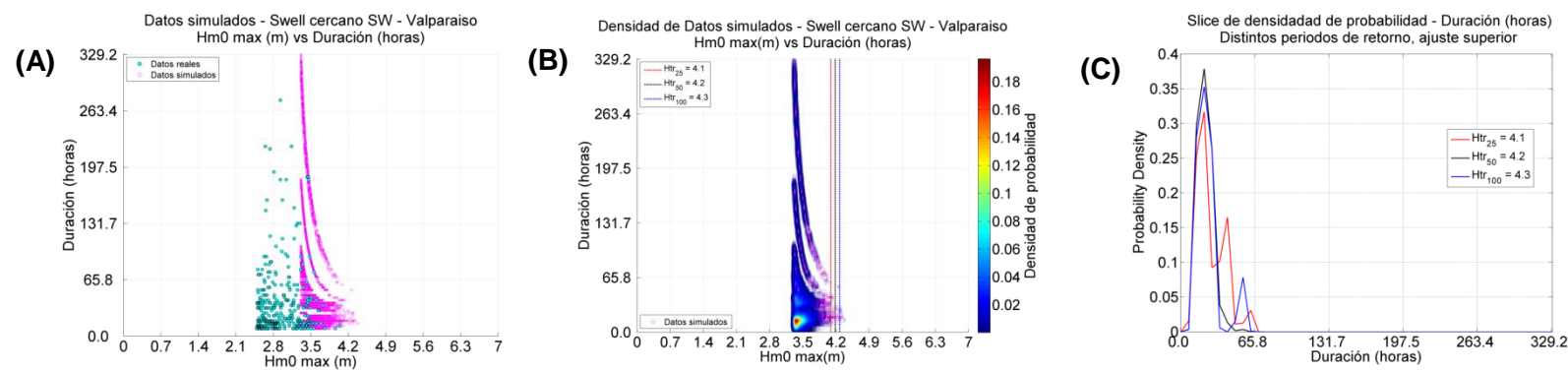


Figura 5.23: Resultados simulación multivariada Hm_0 vs Duración, *Swell* cercano SW – Zona central

Fuente: Elaboración propia

Tabla 5.29: Tabla de percentiles de duraciones, *Swell* cercano SW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	4.1	5	8.2
		50	19.2
		95	44.7
50	4.2	5	8.3
		50	17.8
		95	29.2
100	4.3	5	8.4
		50	18.4
		95	51.2

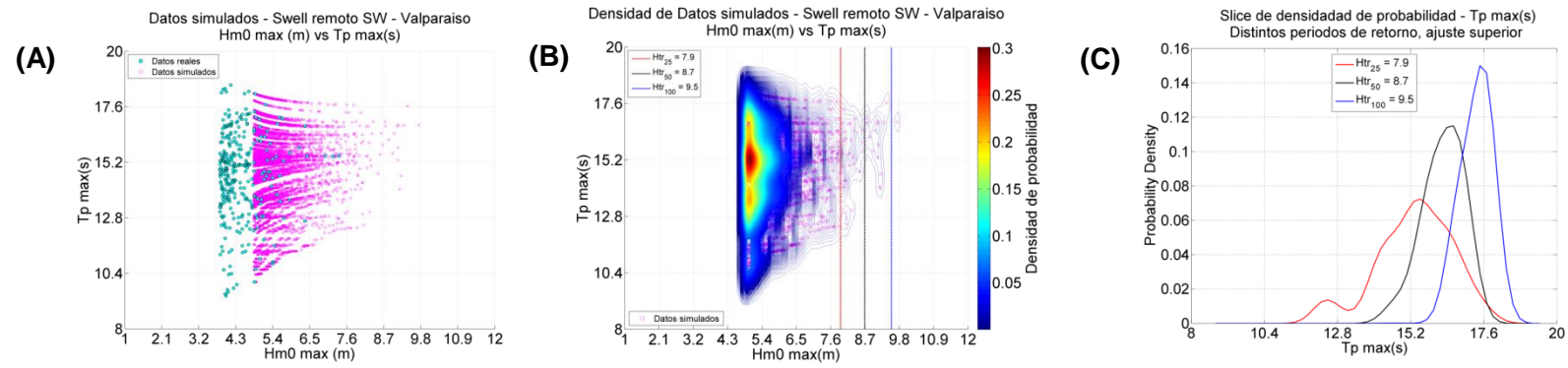


Figura 5.24: Resultados simulación multivariada Hm_0 vs T_p , *Swell* remoto SW – Zona central

Fuente: Elaboración propia

Tabla 5.30: Tabla de percentiles de periodos, *Swell* remoto SW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	7.9	5	12.6
		50	15.3
		95	17.2
50	8.7	5	14.8
		50	16.2
		95	17.2
100	9.5	5	16.6
		50	17.3
		95	18.1

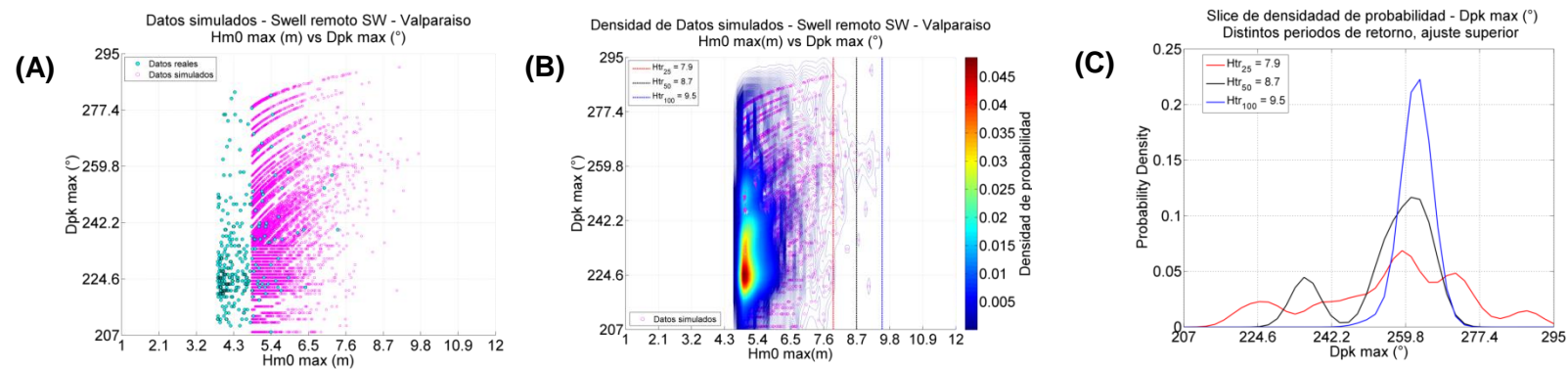


Figura 5.25: Resultados simulación multivariada Hm_0 vs D_{pk} , *Swell* remoto SW – Zona central

Fuente: Elaboración propia

Tabla 5.31: Tabla de percentiles de direcciones, *Swell* remoto SW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	7.9	5	221.9
		50	257.5
		95	284.5
50	8.7	5	232.6
		50	257.6
		95	267.2
100	9.5	5	254.6
		50	261.3
		95	267.3

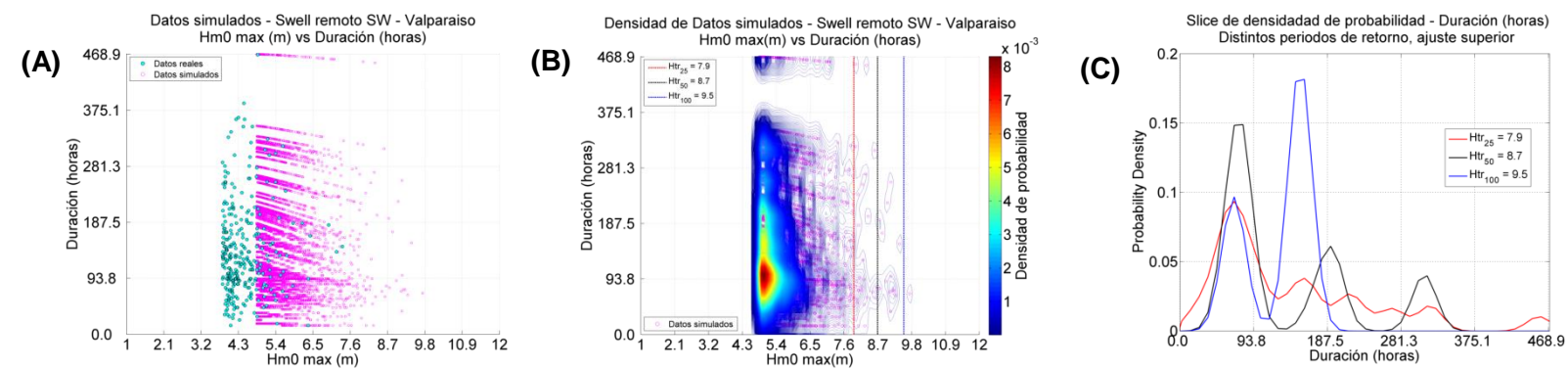


Figura 5.26: Resultados simulación multivariada Hm_0 vs Duración, *Swell* remoto SW – Zona central

Fuente: Elaboración propia

Tabla 5.32: Tabla de percentiles de duraciones, *Swell* remoto SW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	7.9	5	25.0
		50	97.4
		95	324.3
50	8.7	5	44.2
		50	85.3
		95	311.0
100	9.5	5	47.5
		50	137.1
		95	169.2

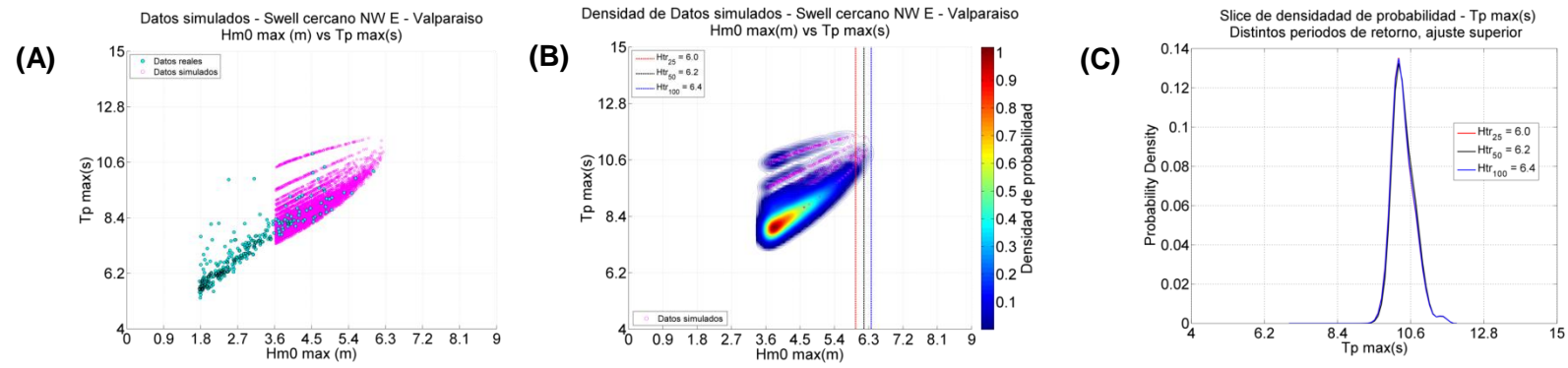


Figura 5.27: Resultados simulación multivariada Hm_0 vs Tp , Swell cercano NE E – Zona central

Fuente: Elaboración propia

Tabla 5.33: Tabla de percentiles de periodos, Swell cercano NE E Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	6.0	5	9.8
		50	10.3
		95	10.9
50	6.2	5	9.9
		50	10.3
		95	11.0
100	6.4	5	9.8
		50	10.3
		95	10.9

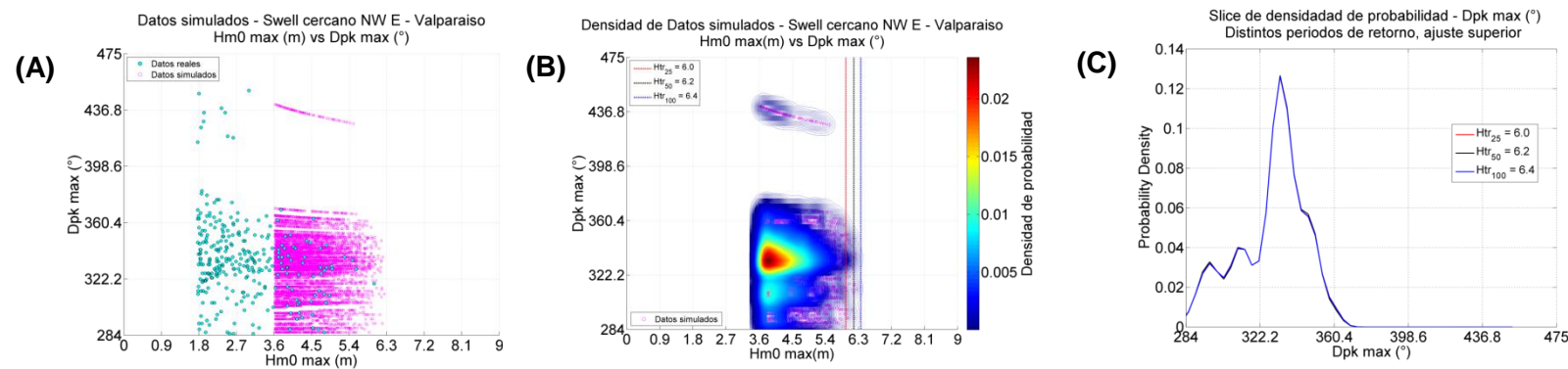


Figura 5.28: Resultados simulación multivariada Hm_0 vs Dpk , Swell cercano NE E – Zona central

Fuente: Elaboración propia

Tabla 5.34: Tabla de percentiles de direcciones, Swell cercano NE E Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	6.0	5	292.1
		50	329.7
		95	351.4
50	6.2	5	292.0
		50	329.7
		95	351.0
100	6.4	5	292.1
		50	329.7
		95	351.4

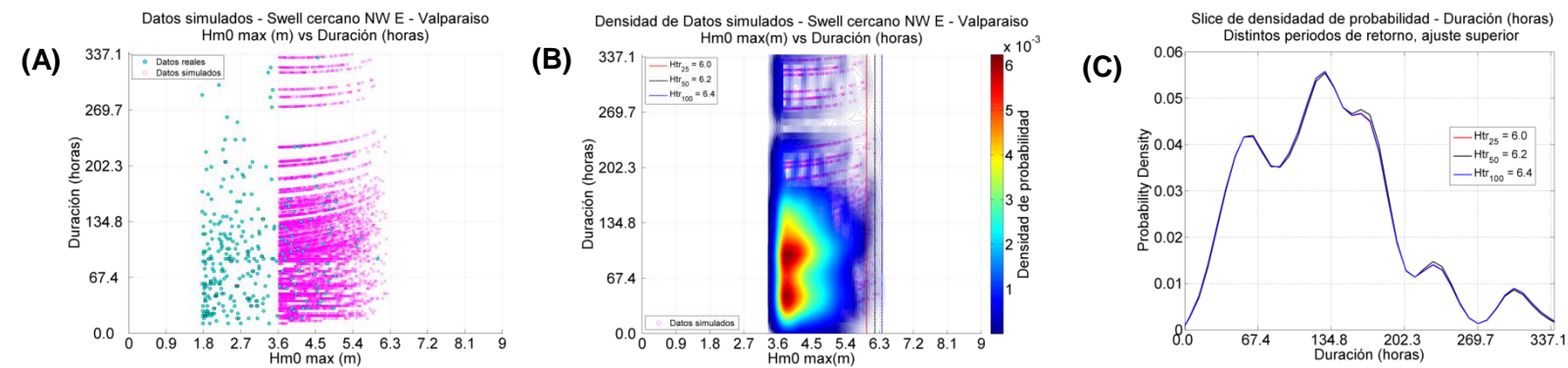


Figura 5.29: Resultados simulación multivariada Hm_0 vs Duración, Swell cercano NW E – Zona central

Fuente: Elaboración propia

Tabla 5.35: Tabla de percentiles de duraciones, Swell cercano NE E Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	6.0	5	30.1
		50	121.9
		95	252.2
50	6.2	5	30.7
		50	122.7
		95	248.6
100	6.4	5	30.2
		50	122.0
		95	251.7

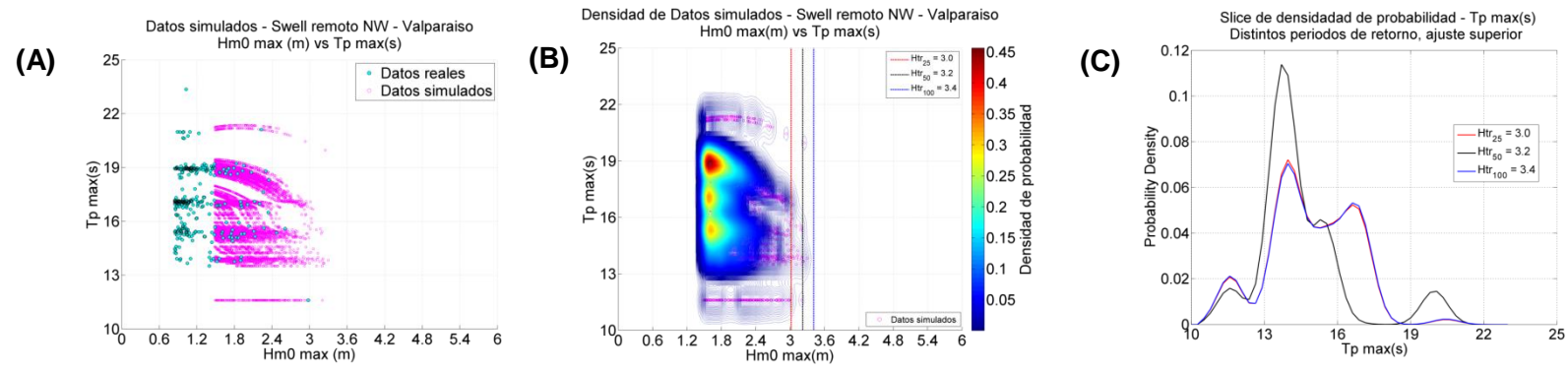


Figura 5.30: Resultados simulación multivariada Hm_0 vs Tp , Swell remoto NW – Zona central

Fuente: Elaboración propia

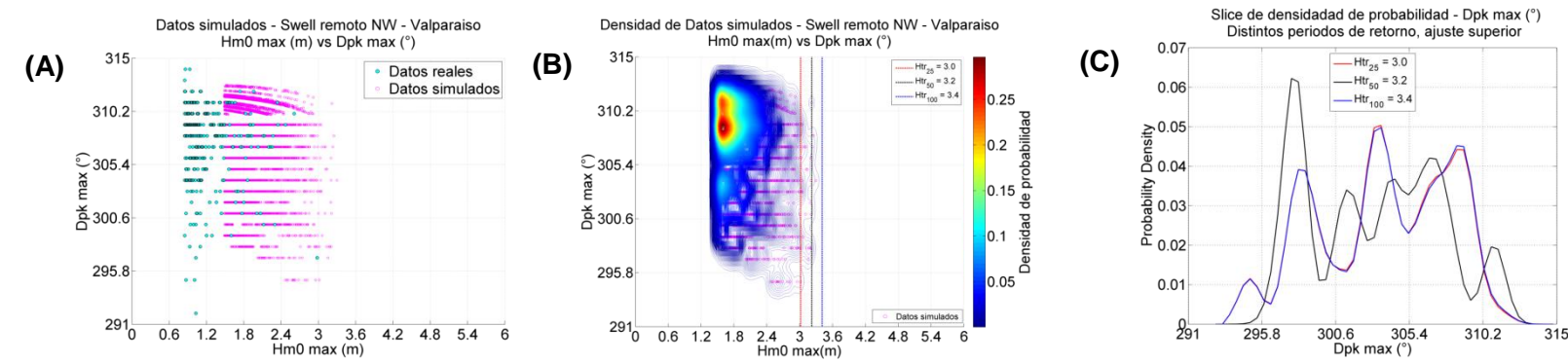


Figura 5.31: Resultados simulación multivariada Hm_0 vs Dpk , Swell remoto NW – Zona central

Fuente: Elaboración propia

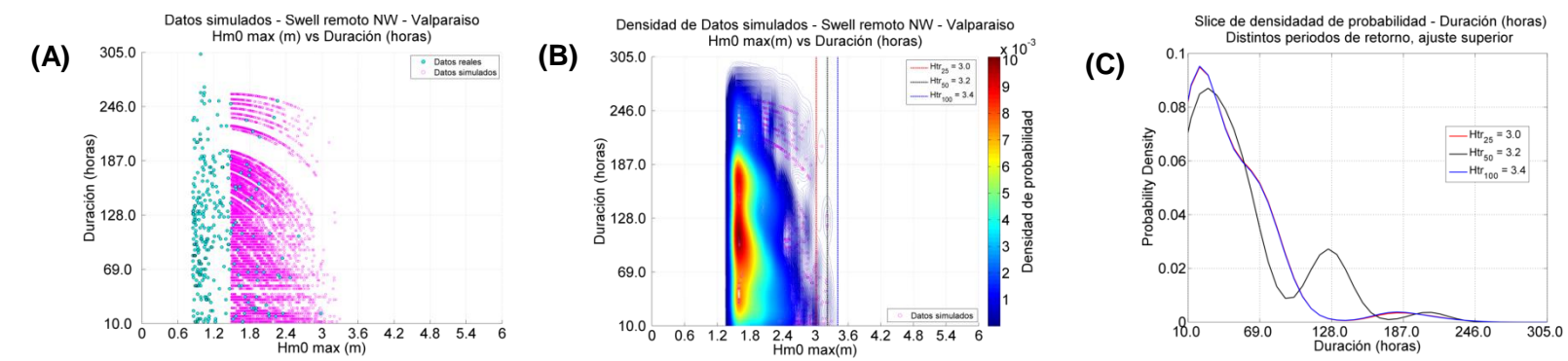


Figura 5.32: Resultados simulación multivariada Hm_0 vs Duración, Swell remoto NW – Zona central

Fuente: Elaboración propia

Tabla 5.36: Tabla de percentiles de periodos, Swell remoto NW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	3.0	5	11.5
		50	14.7
		95	17.31
50	3.2	5	11.7
		50	13.9
		95	19.6
100	3.4	5	11.4
		50	14.7
		95	17.3

Tabla 5.37: Tabla de percentiles de direcciones, Swell remoto NW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	3.0	5	296.4
		50	303.7
		95	309.5
50	3.2	5	296.8
		50	303.2
		95	310.5
100	3.4	5	296.4
		50	303.7
		95	309.5

Tabla 5.38: Tabla de percentiles de duraciones, Swell remoto NW Zona central. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	3.0	5	12.0
		50	31.7
		95	97.1
50	3.2	5	12.0
		50	36.0
		95	138.1
100	3.4	5	12.0
		50	31.5
		95	98.3

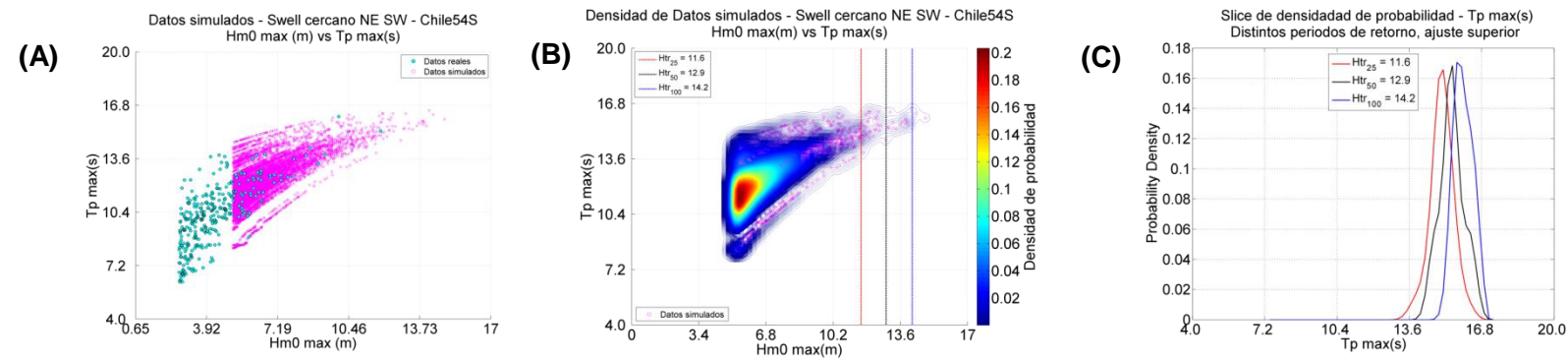


Figura 5.33: Resultados simulación multivariada Hm_0 vs Tp , Swell cercano NE SW – Zona sur

Fuente: Elaboración propia

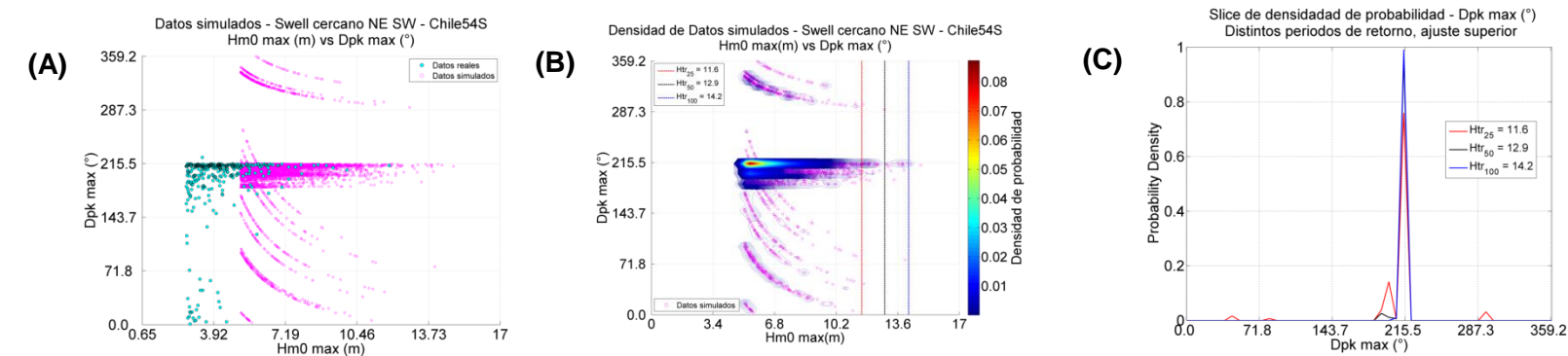


Figura 5.34: Resultados simulación multivariada Hm_0 vs Dpk , Swell cercano NE SW – Zona sur

Fuente: Elaboración propia

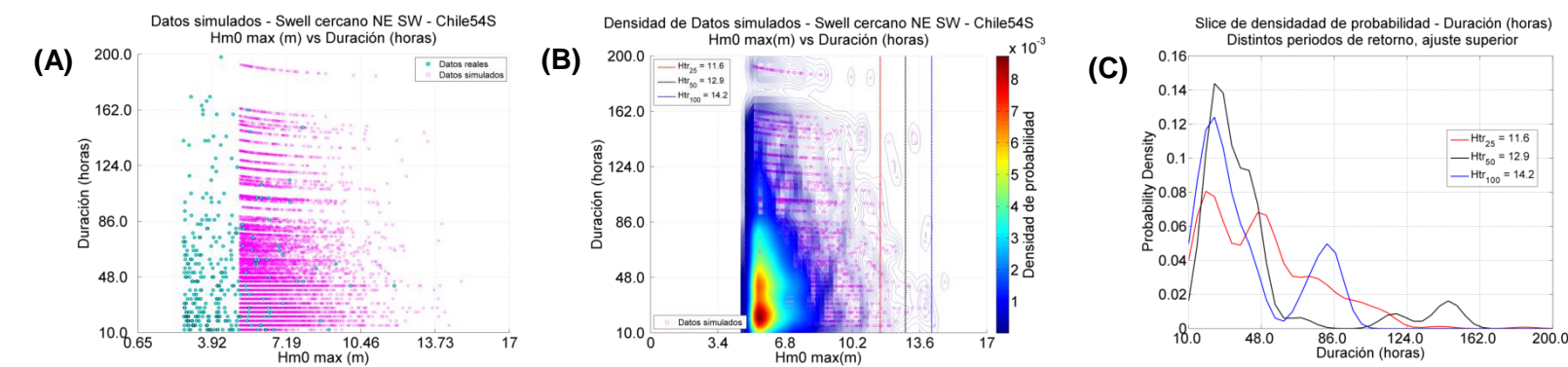


Figura 5.35: Resultados simulación multivariada Hm_0 vs Duración, Swell cercano NE SW – Zona sur

Fuente: Elaboración propia

Tabla 5.39: Tabla de percentiles de periodos, Swell cercano NE SW Zona sur. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	11.6	5	14.0
		50	15.0
		95	15.8
50	12.9	5	14.4
		50	15.4
		95	16.4
100	14.2	5	15.2
		50	15.9
		95	16.6

Tabla 5.40: Tabla de percentiles de direcciones, Swell cercano NE SW Zona sur. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	11.6	5	189.8
		50	209.7
		95	214.0
50	12.9	5	206.9
		50	210.3
		95	213.8
100	14.2	5	207.1
		50	210.5
		95	213.8

Tabla 5.41: Tabla de percentiles de duraciones, Swell cercano NE SW Zona sur. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	11.6	5	8.7
		50	41.1
		95	100.5
50	12.9	5	12.7
		50	29.9
		95	137.6
100	14.2	5	7.6
		50	27.7
		95	85.8

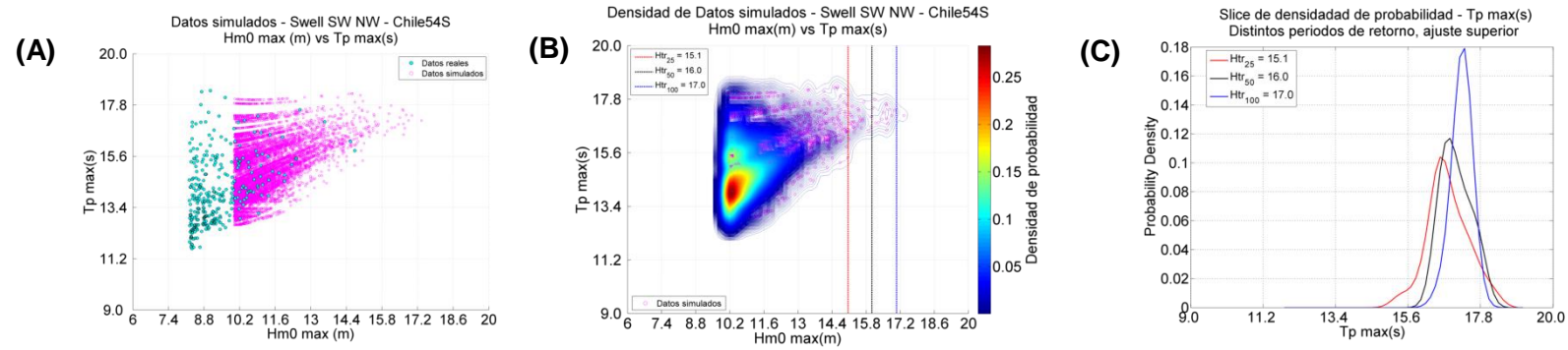


Figura 5.36: Resultados simulación multivariada H_{m0} vs T_p , Swell SW NW – Zona sur
Fuente: Elaboración propia

Tabla 5.42: Tabla de percentiles de periodos, Swell SW NW Zona sur. Ajuste superior

Periodos de retorno (años)	H_{m0} (m)	Percentiles (%)	Periodo (s)
25	15.1	5	15.7
		50	16.9
		95	17.9
50	16.0	5	16.2
		50	16.9
		95	17.8
100	17.0	5	16.6
		50	17.2
		95	17.7

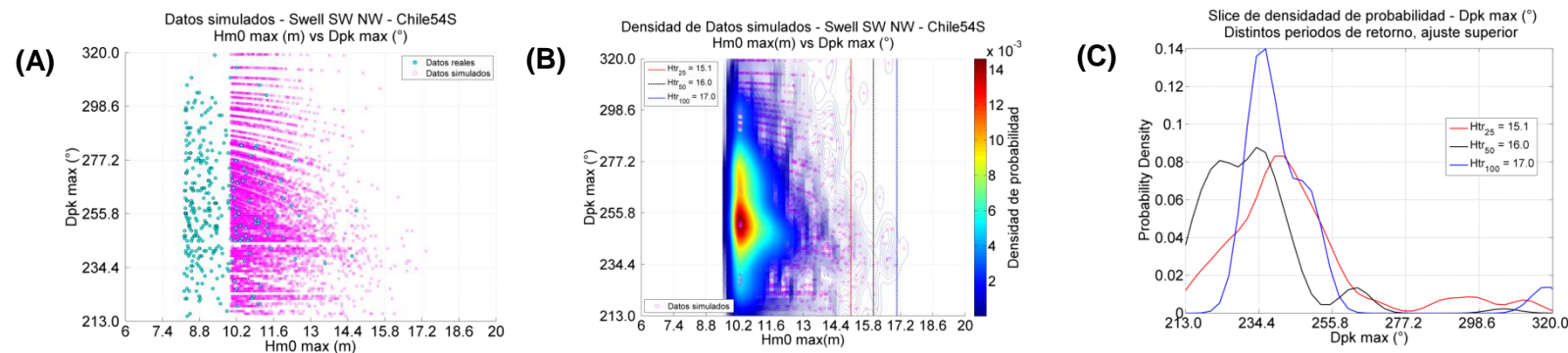


Figura 5.37: Resultados simulación multivariada H_{m0} vs D_{pk} , Swell SW NW – Zona sur
Fuente: Elaboración propia

Tabla 5.43: Tabla de percentiles de direcciones, Swell SW NW Zona sur. Ajuste superior

Periodos de retorno (años)	H_{m0} (m)	Percentiles (%)	Dirección (°)
25	15.1	5	217.0
		50	240.1
		95	295.1
50	16.0	5	212.3
		50	229.1
		95	258.8
100	17.0	5	226.9
		50	237.5
		95	312.1

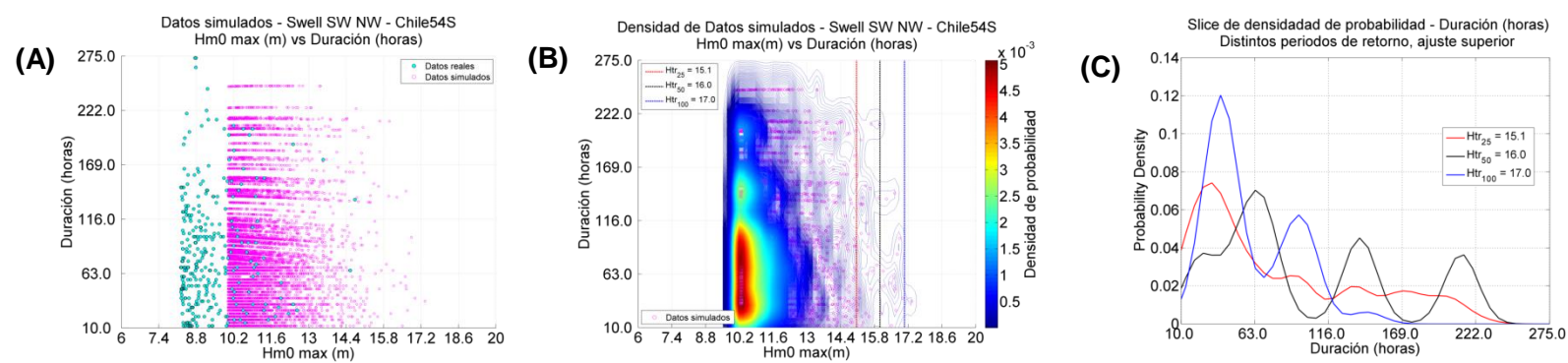


Figura 5.38: Resultados simulación multivariada H_{m0} vs Duración, Swell SW NW – Zona sur
Fuente: Elaboración propia

Tabla 5.44: Tabla de percentiles de duraciones, Swell SW NW Zona sur. Ajuste superior

Periodos de retorno (años)	H_{m0} (m)	Percentiles (%)	Duración (hr)
25	15.1	5	5.3
		50	52.6
		95	195.7
50	16.0	5	14.0
		50	71.4
		95	215.3
100	17.0	5	16.3
		50	46.0
		95	111.6

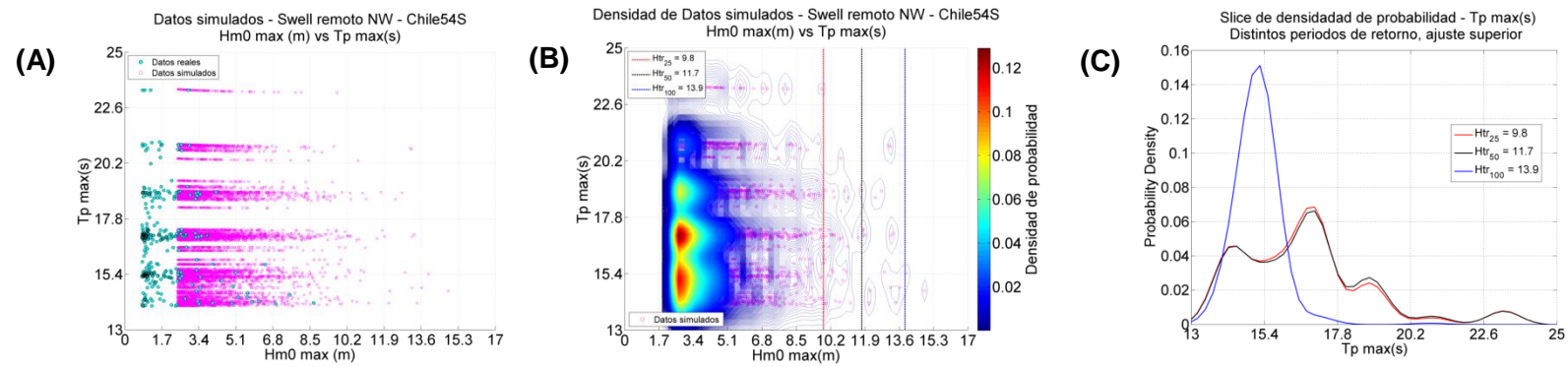


Figura 5.39: Resultados simulación multivariada Hm_0 vs Tp , Swell remoto NW – Zona sur

Fuente: Elaboración propia

Tabla 5.45: Tabla de percentiles de periodos, Swell remoto NW Zona sur. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Periodo (s)
25	9.8	5	13.7
		50	16.4
		95	20.9
50	11.7	5	13.7
		50	16.5
		95	20.9
100	13.9	5	13.9
		50	15.0
		95	16.2

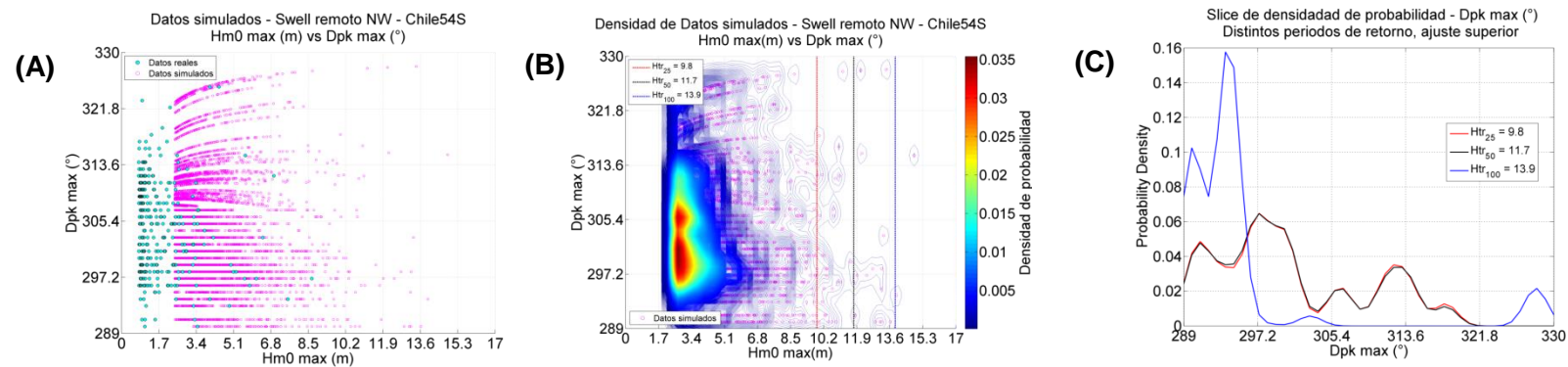


Figura 5.40: Resultados simulación multivariada Hm_0 vs Dpk , Swell remoto NW – Zona sur

Fuente: Elaboración propia

Tabla 5.46: Tabla de percentiles de direcciones, Swell remoto NW Zona sur. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Dirección (°)
25	9.8	5	289.3
		50	298.4
		95	315.1
50	11.7	5	289.4
		50	298.3
		95	314.9
100	13.9	5	288.3
		50	292.8
		95	326.9

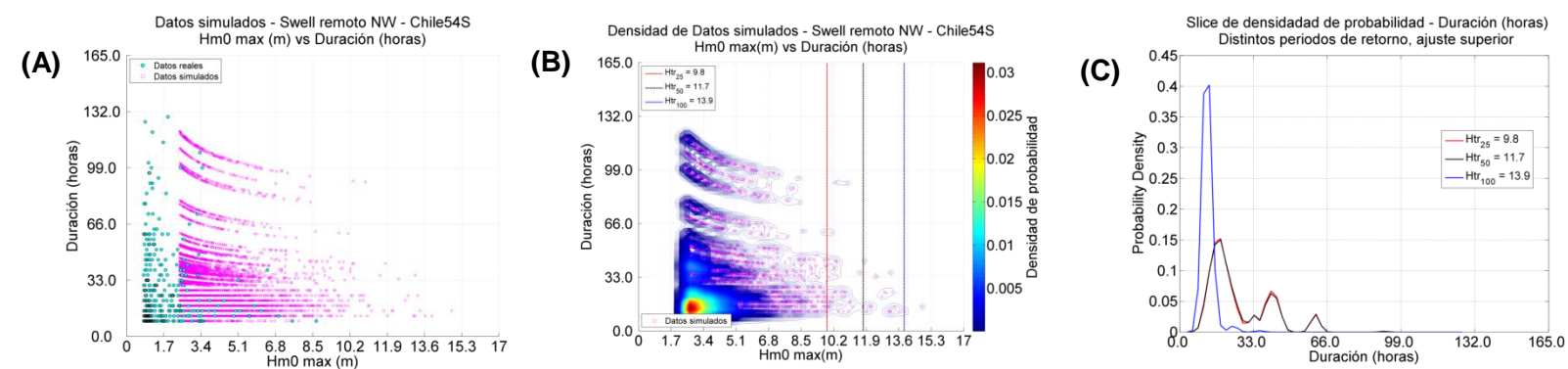


Figura 5.41: Resultados simulación multivariada Hm_0 vs Duración, Swell remoto NW – Zona sur

Fuente: Elaboración propia

Tabla 5.47: Tabla de percentiles de duraciones, Swell remoto NW Zona sur. Ajuste superior

Periodos de retorno (años)	Hm_0 (m)	Percentiles (%)	Duración (hr)
25	9.8	5	10.8
		50	19.6
		95	57.3
50	11.7	5	10.6
		50	19.6
		95	57.0
100	13.9	5	7.5
		50	11.0
		95	14.4

De las figuras expuestas, se desprende que en general, para los sistemas de oleaje *swell* cercanos presentan una tendencia positiva entre alturas y periodos. Esto se representa con la forma cónica ascendente tanto de los datos reales como simulados, donde también es posible notar una disminución en la dispersión de los datos. No obstante, para estos sistemas las direcciones y duraciones presentan dispersiones relativamente constantes. Sin embargo, existen casos específicos como el *swell* cercano SW de la zona norte en el cual las direcciones presentan una asociación positiva con una disminución de la dispersión de los datos a medida que el la altura de oleaje aumenta.

Por otra parte, los casos de oleaje tipo *swell* remotos presentan valores con concentraciones de datos acotadas entre alturas versus direcciones y duraciones, donde la dispersión de estos datos tanto para valores reales como simulados presentan tendencias más bien neutras, con una dispersión casi constante. Así mismo los rangos de direcciones y duraciones se encuentran dentro de los rangos esperados, ya sean para los sistemas asociados a las direcciones SW o NW, debiéndose principalmente a que estos sistemas remotos están completamente desarrollados, por lo cual las olas presentan direcciones más claras y duraciones características de oleaje remoto. En la mayor parte de estos casos, los periodos asociados a las alturas de ola mayormente tienden a tener un comportamiento de asociación positiva entre sí, mostrando dispersiones reducidas hacia las colas.

La Tabla 5.48 expone los resultados de periodo, dirección y duración asociados a la altura de ola con periodo de retorno de 100 años para los percentiles 5 y 95.

Tabla 5.48: Resumen comparativo entre sistemas y nodos. Caso H_{tr} 100 años, ajuste superior

Nodo	Caso	Sistemas comparables	H_{m0} (m)	Periodo (s)		Dirección (°)		Duración (horas)	
				5%	95%	5%	95%	5%	95%
Zona norte	<i>Swell</i> cercano SW	●	3.5	7.9	8.8	212.2	214.8	5.2	110.1
	<i>Swell</i> remoto SW	●	4.8	16.0	17.4	210.6	214.1	51.9	191.0
	<i>Swell</i> cercano NE SE	●	2.6	6.1	7.4	360.0	342.3	62.0	240.6
	<i>Swell</i> remoto NW	●	2.3	17.6	19.0	298.4	302.0	53.9	249.7
Zona centro	<i>Swell</i> cercano SW	●	4.3	8.7	8.9	221.7	226.6	8.4	51.2
	<i>Swell</i> remoto SW	●	9.5	16.6	18.1	254.6	267.3	47.5	169.2
	<i>Swell</i> cercano NW E	●	6.4	9.8	10.9	292.1	351.4	30.2	251.7
	<i>Swell</i> remoto NW	●	3.4	11.4	17.3	296.4	309.5	12.0	98.3
Zona sur	<i>Swell</i> cercano NE SW	●	14.2	15.2	16.6	207.1	213.8	7.6	85.8
	<i>Swell</i> SW NW	●	17.0	16.6	17.7	226.9	312.1	16.3	111.6
	<i>Swell</i> remoto NW	●	13.9	13.9	16.2	288.3	326.9	7.5	14.4

Así mismo, en la Tabla 5.48 se indican con círculos de igual color los sistemas de oleaje en los cuales las variables simuladas tienen un comportamiento semejante y que pueden ser comparables debido a que las particiones espectrales multimodales de las cuales provienen son de rangos similares.

A partir de esta tabla es posible obtener estimaciones sobre las características de los parámetros extremos estimados para cada uno de los sistemas de oleaje identificados. Por ejemplo, el *swell* cercano proveniente del SW, tanto de la zona norte como central presentan rangos muy similares entre los valores asociados en sus percentiles, con un rango de periodos acotados entre 7.9 s y 8.8 s para la zona norte, mientras que para la zona centro los valores van desde los 8.7 s a 8.9 s. Así mismo se aprecia que para las direcciones de la zona norte estas tienden a estar más orientadas al sur, mientras que la zona centro presenta direcciones que tienden hacia el SW con rangos entre 221.7° y 226.6°. Así mismo, las duraciones de la zona norte se encuentran entre valores que van desde los 5.2 a las 110.1 horas y la zona centro con valores entre las 8.4 y 51.2 horas.

Las alturas de oleaje asociadas al *swell* remoto del SW tienden a aumentar con las latitudes más hacia el sur. Los periodos asociados a estas alturas tienen valores similares en los percentiles expuestos. Las direcciones por su parte tienden hacia el W en la zona central, cuyos valores están acotados entre los 254° y 267°. Esto se explica por la posición relativa de los centros de generación. En el caso del sur, la dispersión de direcciones es mucho más amplia en el rango 227° y 312°. Las duraciones, por su parte, son mayores en el norte, situación que se ve reflejada en todos los sistemas de oleaje analizados.

Por otra parte, para el caso *swell* remoto NW, sistema que se encuentra presente en las tres zonas de análisis, las alturas extremas de 100 años de periodo de retorno tienden a incrementarse hacia el sur, los periodos en estos casos, si bien presentan rangos con valores muy similares, es la zona norte la que presenta los periodos de oleaje más altos con valores entre los 17.6 s y 19 s. Las duraciones por su parte presentan un comportamiento inverso a las alturas de oleaje, ya que estas disminuyen hacia el sur para estos sistemas de oleaje. La tendencia muestra que en la zona norte predomina el *swell* remoto que se genera en el hemisferio norte. En la zona central y especialmente en la sur aparecen componentes generadas en zonas más próximas que predominan sobre las olas del hemisferio norte. Para poder aislar las olas del hemisferio norte en estos casos se requiere hacer un análisis estacional o la aplicación de un método más avanzado de clasificación de oleaje que el propuesto en la sección 4.3.2.

6 DISCUSIÓN SOBRE LA VALIDACIÓN DE LA METODOLOGÍA UTILIZADA

En este apartado se presenta una descripción del desarrollo inicial de la metodología de modelación multivariada de valores extremos utilizada en este trabajo para el proceso de cálculo con el fin de justificar el método y sus resultados debido que el procedimiento utilizado no posee una validación directa.

Aquí se indican algunos documentos que se basan en esta metodología de manera de exponer que realizar validaciones sobre este tipo de análisis es un proceso complejo, y que sin embargo, comunidades científicas lo han aceptado en conjunto con sus resultados debido a su robustez matemática y estadística.

La metodología fue desarrollada y propuesta inicialmente por Jeffernan & Tawn (2004) para realizar estimaciones de valores extremos de contaminantes en la ciudad de Leeds, Inglaterra. El documento de su aplicación fue aceptado y publicado en *The Royal Statistical Society* el año 2004. De manera posterior, aplicaciones en diversas áreas de la ciencia utilizaron esta metodología como base. A continuación se presentan dos documentos basados en la metodología desarrollada por Heffernan & Tawn (2004).

1. Una aplicación y análisis sobre estados de mar de aguas profundas fueron desarrolladas utilizando esta metodología, una de ellas y la cual es referencia de este trabajo corresponde al estudio realizado por Philip Jonathan y Kevin Ewans (2010) en su documento *Joint modelling of wave spectral parameters for extreme sea states*, publicado en el *journal* científico *Ocean Engineering*. En este trabajo sólo se realizaron comparaciones entre distintos modelos estadísticos multivariados para la estimación de periodos de oleaje asociados a alturas de ola con 100 años de periodo de retorno. Los datos utilizados en este trabajo correspondían a datos medidos en 4 localidades distintas. En este trabajo los autores no presentan validaciones de los datos ni del procedimiento.
2. Aplicaciones expuestas por Beatriz Vaz de Melo Mendes y Luis Raúl Pericchi en su documento *Assessing Conditional Extremal Risk of Flooding in Puerto Rico* utilizan la metodología propuesta para la estimación multivariada de condiciones extremas de mareas para evaluar el riesgo de inundación. En este trabajo tampoco se han presentado validaciones de la metodología, sin embargo los resultados demuestran la concordancia entre las simulaciones y los datos reales utilizados.

No obstante, aunque las aplicaciones no indiquen validaciones, es posible verificar que los primeros pasos de la metodología, tal como el ajuste marginal de los datos, estén bien realizados contrastando el modelo de ajuste con los datos distribuidos de manera empírica. Realizar este tipo de comparación permite revisar que los datos seleccionados se ajustan de manera correcta al modelo propuesto.

Respecto de lo anterior y de acuerdo a que la teoría propone la utilización de la distribución de Pareto para el ajuste marginal de los datos reales (modelo para valores sobre un umbral), se han realizado pruebas ajustando los datos por medio éste y también una distribución empírica de probabilidad. Estas figuras se presentan a continuación y corresponden a los sistemas de oleaje identificados en nodo de aguas profundas de Valparaíso.

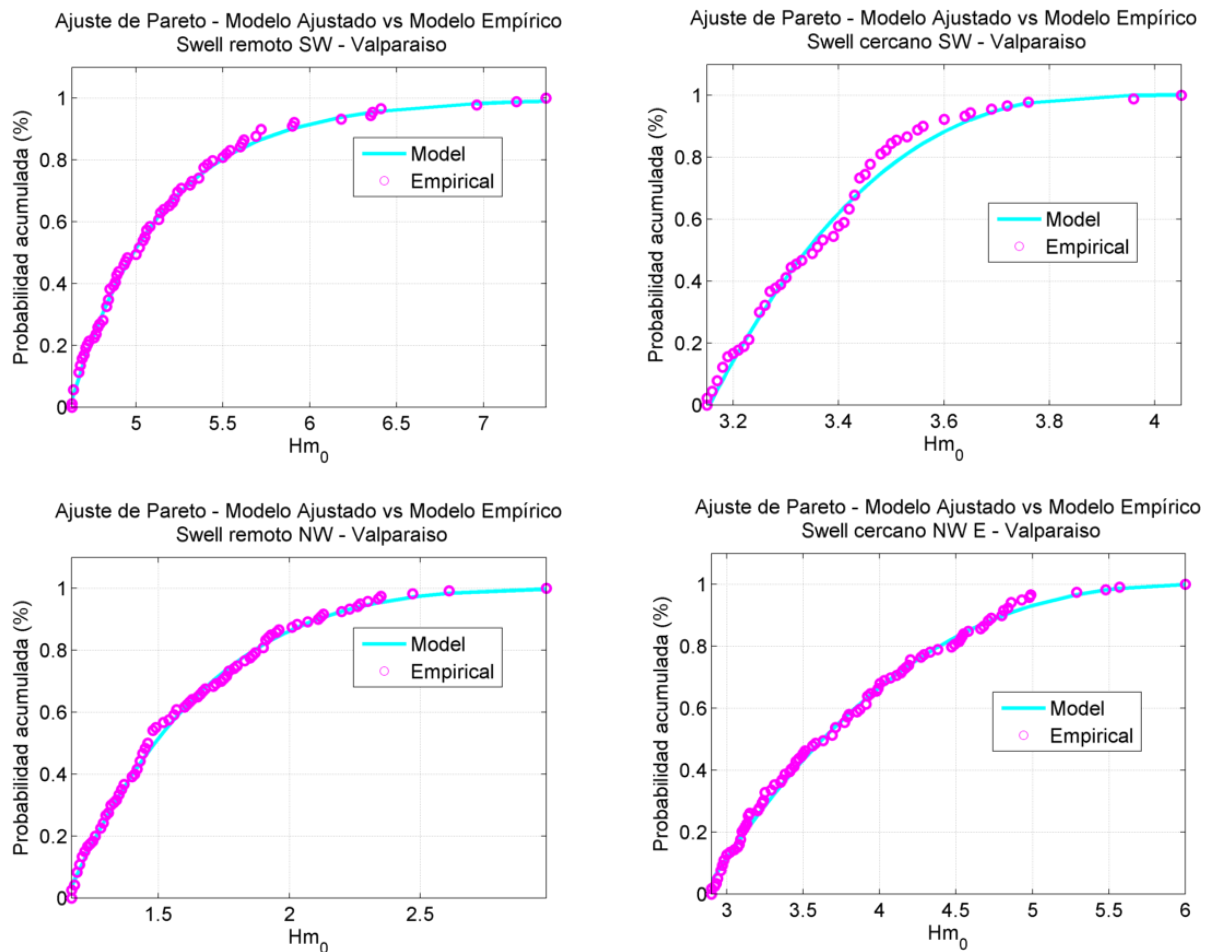


Figura 42: Ajuste de Pareto vs Ajuste empírico sobre datos de oleaje. Nodo Valparaíso

A partir de la Figura 42 es posible apreciar que las estimaciones según el modelo de Pareto se ajustan de manera adecuada a los resultados del modelo empírico utilizado. Esto es una muestra de que las simulaciones univariadas han sido llevadas a cabo de buena manera y que los datos considerados sobre un nivel de umbral son aceptables para las posteriores simulaciones multivariadas. Así mismo, la aceptación y validación del método se basa principalmente en los resultados obtenidos por medio de la simulación multivariada, ya que estos se comportan según la forma de los datos reales considerados. En la Figura 43 se presentan los resultados de las simulaciones multivariadas de alturas de oleaje versus periodos para los sistemas de oleaje identificados en el nodo Valparaíso.

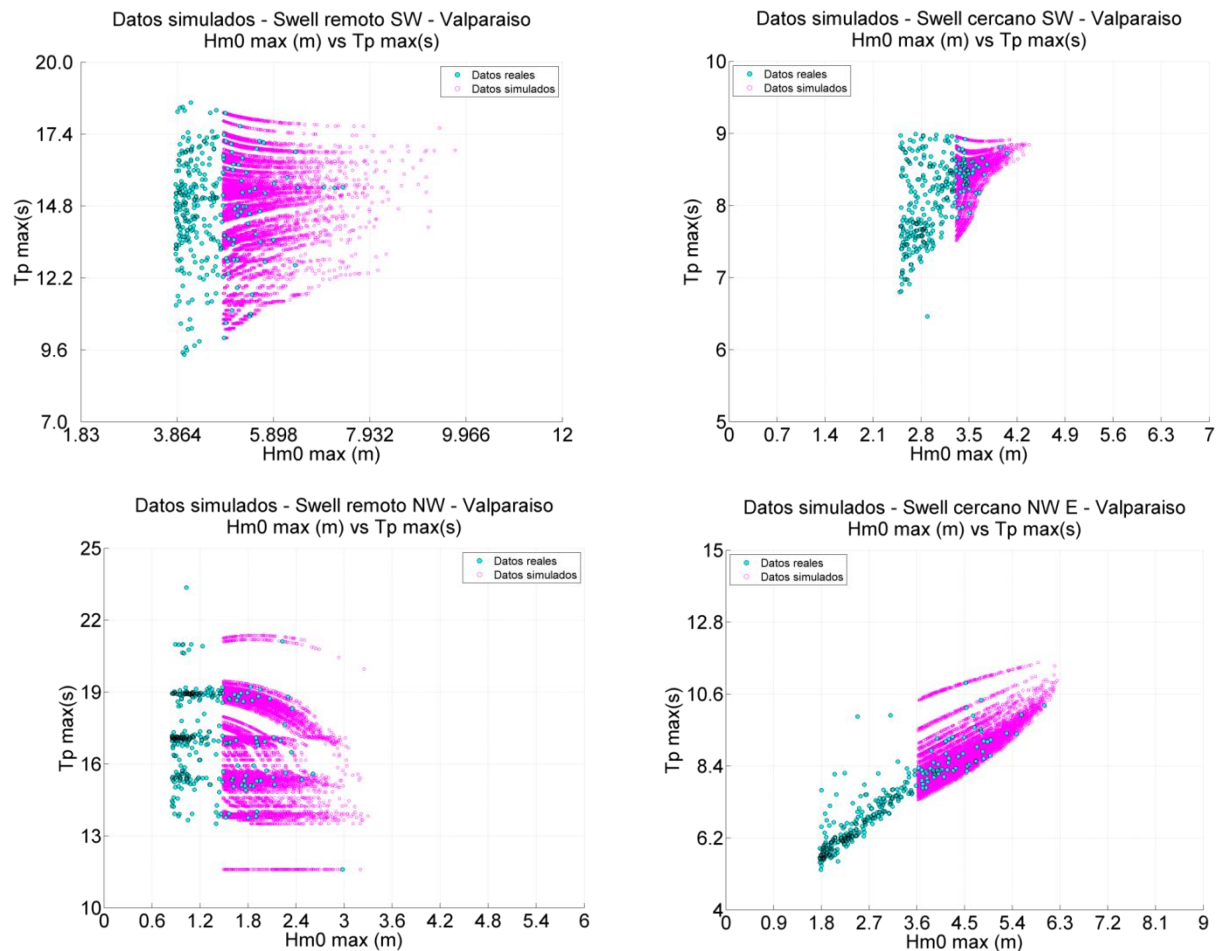


Figura 43: Dispersión de datos reales y simulados - Swell remoto SW, Valparaíso

De acuerdo a que los resultados simulados representan de manera adecuada las tendencias de cada conjunto de variables utilizadas, se muestra que el método es matemáticamente robusto y que los resultados obtenidos han sido suficientes para demostrar que los procesos de cálculo han sido bien realizados.

7 CONCLUSIONES Y RECOMENDACIONES

A partir de los resultados obtenidos, se establece que la metodología propuesta presenta una innovación a los métodos comunes para el cálculo y estimación de valores extremos de oleaje en Chile de acuerdo a lo siguiente:

- La metodología de separación de *swells* permite la obtención de eventos de tormenta dentro de distintos sistemas de oleaje de aguas profundas, separando a cada evento de acuerdo a un criterio físico del oleaje. Esto permite una identificación adecuada de todo el tren de olas que componen una tormenta.
- Para los distintos tipos de oleaje, la utilización de la metodología de separación de *swells* permite una mejor identificación de los valores más altos de altura de oleaje y sus parámetros asociados dentro de un evento de tormenta, ya que se basa sobre una separación de tormentas con criterios físicos precisos y no en aproximaciones, como es el caso del POT sobre una serie de estados de mar.
- Realizar ajustes con datos más precisos en conjunto con una simulación robusta de valores extremos permite obtener extrapolaciones adecuadas de los valores utilizados para distintos periodos de retorno.
- La simulación multivariada de valores extremos se logró con éxito, obteniendo duraciones, periodos y direcciones asociados a valores de oleaje extremas para distintos periodos de retorno.
- Se ha logrado establecer de manera estadística y probabilística la frecuencia de aparición de las variables extremas simuladas para los distintos sistemas de oleaje, pudiendo obtener los valores de estas mismas para diferentes percentiles.
- La metodología propuesta permite mediante métodos estadísticos, matemáticos y físicos, describir de manera cuantitativa el clima extremo, sin dejar valores a criterios o decisiones cualitativas de un calculista o modelador. Dicho criterio se deberá aplicar al momento de utilizar estos valores y definir el riesgo que se requiere asumir de acuerdo a las condiciones específicas del sitio ó en una obra.
- La independencia entre los sistemas de oleaje para realizar el clima extremo es fundamental para aplicar estos métodos, de tal modo que se recomienda no propagar el oleaje hacia aguas somera antes de analizarlo, pues comienza a ser dominado por el fondo y se pierde la independencia estadística entre los sistemas.

Finalmente, se concluye que todo el proceso de simulación realizado en este trabajo representa una metodología robusta para estimar los valores extremos de oleaje en Chile,

basados en procedimientos matemáticos complejos que sustentan los resultados obtenidos y reducen la incertidumbre en la asignación de variables cuando se desean utilizar en conjunto con alturas de ola extremas.

En función de los resultados obtenidos y de las conclusiones efectuadas se recomienda lo siguiente:

- Utilizar partición espectral multimodal sobre los nodos de aguas profundas, de manera de obtener sistemas de oleaje mejor definidos dentro de cada nodo a utilizar.
- Aplicar el método de separación de *swells* para la identificación de eventos extremos, debido a que este procedimiento permite una mejor identificación de los valores de altura de oleaje, además de la obtención de todos los parámetros significativos que componen las tormentas.
- Para la estimación de valores extremos, se recomienda el uso de una metodología de valores extremos multivariados, ya que ésta permite estimar parámetros extremos de manera conjunta, con un sustento matemático y físico, reduciendo la incertidumbre en la estimación de los parámetros. Las variables condicionadas (periodo, dirección y duración) pueden tener una importancia mayor en algunos procesos de ingeniería marítima, tales como la respuesta de playas, diseño de rompeolas, estructuras de borde costero, análisis de sobrepaso, *runup*, entre otros.
- El método podría ser extensible a estimaciones multivariadas de otras variables oceanográficas tales como corrientes, mareas, *steepness* (peralte, escarpe de oleaje en aguas profundas), así también como la interacción entre ellas.

8 REFERENCIAS BIBLIOGRÁFICAS

- Ref. 1:** Heffernan, J., Tawn, J. A., “*A Conditional Approach for Multivariate Extreme Values*”. J. R. Statist. Soc. B. 2004
- Ref. 2:** Jonathan, P., Flynn, J., Ewans, K., “*Joint Modeling of Waves Spectral Parameters for Extreme Sea States*”. Ocean Engineering. 2010
- Ref. 3:** Goda, Y., “*A New Methodology of Extreme Wave Analysis with Regional Frequency Analysis Using L-moments*”. Proc. Special Symposium Celebrating the 60th Anniversary of the Tainan Hydraulic Laboratory, National Cheng-Kung University. 2010
- Ref. 4:** Hanson, J. L, Phillips, O.M., “*Automated Analysis of Ocean Surface Directional Wave Spectra*”. Journal of Atmospheric and Oceanic Technology. 2000
- Ref. 5:** Fougères, A. - L, “*Multivariate Extremes*”. In Extreme Values in Finance, Telecommunications, and the Environment (Eds Finkelstadt, B. & Rootzén, H.). 373-388, Chapman & Hall/CRC, BocaRaton, FL. 2004
- Ref. 6:** Hosking, J. R. M., Wallis J., R., “*Regional Frequency Analysis. An approach based on L-moments*”. Cambridge University Press. 1997
- Ref. 7:** Bortot, P., Gaetan, C., “*Multivariate Extremes*”. Encyclopedia of Environmetrics, 2nd Edition, 1685-1694. John Wiley & Sons, Chichester. 2012
- Ref. 8:** Hanson, J., Lübben, A., Aalberts, P., Lech, M., “*Wave Measurements for the Monitas System*”. Offshore Technology Conference, Houston, Texas, USA, 3 – 6 May, 2010
- Ref. 9:** Hanson, J. L., Tracy, B. A., Tolman, H. L., Scott, D. R., “*Pacific Hindcast Performance of Three Numerical Wave Models*”. Journal of Atmospheric and Oceanic Technology. 2009
- Ref. 10:** Rojas, A., “*Partición espectral multimodal aplicada a la determinación de olas de diseño aplicación en Valparaíso, Chile*”. Universidad de Valparaíso, 2015.

- Ref. 11:** Keef, C., Papastathopoulos, I., Tawn, J. A., “*Estimation of the conditional distribution of a multivariate variable given that one of its components is large: Additional constraints for the Heffernan and Tawn model*”. *Journal of Multivariate Analysis* 115, 2013, 396 - 404.
- Ref. 12:** Wyncoll, D., Gouldby, B., “*Application of a multivariate extreme value approach to system flood risk analysis*”. *Journal of Flood Risk Management*, 2013.
- Ref. 13:** Li, F., Bicknell, C., Lowry, R., Li, Y., “*A comparison of extreme wave analysis methods with 1994-2010 offshore Perth dataset*”. *Coastal Engineering*, 2012.
- Ref. 14:** Pantoja, C., Nicolau del Roure, F., Scott, D., “*Proyecto Olas Chile: The development of a reliable long term wave climate for entire Chilean coastline*”. *Waves 2005*, Madrid, España.