

Universidad Nacional Autónoma de México

Facultad de Ciencias

VALIDACIÓN DEL NUEVO SISTEMA DE PRONÓSTICO OPERATIVO DE OLEAJE DEL CENTRO DE CIENCIAS DE LA ATMÓSFERA

T E S I S

QUE PARA OBTENER EL TÍTULO DE

LICENCIADA EN CIENCIAS DE LA TIERRA

PRESENTA

ALMA TATIANA HERNÁNDEZ SIADE

TUTOR:

OCTAVIO GÓMEZ RAMOS

Ciudad Universitaria, Cd.Mx. Marzo 2020





Universidad Nacional Autónoma de México



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Datos del alumno

Hernández Siade Alma Tatiana 4423672509 Universidad Nacional Autónoma de México Facultad de Ciencias Ciencias de la Tierra 414039780

Datos del tutor

M. en C. Octavio Gómez Ramos

Datos del sinodal 1 Dr. Jorge Zavala Hidalgo

- Datos del sinodal 2 Dr. Arturo Ignacio Quintanar Isaías
- Datos del sinodal 3 Dr. José Antonio Salinas Prieto

• Datos del sinodal 4

Dra. Elsa Arellano Torres

Datos del trabajo escrito Validación del nuevo sistema de pronóstico operativo de oleaje del Centro de Ciencias de la Atmósfera 90 p 2020

Agradecimientos

Agradecimientos académicos

A los investigadores que fueron sinodales de esta tesis: Dr. Jorge Zavala Hidalgo, Dr. Arturo Quintanar Isaías, Dr. José Antonio Salinas Prieto y Dra. Elsa Arellano Torres, gracias por enriquecer este trabajo con sus observaciones. Pero sobre todo, gracias por hacerme ver detalles que en el futuro me ayudarán a mejorar la forma en que desarrolle cualquier proyecto académico al que me enfrente. Gracias también por su disposición y entrega para facilitarme terminar con este proceso.

Agradezco los recursos de supercómputo del Centro de Ciencias de la Atmósfera por el tiempo de uso del Clúster Ometéotl.

El desarrollo de este trabajo fue posible gracias al Centro Nacional de Prevención de Desastres (CE-NAPRED).

Agradecimientos personales

A mis papás, por tantas cosas que ni siquiera las puedo enumerar, pero que me ayudaron a llegar hasta aquí. Por esos esfuerzos enormes que hicieron para que, al llegar a la Ciudad de México, yo solo tuviera que preocuparme por hacer mis tareas y ver doramas. Por la paciencia y la confianza en mis decisiones, incluso cuando no estaban de acuerdo o les atemorizaban. No alcanza, pero los amo.

A mis hermanas, por el cariño incondicional. Cirenia, gracias por escuchar las cosas que me preocupaban y asustaban y que no podría haber compartido con nadie más. Por darme una perspectiva más ligera ante ellas y por ponerte la camiseta de hermana mayor cada vez que lo necesito. María, gracias por todas las veces que me has dado palabras de aliento y me has hecho ver que crees en mi. En los momentos difíciles han sido una fuente de luz. Las amo.

A mis madrinas: Conchita, Rosita y Guille. Casi parecen hadas. Gracias por el apoyo y cariño que me han dado en tantas formas. Por estar en momentos difíciles y felices.

A las amigas que han estado toda la vida: Andrea, Ale, Majo y Abril. Gracias por todos los recuerdos memorables. Ale, gracias infinitas por el apoyo y las palabras que me diste en meses recientes. Los atesoro mucho.

A Claudia, mi NHPMQA. Me has hecho saber que no hay nada que no pueda confiar en ti. En los momentos díficiles durante la carrera, durante la escritura de esta tesis, durante la vida, me has recordado que no estoy sola y que puedo. Gracias eternas. Te quiero muchísimo. A las "niñas": Ale, Ari, Bren, Diana, Mayra, Sam y Regina. Me dieron los días más felices y tranquilos que he tenido, que aún hoy me hacen sonreír cuando las cosas están difíciles. Y cuando las veo ahora (a las que logro ver) son un refugio. Ale, Ari y Diana, la verdad es que los últimos meses no los hubiera podido recorrer sin ustedes. Gracias. Las amodoro.

A Aura, Lulú, Mali, Norma y Charlie. Probablemente hubiera abandonado la carrera en tercer semestre si no los hubiera tenido. Valía la pena ir cada día a la UMDI por compartir momentos con ustedes. Son personas únicas a las que admiro muchísimo y de las que he aprendido tantas cosas. Aura, Mali y Norma, en los momentos difíciles de los últimos veces sus palabras me han sanado tanto. Las quiero muchísimo.

A Aldo. Las palabras que me dijiste antes de que me viniera a la Ciudad de México las tengo guardadas con mucho cariño. En los momentos en los que he sentido que no puedo con nada, que no han sido pocos, las recuerdo y las cosas mejoran. Gracias.

A Meli y a Dani. No sé cómo hubiera terminado la carrera sin ustedes. Hicieron que mis semestres en CU fueran realmente bonitos. Meli, gracias por las noches de desvelo en los que acababámos tareas de 15 páginas y que siempre recordábamos para auto-animarnos a terminar la tesis. Dani, gracias por ser un gran roomie y por todas esas pláticas que, inadvertidamente, se han vuelto algunas de las más honestas que he tenido. Gracias por las acuapijamadas. Los quiero. Gracias.

A Mar. Te admiro mucho por lo fuerte y entregada que eres. Gracias por todas las veces en las que me dijiste las cosas chidas que veías en mi. No sabes cómo me ha ayudado tener eso presente para seguir adelante.

A Samy: gracias por muchas cosas. Tienes mucha energía que no siempre puedo manejar, pero justamente por eso ha habido muchas veces en las que me has levantado el ánimo como nadie. Gracias por ser una gran roomie, pero más por ser una verdadera amiga. Pero sobre todo, gracias porque me has hecho sentir acompañada cuando lo necesitaba. Te quiero.

Al Dr. Jorge Zavala. Muchas gracias por el apoyo y la confianza que me ha dado en mis momentos personales difíciles. Y gracias por las palabras que me recuerdan que puedo con retos más grandes.

Y a Octavio. Perdona porque seguro este proceso fue mucho menos fácil y mucho más largo de lo que pensabas. Tampoco yo imaginé que sería así. Muchas gracias por tu paciencia. Pero sobre todo, gracias por ser mi amigo. Gracias por hacerme saber que lo más importante era yo y que la tesis venía después.

Hay muchas otras personas y amigos que han impactado mi vida de un modo positivo, ayudándome a llegar hasta aquí. Gracias a ustedes también.

Gracias a todos, por hacer de esta una lista larga.

Resumen

El oleaje generado por viento es un proceso importante para acoplamiento del sistema océano-atmósfera, así como para el desarrollo de las actividades marinas que realiza el hombre. México posee mares de gran extensión en cuyas costas habita una importante parte de la población y en donde se desarrolla una de sus actividades económicas más importantes: la extracción de petróleo. En este contexto, el Grupo de Interacción Océano-Atmósfera del Centro de Ciencias de la Atmósfera implementó un sistema de pronóstico operativo de oleaje como parte de un proyecto realizado para el Centro Nacional de Prevención de Desastres. Esta tesis tuvo como objetivo desarrollar dicho sistema, haciendo uso del modelo numérico WAVEWATCH III (WWIII), para el cual se implementaron tres dominios: un dominio global, un dominio que cubre el Pacífico mexicano y un dominio que cubre el Golfo de México. El modelo fue forzado con vientos obtenidos del modelo meteorológico WRF, que también se corre en modo pronóstico en la UNAM. Cada día, este sistema pronostica las condiciones de oleaje para las siguientes 120 horas y ha trabajado de manera diaria desde octubre de 2017. Para el dominio del Golfo de México se realizó una validación estadística sobre tres variables de salida: la altura significante, el periodo de oleaje y la dirección de oleaje. Para las primeras dos variables, se utilizaron cinco métricas de validación: el sesgo (EM), la raíz del error cuadrático medio (RMSE), el error no sistemático (RMSEb), el coeficiente de correlación lineal (r) y el índice de concordancia (IOA). Para la dirección de oleaje se utilizó el coeficiente de correlación compleja (r_{ang}) y el sesgo angular (EM_{ang}) . Se determinó que el modelo se desempeña muy bien para pronosticar la dirección de oleaje y de altura significante, pero no tiene tan buen desempeño para el periodo del oleaje. No se identificó un patrón estacional claro en el desempeño del modelo, pero se confirmó que el pronostico va perdiendo calidad conforme pasa el tiempo de simulación, teniendo mejor desempeño en las primeras 48 horas y siendo más pobre entre las 48 y 120 horas de simuñación. Para la altura significante se observó un patrón espacial en el desempeño del modelo: para la zona del Caribe se observó un sesgo positivo muy alto, para las zonas más someras se observó un sesgo negativo persistente y para las boyas ubicadas en otras zonas se observó un sesgo positivo al pronosticar alturas significantes pequeñas y un sesgo positivo al pronosticar alturas significantes altas. En general, los resultados coinciden con lo reportado en la literatura, la cual indica que los modelos de oleaje pronostican muy bien las variables de altura significante y dirección del oleaje; que suelen reproducir los patrones de error observados en los vientos utilizados como forzantes y que el WWIII presenta una tendencia a sobreestimar alturas significantes pequeñas y subestimar alturas significantes altas.

Índice general

1.	1. Introducción		6	
2 .	2. Marco teórico	Marco teórico		
	2.1. Conformación del campo de olea	je	9	
	2.2. Teorías sobre la formación de las	3 olas	11	
	2.3. Modelación numérica del oleaje		13	
3.	3. Revisión de trabajos de validació	n de modelos de oleaje	15	
	3.1. Métodos de validación	· · · · · · · · · · · · · · · · · · ·	15	
	3.1.1. Evaluación de los paráme	etros integrales	15	
	3.1.2. Evaluación de las partici	ones del espectro	16	
	3.2. Trabajos de validación realizado	s previamente	17	
	3.2.1. Validación de implementa	aciones operativas de modelos de oleaje	17	
	3.2.2. Validación del modelo co	mo un análisis de sensibilidad: vientos forzantes y para-		
	metrizaciones físicas		21	
	3.3. Conclusiones del capítulo		24	
4.	4. Descripción del modelo y su imp	lementación	26	
	4.1. Ecuaciones gobernantes		26	
	4.1.1. Propagación en el model	• WWIII	28	
	4.1.2. Términos fuente		28	
	4.2. Consideraciones numéricas		29	
	4.3. Implementación del modelo WW	/III	30	
	4.3.1. Parametrizaciones		30	
	4.3.2. Flujo de ejecución del me	odelo	31	
	4.3.3. Descripción de los domin	ios simulados	32	
	4.3.4. Datos de entrada para el	modelo	33	
	4.3.5. Automatización de la eje	cución modelo	35	
	4.4. Consideraciones para la validació	5n	38	
5.	5. Método de la validación		39	
	5.1. Observaciones utilizadas		39	
	5.1.1. Variables medidas en las	boyas	39	
	5.1.2. Selección de boyas	·	40	
	5.1.3. Preprocesamiento de las	observaciones	42	
	5.2. Procesamiento de los datos del r	nodelo	43	
	5.2.1. Interpolación a la posició	n espacio-temporal de las boyas	43	
	5.2.2. Separación de las salidas	del modelo por periodo de simulación	44	
	5.3. Métricas estadísticas utilizadas ϵ	en la evaluación del modelo	45	

		5.3.1.	Sesgo o error medio (EM)	45					
		5.3.2.	Raíz del error cuadrático medio (RMSE)	46					
		5.3.3.	Error no sistemático (RMSEb)	46					
		5.3.4.	Coeficiente de correlación de Pearson (r)	47					
		5.3.5.	Índice de concordancia (IOA)	47					
		5.3.6.	Sesgo angular (EM_{ang})	48					
		5.3.7.	Coeficiente de correlación compleja (r_{ang})	48					
	5.4.	Conclu	siones del capítulo	49					
6.	Resultados y discusión								
	6.1.	Result	ados de la implementación	50					
	6.2.	Result	ados de la validación	56					
		6.2.1.	Magnitud del viento	57					
		6.2.2.	Altura significante	60					
		6.2.3.	Periodo del oleaje	64					
		6.2.4.	Dirección del oleaje	68					
		6.2.5.	Distribución direccional de la energía	71					
	6.3.	Discus	ión de resultados	71					
7.	Con	clusio	les	75					
	7.1.	Trabaj	o a futuro	76					
	7.2.	Observ	vaciones finales	77					
A. Registros de boyas utilizados en la validación de las variables de interés: rapidez del									
	viento, altura significante, periodo del oleaje y dirección del oleaje 7								
Bi	Bibliografía 83								

Índice de figuras

1.1.	Esquema de las capas límites en la interfase océano-atmósfera. Modificado de Konda (sf) y Brumer (2017).	7
2.1.	Esquema de formación de olas por viento. Cuando las olas escapan de la región de in- fluencia del viento, pasan a formar parte del <i>swell</i> también llamado <i>mar de fondo</i>	9
2.2.	Clasificación de ondas oceánicas superficiales. En rojo se indica la fuerza de restitución principal para cada intervalo. Modificado de Massel (1996) y Li (2010)	10
2.3. 2.4.	Atributos utilizados para describir una onda armonica	$\frac{11}{12}$
4.1. 4.2. 4.3.	Diagrama de ejecución del modelo WWIII	32 33 37
5.1. 5.2. 5.3.	Registros de la boya 42057 para el año 2015	41 42
5.4.	filtradas	43
5.5.	Ejemplo del agrupamiento de las simulaciones diarias en cinco series mensuales armadas con periodos de simulación de 24 horas.	44 45
6.1.	Sección de oleaje de la página de pronósticos numéricos del Grupo IOA (http://grupo-ioa. atmosfera.unam.mx/pronosticos/index.php/oleaje)	52
6.2.	Productos gráficos generados con el sistema de pronóstico de oleaje para la variable de altura significante. En la parte superior se aprecia una imagen para el dominio del Océano mundial, mientras que en la parte de abajo se aprecia una imagen para el dominio del Pacífico mexicano (izquierda) y una para Golfo de México (derecha). Las tres gráficas fueron generadas con las salidas correspondientes a la hora 05 del pronóstico iniciado el	~0
6.3.	día 03 de diciembre de 2019 a las 00:00 GMT	53
6.4.	2019 a las 00:00 GMT	54
	$00:00 \mathrm{GMT}$	55

6.5.	Productos gráficos generados con el sistema de pronóstico de oleaje para la variable de periodo de marejada. Las gráficas para los tres dominios fueron generadas con las salidas correspondientes a la hora 05 del pronóstico iniciado el día 03 de diciembre de 2019 a las 00:00 GMT.
6.6.	Diagramas anuales de dispersión para las ocho boyas analizadas para la magitud del vien- to. El color de los puntos representa el intervalo de horas de simulación al que pertenece el dato, la línea roja indica la recta que mejor se ajusta a los datos y la línea gris indica la función identidad $(y = x)$
6.7.	Índices estadísticos para WSP como función del mes para todas las boyas analizadas. 59
6.8.	Índices estadísticos para WSP como función del intervalo de simulación 60
6.10. 6.9.	Índices estadísticos para H_S como función del mes para todas las boyas analizadas 61 Diagramas anuales de dispersión para las ocho boyas analizadas para la altura significante. El color de los puntos representa el intervalo de horas de simulación al que pertenece el dato, la línea roja indica la recta que mejor se ajusta a los datos y la línea gris indica la función identidad $(u - x)$
6 11	Índices estadísticos para H_c como función del intervalo de simulación 63
6.12.	Serie de tiempo y diagrama de dispersión para HS en la boya 42055 para el periodo de simulación de 24 horas (arriba) y 120 horas (abajo).
6.13.	Diagramas anuales de dispersión para las ocho boyas analizadas para el periodo de oleaje. El color de los puntos representa el intervalo de horas de simulación al que pertenece el dato, la línea roja indica la recta que mejor se ajusta a los datos y la línea gris indica la función identidad (u m)
6 14	Índicos estadísticos para T como función del mos para todas las hovas analizadas 66
6.14. 6.15. 6.16.	Índices estadísticos para T como función del mes para todas las boyas analizadas 60 Índices estadísticos para T como función del mes para todas las boyas analizadas 67 Serie de tiempo y diagrama de dispersión para T en la boya 42055 para el periodo de
	simulación de 24 horas (arriba) y 120 horas (abajo)
6.17.	Series anuales de dirección de oleaje
6.18.	Indices estadísticos para DIR como función del mes para todas las boyas analizadas 69
6.19. 6.20.	Indices estadísticos para DIR como función del intervalo de simulación
6.21.	(arriba) y 120 horas (abajo) en el mes de marzo
A.1. A.2. A.3. A.4. A.5. A.6. A.7. A.8.	Registros de la boya 42001 para el año 2015. .

Índice de cuadros

4.1.	Descripción de los dominios utilizados en el sistema de pronóstico de oleaje implementado con el modelo WWIII	32
5.1.	Descripción de las variables de oleaje monitoreadas por las boyas en el Golfo de México.	40
5.2.	Ubicación y porcentaje de registros de las boyas para cada variable de interés	41

Capítulo 1 Introducción

Las olas generadas por el viento son una de las manifestaciones más claras del estado del mar y han sido un fenómeno muy estudiado desde hace mucho tiempo. Se tiene registro de su observación desde la época de los antiguos griegos. Aristóteles (384-322 AC), por ejemplo, estaba al tanto de que la atmósfera interactuaba con el océano a través del viento y que el resultado de esa interacción eran las olas (Massel, 1996).

El oleaje es un proceso que contirubuye al funcionamiento del medio marino y el sistema climático. Actúa como generador de turbulencia, la cual estimula los flujos de energía y gases entre la atmósfera y el océano, acoplando ambos sistemas y ayudándoles a mantener su dinámica. Esta turbulencia es un importante agente de mezcla que mantiene la dinámica de la capa límite océanica. Además, el oleaje es el mecanismo intermediario en la transferencia de momento de la atmósfera al océano, pues la mayor parte de la energía del viento es transferida a las olas y luego al océano cuando estas rompen (Figura 1.1). Por otro lado, la presencia de rugosidad aumenta la superficie efectiva de intercambio de energía entre el océano y la atmósfera y altera los patrones de presión a escala sinóptica, modificando los flujos de viento sobre el océano (Affholder and Valiron, 2001; Cavaleri et al., 2012). El oleaje también funge como modelador del entorno, influyendo en la evolución de las líneas de costa (Lionello, 2012) y en la distribución y estrategias que adoptan los organismos que viven sometidos a condiciones de oleaje intenso (Blanchette et al., 2008; Lionello, 2012).

Dados los procesos en los que interfiere el oleaje y su capacidad para afectar severamente las actividades marítimas de importancia socio-económica, se ha convertido en un objeto de estudio de relevancia para el hombre. Su conocimiento tiene aplicación en campos que van desde la seguridad de la navegación hasta proyectos de ingeniería costera (Ardhuin and Orfila, 2018). Además, entre las razones de su importancia destaca su potencial destructivo, ya que la energía que transportan las olas es tal que puede causar daños a embarcaciones y a algunas estructuras fijas (Lefèvre and Cotton, 2001). Esto fue de particular interés durante la Segunda Guerra Mundial (1939-1945), cuando el conocimiento del estado del mar era esencial para llevar a cabo operaciones de aterrizaje en costas (Komen et al., 1994). El oleaje también modula actividades como la pesca: su intensidad puede limitar la navegación de embarcaciones pequeñas e influir en la profundidad local de la capa de mezcla, factores que de forma conjunta determinan las especies que se pueden capturar en una cierta temporada (Affholder and Valiron, 2001).

Por otro lado, el oleaje también es una fuente de energía que se ha buscado aprovechar desde hace mucho tiempo. La primera patente conocida enfocada al aprovechamiento de la energía del oleaje data de 1799, en París, y era de un dispositivo que se colocaría en una costa con el fin de bombear agua dulce a un pueblo cercano. Desde entonces, múltiples esfuerzos se han dado en este sentido (Alcorn, 2014), a pesar de los cuales sigue siendo un recurso potencial cuyas formas de explotación no han sido bien



Figura 1.1: Esquema de las capas límites en la interfase océano-atmósfera. Modificado de Konda (sf) y Brumer (2017).

exploradas (Guiberteau et al., 2012). Sin embargo, ante la reciente preocupación causada por el cambio climático y el consumo de combustibles fósiles, ha habido un interés renovado en el oleaje como fuente de energía. La satisfacción de este interés demanda un análisis del régimen climático del oleaje, el cual se ha buscado satisfacer a través del uso de modelos numéricos (Arinaga and Cheung, 2012).

Además, también se ha postulado utilizar el oleaje como una solución para las demandas energéticas de las plataformas petroleras que operan en el mar (Guiberteau et al., 2012). Para un país como México, que depende en gran medida de la explotación del petróleo, esto abre la perspectiva de la importancia de contar con información certera de las condiciones de oleaje en aguas nacionales, en particular en el Golfo de México. Esto no solo por el potencial que tiene el oleaje como fuente de energía, sino también porque puede representar un riesgo para las operaciones de dichas plataformas, en particular en condiciones de eventos extremos. De hecho, Panchang et al. (2013) reportan que durante las temporadas de huracanes de 2004 a 2008, se presentaron alturas significantes de olas de hasta casi 18 m, muy por encima de los 11 m caracterizados para un periodo de retorno de 100 años. Ya que el diseño de las plataformas petroleras oceánicas se basa en las alturas significantes de este periodo de retorno, múltiples instalaciones sufrieron daños ante la presencia del oleaje extraordinariamente alto.

Por todo lo anterior, el monitoreo y la predicción de las condiciones del estado del mar son de vital importancia para el país, tanto desde una perspectiva económica como de protección civil. Con esto como motivación, en 2016, el Grupo de Interacción Océano-Atmósfera (Grupo IOA) del Centro de Ciencias de la Atmósfera (CCA) comenzó un proyecto para el Centro Nacional de Prevención de Desastres (CENAPRED), para desarrollar un Sistema Operacional de Pronóstico Numérico de las condiciones meteorológicas, de oleaje y de marea de tormenta para el territorio nacional. Para este proyecto, el Grupo IOA implementó tres modelos numéricos: el modelo meteorológico Weather Research and Forecasting (WRF), el modelo de marea de tormenta Advanced Circulation (ADCIRC) y el modelo de oleaje WAVEWATCH III (WWIII).

Es en el marco de este proyecto que se desarrolló la presente tesis, la cual consistió en implementar

de forma operativa el modelo numérico de oleaje y, posteriormente, realizar la evaluación de la calidad de su desempeño.

Objetivos:

- 1. Implementar el modelo numérico de oleaje WWIII en un equipo de supercómputo para que se ejecute diariamente en modo pronóstico de manera operativa.
- 2. Evaluar el desempeño del modelo numérico de oleaje WWIII para pronosticar el campo de oleaje en el Golfo de México mediante una comparación con datos de boyas.

Dentro de estos objetivos generales, se enmarcan los siguientes objetivos particulares:

- a) Diseñar una configuración del modelo adecuada para el proyecto e implementarla.
- b) Generar un conjunto de herramientas de software (scripts) que permitan realizar la ejecución automática del modelo WWIII.
- c) Generar un conjunto de herramientas de software (scripts) que permitan realizar la comparación estadística del modelo WWIII con observaciones de boyas y presentar los resultados de forma gráfica.
- d) Determinar la existencia de patrones espaciales en la calidad del desempeño del modelo a través de la comparación de los resultados para distintas boyas.
- e) Determinar la existencia de patrones temporales en el desempeño del modelo como función del tiempo de pronóstico.
- f) Determinar la existencia de patrones estacionales en la calidad del desepeño del modelo.
- g) Discutir las posibles fuentes del error del modelo.

Para cumplir con estos objetivos, en el Capítulo 2 se presentan las bases teóricas del estudio y modelado numérico del oleaje; en el Capítulo 3 se hace una revisión bibliográfica de los principales métodos implementados y resultados obtenidos de la validación de diversos modelos de oleaje; en el Capítulo 4 se describe con detalle el modelo numérico WWIII y la implementación que se hizo; en el Capítulo 5 se describe el método seguido para realizar la validación, así como la interpretación de las métricas utilizadas; en el Capítulo 6 se muestran y discuten los resultados de la validación; y finalmente, en el Capítulo 7 se sintetizan las principales contribuciones y conclusiones de este trabajo.

Capítulo 2 Marco teórico

La superficie del mar está sometida a la constante influencia de fuerzas externas que alteran su aspecto al generar ondas superficiales. La principal causa de estas perturbaciones es la acción del viento, el cual le transfiere energía al agua al soplar sobre su superficie (Massel, 1996). Durante este proceso de intercambio de energía, se generan ondas con diferentes características, las cuales se combinan para formar un campo de oleaje más complejo. En adición a estas perturbaciones formadas localmente, sistemas de olas provenientes de regiones lejanas pueden también hacer una aportación al oleaje observado en un punto en un cierto momento.

En este capítulo se revisa la clasificación de las perturbaciones de la superficie del mar, los mecanismos de formación de las olas y la incorporación de estos conceptos en la evolución histórica que siguió el desarrollo de los modelos numéricos del oleaje.

2.1. Conformación del campo de oleaje

Aunque existen diferentes forzantes externos al océano que pueden generar olas superficiales, el más importante de ellos es el viento. La generación de olas asociadas a él se da en la capa límite océanica y depende básicamente de tres aspectos: la intensidad del viento (entendida como su rápidez), la duración y el *fetch* (es decir, la máxima longitud de la superficie en la que el viento sopla libremente). Cuanto más intenso, duradero y mayor *fetch* tenga el viento, más oportunidad tendrá de desarrollar un campo de oleaje (ver Figura 2.1).



Figura 2.1: Esquema de formación de olas por viento. Cuando las olas escapan de la región de influencia del viento, pasan a formar parte del *swell* también llamado *mar de fondo*.

Las olas generadas por viento se pueden clasificar en tres tipos, de acuerdo a la fuerza de restitución a la que están asociadas:

- 1. Ondas capilares. Las fluctuaciones turbulentas del campo de presión que se encuentra por encima de la superficie del agua permiten que, por momentos, haya flujos de aire en dirección vertical. Dichos flujos generan pequeños abultamientos de agua, los cuales reciben nombre de olas capilares (Massel, 1996). Este tipo de ondas tienen longitudes de onda de menos de 1 cm y la fuerza de restitución que las mantiene es la tensión superficial del agua (Lefèvre and Cotton, 2001).
- 2. Ondas capilares-de gravedad. Son ondas cuya longitud de onda oscila entre 1 y 10 cm, y pueden ser moduladas por ondas más largas. Por esto, tanto la tensión superficial como la gravedad juegan un papel como fuerza de restitución (Lefèvre and Cotton, 2001).
- 3. Ondas de gravedad. Estas ondas presentan longitudes de onda que varían entre 10 cm y 1 km. Suelen ser las que dominan el aspecto del mar y su fuerza de restitución es la gravedad (Lefèvre and Cotton, 2001).



Figura 2.2: Clasificación de ondas oceánicas superficiales. En rojo se indica la fuerza de restitución principal para cada intervalo. Modificado de Massel (1996) y Li (2010).

Estos tres tipos de ola constituyen el oleaje local (también conocido como *wind sea*) y la principal diferencia entre ellos es su duración: mientras que las olas capilares se disipan rápidamente cuando el viento deja de soplar, las olas de gravedad se pueden mantener por periodos más prolongados (Kinsman, 1984). Gracias a su duración, las olas de gravedad pueden viajar hacia los alrededores de su zona de formación.

Conforme viajan, un conjunto de olas producidas en una misma región se irá transformando, pues se verá afectado por diferentes eventos de interferencia. A pesar de esto, los paquetes de onda siempre conservarán una distribución en la que las ondas más largas viajen en la parte frontal del tren y las más cortas en la parte de atrás. Este arreglo obedece a que, en aguas profundas, la velocidad de fase es mayor para ondas con mayor longitud de onda. Si las olas llegan a escapar de su región de formación, dejan de ser parte del oleaje local y pasan a denominarse *swell*, que comunmente se traduce como *mar de fondo*. Estas olas pueden viajar grandes distacias, hasta encontrar un obstáculo que disipe su energía (Kinsman, 1984). De este modo, el campo de oleaje en un punto del océano es una combinación cuasi lineal de los distintos tipos de olas formadas localmente y una componente de *swell* (Lefèvre and Cotton, 2001). Esta componente tiene la característica de que puede moverse en direcciones arbitrarias con respecto a las componentes locales, e incluso dominar el campo de oleaje si el viento es muy débil (Sullivan and McWilliams, 2010).

Esta composición del campo de oleaje permite que la superficie del océano sea modelada como la suma de un gran número de armónicos senoidales, cada uno de ellos descrito por:

$$\eta(x, y, t) = a \sin\left(\frac{2\pi}{T}t - \frac{2\pi}{L}x\cos\theta - \frac{2\pi}{L}y\cos\theta\right)$$
(2.1)

donde t es el tiempo, a = H/2 es la amplitud, L es la longitud de onda, T es el periodo y θ es la dirección de propagación de la onda (ver Figura 2.3). Considerando que la frecuencia, amplitud y fase de cada armónico son aleatorias, los cambios en la elevación de la superficie del mar (η) quedan dados por:

$$\eta(x, y, t) = \sum_{i=1}^{N} \sum_{j=1}^{M} a_{i,j} \sin\left(\omega_i t - k_i x \cos \theta_j - k_i y \sin \theta_j + \phi_{i,j}\right)$$
(2.2)

donde N y M son números muy grandes, $\omega = 2\pi/T$ es la frecuencia radial y $k = 2\pi/L$ es el número de onda. Tanto la amplitud $a_{i,j}$ como la fase $\phi_{i,j}$ ($0 < \phi_{i,j} \le 2\pi$) son variables aleatorias (Holthuijsen, 2007). Así, cada componente de la ecuación anterior queda descrita por el índice *i* para el número de onda y por el índice *j* para la dirección.



Figura 2.3: Atributos utilizados para describir una onda armónica.

2.2. Teorías sobre la formación de las olas

A pesar de que el estudio del oleaje se remonta a mucho tiempo atrás, no fue sino a hasta el siglo XX que se hicieron los avances más importantes en el desarrollo de la descripción teórica-matemática del fenómeno. Específicamente en 1957, John W. Miles y Owen M. Phillips publicaron (de forma independiente) sus teorías sobre la generación y crecimiento del oleaje generado por viento (Massel, 1996). Ambas teorías han sido unificadas en la actual teoría de oleaje.

La teoría de Phillips propone que las variaciones turbulentas del campo de presión que se encuentra encima de la superficie del mar generan las primeras olas. Una vez formadas estas pequeñas perturbaciones, habrá una parte de ellas que viaje a una velocidad igual a la velocidad media del campo de viento que se encuentra por encima. Esto generará un efecto de resonancia por el cual el viento seguirá transmitiendo energía al agua, permitiendo que las olas crezcan (ver Figura 2.4 paneles A y B). Dicho efecto es posible porque esta teoría supone que el campo de presión por encima de la superficie del mar no se ve afectado por la formación de las olas (Affholder and Valiron, 2001). El desarrollo matemático de este proceso lleva a determinar que el creciemiento de las olas es lineal y, aunque es útil para describir la fase de formación de las olas, resulta insuficiente para predecir el crecimiento posterior de las mismas (Massel, 1996).

De manera contraria a Phillips, Miles supuso en su teoría que, una vez que se han formado las primeras olas, su presencia modifica los flujos de aire por encima de ellas. Se puede pensar que la modificación de este flujo se da en dos etapas. Primero, por encima de las crestas se generan mínimos locales de presión, mientras que por encima de los valles se forman máximos locales. Estas diferencias tienen efecto hasta una cierta altura crítica. En dicha altura, el flujo de aire viaja más rápido encima de las crestas, adelantándose a ellas, y luego retrocendiendo debido a la alta presión de los valles. Así pues, en la segunda etapa, en la parte frontal de la cresta existe un flujo ascendente, mientras que en la parte posterior habrá un flujo descendente (Figura 2.4 paneles B y C). Lo anterior provoca una transferencia de energía neta que va desde el viento hacia el agua. Este mecanismo lleva a un crecimiento exponencial, consistente con las observaciones hechas hasta la fecha (Lighthill, 1962).



Figura 2.4: Mecanismo de formación de olas de Miles y Phillips. Modificado de Sawan (2003).

Una vez que las olas se han formado, comienza a haber interacción entre ellas. A través de esta interacción, las olas de menor longitud de onda le transfieren energía a las de mayor longitud. Este proceso lleva a que las olas aumenten su velocidad hasta superar la del viento (Li, 2010). A partir del momento en el que las velocidades se igualan, deja de haber transferencia neta de energía desde el viento hacia el agua y se dice que el mar está completamente desarrollado (*fully developed sea*, Holmes 2001). Dicho estado se alcanza cuando el viento ha soplado por un tiempo y sobre un área lo suficientemente grandes. En este punto, la distribución de energía en el campo de oleaje (*espectro*) solo depende de la velocidad del viento (Folley, 2017).

2.3. Modelación numérica del oleaje

Incluso antes de que fueran desarrollados los modelos teóricos de oleaje, ya existía un gran interés por realizar la predicción del estado de oleaje. Dicho interés surgió durante la Segunda Guerra Mundial como resultado de la necesidad de conocer el estado del mar durante operaciones de aterrizaje en el océano. En aquel momento, las predicciones estaban basadas en leyes empíricas propuestas por H.U. Sverdrup y W.H. Munk (Janssen, 2008).

El primer modelo basado en una ecuación de transporte fue hecho por R. Gelci y colaboradores y fue publicado en 1957. En este trabajo propuesieron la siguiente ecuación de balance de energía:

$$\frac{\partial F}{\partial t} + v \cdot \nabla F = S, \qquad \text{con } S = S_{in} + S_{nl} + S_{ds}$$
(2.3)

donde $F(f, \theta; x, t)$ es la función que describe las variaciones de la superficie del mar, v es la velocidad de grupo, S_{in} son las entradas de energía por viento, S_{nl} son las transferencias de energía por mecanismos no lineales y S_{ds} es el término de disipación (The SWAMP Group, 2013). Esta ecuación indica que la evolución del espectro de oleaje, F, está determinada por los efectos de la advección y la presencia de los llamados términos fuente, S, los cuales son funciones escalares que se añaden a las ecuaciones conservativas para describir efectos no lineales (Janssen, 2008).

Aunque no se entendían a detalle los términos fuente, en especial los no lineales, los modelos de las décadas de los cincuenta y sesenta buscaron la manera de incorporarlos. La mayoría de esos modelos, que ahora se denominan de **primera generación**, ignoraban completamente los términos no lineales, trataban el término de dispersión como un parámetro de ajuste que evitaba un crecimiento desmedido de las olas y aproximaban el término de viento como $S_{in} = A + B \cdot F$, donde A representaba el forzamiento turbulento de Phillips y $B \cdot F$ el mecanismo de retroalimentación de Miles. Estos modelos eran forzados a funcionar, pues para que lograran representar los valores medidos en las observaciones, los parámetros A y B eran escogidos órdenes de magnitud más grandes que los estimados teóricamente por los autores (The SWAMP Group, 2013).

Fue hasta la década de los setenta que se encontró la clave de la carencia de estos primeros modelos. En aquel periodo se llevaron a cabo una importante cantidad de campañas de medición del crecimiento de olas en sitios con condiciones de *fetch* e intensidad de viento bien acotadas, así como experimentos de simulación de oleaje en tanques. Con la información obtenida, se pudo determinar que la principal fuente de energía durante la fase de crecimiento de las olas de baja frecuencia es la transmitida de forma no lineal desde las olas de mayor frecuencia (The SWAMP Group, 2013).

A pesar de este avance conceptual, la capacidad computacional de aquel momento limitaba la posibilidad de representar interacciones no lineales de manera explícita. Sin embargo, hubo una propiedad de estas interacciones que permitió incluirlas. Se determinó que cuando el oleaje se está desarrollando, las transferencias de energía de forma no lineal son los procesos más rápidos, de modo que controlan la forma que toma el espectro F. Esto permitió crear una función de distribución que se podía aplicar de manera prácticamente universal, la cual representaba a la mayoría de los sistemas en cuasi-equilibrio. Dicha función dependía, al final, de un sólo parámetro que podía ser la energía total del sistema o la frecuencia en la que se da el pico de energía (Janssen, 2008). Los modelos que utilizan este tipo de parametrizaciones implícitas se denominan de **segunda generación**.

Los modelos de segunda generación fueron capaces de representar muy bien la mayoría de los escenarios de oleaje, tras un importante esfuerzo en la realización de calibraciones. Sin embargo, seguía habiendo interacciones que no podían resolver (Lefèvre and Cotton, 2001). Particularmente, cuando los campos de viento eran demasiado intensos, los términos advectivos ganaban importancia frente a los no lineales, de modo que el campo se desviaba de las distribuciones propuestas. Esto dejó abierta la necesidad de mejorar la forma en que se parametrizaban los efectos no lineales (Janssen, 2008).

En 1985, S. Hasselmann y colaboradores propusieron la primera parametrización explícita para las interacciones no lineales, denominada *Aproximación de Interacciones Discretas* (DIA, por sus siglas en inglés). Esta parametrización, eficiente y capaz de reproducir adecuadamente la evolución del espectro, junto con los avances en la capacidad de cómputo, llevaron a que en 1988 se presentara el primer modelo de **tercera generación**: el WAM (Janssen, 2008).

Hasta la fecha, son los modelos de tercera generación los que se utilizan en el pronóstico de oleaje. Por supuesto, desde la propuesta de Hasselmann y sus colaboradores hasta ahora, se ha trabajado en proponer nuevas y mejores parametrizaciones. Además, han surgido más modelos de tercera generación. Uno de ellos es el modelo Wave Watch III, desarrollado por los National Centers for Environmental Prediction de la NOAA, el cual fue utilizado y evaluado para este trabajo.

Capítulo 3

Revisión de trabajos de validación de modelos de oleaje

Los resultados obtenidos por un modelo numérico deben ser validados con el fin de determinar su utilidad en la predicción o reconstrucción de las variables que se simulan. El proceso de validación se entiende como una evaluación de la concordancia entre estimaciones generadas por un modelo y observaciones (Lee, 2015). Ya que esta concordancia depende de las condiciones particulares de la región modelada, así como de la forma en que se hace la implementación del modelo, diversos trabajos relacionados con la validación del modelo WWIII para diferentes regiones y escenarios han sido publicados. En este capítulo se hace una breve revisión de algunos de estos trabajos, con el fin de ilustrar los distintos enfoques que se han utilizado en la literatura.

3.1. Métodos de validación

Aunque el principal objetivo de los trabajos de validación es determinar de forma cuantitativa el desempeño de un modelo, también pueden ser utilizados para otros fines como identificar cuáles son las principales fuentes de error (Bidlot et al. 2002; Chawla et al. 2009; Chu et al. 2004; Tolman 2002, etc.), determinar si existen patrones estacionales o espaciales en la distribución del error (Chawla et al. 2013, 2009; Chu et al. 2004; Mentaschi et al. 2015; Tolman 2002, etc.) e incluso detectar irregularidades en las observaciones contra las que se hace la comparación del modelo (Chawla et al. 2013, 2009). De este modo, el fin del estudio es el que determina el método a seguir para realizar la validación.

Para realizar la evaluación de los modelos de oleaje existen dos enfoques principales: el primero consiste en comparar los denominados parámetros integrales (que se definen en la siguiente sección) contra observaciones a través de diferentes métricas estadísticas, y el segundo en aplicar un algoritmo de partición que permita separar el espectro de oleaje en diferentes eventos caracterizados por su nivel energético, que después serán comparados con las observaciones espectrales que les correspondan.

3.1.1. Evaluación de los parámetros integrales

Los parámetros integrales son aquellas variables que describen el estado del campo de oleaje que resulta de combinar todas las fuentes y eventos que generan olas. Esto significa que representan la suma de los efectos del viento local con los de eventos lejanos que producen *swell*. Estos parámetros están definidos en el dominio espacial F(x,y) y, en este sentido, son los que mejor describen lo que un observador percibe como el estado del oleaje. En las salidas del modelo implementado, la altura significante (HS), el periodo (T01) y la dirección de oleaje (DIR) corresponden a este tipo de parámetros.

Las evaluaciones más comúnes consisten en comparar directamente las observaciones que se tienen del campo de oleaje contra las salidas del modelo para cada una de las variables de interés. Dicha comparación se lleva a cabo a través de correlaciones temporales o de análisis cuantil-cuantil (Q-Q). La correlación temporal consiste en la comparación de series de tiempo y permite identificar qué tanto se asemejan las predicciones a las observaciones a lo largo del periodo simulado. Por otro lado, los análisis Q-Q permiten identificar si la distribución de las magnitudes de las predicciones es correcta al compararla con la de las observaciones (Hanson et al., 2009).

En ambos tipos de análisis, la comparación cuantitativa se hace a través de métricas estadísticas, tales como el sesgo, la raíz del error cuadrático medio (*Root mean square error*, RMSE), el coeficiente de correlación (lineal o compleja) u otros índices que resuman la información dada por varios estadísticos, como el índice de concordancia (*Index of agreement, IoA*) o el índice de dispersión (*Scatter index, SI*).Una descripción detallada de estas métricas se dará más adelante en el Capítulo 5.

3.1.2. Evaluación de las particiones del espectro

Como se describió en el capítulo anterior, el campo de oleaje en un punto del espacio es el resultado de la interacción del denominado *wind sea* con diferentes eventos de *swell* generados en regiones lejanas. Ser capaces de identificar y aislar cada uno de estos componentes es a lo que se denomina "particionar el espectro de oleaje".

El procedimiento más recurrente en la literatura para realizar esta identificación es el descrito por Hanson and Phillips (2001), mejorado en Hanson et al. (2009), el cual consta de manera general de cinco etapas:

- 1. En un cierto dominio del espectro de oleaje, $S(\theta, f)$, se identifican máximos de energía locales.
- 2. Usando un criterio de *edad de ola*, se identifican los eventos que se consideran como parte del oleaje por viento local. Este criterio está dado por la siguiente relación:

$$f_p \ge \frac{g}{2\pi} [1.5U_{10}\cos\delta], \qquad 0 \le \delta \le \frac{\pi}{2}$$
 (3.1)

donde U_{10} es la magnitud de la velocidad del viento a 10 metros, δ es el ángulo entre el viento y las olas y f_p es la frecuencia pico. El factor de 1.5 se utiliza con la finalidad de asegurar que todos los eventos locales se incluyan. Esta relación define una parábola en el espacio $S(\theta, f)$ y se considera que cualquier partición identificada en el paso anterior que cae en ella, corresponde a un evento de oleaje local.

3. A continuación, se usan dos criterios para determinar los máximos adyacentes que corresponden a un mismo evento de *swell*:

a) Separación entre picos: se calcula la distancia euclidiana (en el plano espectral) entre los puntos y luego se compara con la dispersión (*spread*) de cada pico. De forma general, la dispersión es una medida del área de influencia de cada pico de energía. De este modo, si se cumple que la distancia entre los picos es menor que la dispersión de alguno de ellos, se considera que son parte del mismo evento y se combinan en una sola partición.

b) Mínimo entre picos: el menor de los picos es multiplicado por un factor, denominado factor de pico mínimo (ζ). Si el resultado de esta multiplicación es menor que el valor mínimo entre los picos, se considera que estos últimos corresponden al mismo evento.

4. Por último, se define un límite de energía mínimo, de modo que cualquier partición resultante que esté por debajo de dicho umbral es eliminada, pues se considera que representa ruido.

A partir del algoritmo anterior, se obtiene una matriz plantilla, $T(f, \theta)$, que asigna un número de partición a cada punto del espectro S. En su momento, el algoritmo fue creado para ser aplicado a las observaciones y a partir de ellas obtener la matriz T, que luego se empleaba para identificar en las salidas del modelo las particiones correspondientes (Hanson et al., 2009). En el caso de la versión de WWIII utilizada en este trabajo, ya viene implementado un algoritmo de partición, de modo que los parámetros de las particiones son también parte de las salidas del modelo.

Una vez determinadas las particiones, tanto de las observaciones como del modelo, la comparación de la información puede realizarse con las mismas métricas que se utilizan para los parámetros integrales.

Desde su publicación, este algoritmo se ha considerado de gran utilidad porque se ha visto que los parámetros integrales del oleaje no proveen suficiente información de las características de los procesos físicos a diferentes escalas espectrales (Bi et al., 2015). Además, permite identificar las fuentes de error en el modelo, pues se pueden rastrear las partes del espectro que no se están representando adecuadamente (Hanson et al., 2009).

En virtud de que el sistema de pronóstico evaluado en este trabajo tiene como principal objetivo el alertamiento ante la posibilidad de eventos extremos, es de mayor interés el estado general del campo de oleaje, por lo que se decidió hacer la validación directamente con los parámetros integrales.

3.2. Trabajos de validación realizados previamente

Teniendo en cuenta las diferencias conceptuales que implican las formas de validar antes descritas, a continuación se presenta un resumen de algunos trabajos hechos sobre el tema. La mayoría de ellos son específicamente de validación del modelo WWIII, mientras que otros son comparaciones de varios modelos. Destaca que únicamente los trabajos de Chawla et al. (2009) y Bi et al. (2015) utilizan el método de partición del espectro. El objetivo de esta revisión es presentar los antecedentes necesarios para realizar la discusión de los resultados obtenidos en este trabajo.

Los artículos y reportes han sido clasificados en dos categorías de acuerdo a su objetivo principal: validación de implentaciones operativas y validación como un análisis de sensibilidad.

3.2.1. Validación de implementaciones operativas de modelos de oleaje.

El primero de los trabajos de esta categoría es un caso particular en el que se compararon modelos de varios centros meteorológicos del mundo para identificar las fortalezas y las fuentes de error de diferentes implementaciones. Los demás son, en su mayoría, una comparación entre los modelos WAM y WWIII. Estos trabajos se dieron en el marco de la transición del WAM al WWIII como modelo de oleaje operativo del NCEP (Rogers et al., 2012).

1. Intercomparison of the Performance of Operational Ocean Wave Forecasting Systems with Buoy Data, Bidlot et al. (2002)

En 1995, cinco centros meteorológicos comenzaron un acuerdo para intercambiar los resultados de sus modelos de oleaje en determinadas ubicaciones. Este acuerdo permitía a los participantes evaluar el desempeño de sus sistemas de pronóstico y compararlo con el de los demás, de modo que se pudieran determinar los puntos fuertes de cada sistema. Los centros participantes fueron el *European Centre for Medium-Range Weather Forecasts* (ECMWF), la *Met Office* de Reino Unido, el *Fleet Numerical Meteorology and Oceanography Center* (FNMOC) de Estados Unidos, el *Atmospheric Environment Service* (AES) de Canadá y el *National Center for Environmental Prediction* (NCEP) de Estados Unidos.

Describir los detalles de la implementación de cada modelo está más allá del alcance de este trabajo, pero hay tres puntos que deben resaltarse: 1) todos los centros utilizaban el modelo WAM en su versión 4.0, salvo la Met Office, que utilizaba un modelo de segunda generación; 2) el ECMWF utilizaba la resolución más fina, siendo su malla de 55 km x 55 km; 3) el ECMWF era el único en utilizar física de aguas someras.

De manera general, los autores de este trabajo realizaron la validación de cada modelo utilizandos los parámetros integrales, pero con tres enfoques distintos: primero se hicieron correlaciones temporales, luego se estudiaron los estadísticos obtenidos como función de las cantidades observadas (por ejemplo gráficas de coeficiente de correlación lineal como función de la altura significante). Por último se hizo una comparación por zonas geográficas, para buscar patrones regionales en el desempeño del modelo. Las métricas estadísticas utilizadas fueron el sesgo, la raíz del error cuadrático medio (RMSE), el coeficiente de correlación lineal (ρ) y el índice de dispersión (SI, por sus siglas en inglés).

Los resultados mostraron que todas las implementaciones de modelos de tercera generación (WAM v.4.0) tenían una tendencia a subestimar la altura de oleaje. Sin embargo, al comparar los vientos, se encontró que los utilizados por la *Met Office* presentaban una sobreestimación de la magnitud, lo que podría haber jugado como factor de correción en la predicción de las alturas de oleaje. De manera general, el sistema del ECMWF mostraba el mejor desempeño, lo que parecía estar ligado a que presentaba también el menor error en el pronóstico de viento usado para forzar el modelo WAM.

Se concluyó que la incorporación de datos de altimetría, el incremento de la resolución angular de la partición del espectro y un ajuste de los datos del viento previo a que sean usados como forzantes, representarían una importante mejora en el desempeño de los pronósticos.

2. Validation of WAVEWATCH III version 1.15 for a global domain, Tolman (2002)

El objetivo de este trabajo era evaluar el desempeño del WWIII v.1.15 en el entorno de un sistema funcionando de manera operativa, para después compararlo con el desempeño del WAM. Ambas implementaciones utilizaban las mismas resoluciones espaciales y espectrales, así como los mismos forzamientos de viento, los cuales eran obtenidos del *Global Data Assimilation System* (GDAS) del NCEP. La única diferencia entre las simulaciones era que en las de WAM no se incluyó la

presencia de hielo. Sin embargo, los análisis se limitaron hasta la latitud de 60°S, de modo que se consideró que esta diferencia no afectaría de manera significativa la comparación de los modelos.

Los resultados de ambos modelos fueron comparados con las observaciones de varias boyas de la *World Meteorological Organization* (WMO), así como con datos del satélite ERS-1. Las métricas que se utilizaron para la evaluación fueron el sesgo, el error cuadrático medio y el SI.

En general, los resultados mostraron que el WWIII se desempeña mejor que el WAM, salvo en latitudes altas. Para ambos modelos, se obtuvieron índices de dispersión de cerca del 15 %, lo que se considera aceptable para un sistema de pronóstico. Sin embargo, el error cuadrático medio del WWIII fue menor que el de WAM, estando por debajo del 15 % en casi todo el dominio. El sesgo del WWIII fue también menor que el de WAM y no mostró ningún patrón de distribución espacial. El WWIII mostró ser mejor que el WAM para predecir el *swell* en zonas tropicales, así como para describir la distribución de eventos extremos. A pesar de este buen desempeño general, el modelo mostró una tendencia a sobreestimar las olas más bajas y a subestimar las más altas.

Finalmente, resaltó que en las regiones de altas latitudes del dominio (en el NE del Atlántico), el WAM mostró valores menores de RMSE, así como menor sesgo durante eventos extremos al ser comparado con boyas. Este resultado parece ser debido a que el WAM era calibrado utilizando datos de boyas, mientras que el WWIII lo era principalmente con datos satelitales.

3. A review of operational forecasing of wind generated waves by hurricane Isabel at NCEP, Tolman et al. (2004)

El objetivo de este estudio fue evaluar la habilidad del modelo WWIII, implementado de manera operativa, para predecir las condiciones de un evento extremo. El modelo estaba implementado en su versión 2.2 y tenía un dominio global con resolución de $1^{\circ}x1.25^{\circ}$, con diversos dominios anidados con resolución de 0.25° , los cuales eran activados según la demanda. Ya que el evento estudiado fue el huracán Isabel, que tocó la costa este de Estados Unidos en septiembre de 2003, los anidamientos utilizados para la validación fueron el denominado North Atlantic Hurricane (NAH) y el Western North Atlantic (WNA). Las simulaciones fueron forzadas con vientos obtenidos del modelo GFS.

Los resultados fueron validados contra datos de 15 boyas del *National Data Buoy Center*, así como con los del satélite Jason-1, considerando el RMSE, el sesgo y el SI.

Ambos anidamientos tuvieron buen desempeño hasta antes de que el huracán Isabel tocara tierra, lo que parece estar asociado al error en los vientos utilizados: las velocidades máximas de viento habían sido subestimadas por el modelo GFS, de modo que cuando el huracán se intensificó (lo cual ocurrió durante su llegada a tierra), la altura de las olas fue subestimada. Aún así, el NAH representó mejor las condiciones cuando Isabel alcanzó categoría 5, así como el *swell* asociado al evento. Los índices de dispersión del NAH también fueron menores al compararse con boyas. En el caso de la comparación con datos de altimetría, WNA presentó importantes subestimaciones de las alturas, mientras que el NAH presentó ligeras sobreestimaciones.

Estos resultados llevaron a concluir que la versión 2.22 del modelo necesitaba implementar mejoras en la forma en que resuelve los procesos de somerización y, en general, en las parametrizaciones que utiliza para eventos extremos.

4. Validation of a Multi-Grid WAVEWATCH III TM Modeling System, Chawla et al. (2009)

En este reporte se validó el pronóstico operativo del NCEP con WWIII que a partir del 2007 incorporó la posibilidad de usar múltiples mallas. El sistema implementado consistía en un arreglo de ocho mallas de diferente resolución, incluyendo un dominio global de 0.5° de resolución, todas forzadas por vientos GFS de resolución de 0.5° , que eran internamente interpolados a la resolución de cada malla.

El análisis estadístico se hizo utilizando datos del satélite Jason-1 y de las boyas del NBDC para los parámetros integrales y con una herramienta denominada *Interactive Model Evaluation and Diagnostics System* (IMEDS) para las particiones espectrales. El IMEDS consitía en una automatización del método de partición y evaluación de Hanson and Phillips (2001) descrito en la sección 3.1.2, y utilizaba datos de boyas para realizar las evaluaciones.

Para la altura significante, los promedios espaciales de las métricas de error no mostraron una variabilidad estacional. Sin embargo, al realizar mapas de sesgo para esta variable, encontraron que había un patrón estacional y espacial para el hemisferio norte. Observaron que en la parte oeste de las cuencas oceánicas, las cuales en general están dominadas por la formación de olas por vientos locales, se presentó un sesgo negativo. Por otro lado, en la parte este de las cuencas, la cual está dominada por eventos de *swell*, se presentó un sesgo positivo. Este patrón parece tener origen en el esquema de formación de olas que utiliza el modelo, el cual subestima el creciemiento inicial de las olas, pero una vez que la energía se transfiere a las particiones de *swell*, la disipación no es lo suficientemente rápida, generando un crecimiento mayor al real. Este patrón es, además, estacional, pues sólo se observa en los meses de invierno.

En el hemisferio sur, para el pronóstico de la altura significante se observó un sesgo positivo persistente a lo largo de varios años. La fuente más probable para este patrón fue identificada como un cambio de configuración hecho en el modelo GFS. Cuando se hizo un análisis estadístico para una serie de tiempo de 10 años (1999-2009) del viento en el hemisferio sur, se encontró que a partir de 2005 el modelo GFS presentaba un aumento en las magnitudes de las velocidades mayores. El incremento en el sesgo positivo de la altura significante en la región del HS coincide con este cambio en los vientos.

Estos resultados permitieron concluir que los términos de crecimiento de ola y disipación de la energía debían ser utilizados con mayor cuidado y no únicamente como parámetros de *ajuste*.

5. Validation Test Report for WAVEWATCH III, Rogers et al. (2012)

En 2012 se publicó un reporte de la validación del modelo WWIII versión 3.14 que fue realizada por el equipo del *Naval Research Laboratory*, de la U.S.NAVY, como un documento para respaldar el cambio del modelo de oleaje utilizado en el pronóstico operativo de la Naval Oceanographic

Office, que pasó de ser el WAM al WWIII.

Ambos modelos fueron implementados con nueve mallas de diferentes resoluciones, anidadas en una malla global de 0.5°. Algunas de las mallas, aunque representaban el mismo dominio en ambos modelos, tenían mayor resolución en el WWIII. Las salidas de los modelos fueron comparadas con datos de boyas del NBDC haciendo uso del sesgo, el RMSE y del SI, así como el índice de correlación lineal.

Los resultados mostraron que el desempeño del WWIII es, en general, mejor que el del WAM, pues el sesgo, el RMSE y el SI presentaron valores más pequeños para el WWIII. Sin embargo, los coeficientes de correlación fueron ligeramente mayores para el WAM.

En respuesta a los resultados de la comparación, se emitieron recomendaciones para mejorar el desempeño del WWIII:

- Implementar hielo en el pronóstico para mejorar el desempeño cerca de la Antártica.
- Asimilar datos modificando los archivos de restart, de manera que se incorpore a ellos información de altimetría.
- Incluir datos espectrales en la asimilación, para evitar que las modificaciones se propaguen en la dirección y a la velocidad equivocada.
- Recalibrar los términos fuente para que se ajusten a un tipo específico de datos (ya sea de altimetría o de boyas).
- Utilizar como término fuente de entradas de energía por viento (S_{in}) el esquema descrito por Ardhuin y colaboradores en 2010 (ST4 en el modelo) cuando se implemente un sistema con anidamientos.

3.2.2. Validación del modelo como un análisis de sensibilidad: vientos forzantes y parametrizaciones físicas

Muchos trabajos de validación comparan el desempeño del modelo cuando las condiciones en las que se ejecuta son ligeramente distintas, por ejemplo cuando se cambia la fuente de vientos o bien cuando se utilizan diferentes paquetes de parametrización de procesos físicos. En este sentido, estos trabajos pueden ser vistos también como un análisis de sensibilidad del modelo ante diferentes factores. En esta sección se revisan trabajos que hacen este tipo de análisis de sensibilidad.

1. The US Navy's Global wind-wave models: an investigation into sources of errors in low frequency energy predictions, Rogers (2002)

Este estudio se enfocó en buscar los errores de los dos modelos implementados en el pronóstico operativo de la U.S. Navy (WAM y WWIII) asociados a los vientos de entrada. En particular, se centró en las zonas del espectro de baja frecuencia, es decir, en zonas de *swell*, de las que se sabía presentaban una tendencia a subestimar las alturas de oleaje.

Para ambos modelos se utilizó una resolución de 1°, pues se consideró suficiente para ver el efecto del viento. Se utilizaron dos fuentes de viento en cada modelo: los GDAS del NCEP y los obtenidos

por el Navy Operational Global Atmospheric Prediction System (NOGAPS).

Los resultados fueron comparados contra datos de boyas, datos del satélite ERS-2 y datos del satélite TOPEX, a través del sesgo y el RMSE. Los posibles errores se clasificaron en tres categorías: numéricos y de resolución, de parametrizaciones físicas y de forzamientos.

Se determinó que los errores numéricos jugaban un papel despreciable en los errores de predicción a bajas frecuencias. En el caso de las parametrizaciones, aunque importantes, no cumplieron con un papel dominante. Así, la principal fuente de error de ambos modelos se debió a los errores heredados del modelo de viento, siendo las simulaciones forzadas con los vientos NOGAPS las más afectadas. Este resultado fue consistente con la validación realizada a los datos de viento, en la que se determinó que los vientos NOGAPS presentaban un marcado sesgo negativo a altas velocidades. Estos resultados eran esperados, pues trabajos previos ya habían determinado que el error sistemático de los modelos de oleaje era pequeño, y que la mayor parte del error se debe al campo forzante.

2. Wave modeling performance in the Gulf of Mexico and Western Caribbean: Wind reanalyses assessment, Appendini et al. (2013)

El objetivo de este trabajo era comparar el desempeño del modelo WWIII al utilizar diferentes fuentes de vientos, provenientes de datos de reanálisis: los North American Regional Reanalysis (NARR) de NCEP, los *Global Reanalysis* de NCEP/NCAR y los del *European Centre for Medium-Range Weather Forecasts* (ERA).

Las simulaciones se hicieron para los años 2005 y 2006 con un dominio que contenía diferentes anidamientos, cuyas resoluciones iban desde los 0.005° hasta los 0.06° . Por su lado, cada fuente de datos de viento tenía su propia resolución: los ERA tenían una resolución de 1.5°, los NARR de 0.375° , y los NCEP, de $1.8^{\circ}x1.9^{\circ}$.

Para cada simulación realizada, el análisis de error se llevó a cabo evaluando el sesgo, el RMSE, el SI y el coeficiente de correlación. Los resultados se compararon con datos de boyas de NDBC, separando el desempeño en condiciones normales del desempeño en condiciones de eventos extremos.

Los valores de correlación mostraron que los datos de ERA y NARR proveen una mejor representación promedio de las condiciones de oleaje, obteniendo valores de r mayores a 0.9. En particular, al analizar eventos extremos se encontró que los ciclones son mejor representados al usar datos de NARR. Estos resultados muestran que la resolución en los datos de los vientos juega un papel importante en el desempeño del modelo de oleaje, pues los mejores resultados se obtuvieron al forzar con los vientos de mayor resolución.

3. Evaluation of the simulation capability of the Wavewatch III model for Pacific Ocean wave, Bi et al. (2015)

Los autores trabajaron en la evaluación del desempeño del WWIII en la cuenca del Océano Pacífico, enfocándose en comparar cinco paquetes de términos fuentes disponibles en las versiones 3.14 y 4.18 del modelo: el Tolman y Chalikov (TC), el Babanin/Young/Donelan/Rogers/Zieger (ST6), dos formas de la parametrización propuesa por Ardhuin y colaboradores en 2010 (ST4 y ACC350) y el de Bidlot, Jassen y Abdallah (BJA).

El estudio se hizo siguiendo una versión modificada del método propuesto por Hanson et al. (2009) y las salidas del modelo fueron comparadas con tres fuentes de observaciones: datos de altimetría del satélite JASON-1, datos de las boyas del NDBC y datos del radar de apertura sintética que va a bordo del satélite ENVISAT.

Para las variables no direccionales, el análisis de error fue hecho a través del sesgo medio, el sesgo medio relativo, el RMSE y el SI. Por su lado, para la dirección del oleaje se utilizó el sesgo angular y la correlación circular. Todas la métricas anteriores fueron normalizadas y luego promediadas para generar indicadores generales del desempeño del modelo.

Entre sus resultados principales están que, para todos los esquemas utilizados, la altura significante de oleaje por viento local es mejor reproducida que la altura de particiones de *swell*. Sin embargo, cuando se implementan esquemas de disipación de *swell* (ST4 y ST6), se mejora el desempeño del modelo para simular estos sistemas, lo cual se ve reflejado en una mejora en la simulación de las latitudes bajas de la cuenca, pues son las zonas donde la influencia de los swell es más notoria. Además, obtuvieron que la dirección y el periodo fueron mejor representados que la altura significante en todos sus casos experimentales.

4. Performance evaluation of Wavewatch III in the Mediterranean Sea, Mentaschi et al. (2015)

El trabajo realizado por Mentaschi y colaboradores tuvo como principal objetivo evaluar cómo las características de la cuenca del Mediterráneo (compleja batimetría, área de fecth limitada y marcados procesos meteorológicos de mesoescala) afectan el desempeño del modelo WWIII en la región, así como determinar el paquete de parametrizaciones físicas que mejor reproduce las caracteríticas del lugar.

Para esto, llevaron a cabo 17 casos de estudio correspondientes a eventos de tormentas en la región, en los que evaluaron el desempeño del modelo utilizando tres esquemas de crecimiento-disipación de oleaje: el propuesto por Arduhin y colaboradores en 2010 (ACC350), el predeterminado del modelo que fue propuesto por Tolman y Chalikov en 1996 (TC) y el descrito por Bidlot y colaboradores en 2007 (BJA). Además de comparar estos esquemas, también evaluaron el efecto de la resolución del campo forzante (vientos) y de la batimetría en los resultados.

Las salidas del modelo fueron comparadas con datos de boyas haciendo uso de indicadores estadísticos puntuales que fueron el sesgo normalizado, el coeficiente de correlación, el error cuadrático medio simétricamente normalizado, así como el sesgo angular normalizado y el error cuadrático medio normalizado para la variable de dirección.

Sus resultados muestran que el modelo tiene el mejor desempeño en la cuenca del Mediterráneo cuando se utiliza el esquema de Ardhuin y colaboradores. También encontraron una tendencia en el sesgo de la altura significante, el cual decrece conforme dicha altura crece, volviéndose negativo para alturas mayores a 1.7 m. En general, se apreció una tendencia del modelo a sobreestimar

la altura significante en condiciones de calma, mientras que subestima en condiciones de tormenta.

En cuanto a las pruebas de resolución, encontraron que es suficiente tener vientos en una malla de 10 km para generar un pronóstico adecuado, pero para el caso de la batimetría es necesaria una resolución mayor que pueda representar adecuadamente las condiciones de la zona (como la presencia de pequeños archipiélagos) que pueden causar efectos locales de refracción y somerización, los cuales se ven reflejados en las boyas cercanas.

5. Evaluation of WAVEWATCH III performance with wind input and dissipation source terms using wave buoy measurements for October 2006 along the east Korean coast in the East Sea, Lee (2015)

El autor se enfocó en evaluar el desempeño del modelo WWIII para representar el oleaje de tormenta generado por sistemas de baja presión en movimiento, que se desarrollan durante el invierno en la costa este de la península coreana. Para encontrar la implementación que mejor simulara el oleaje bajo esas condiciones, probó tres esquemas de disipación para simular el proceso generado por el rompimiento de olas: el WAM3, el WAM4 y el de Tolman y Chalikov (TC).

Los salidas del modelo fueron comparadas puntualmente con cuatro boyas y los resultados fueron presentados a través de diagramas de Taylor, para lo que se calculó el RMSE, la desviación estándar y el coeficiente de correlación lineal. Además se presentaron datos del sesgo global y el índice de concordancia (IoA).

Aunque en general todos los paquetes generaron buenos resultados, el esquema WAM4 mostró el mejor desempeño de acuerdo con el índice de concordancia y el diagrama de Taylor. Además se observó que el esquema TC genera pérdidas de energía en frecuencias altas (mayores a 0.25 Hz), llevando a una subestimación de la altura significante en esa frecuencia, mientras que el WAM3 tiene el comportamiento contrario. Los resultados obtenidos con el esquema WAM4 son intermedios, lo cual es consistente con que haya mostrado mejor desempeño, pues genera un comportamiento más amortiguado en el proceso de crecimiento y disipación del oleaje.

3.3. Conclusiones del capítulo.

El objetivo principal de la validación de modelos en general es responder a la pregunta ¿qué tan bien reproduce el modelo a la realidad?. Sin embargo, en el proceso se obtiene mucha información útil, que responde a las preguntas de ¿qué factores afectan el desempeño del modelo? y ¿cómo corregir o mejorar un modelo y/o su implementación? Además, una vez que se han comparado resultados de diferentes estudios, se responde también a ¿qué se puede esperar de un modelo en particular?. En este sentido, entre los resultados obtenidos por los trabajos antes expuestos, y que son de interés para esta tesis, resaltan los siguientes:

• La calidad de los resultados de los modelos de oleaje depende directamente de la calidad de los datos de viento utilizados como forzantes, pues los cálculos heredan los errores asociados a estos últimos.

- El error sistemático del WWIII es muy pequeño, mostrando una tendencia a subestimar valores máximos de altura significante y a sobreestimar valores mínimos;
- El paquete de física ST4 es el más recomendado para usar, en particular para sistemas con mallas anidadas.

En cuanto al tipo de datos que se validan, si son parámetros integrales o de particiones, podemos decir que la práctica más extendida es la primera. Esto porque si bien, como se vio en el trabajo de Bi et al. (2015), validar cada partición da información más detallada sobre las áreas de oportunidad del modelo, validar con parámetros integrales da suficiente información para determinar qué tan bueno es el desempeño del modelo. Esto último respaldó la decisión de que en este trabajo se hiciera una validación con parámetros integrales.

Capítulo 4

Descripción del modelo y su implementación

El sistema de pronóstico de oleaje que se valida en esta tesis se desarrolló con el modelo de oleaje WaveWatch III (WWIII), el cual es un modelo de tercera generación (ver Sec. 2.3). Fue desarrollado por el NCEP de la NOAA para ser el sucesor del modelo WAM como su modelo operativo de oleaje. Su primera versión fue el WWI, que fue desarrollado por la Delft University of Technology y más tarde evolucionó en el WWII desarrollado por el Goddard Space Flight Center de la NASA. Sin embargo, WWIII difiere de ellos tanto en la estructura de su programación como en la forma de resolver las ecuaciones gobernantes (WAVEWATCH III Development Group, 2016).

En este capítulo se describen de manera breve las ecuaciones fundamentales del modelo, sus consideraciones numéricas y la configuración que se hizo del modelo para el pronóstico operativo.

4.1. Ecuaciones gobernantes

Como se describió en la sección 2.1 del capítulo 2, la varianza del desplazamiento vertical de la superficie del océano (η) puede describirse como el resultado de la superposición de varias ondas con diferentes longitudes de onda y frecuencias, viajando en todas direcciones (Eq.2.2). La energía asociada a estas ondas (E) en unidades de J/m^2 se relaciona con η a través de:

$$E = \rho_w g \langle \eta^2 \rangle \tag{4.1}$$

donde ρ_w es la densidad del agua, g es la acelaración de la gravedad y los paréntesis denotan un promedio espacial o temporal (Stewart, 2008).

Resulta de interés describir cómo se distribuye esta energía entre las diferentes componentes del campo de oleaje, es decir, cuál es la energía asociada a cada frecuencia y a cada dirección. Para hacerlo, se utiliza el denominado *espectro de densidad de la varianza de elevación*, el cual está dado por:

$$F(f,\theta) = Z \cdot Z^* \tag{4.2}$$

donde Z es la transformada de Fourier de $\eta(t)$ y Z^{*} es su complejo conjugado (Stewart, 2008). Debido a esto, a F se le conoce también como el *espectro de energía*. Al integrar F sobre todo el espectro se obtiene la energía total E (WAVEWATCH III Development Group, 2016).

En principio, F es función de todos los parámetros de fase que describen a una onda, así como del tiempo y la posición, tal que $F(\vec{k}, \omega, \theta; \vec{x}, t)$, donde $\vec{k} = (k_x, k_y)$ es el vector de número de onda y $\vec{x} = (x, y)$ es el vector de posición. Sin embargo, el espectro local e instantáneo se puede reducir a una función bidimensional $F(k, \theta)$ al asumir que las ondas individuales (componentes del espectro) cumplen localmente con la teoría lineal de olas (WAVEWATCH III Development Group, 2016). Bajo esta teoría, la frecuencia radial ω se relaciona con el número de onda k a través de la relación de dispersión:

$$\omega^2 = gk \tanh(kd) \tag{4.3}$$

donde d es la profundidad media del agua (Stewart, 2008). Considerando un caso más general en donde hay presencia de corrientes, la frecuencia angular se describe con un corrimiento tipo Doppler:

$$\sigma^2 = gk \tanh(kd) \tag{4.4}$$

$$\omega = \sigma + \vec{k} \cdot \vec{U} \tag{4.5}$$

donde $\sigma = 2\pi/T$ es la frecuencia intrínseca, la cual se observa desde un sistema de referencia que se mueve con la corriente y corresponde con la definición de ω que se había usado hasta este punto; ω es ahora la frecuencia absoluta, la cual se observa desde un sistema de referencia fijo; y \vec{U} es la velocidad de la corriente promediada en tiempo y espacio. Esta aproximación lineal se puede aplicar si suponemos que la profundidad d y la corriente \vec{U} varían a una escala temporal y espacial mucho menor que las escalas características de las olas. Al realizar estas suposiciones, los fenómenos de difracción, dispersión e interferencia son despreciados, pero pueden ser añadidos después como términos fuente (WAVEWATCH III Development Group, 2016).

El modelo WWIII tiene como base el espectro $F(k, \theta)$ para resolver la evolución del campo de oleaje. Sin embargo, las salidas que genera usan el dominio de frecuencia-dirección, $F(f, \theta)$, pues es la forma más tradicional de presentar la información espectral. El espectro $F(f, \theta)$ puede ser obtenido a partir de $F(k, \theta)$ a través de un jacobiano (WAVEWATCH III Development Group, 2016).

Además de la información que da F en el dominio espectral, también resulta de interés caracterizar el estado de la superficie del mar a través de los parámetros con los que regularmente se describe una onda, tales como su altura, periodo y dirección de propagación. Para obtener parámetros de este tipo que representen a todo el espectro, se tiene que integrar F completamente, es decir, se tiene que obtener E. Por esta razón, estos parámetros se conocen como *parámetros integrales*. Los parámetros integrales de mayor interés para este trabajo y para el proyecto para el cual fue desarrollado el sistema de pronóstico de oleaje son:

1. Altura significante (H_s) , dada por:

$$H_s = 4\sqrt{E} \tag{4.6}$$

2. Periodo medio de oleaje (T_m) , obtenido como:

$$T_m = 2\pi\overline{\sigma} \tag{4.7}$$

donde $\overline{\sigma}$ es un promedio hecho sobre todo el espectro;

3. Dirección media de oleaje (θ_m) , que se calcula con:

$$\theta_m = atan\left(\frac{b}{a}\right), \tag{4.8}$$
$$a = \int_0^{2\pi} \int_0^\infty \cos(\theta) F(\sigma, \theta) \mathrm{d}\sigma \mathrm{d}\theta,$$

$$b = \int_0^{2\pi} \int_0^\infty \sin(\theta) F(\sigma, \theta) d\sigma d\theta$$

Las expresiones mostradas para estos parámetros son las que utiliza WWIII para su cálculo (WA-VEWATCH III Development Group, 2016). En el caso de la altura significante, su definición tiene un principio histórico. Cuando el estado del mar era descrito únicamente por las observaciones (visuales, no mediciones), se encontró que la altura promedio observada correspondía con el promedio de alturas del 30 % de las olas más altas ($H_{1/3}$) y se puede mostrar que $H_s \approx H_{1/3}$ (Janssen, 2004).

4.1.1. Propagación en el modelo WWIII

Cuando no hay corrientes, la energía de un paquete de olas es conservativa, pero bajo la influencia de una corriente, esta propiedad se pierde, pues existe una transferencia de momento entre la corriente y las olas. Sin embargo, la *acción de ola* que se define como $A \equiv E/\sigma$ siempre es conservada. De forma análoga, el *espectro de densidad de la acción de ola*, definido como $N(k, \theta) \equiv F(k, \theta)/\sigma$, es conservativo y es el espectro que resuelve el modelo WWIII. De forma general, la propagación de las olas queda entonces descrita por:

$$\frac{DN}{Dt} = \frac{S}{\sigma} \tag{4.9}$$

donde D/Dt representa la derivada material y S representa el efecto neto de fuentes y sumideros $(t\acute{erminos fuente})$ del espectro F. El lado izquierdo de la ecuación anterior surge de la aproximación lineal, por lo que todos los efectos no lineales, así como los fenómenos de difracción, dispersión e interferencia están contenidos en S.

El modelo utiliza una forma Euleriana de la ecuación de balance 4.9 descrita de manera conservativa, la cual está dada por:

$$\frac{\partial N}{\partial t} + \nabla_x \cdot \dot{x}N + \frac{\partial}{\partial k} \dot{k}N + \frac{\partial}{\partial \theta} \dot{\theta}N = \frac{S}{\sigma}$$
(4.10)

En la expresión anterior, \dot{x} , \dot{k} y $\dot{\theta}$ son derivadas, definidas como:

$$\dot{x} = \vec{c_g} + \vec{U},\tag{4.11}$$

$$\dot{k} = -\frac{\partial\sigma}{\partial d}\frac{\partial d}{\partial s} - \vec{k}\frac{\partial\vec{U}}{\partial s},\tag{4.12}$$

$$\dot{\theta} = -\frac{1}{k} \left[\frac{\partial \sigma}{\partial d} \frac{\partial d}{\partial m} + \vec{k} \frac{\partial \vec{U}}{\partial m} \right], \qquad (4.13)$$

donde $\vec{c_g} = (c_g \sin \theta, c_g \cos \theta)$ es la velocidad de grupo, *s* es una coordenada en la dirección de θ y *m* es una coordenada perpendicular a *s*. Estas ecuaciones se encuentran en coordenadas cartesianas, pero pueden ser transformadas a coordenadas esféricas, definidas por la longitud y latitud, forma en la cual son más adecuadas para simulaciones de dominios de gran escala (WAVEWATCH III Development Group, 2016). En esta tesis se utilizó un dominio descrito en coordenadas cartesianas.

4.1.2. Términos fuente

De forma general, S está conformada por tres términos como se describió en la Eq.2.3: el término S_{in} que describe el intercambio de energía entre la atmósfera y las olas, el cual es positivo para la interacción con windsea, pero puede ser negativo para el swell; el término S_{nl} que describe las interacciones no lineales que se dan entre olas y el término S_{ds} que describe la interacción entre las olas y el océano, el cual describe la disipación. El término S_{in} está dominado por el crecimiento exponencial descrito por el mecanismo de Miles, por lo que para hacer más realista el crecimiento de las olas, se puede agregar un término lineal S_{ln} que describa el mecanismo de Phillips (WAVEWATCH III Development Group, 2016).

Estos términos son suficientes para describir los procesos de intercambio de energía principales a los que se ven sometidas las olas en aguas profundas. Sin embargo, en aguas someras, se tienen que considerar interacciones con el fondo, las cuales se describen con el término S_{bot} (WAVEWATCH III Development Group, 2016).

Lo anterior define el término fuente general, S, usado en el modelo como:

$$S = S_{ln} + S_{in} + S_{nl} + S_{ds} + S_{bot}$$
(4.14)

En realidad, WWIII permite añadir algunos términos fuente para otros procesos, así como términos fuente personalizados por el usuario, pero la Eq. 4.14 describe la configuración del modelo usada en esta tesis.

4.2. Consideraciones numéricas

La ecuación 4.10 es la ecuación fundamental del modelo. Sin embargo, su resolución directa genera problemas numéricos, disminuyendo la resolución espectral en aguas someras. Este problema se resuelve utilizando una malla cuya resolución espacial (dada por el número de onda, k) cambie en función de la profundidad. Introduciendo esta malla de número de onda variable, denotada con κ , las ecuaciones 4.10 y 4.12 se vuelven:

$$\frac{\partial}{\partial t}\frac{N}{c_g} + \frac{\partial}{\partial x}\frac{\dot{x}N}{c_g} + \frac{\partial}{\partial y}\frac{\dot{y}N}{c_g} + \frac{\partial}{\partial \kappa}\frac{\dot{\kappa}N}{c_g} + \frac{\partial}{\partial\theta}\frac{\dot{\theta}N}{c_g} = \frac{S}{\sigma c_g},\tag{4.15}$$

$$\dot{\kappa}\frac{\partial k}{\partial \kappa} = c_g^{-1}\frac{\partial \sigma}{\partial d} \left(\frac{\partial d}{\partial t} + \vec{U} \cdot \nabla_x d\right) - \vec{k}\frac{\partial \vec{U}}{\partial s}$$
(4.16)

Una vez incorporada κ , WWIII resuelve la Ec. 4.15 diviéndola en varias partes que son integradas por separado. Este enfoque se denomina *fractional step method* y consiste de seis etapas:

- 1. La primera etapa consiste en tratar las variaciones temporales de la profundidad, de modo que se determine la malla para el tiempo que se está integrando en ese momento y ésta permanezca invariante para las siguientes etapas de la integración.
- 2. En la segunda etapa se resuelve la propagación espectral, integrando sobre un paso de tiempo $\Delta t_g/2$ la parte espectral de la Ec. 4.15, es decir:

$$\frac{\partial}{\partial t}\frac{N}{c_g} + \frac{\partial}{\partial \kappa}\frac{\dot{\kappa}N}{c_g} + \frac{\partial}{\partial \theta}\frac{\theta N}{c_g} = 0$$

3. La tercera etapa resuelve la propagación espacial, integrando sobre un paso de tiempo Δt_g la parte correspondiente de la Ec. 4.15, es decir:

$$\frac{\partial}{\partial t}\frac{N}{c_g} + \frac{\partial}{\partial x}\frac{\dot{x}N}{c_g} + \frac{\partial}{\partial y}\frac{\dot{y}N}{c_g} = 0$$

- 4. En la cuarta etapa se hace la segunda parte de la propagación espectral, integrando nuevamente sobre un paso de tiempo $\Delta_q/2$ la parte espectral de la Ec. 4.15.
- 5. La quinta etapa integra sobre Δt_g todos los términos fuente que no estén asociados a interacciones con hielo, es decir:

$$\frac{\partial}{\partial t}\frac{N}{c_g} = \frac{S_{noice}}{\sigma c_g}$$

6. Finalmente, en la sexta etapa se integran sobre Δt_g los términos fuente asociados a interacciones con hielo:

$$\frac{\partial}{\partial t}\frac{N}{c_g} = \frac{S_{ice}}{\sigma c_g}$$

En el límite cuando $\Delta t_g \rightarrow 0$, la sucesión de las etapas antes descritas es equivalente a la integración de la Ec. 4.15 sobre un paso de tiempo general Δt_g . La separación por etapas de la integración permite que la paralelización del modelo sea más eficiente.

4.3. Implementación del modelo WWIII

La versión 5.16 del modelo (la más actual disponible al momento de la instalación) fue implementada en el clúster de supercómputo del Centro de Ciencias de la Atmósfera, Ometéotl. Este clúster está integrado por 1360 núcleos, distribuidos en 20 nodos de 24 núcleos cada uno más 20 nodos de 44 núcleos cada uno. En conjunto, estos procesadores equivalen a una capacidad de cómputo de 42 Tflops. Ometéotl cuenta con un sistema de almacenamiento LUSTRE con una capacidad de aproximadamente 200TB. Ometéotl cuenta con Linux Centos 6.9 como sistema operativo y con *Simple Linux Utility Resource Management*(SLURM) como administrador de recursos. La función esencial de SLURM es programar la ejecución de los trabajos en el clúster, gestionando los nodos de cálculo y el tiempo de procesador que se le asignan a cada tarea.

El modelo fue compilado con compiladores Intel para C, C++ y FORTRAN, tanto en sus versiones en serie como en sus versiones en paralelo. Además, se requirió de la compilación de las librerías NetCDF y NCL-Ncarg. El proceso de compilación se llevó a cabo en dos etapas. En la primera etapa se generaron las librerías base y el archivo de configuración inicial, el cual contiene las banderas y variables de ambiente generales del modelo. En esta etapa se usó el compilador ifort de Intel. En la segunda etapa de compilación es donde se deben generar, uno a uno, los archivos ejecutables de los cuales consta el modelo, para verificar su correcta creación. Esto se hace mediante el programa $ww3_make$, el cual fue generado en la primera etapa. Es en esta etapa en la que se definen las banderas de compilación, las cuales determinan aspectos de la configuración del modelo, tales como si el programa será ejecutado en serie o en paralelo, así como qué parametrizaciones se van a utilizar para resolver los procesos físicos simulados. Una breve descripción de los módulos (programas) compilados se puede encontrar en la sección 4.3.2.

4.3.1. Parametrizaciones

Como se mencionó anteriormente, durante la segunda etapa del proceso de compilación es cuando se debe elegir la forma en que se resolverán los términos fuente descritos en la sección 4.1.2. Para cada término fuente, WWIII tiene disponibles varias parametrizaciones, las cuales representan diferentes
formas de modelar el proceso que describe ese término. Las parametrizaciones son elegidas a través de un archivo que contiene los denominados *switches*, que son claves alfanuméricas que identifican las diferentes ecuaciones programadas para cada término fuente. Las parametrizaciones con las que fue configurado el modelo fueron:

- LN1 (S_{ln}) . Para representar el crecimiento lineal de las olas se incluyó el término S_{nl} de Cavaleri and Rizzoli (1981).
- ST2 (S_{in} y S_{ds}). Aunque los términos S_{in} y S_{ds} representan procesos distintos, están estrechamente relacionados, pues gobiernan el crecimiento integral de las olas. Por tanto, se eligió una parametrización, descrita por Tolman and Chalikov (1996), la cual integra ambos procesos. Esta parametrización es también referida como TC en la literatura.
- **NL1** (S_{nl}) . Para representar los procesos no lineales se utilizó la Aproximación de Interaciones Discretas (DIA) descrita por Hasselmann et al. (1985).
- **BT1** (S_{bot}) . La fricción de fondo fue representada con una función empírica lineal que fue desarrollada como resultado del proyecto JONSWAP y es descrita en Hasselmann et al. (1973).

4.3.2. Flujo de ejecución del modelo

Una vez compilado el modelo, se obtienen varios archivos ejecutables que se denominan módulos, cada uno de los cuales cubre una función dentro del proceso de simulación. Cada módulo cumple, esencialmente, con la tarea de transformar un conjunto de datos de entrada (que se conocen como *campos forzantes*), que se encuentran en un formato de texto, a datos en un formato binario entendibles por el modelo y la computadora. Los módulos más importantes son los siguientes:

- ww3_grid Se encarga de procesar la malla que se utilizará en la simulación, incluyendo la batimetría.
- ww3_strt Genera un archivo con las condiciones iniciales del océano. Si no se cuenta con condiciones que aproximen la realidad (salidas de otro simulación o un campo observado), el modelo puede comenzar desde la calma o con un espectro idealizado basado en el campo de viento de entrada.
- ww3_prep Procesa los datos de viento que se utilizarán como forzantes para convertirlos a un formato asimilable por el modelo.
- ww3_multi Este es el módulo principal, pues es donde se resuelven las ecuaciones del modelo. Se llama multi ya que está compilado para ejecutarse en paralelo, esto es, en múltiples núcleos a la vez, reduciendo así el tiempo de cálculo.
- ww3_ounf Las salidas generadas por ww3_multi están en un formato binario y en este módulo se transforman a formato NetCDF (Rew et al., 1997).

El diagrama de la Figura 4.1 muestra cómo se da el flujo de información entre los módulos.



Figura 4.1: Diagrama de ejecución del modelo WWIII.

4.3.3. Descripción de los dominios simulados

El modelo se implementó para tres dominios: Océano mundial, Pacífico mexicano y Golfo de México. Los detalles se pueden ver en el Cuadro 4.1. Los dominios del Pacífico y el Golfo están anidados dentro del dominio del Océano mundial. Esto significa que mientras se ejecuta la simulación, el modelo actualiza automáticamente las condiciones de frontera para los dominios del Pacífico y el Golfo a partir de las condiciones generadas por el dominio mundial. Este dominio tiene la característica de que su malla es periódica, es decir, sus fronteras meridionales están conectadas, de modo que los sistemas de oleaje que salen por una de estas fronteras, entran por la otra. Debido a que el corte de la malla se hace en medio del Pacífico (Figura 4.2a), esta característica es importante, pues asegura la continuidad de sistemas de *swell* que pueden afectar el dominio del Pacífico mexicano. Para los tres dominios se generan salidas horarias para un periodo de 120 horas.

Dominio	Latitud	Longitud	Resolución	Batimetría Forzamie:	
				base	de viento
Océano	75°S-75°N	180°W-	0.5°	ETOPO1	GFS
mundial		180°E			
Pacífico	10°S-33°N	120°W-	0.25°	ETOPO1	WRF
mexicano		$85^{\circ}W$			
Golfo de	15°S-33°N	100°W-77°E	0.25°	ETOPO1	WRF
México					

Cuadro 4.1: Descripción de los dominios utilizados en el sistema de pronóstico de oleaje implementado con el modelo WWIII.

Como ya se ha mencionado, el sistema de pronóstico de oleaje descrito y evaluado en esta tesis, forma parte de un proyecto con fines de alertamiento, por lo que resultaba de interés evaluar la calidad de las salidas generadas para los dominios de los mares mexicanos. Sin embargo, ya que hay una muy pequeña cantidad de boyas disponibles en el Pacífico Mexicano, la validación se realizó únicamente para el dominio del Golfo de México (Figura 4.2c).



Figura 4.2: Dominios del sistema de pronóstico de oleaje.

4.3.4. Datos de entrada para el modelo

El modelo WWIII puede recibir varios campos de entrada, tales como la batimetría, la velocidad media de las corrientes, el campo de velocidad del viento, la diferencia de temperatura entre el agua y el aire, la elevación del nivel medio del mar, la concentración de hielo, etc. Cuanto más campos de entrada se ingresen, más realista es la configuración del modelo. Sin embargo, esto genera un mayor costo computacional, pues requiere mayor memoria para alojar la información, así como más tiempo de cómputo para resolver todos los procesos físicos asociados a cada campo ingresado. Por esto, generalmente se ingresan sólo los campos más relevantes para la zona de estudio, siendo la batimetría y el campo de velocidad de viento los forzantes más esenciales. Estos dos campos, junto con un archivo de condiciones iniciales, son los datos de entrada que se utilizan para realizar la ejecución del sistema de pronóstico de oleaje que se valida en esta tesis.

Modelo global de relieve ETOPO1

El Modelo Global de Relieve ETOPO1, como su nombre lo indica, es un modelo de relieve de todo el planeta con resolución espacial de un arcominuto. Este modelo provee información topográfica y batimétrica con cobertura completa (de 90°S a 90°N y de 180°W a 180°E), resultado de la integración y procesamiento de datos obtenidos de varias agencias gubernamentales de EE. UU., agencias internacionales e instituciones académicas de todo el mundo. Fue diseñado con la finalidad de ser útil en el desarrollo de modelos de tsunamis, modelos de circulación oceánica y visualización de la Tierra (Amante and Eakins, 2009).

La batimetría de ETOPO1 es usada por el módulo $ww3_grid$ al momento de crear las mallas de cada dominio, pues a cada uno de los puntos de malla se le asocia la profundidad correspondiente. Esta información es necesaria tanto para resolver los procesos de interacciones con el fondo, como para determinar los puntos que corresponden a mar y los que corresponden a tierra.

Los datos de ETOPO1 se pueden descargar de forma libre en la página de National Centers for Environmental Information de la NOAA (https://www.ngdc.noaa.gov/mgg/global/).

Modelo GFS

El dominio del Océano mundial es forzado con datos provenientes del modelo meteorológico *Global Forecast System* (GFS), desarrollado y operado por el NCEP. El modelo GFS es un modelo acoplado compuesto por cuatro modelos independientes: uno atmosférico, uno oceánico, uno de suelo y uno de hielo marino. Tiene cobertura global y produce pronósticos de hasta 16 días. La resolución horizontal del modelo es de aproximadamente 28 km durante la primera semana de pronóstico y de cerca de 70 km para la segunda semana. Las salidas de este modelo incluyen un gran número de variables, como temperatura, componentes de viento, precipitación, humedad del suelo, etc.(National Centers for Environmental Prediction, sf).

Para forzar al WWIII, se usan las componentes horizontales del viento ($\mathbf{u} \ \mathbf{y} \ \mathbf{v}$) y la proporción de cobertura de hielo marino. Estas variables se toman de las salidas con resolución espacial de 0.25° (28 km) y temporal de 3 horas, que se toman de los pronósticos con hora de inicio a las 00 Z de cada día y cubriendo un periodo de 120 horas.

Los datos del pronóstico operativo de GFS pueden consultarse y descargarse libremente mediante varios protocolos (HTTP, FTP, etc.) en la dirección electrónica https://www.ncdc.noaa.gov/ data-access/model-data/model-datasets/global-forcast-system-gfs.

Modelo WRF

Los dominios del Pacífico y el Golfo son forzados con salidas del modelo Weather Research and Forecasting (WRF) que es un modelo meteorológico de mesoescala desarollado de manera conjunta por varios centros de investigación de EE. UU., encabezados por el National Center for Atmospheric Research (NCAR). Es un modelo de código abierto y puede ser descargado de forma libre en http://www2.mmm.ucar.edu/wrf/users/downloads.html (National Center for Atmospheric Research, sf).

El modelo WRF también fue implementado por el Grupo IOA como parte del proyecto del cual forma parte esta tesis. De este modelo se configuraron dos dominios:

- Dominio 1, acotado entre $68^{\circ}W$ y $111^{\circ}W$ en longitud y entre $10^{\circ}N$ y $35^{\circ}N$ en latitud, cubre todo el territorio nacional. Tiene una resolución espacial de 15 km y genera salidas cada hora por 120 horas.
- Dominio 2, acotado entre $89^{\circ}W$ y $98.5^{\circ}W$ en longitud y entre $17^{\circ}N$ y $30^{\circ}N$ en latitud, cubre la parte central del país. Tiene una resolución espacial de 5 km y genera salidas cada hora por 120 horas.

Al igual que en el caso del modelo GFS, el modelo WRF genera una amplia variedad de variables meteorológicas, de las cuales solo se utilizan las componentes del viento como forzantes para el WWIII. Estas variables se extraen de las salidas del Dominio 1, que es el que cubre tanto el Pacífico mexicano como el Golfo de México.

Condiciones iniciales de oleaje

Además de los datos de batimetría y de viento, los cuales son procesados para generar los archivos $mod_def.ww3$ y wind.ww3, respectivamente, el modelo necesita el archivo restart.ww3. Este archivo contiene las condiciones iniciales del campo de oleaje correspondientes a la fecha en la que inicia la simulación, lo que permite partir de una configuración más realista. El archivo restart.ww3 puede ser generado de dos formas: por el modelo mismo o con el módulo de condiciones iniciales, $ww3_strt$. Cuando el modelo corre, genera archivos restart.ww3 cada cierto número de pasos de tiempo. Estos archivos contienen las condiciones de oleaje simuladas para la fecha correspondiente al paso de tiempo en el que fueron creados. Si una simulación se interrumpe por cualquier razón, el archivo restart.ww3 permite continuar con la simulación desde la fecha para la cual fue creado, de modo que no sea necesario iniciar desde cero la simulación. Si no se cuenta con un archivo restart.ww3 generado por una simulación previa del modelo, se puede generar con el programa $ww3_strt$. Dependiendo de la configuración que se haga para el programa de condiciones iniciales, el modelo puede iniciar de la calma ($H_s = 0$), de un espectro de oleaje proporcionado por el usuario, de un espectro idealizado o de un espectro calculado a partir de los datos de viento de entrada.

En el caso del sistema de pronóstico que se implementó, el modelo genera un archivo **restart.ww3** para la hora cero del día siguiente al de la fecha de inicio de la simulación. El campo de oleaje contenido en este archivo es utilizado como condición inicial para la ejecución del sistema de pronóstico del día siguiente. En caso de que no se cuente con este archivo, el módulo $ww3_strt$ está configurado para que genere las condiciones iniciales haciendo uso de un espectro denominado "JONSWAP con fetch limitado", el cual estima el campo de oleaje a partir del campo de viento ingresado.

4.3.5. Automatización de la ejecución modelo

El sistema de pronóstico operativo del cual forma parte el pronóstico de oleaje fue automatizado con scripts programados en el lenguaje BASH, mientras que la generación de los productos gráficos de cada modelo se hizo con el lenguaje NCL. Como parte de esta tesis, se trabajó tanto en la automatización de la ejecución del modelo WWII en BASH como en la paralelización del graficado. Sin embargo, los scripts que generan los productos gráficos finales (que se muestran en el capítulo de Resultados) fueron desarrollados por el M. en C. Fernando Nicolás Arellano Guerrero.

El proceso de ejecución del modelo consiste en general de cuatro etapas:

1. Generación de condiciones iniciales y de frontera. En esta etapa se interpolan los vientos descargados del modelo GFS a la malla correspondiente al dominio del Océano mundial, mientras

que los vientos generados con el modelo WRF se interpolan a las mallas del Pacífico mexicano y el Golfo de México. La interpolación se realiza con una subrutina programada en el lenguaje NCL.

- 2. Preparación de archivos de configuración. En este paso se preparan los archivos de configuración necesarios para ejecutar cada uno de los módulos descritos en la Sección 4.3.2. Para cada módulo se genera un archivo de configuración con el mismo nombre del módulo, pero con extensión .inp (por ejemplo, ww3_grid.inp es el archivo de configuración del módulo ww3_grid). Estos archivos contienen una descripción del dominio o campo de entrada que su ejecutable correspondiente procesará.
- 3. Ejecución del modelo. Esta etapa consiste en ejecutar cada uno de los módulos descritos en la Sección 4.3.2 de manera secuencial, hasta obtener las salidas en formato NetCDF.
- 4. Generación de productos gráficos. Finalmente, una vez que se tienen las simulaciones en NetCDF, estas son utilizadas para graficar cuatro variables de interés: la altura significante, el periodo medio de oleaje, la altura de marejada y el periodo de marejada. La marejada corresponde al *swell* más energético identificado en el modelo. La generación de gráficas se realiza con un script de graficado desarrollado en NCL, pero se controla desde un script de BASH que se encarga de paralelizar el proceso. Una vez que terminan de generarse las salidas, el script de BASH copia las imágenes al portal web de visualización del sistema de pronóstico (http://grupo-ioa.atmosfera.unam.mx/pronosticos/index.php/oleaje).

Las primeras tres etapas son ejecutadas desde el script maestro modulo_ejecuta_wwiii.sh, mientras que la cuarta etapa es llevada a cabo por el script modulo_grafica_wwiii.sh, ambos desarrollados en BASH. En la Figura 4.3 se muestra el diagrama de flujo de la ejecución del sistema pronóstico de oleaje y se pueden identificar dichos scripts. En este diagrama se consideran todos los pasos de le ejecución del sistema operativo completo que son relevantes para la ejecución del WWIII, como la descarga de datos GFS y la ejecución del modelo WRF, pero únicamente se detallan las subrutinas de ejecución para el modelo WWIII. No se considera en el esquema la ejecución del modelo ADCIRC ni los detalles de la ejecución y graficado del modelo WRF. Se aprecia que el proceso de ejecución del sistema de pronóstico inicia con el comando CRON, el cual es un comando de los sistemas operativos tipo Unix (tal como es el caso de Linux Centos) que permite programar la ejecución de tareas a una cierta hora del día. El sistema de pronóstico operativo completo se ejecuta todos los días a la media noche y su ejecución completa toma de 4 a 6 horas.



Figura 4.3: Diagrama de ejecución del sistema de pronóstico de oleaje.

4.4. Consideraciones para la validación

La implementación antes descrita corresponde con la configuración en modo pronóstico que se tiene del modelo de oleaje, la cual funciona de manera operativa hasta la fecha. Esta configuración y toda su arquitectura (ambiente de ejecución) fue utilizada para realizar un año completo de simulaciones (2015) con el objetivo de realizar la validación del modelo al comparar sus salidas con observaciones. La única diferencia entre las simulaciones para la validación y el sistema de pronóstico, es que para las primeras se utilizaron datos históricos de GFS, los cuales se descaragaron de https://rda.ucar.edu/datasets/ds084.1/#!access.

El año 2015 fue elegido para la validación debido a que como parte del proyecto en el que se enmarca esta tesis, el modelo WRF se validó con este año. Así, teniendo los datos de GFS para 2015 (que también son usados como forzantes para el modelo WRF) y las simulaciones de WRF para el mismo año, se procedió a realizar las corridas del modelo WWIII. Son estos resultados del año 2015 los que se compararon con observaciones y que constituyen el objeto de estudio de esta tesis.

Capítulo 5

Método de la validación

En este capítulo se describen las herramientas y métodos utilizados para realizar la evaluación de la calidad del pronóstico de oleaje generado con el modelo WWIII para el Golfo de México. De manera general, el proceso de validación consistió de los siguientes pasos:

- 1. Se seleccionaron boyas distribuidas en el dominio del Golfo de México que abarca el modelo. Sus registros fueron filtrados para eliminar las frecuencias correspondientes a ruido.
- 2. Las 120 horas de simulación se divieron en periodos de 24 horas y con ellos se armaron cinco series para cada mes.
- 3. Las series mensuales fueron comparadas con los datos de oleaje obtenidos de las boyas, haciendo uso tanto de métodos gráficos como de métricas estadísticas. Se validaron las variables de magnitud de la velocidad del viento (WSP), altura significante (H_s), periodo del oleaje (T) y dirección del oleaje (DIR).

Los detalles de este procedimiento se detallan a continuación.

5.1. Observaciones utilizadas

Las boyas utilizadas pertenecen a la red de boyas del National Data Buoy Center (NDBC), agencia perteneciente al National Weather Service de la NOAA.

Las boyas contienen un conjunto de sensores que registran y transmiten información meteorológica y oceanográfica de los sitios en donde se encuentran instaladas. La frecuencia de registro de dicha información varía dependiendo de la variable monitoreada. En el caso de los datos de oleaje, se hace de manera continua por periodos de 20 a 40 minutos, dependiendo de cada boya. Los datos se reportan redondeados a la hora más cercana para sistemas de adquisición de 40 minutos o a la media hora más cercana para sistemas de adquisición de 20 minutos, siempre en horario UTC. Para las boyas utilizadas, todos los registros se reportan para el minuto 50 de cada hora (National Data Buoy Center, 2018a).

5.1.1. Variables medidas en las boyas

Los datos de oleaje reportados por las boyas no se miden con los sensores montados a bordo, sino que son derivados de manera indirecta. Las boyas tienen a bordo acelerómetros o inclinómetros, los cuales miden la aceleración de elevación o el desplazamiento vertical del casco de la boya durante el tiempo de adquisición de datos. A estas mediciones se les aplica una Transformada Rápida de Fourier (FFT), llevándolas del dominio temporal al dominio de la frecuencia. A continuación, se aplica un *Respond amplitude operator* (RAO), que es una función de transferencia utilizada para determinar el efecto que tiene el estado del mar en objetos flotantes y que, por tanto, permite determinar la parte de la señal que representa ruido del casco o ruido electrónico. Después de aplicar ambas funciones, se pueden obtener las mediciones espectrales del campo de oleaje (energías de onda con sus frecuencias asociadas) y posteriormente la altura de ola significante, el periodo de oleaje promedio y el período dominante (National Data Buoy Center, 2018b).

Las variables reportadas por las boyas que fueron utilizadas en este trabajo se muestran en el Cuadro 5.1.

Variable	Abreviación	Descripción	
Rapidez del viento	WSPD	Es el promedio aritmético de las mediciones	
		de la magnitud del viento realizadas en un	
		lapso de 8 minutos. Su unidad de medición	
		es en m/s.	
Altura significante	WVHT	Es el promedio aritmético del tercio más alto	
		de todas las alturas de las olas medidas en un	
		lapso de muestreo de 20 minutos. Su unidad	
		de medición es en m.	
Dirección de la ola	MWD	Es la dirección de donde vienen las olas en el	
		periodo dominante. Se mide en grados Dex-	
		trorsum (giro en sentido de las manecillas del	
		reloj), donde 0° indica el norte verdadero.	
Periodo de la ola	APD	Es el promedio aritmético del periodo de to-	
		das las olas calculado en un lapso de 20 mi-	
		nutos. Su unidad de medición es en segundos.	

Cuadro 5.1: Descripción de las variables de oleaje monitoreadas por las boyas en el Golfo de México.

5.1.2. Selección de boyas

Para realizar la selección de boyas a utilizar en la validación, se tomó en cuenta su ubicación, tratando de que estuvieran distribuidas en todo el dominio, y la completitud de su registro para las variables de interés para el año 2015, de modo que tuvieran más del 90 % de los datos del periodo. La revisión de los registros se hizo generando series de tiempo para cada variable en cada boya (ver Figura 5.1).



Figura 5.1: Registros de la boya 42057 para el año 2015.

Con los criterios anteriores se seleccionaron cinco boyas para la validación. Además, se incluyeron dos boyas más que, si bien no cumplen con tener más del 90 % de registro de datos, se consideraron de interés debido a que se usaron en la validación de las salidas del modelo WRF (hecha dentro del proyecto del cual forma parte esta tesis), con las cuales se forzó el modelo WWIII. Las siete boyas seleccionadas, identificadas por un número (ID), así como el porcentaje de registros que presentan para cada variable, se enlistan en el Cuadro 5.2, y su distribución se puede ver en el mapa de la Figura 5.2.

ID	Latitud	Longitud	Profundidad	Porcentaje de observaciones			
				WSPD	WVHT	APD	MWD
42001	25.897	-89.668	3194 m	66.35	57.92	57.92	0
42002	26.091	-93.758	3088 m	74.62	74.62	74.62	73.34
42020	26.968	-96.694	84.1 m	99.33	99.33	99.33	98.3
42035	29.232	-94.413	$16.2 \mathrm{~m}$	99.63	96.11	96.11	90.89
42040	29.208	-88.226	$183 \mathrm{~m}$	99.95	99.25	99.25	62.85
42055	22.120	-93.960	$3624~\mathrm{m}$	99.77	99.77	99.77	98.58
42056	19.769	-84.973	$4554~\mathrm{m}$	99.46	99.46	99.46	99.42
42057	16.908	-81.422	$377 \mathrm{m}$	99.78	99.78	99.78	99.78

Cuadro 5.2: Ubicación y porcentaje de registros de las boyas para cada variable de interés.



Figura 5.2: Boyas utilizadas para la validación del modelo WWIII.

5.1.3. Preprocesamiento de las observaciones

Antes de realizar la comparación estadística entre los datos de boyas y el pronóstico, se aplicó un filtro paso bajo de media móvil ponderada a las observaciones. El uso de este tipo de filtro tiene base en la suposición de que la señal con la que se trabaja es localmente suave, es decir, que los datos vecinos son espectralmente similares (Park and Lu, 2015). De este modo, los filtros paso bajo suprimen en su mayoría los valores provenientes de señales de frecuencias altas (que se asume que son ruido), pero conservan las oscilaciones de baja frecuencia (Kubben et al., 2019).

El grado de preservación de la señal depende de las características de la media móvil ponderada que se emplea. La forma más básica de una media móvil ponderada es la media móvil uniforme, donde el *i-ésimo* punto y los N-1 puntos anteriores son promediados. Esto equivale a que cada punto sea multiplicado por un coeficiente de 1/N. Al cambiar estos coeficientes y la longitud del filtro (N), se pueden ajustar las características de la señal de salida. Es importante notar que estos filtros generan un desfase temporal de los datos. Cuando el filtro tiene pesos simétricos alrededor del centro, el desfase temporal es igual a la longitud del filtro dividida entre 2 (Kubben et al., 2019).

El suavizado de una señal hecho con un filtro de este tipo puede ser generalizado por una convolución unidimensional de la siguiente manera:

$$y[i] = x[i] * h[i] = \sum_{m=0}^{N-1} x[m]h[i+m], \qquad (5.1)$$

donde x es el espectro medido, h es un filtro y y es el espectro suavizado (Park and Lu, 2015).

El filtro aplicado a las observaciones sigue la forma de la expansión binomial de $[1/2 \ 1/2]^n$, que para valores altos de *n* se aproxima a una curva normal. Los coeficientes del filtro son determinados convolucionando la matriz $[1/2 \ 1/2]$ consigo misma y luego, iterativamente, convolucionando el resultado con $[1/2 \ 1/2]$ un cierto número de veces (The MathWorks Inc., sf). En nuestro caso, se utilizó n=4. La Figura 5.3 muestra como ejemplo una comparación de los datos crudos contra los datos filtrados para la boya 42020 en el mes de febrero.



Figura 5.3: Ejemplo de datos filtrados. Se muestra la serie de tiempo de febrero de la altura significante en la boya 42020. En negro las observaciones crudas y en rojo las observaciones filtradas.

5.2. Procesamiento de los datos del modelo

Antes de realizar la comparación del modelo con las observaciones de las boyas, los datos de las simulaciones fueron procesados con dos objetivos: el primero, que representaran adecuadamente la información generada por el modelo en la posición de cada boya y el segundo, que las comparaciones dieran oportunidad de disctutir las variaciones temporales del desempeño del modelo, tanto en función del periodo del año como del tiempo de simulación. Para el primer objetivo, se realizó una interpolación espacial y una temporal de las salidas del modelo a la posición de cada boya. Por su lado, para el segundo objetivo las simulaciones se divieron en periodos de 24 horas y se agruparon mensualmente.

A continuación se describen a mayor detalle estos procedimientos.

5.2.1. Interpolación a la posición espacio-temporal de las boyas

La malla utilizada en las simulaciones es de tipo regular, específicamente, rectilínea. Ya que las boyas en general no caen en un punto de malla, para generar las series de tiempo que se utilizaron para comparar con cada boya, se hicieron interpolaciones de los puntos de malla del modelo a la posición de cada boya. El esquema de la Figura 5.4 ejemplifica una distribución hipotética de puntos de malla y, dada la posición de una boya, los puntos que se toman para realizar la interpolación. Así, para una boya cualquiera se localizaron los cuatro puntos de malla con menor distancia euclidiana a ésta y su información se interpoló bilinealmente a la posición de la boya.



Figura 5.4: Ejemplo de la distribución de los puntos de malla usados para validar el modelo contra un boya.

Una vez que se obtuvieron las series de tiempo en la posición de cada boya, se realizaron interpolaciones lineales en tiempo. Esto debido a que los datos del modelo se reportan en el minuto 00 de cada hora, mientras que los datos de las boyas están en el minuto 50 de cada hora.

5.2.2. Separación de las salidas del modelo por periodo de simulación

Como se describió en el Capítulo 4, el sistema de pronóstico de oleaje que se valida en esta tesis genera para cada día el pronóstico de las siguientes 120 horas. Por lo anterior, para el año simulado para realizar la validación (2015), también se generaron 120 horas de simulación para cada día.

Estas 120 horas de simulación fueron segmentadas en lapsos de 24 horas, pues uno de los objetivos de la validación es determinar el cambio en la calidad del pronóstico de acuerdo al tiempo de simulación. De este modo, para cada día simulado se obtuvieron cinco series de tiempo (*ver Figura 5.5a*): serie a 24 horas, que va de la hora 1 a la 24 (azul); la serie a 48 horas, que va de la hora 25 a la 48 (rojo); la serie a 72 horas, que va de la hora 49 a la 72 (amarillo); la serie a 96 horas, que va de la hora 73 a la 96 (morado) y la serie a 120 horas, que va de la hora 97 a la 120 (verde).

Estas series diarias fueron agrupadas mensualmente, de modo que para cada mes, se obtuvieron cinco series (ver Figura 5.5b): la serie mensual a 24 horas (que agrupa las series a 24 horas de cada día del mes), la serie mensual a 48 horas (que agrupa las series a 48 horas de cada día del mes), la serie mensual a 72 horas (que agrupa las series a 72 horas de cada día del mes), la serie mensual a 96 horas (que agrupa las series a 96 horas de cada día del mes) y la serie mensual a 120 horas (que agrupa las series a 120 horas de cada día del mes).



Figura 5.5: Ejemplo del agrupamiento de las simulaciones diarias en cinco series mensuales armadas con periodos de simulación de 24 horas.

Las cinco series mensuales obtenidas con este procedimiento para cada variablea validar fueron las que se compararon con los datos de las boyas para los periodos correspondientes.

5.3. Métricas estadísticas utilizadas en la evaluación del modelo

Una vez obtenidas las series de tiempo mensuales, tanto las de observaciones como las del modelo, se procedió a realizar su comparación cuantitativa. Esta evaluación cuantitativa fue hecha utilizando siete métricas estadísticas distintas, de acuerdo a cada tipo de variable analizada. La altura significante, el periodo de oleaje y la rapidez del viento fueron evaluadas utilizando cinco de estas métricas: el sesgo, la raíz del error cuadrático medio, el coeficiente de correlación de Pearson, el error no sistemático y el índice de concordancia. Por su lado, para la dirección de oleaje se utilizó el sesgo angular y el coeficiente de correlación compleja. En este capítulo se describen las formulaciones e interpretaciones de cada métrica utilizada.

5.3.1. Sesgo o error medio (EM)

El sesgo está definido como la sobreestimación o subestimación sistemática hecha con respecto al valor real de una variable (Rumsey, 2007), por lo que también se le conoce como error sistemático o error medio. Es calculado como el promedio de las diferencias entre los valores pronósticados y los observados:

$$EM = \frac{1}{N} \sum_{n=1}^{N} (p_n - o_n)$$
(5.2)

donde N es el total de pares de datos (es decir, cada valor pronosticado con su observación correspondiente), p_n es el n-ésimo valor pronósticado y o_n la n-ésima observación.

Si los valores del pronóstico son en promedio mayores a los reales, presentará un EM mayor a cero, mientras que valores del pronóstico menores a los reales presentará un EM menor a cero (Wilks, 2005). Un pronóstico perfecto debería exhibir un EM igual a cero; sin embargo, errores de diferente signo pero misma magnitud pueden cancelarse entre sí y producir dicho valor.

Es por lo anterior que el sesgo indica el sentido medio en el que el pronóstico se desvía de los valores observados, pero no necesariamente representa la magnitud del error. Así pues, esta no es una métrica que pueda, por sí sola, describir la exactitud en el desempeño de un modelo.

5.3.2. Raíz del error cuadrático medio (RMSE).

Para definir la raíz del error cuadrático medio (RMSE, por sus siglas en inglés), es importante antes definir al error cuadrático medio (MSE). El MSE es un promedio espacial de las diferencias cuadradas individuales entre cada valor pronosticado por el modelo y su correspondiente observación (Wilks, 2005):

$$MSE = \frac{1}{N} \sum_{n=1}^{N} (p_n - o_n)^2$$
(5.3)

Esta métrica incorpora tanto el concepto de sesgo como el de varianza, pues se puede demostrar que el MSE es igual a la suma de la varianza más el EM al cuadrado (Walther and Moore, 2005). Ya que la varianza describe la precisión de un arreglo de datos, al incorporarla el MSE se convierte en una buena herramienta para medir la exactitud de un modelo.

Ahora bien, ya que es deseable trabajar con la escala y unidades originales de las variables, a menudo se toma la raíz cuadrada del MSE,

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (p_n - o_n)^2}$$
(5.4)

que es la denominada RMSE, y tiene la ventaja de que puede ser fácilmente interpretada como la magnitud del error del modelo. La RMSE representa el error total del pronóstico, conteniendo tanto los errores sistemáticos como los aleatorios, de modo que un desempeño perfecto del modelo estaría dado por un RMSE = 0.

Esta métrica, sin embargo, tiene la desventaja de que no indica la dirección de los errores y tiende a ser dominada por los valores de error extremos (valores en los cuales la diferencia entre la medición y el pronóstico es muy grande).

5.3.3. Error no sistemático (RMSEb)

El error total que tiene el pronóstico generado por un modelo es la suma de errores sistemáticos y errores aleatorios. Como ya se mencionó, el EM es una medida del error sistemático, mientras que el RMSE es una medida del error total. Así pues, la forma más sencilla de calcular la magnitud de los errores aleatorios presentes en un pronóstico es removiendo el error sistemático del error total. La métrica utilizada en este trabajo para dicho fin es la siguiente:

$$RMSEb = \sqrt{\frac{1}{N} \sum_{n=1}^{N} [(p_n - o_n) - EM]^2}$$
(5.5)

que es equivalente al RMSE una vez que se ha removido el EM.

5.3.4. Coeficiente de correlación de Pearson (r)

El coeficiente de correlación de Pearson es una medida de la fuerza de la relación lineal entre dos variables (Isotalo, 2001), representada como la proporción entre la covarianza de las variables con respecto al producto de sus desviaciones estándar:

$$r_{xy} = \frac{Cov(x,y)}{S_x S_y}$$

Los valores de este coeficiente están acotados entre -1 y 1, tomando el valor de 1 cuando las dos variables tienen una relación lineal positiva perfecta. Por otro lado, si el coeficiente toma un valor de -1, las variables tienen una relación lineal negativa perfecta. Cuando las variables no tienen relación lineal alguna, el coeficiente toma un valor de cero (Wilks, 2005).

En el caso de la evaluación del desempeño de un modelo, más que una correlación, el cálculo de este coeficente ofrece una medida intuitiva de qué tan parecidas son las predicciones a las observaciones. Su cálculo queda dado por la siguiente expresión:

$$r = \frac{\sum_{n=1}^{N} [(p_n - \bar{p})(o_n - \bar{o})]}{\sqrt{\sum_{n=1}^{N} (p_n - \bar{p})^2} \sqrt{\sum_{n=1}^{N} (o_n - \bar{o})^2}}$$
(5.6)

donde la barra denota la media aritmética de las variables. En este sentido, un desempeño perfecto del modelo estaría representado por un coeficiente de correlación r = 1, además de cumplir que en un diagrama de dispersión la recta que mejor se ajuste sea la de pendiente igual a 1.

Aunque el uso de esta métrica para la evaluación de modelos ha sido controvertido, principalmente porque su formulación no está relacionada con el concepto de exactitud de predicción (Willmott, 1982), su sencilla interpretación es útil para tener una idea general del desempeño de un modelo.

5.3.5. Indice de concordancia (IOA)

El Index of agreement (IOA), traducido en esta tesis como *índice de concordancia*, fue originalmente desarrollado por Cort J. Willmott y Donald E. Wicks en un trabajo de 1980, y modificado en varios trabajos posteriores. La versión utilizada en esta tesis corresponde a la publicación más reciente, que se realizó en el año 2012.

Este índice pondera la magnitud del error del pronóstico (entendido como la diferencia entre cada valor pronosticado con su observación correspondiente) en relación con la desviación que presentan las observaciones respecto a su media,

$$IOA = \begin{cases} 1 - \frac{\sum\limits_{n=1}^{N} |p_n - o_n|}{c \sum\limits_{n=1}^{N} |o_n - \bar{o}|}, \text{si } \sum\limits_{n=1}^{N} |p_n - o_n| \le c \sum\limits_{n=1}^{N} |o_n - \bar{o}| \\ \frac{c \sum\limits_{n=1}^{N} |o_n - \bar{o}|}{\sum\limits_{n=1}^{N} |p_n - o_n|} - 1, \text{si } \sum\limits_{n=1}^{N} |p_n - o_n| > c \sum\limits_{n=1}^{N} |o_n - \bar{o}| \end{cases}$$
(5.7)

donde c es una constante de valor 2, que surge de la forma en que fue desarrollado el índice (Willmott et al., 2012).

El valor del IOA está acotado entre -1 y 1, tomando este último cuando el pronóstico tiene un desempeño perfecto. Por otro lado, cuando el índice toma un valor de cero significa que la suma de los errores del modelo es equivalente a la magnitud de la desviación de las observaciones. Finalmente, un valor de -1 indica que existe poca representación entre el modelo y las observaciones, aunque también puede representar que existe una muy baja variabilidad en las observaciones.

5.3.6. Sesgo angular (EM_{ang})

Al igual que el EM, el sesgo angular se define como la sobreestimación o subestimación sistemática hecha con respecto al valor real de un ángulo (Bi et al., 2015):

$$EM_{ang} = \begin{cases} \arctan(\frac{S}{C}), & \text{si } S > 0, C > 0\\ \arctan(\frac{S}{C}) + \pi, & \text{si } C < 0\\ \arctan(\frac{S}{C}) + 2\pi, & \text{si } S < 0, C > 0 \end{cases}$$
(5.8)

donde S y C representan la suma del seno y el coseno, respectivamente, de las diferencias entre el ángulo pronosticado por el modelo y el registrado en las observaciones:

$$S = \sum_{n=1}^{N} sen |\theta_p - \theta_o|_n$$
$$iC = \sum_{n=1}^{N} cos |\theta_p - \theta_o|_n$$

El valor máximo absoluto que puede tomar el sesgo angular es de 180° , que representaría el caso en el que el valor pronósticado está exactamente en dirección contraria al valor medido. Por su lado, un pronóstico perfecto tendría un sesgo de 0° .

5.3.7. Coeficiente de correlación compleja (r_{anq})

Kundu (1976) propuso que una forma de encontrar el desplazamiento angular promedio entre dos series de vectores es a través del coeficiente de correlación compleja. Para calcular el coeficiente, cada dato debe ser representando como un número complejo. En el caso de la dirección de oleaje (reportada en grados desde el norte) tenemos que:

$$\omega_i = sen(\theta_i) + i \cdot cos(\theta_i)$$

donde θ representa el ángulo. Una vez que las dos series de datos (valores pronosticados y valores observados) son representadas de forma compleja, el coeficiente de correlación compleja se calcula como el producto interno normalizado:

$$r_{ang} = \frac{\langle \tilde{\omega_o} \cdot \tilde{\omega_p} \rangle}{\sqrt{\langle \tilde{\omega_o} \cdot \omega_o \rangle \langle \tilde{\omega_p} \cdot \omega_p \rangle}}$$
(5.9)

donde la tilde representa el complejo conjugado.

El coeficiente ρ es un número complejo cuya magnitud, que es siempre menor a 1, indica la correlación (entendida de la misma manera que la correlación lineal) entre las dos series de datos.

5.4. Conclusiones del capítulo

Las observaciones utilizadas para realizar la validación fueron seleccionadas de modo que se pueda considerar que los resultados obtenidos a partir de este análisis son representativos del desempeño del modelo.

Las métricas descritas en este capítulo han sido desarrolladas para diferentes fines. En cada caso, sus usos tienen tanto ventajas como desventajas, por lo que las interpretaciones que se hagan de sus resultados deben tomar en cuenta las limitaciones de cada una de ellas. Es por lo anterior que la evaluación del desempeño del modelo se hizo utilizando varias métricas y analizando de manera conjunta sus resultados.

A través de las métricas de error utilizadas $(EM/EM_{ang}, RMSE \ y \ RMSEb)$ no solo se puede cuantificar el error total del modelo, sino también conocer el sentido predominante en el que se da ese error y determinar si tiene un origen sistemático o aleatorio. Por su lado, las métricas de bondad $(r/r_{ang}$ e IOA) cuantifican qué tanto el modelo reproduce la magnitud y el sentido de las observaciones. El coeficiente de correlación permite realizar esta interpretación de manera sencilla e intuitiva. Sin embargo, como se discutió en este capítulo, no fue creado con el fin de ser una métrica de validación, por lo que puede resultar un poco laxo. Por lo tanto, el uso del IOA, que sí es una métrica desarrollada con fines de validación y toma valores en el mismo intervalo que r (-1 a 1), permite interpretar los resultados con mayor reserva. De este modo, las métricas seleccionadas ayudan a tener una visión integral del desempeño del modelo.

Capítulo 6 Resultados y discusión

Este capítulo resume los resultados de esta tesis en dos secciones principales: una enfocada en los resultados de la implementación del modelo como tal y otra enfocada en la validación. En la primera parte de este capítulo se describen los productos que se obtuvieron como resultado de la implementación operativa del modelo de oleaje. Estos productos incluyen varias piezas de software (scripts) que se generaron como parte del desarrollo de esta tesis para la implementación del sistema de pronóstico de oleaje, los archivos NetCDF de las simulaciones generadas diariamente por el modelo WWIII y los productos gráficos que se generan y publican a partir de dichas simulaciones. Para ejemplificar los productos gráficos generados se muestran figuras recientes, correspondientes al 03 de diciembre de 2019. En la segunda parte de este capítulo se exponen los resultados estadísticos de la validación del modelo de oleaje que se hizo para el año 2015 y se discuten sus posibles causas, así como su concordancia con estudios previos.

6.1. Resultados de la implementación

El modelo de oleaje WWIII quedó implementado de forma operativa desde octubre de 2017 y hasta la fecha sigue ejecutándose de manera diaria. Resultado de esta implementación, se generaron los siguientes programas:

- 1. Un script maestro desarrollado en BASH que se encarga de gestionar el preprocesamiento de los datos de entrada del WWIII, crear los archivos de configuración necesarios y lanzar la ejecución del modelo.
- 2. Cuatro scripts desarrollados en NCL para realizar la interpolación de los campos forzantes: uno para el campo de proporción de cobertura de hielo y tres para los campos de viento de cada dominio simulado.
- 3. Un script desarrollado en BASH para la gestión del graficado de las variables de salida del modelo.
- 4. Cuatro scripts desarrollados en NCL para generar los productos gráficos, uno para cada variable: la altura significante, el periodo de oleaje, la altura de la marejada y el periodo de la marejada.
- 5. Un script desarrollado en BASH que se encarga de respaldar tanto las salidas en NetCDF como los productos gráficos que se generan cada día con el sistema de pronóstico de oleaje.

Como se ha mencionado varias veces, la implementación del modelo de oleaje es parte de un sistema de pronóstico numérico más grande, por lo que los algunos de estos scripts fueron trabajados por varias

personas que tuvieron participación en el desarrollo del proyecto en el que se enmarca esta tesis.

La ejecución del modelo WWIII genera 15 archivos NetCDF diarios, 5 para cada dominio (uno para cada 24 horas de simulación). Los archivos del dominio del Oceáno mundial pesan 735 Mb cada uno, los del dominio del Pacífico mexicano 45 Mb y los del dominio del Golfo de México 24 Mb, de modo que de manera conjunta se generan 4 Gb de información cada día. Cada uno de estos archivos contiene las siguientes variables de salida:

- 1. Parámetros integrales de oleaje
 - Altura significante (HS)
 - Longitud de onda promedio (LM)
 - Periodo de oleaje promedio (T01)
 - Periodo de oleaje promedio (T02)
 - Periodo de oleaje promedio (T0M1)
 - Frecuencia pico (FP)
 - Dirección promedio (DIR)
 - Dirección pico (DP)
- 2. Parámetros de seis particiones del espectro
 - Altura de ola (PHS_n)
 - Periodo pico (PTP_n)
 - Longitud de onda pico (PLP_n)
 - Dirección promedio $(PDIR_n)$

En la lista anterior, el subíndice indica el número de partición que va de 0 a 5. Las particiones que hace el modelo corresponden al oleaje local (n = 0) y a los cinco sistemas de *swell* más energéticos que se identifiquen, con n = 1 para el sistema más energético.

De todas las variables anteriores, se generan productos gráficos para cuatro: la altura significante HS (Figura 6.2), el periodo de oleaje promedio T01 (Figura 6.3), la altura de ola del primer *swell* PHS_1 (Figura 6.4) y el periodo pico del primer *swell* PTP_1 (Figura 6.5). Estas cuatro variables se grafican para cada dominio. Todos los productos gráficos pueden consultarse en la sección de *pronósticos* de la página web del Grupo IOA, en la pestaña de *Oleaje*: http://grupo-ioa.atmosfera.unam.mx/pronosticos/ index.php/oleaje (Figura 6.1).

En la Figura 6.2, la barra de color indica la altura significante en metros, mientras que las flechas indican la dirección media del oleaje. La barra de color empleada en el dominio global difiere de la barra utilizada para los dominios del Pacífico mexicano y el Golfo de México ya que, por fines de protección civil, resulta de interés resaltar alturas significantes a partir de 2 m en los dominios nacionales.



Figura 6.1: Sección de oleaje de la página de pronósticos numéricos del Grupo IOA (http://grupo-ioa. atmosfera.unam.mx/pronosticos/index.php/oleaje).



Figura 6.2: Productos gráficos generados con el sistema de pronóstico de oleaje para la variable de altura significante. En la parte superior se aprecia una imagen para el dominio del Océano mundial, mientras que en la parte de abajo se aprecia una imagen para el dominio del Pacífico mexicano (izquierda) y una para Golfo de México (derecha). Las tres gráficas fueron generadas con las salidas correspondientes a la hora 05 del pronóstico iniciado el día 03 de diciembre de 2019 a las 00:00 GMT.

En la Figura 6.3, la barra de color indica el periodo medio de oleaje en segundos, mientras que las flechas indican la dirección media del oleaje. A diferencia del caso de los productos gráficos para la variable de altura significante, para el periodo de oleaje la barra de color del dominio global es igual a la utilizada para el dominio del Pacífico mexicano, pero diferente a la utilizada en el Golfo de México. En este caso, la diferencia acata al hecho de que tanto el Océano mundial como el Pacífico mexicano están constantemente sometidos a la presencia de sistemas de *swell*, los cuales se caracterizan por tener periodos más largos. Por su lado, el Golfo de México no suele estar bajo la influencia de grandes sistemas de *swell*, por lo que en este dominio los periodos son más cortos.

Modelo WWIII 03/12/2019 00:00 GMT Periodo (s) y dirección



Figura 6.3: Productos gráficos generados con el sistema de pronóstico de oleaje para la variable de periodo medio de oleaje. Las gráficas para los tres dominios fueron generadas con las salidas correspondientes a la hora 05 del pronóstico iniciado el día 03 de diciembre de 2019 a las 00:00 GMT.

Para la Figura 6.4, la barra de color indica la altura del sistema de *swell* más energético mientras que las flechas indican la dirección promedio para dicho sistema. Las barras de color utilizadas para esta variable son iguales que las utilizadas para la altura significante. Pese a que esta figura muestra un ejemplo en particular, captura dos características importantes de los dominios: en primer lugar, se aprecia la presencia de sistemas de *swell* en casi todo el dominio del Océano mundial y del Pacífico mexicano, pero no así en el Golfo de México; en segundo lugar, se aprecia que las alturas de oleaje



asociadas al *swell* son mayores en el Pacífico mexicano que en el Golfo de México.

Figura 6.4: Productos gráficos generados con el sistema de pronóstico de oleaje para la variable de altura de marejada. Las gráficas para los tres dominios fueron generadas con las salidas correspondientes a la hora 05 del pronóstico iniciado el día 03 de diciembre de 2019 a las 00:00 GMT.

Finalmente, en la Figura 6.5 la barra de color indica el periodo pico del primer sistema de *swell*, mientras que las flechas indican la dirección promedio de esta partición. Las barras de color empleadas en esta figura son iguales que las utilizadas para la variable de periodo medio de oleaje. Ya que los ejemplos mostrados en esta figura corresponden a la misma fecha y hora que los ejemplos de las figuras anteriores, las zonas con presencia de *swell* son las mismas que en la Figura 6.4. Destaca que la variable de periodo pico representa muy claramente los diferentes frentes de onda, coincidiendo muy bien con la

Pronóstico a 005 hrs Modelo WWIII 03/12/2019 00:00 GMT Periodo Swell (s) y dirección Hora Local CDMX 23:00 02/12/2019 305 150W 120W 60W 30W 30E 120E 90W 0 60E 90E 150E Periodo Swell (s) 5 6 8 16 18 2 3 4 9 10 11 12 13 14 15 17 19 20 Pronóstico a 005 hrs Hora Local CDMX 23:00 02/12/2019 Modelo WWIII 03/12/2019 00:00 GMT Periodo Swell (s) y dirección Modelo WWIII 03/12/2019 Periodo Swell (s) y dirección 03/12/2019 00:00 GMT Pronóstico a 005 hrs Hora Local CDMX 23:00 02/12/2019 Periodo Swell (s) 11 12 13 14 15 16 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0 12.0 10 17 18

dirección en la que se propagan estos sistemas. Además, se observa que los periodos asociados a *swell* son, en general, mayores a 5 segundos.

Figura 6.5: Productos gráficos generados con el sistema de pronóstico de oleaje para la variable de periodo de marejada. Las gráficas para los tres dominios fueron generadas con las salidas correspondientes a la hora 05 del pronóstico iniciado el día 03 de diciembre de 2019 a las 00:00 GMT.

6.2. Resultados de la validación

Como se mencionó en el capítulo de Método de validación, las 120 horas de simulaciones diarias fueron separadas en cinco periodos de 24 horas y agrupadas mensualmente para construir cinco series

mensuales. Cada una de estas series se interpoló a la posición de la boya correspondiente, obteniendo métricas para cada periodo de simulación y cada mes, esto es, 480 valores de estadísticos (8 boyas x 12 meses x 5 periodos). Estos resultados se resumen en varios tipos de gráficos mostrados en esta sección.

En esta sección se presentan diagramas de dispersión con recta de ajuste para las variables de altura significante y periodo de oleaje (identificadas en las gráficas como HS y T, respectivamente), mientras que para la dirección de oleaje (identificada como DIR en las gráficas) se presentan series de tiempo. Para las tres variables se presentan gráficas de las métricas estadísticas como función del tiempo (mes del año) y como función del periodo de simulación. Finalmente, se presentan rosas de oleaje para comparar la distribución (frecuencia de datos en determinadas direcciones y de determinadas alturas significantes) que sigue el modelo con respecto a las boyas.

Además de la validación de las variables del modelo de oleaje, se presentan los resultados para la validación de la magnitud del viento utilizado como forzante (proveniente del modelo WRF). La variable es identificada en las gráficas como WSP. Estos resultados se presentan con el fin de determinar si el comportamiento presentado en el desempeño del modelo de oleaje sigue al del campo forzante.

6.2.1. Magnitud del viento

Nota: únicamente en esta subsección de resultados de validación de magnitud del viento, al hablar de el "modelo", se hace referencia al modelo WRF.

Los diagramas de dispersión son útiles para la interpretación de la calidad de los resultados de un modelo numérico, pues en el caso de un pronóstico perfecto, todos los puntos tendrían que caer en la recta de pendiente igual a uno. Así pues, estos gráficos muestran, a través de la dispersión que presentan los puntos y de qué tanto se aleja la recta ajustada de la recta identidad, el parecido de los valores pronosticados con las observaciones.

En la Figura 6.6 se muestran diagramas de dispersión generados con los datos del todo el año y para las 120 horas de simulación de la magnitud del viento usado como forzante (proveniente del modelo WRF). Es importante mencionar que estos diagramas se construyeron con los datos interpolados a la malla del modelo WWIII y luego a la posición de la boya.

Se puede observar que aquellos puntos correspondientes a las primeras 24 horas de simulación (azul más oscuro) son los que presentan la menor dispersión y se acercan más a la recta de la función identidad (y = x, identificada en gris). Si el modelo pronosticara perfectamente las observaciones de las boyas, entonces todos los puntos tendrían que caer sobre dicha recta. Por su lado, los puntos correspondientes al periodo de las 120 horas, es decir, las últimas 24 horas de simulación (puntos amarillos) son los que presentan mayor dispersión, alejándose de la recta identidad, lo cual indica que el modelo WRF reproduce mejor a las observaciones en las primeras 24 horas de simulación. Por su lado, la recta de mejor ajuste (rojo), cuya ecuación se muestra en la esquina superior izquierda de cada diagrama, nos brinda información de la tendencia general del sesgo del modelo WRF (que se discutirá más adelante).



Figura 6.6: Diagramas anuales de dispersión para las ocho boyas analizadas para la magitud del viento. El color de los puntos representa el intervalo de horas de simulación al que pertenece el dato, la línea roja indica la recta que mejor se ajusta a los datos y la línea gris indica la función identidad (y = x).

En general, el desempeño del WRF parece ser muy similar en las ubicaciones de todas las boyas, teniendo coeficientes de correlación de más de 0.7 para todos los casos. Sin embargo, se aprecia una tendencia del modelo a sobreestimar los vientos de menor magnitud y a subestimar magnitudes mayores. Esto se observa porque la recta de mejor ajuste inicia por encima de la recta identidad, posteriormente la cruza y termina debajo de ella. Aunque el punto de cruce es distinto para todos los casos, es cercano a los 5 m/s, que es aproximadamente la mitad del intervalo de velocidades presentadas. La excepción a esta tendencia se da en la boya 42057, ubicada en el Caribe, en la cual el WRF sobreestima los vientos de hasta 7 m/s. A pesar de esto, junto con la boya 42002, tiene el coeficiende de correlación más alto (r = 0.76).

En la Figura 6.7 se muestran los estadísticos de cada boya como función del mes, para lo cual fueron promediados los resultados de los cinco periodos de simulación para cada mes. En la gráfica del EM se puede ver que la boya 42057 presenta un sesgo positivo claramente superior al resto de las boyas, consistente con lo observado en su diagrama de dispersión. No obstante, en la gráfica del error no sistemático,

esta boya presenta un error menor que las demás. Esto indica que el modelo WRF no está simulando adecuadamente las condiciones de viento en la ubicación de esta boya. Dicho error no es capturado en el coeficiente de correlación lineal ni en el *IOA*, pues se observa que los valores de estas métricas en esta boya son muy semejantes a los que se presentan en el resto de las boyas.



Figura 6.7: Índices estadísticos para WSP como función del mes para todas las boyas analizadas.

Aunque no se aprecia una tendencia estacional clara, en las gráficas de $r \in IOA$ se observa que en abril y septiembre disminuye la calidad del desempeño del pronóstico de viento. Además, resalta que en la boya 42001 se tienen un desempeño particularmente malo en agosto. Al observar la serie de tiempo de datos de esta boya (ver Figura A.1), se puede ver que durante este periodo hay muy pocos datos, de modo que es probable que ese sea el origen del repentino declive del desempeño del modelo WRF en la ubicación de esa boya.

Más allá de las observaciones anteriores, el desempeño general del modelo WRF se puede considerar bueno. La media de los errores sistemáticos (EM) oscila al rededor de los 0.5 m/s, que es poco, mientras que las medias del coeficiente de correlación lineal y del *IOA* en general es mayor a 0.6 y 0.5, respectivamente.



Figura 6.8: Índices estadísticos para WSP como función del intervalo de simulación.

En la Figura 6.8 se muestran los índices como función del intervalo de simulación. La dependencia es muy clara para todos los estadísticos: el modelo WRF se desempeña mejor en las primeras horas de simulación y va perdiendo precisión hacia el final del periodo simulado. En esta Figura es aún más claro que el desempeño del modelo en la posición de la boya 42057 es el peor, pues para todos los periodos de simulación presenta el mayor sesgo. Nuevamente, llama la atención que ni el valor del coeficiente de correlación ni el del *IOA* indican que para esta boya el modelo WRF tenga un mal desempeño en comparación con las otras.

6.2.2. Altura significante

En la Figura 6.9 se muestran los diagramas de dispersión de la altura significante para las ocho boyas analizadas, incluyendo todos los datos del año y para las 120 horas de simulaciones diarias. Se aprecia que para las boyas 42001, 42002, 42020 y 42055 la recta de mejor ajuste está por encima de la recta identidad para los valores más bajos de *HS*, pero hacia la mitad del intervalo la atraviesa, quedando por debajo. Esto indica que de manera general el modelo está sobreestimando los valores más bajos de las olas y subestimando los valores más altos. Las boyas 42035 y 42040 muestran siempre una tendencia a ser subestimadas por el modelo, pues la recta de ajuste está siempre por debajo de la recta identidad (para la boya 42040 en realidad la ordenada al origen es positiva, pero es muy cercana a cero). Estas son dos de las boyas que se localización en aguas más someras (ver Cuadro 5.2). Finalmente, las boyas 42056 y 42057 presentan una tendencia contraria, siempre siendo sobreestimadas por el modelo y con rectas de ajuste más alejadas de la recta identidad. Ambas boyas se encuentran en realidad en el Caribe. Lo

anterior podría indicar la existencia de un patrón espacial en el desempeño del modelo para esta variable.

A pesar de las diferentes tendencias mencionadas, todas las boyas tienen un coeficiente de correlación muy alto (todos mayores a 0.8), siendo la boya 42002 la que presenta el coeficiente mayor (0.89). Los coeficientes de correlación mostrados en la Fig 6.9 corresponden a la correlación de la serie anual completa (es decir, incluyendo las cinco series de cada mes) para cada boya. Los coeficientes obtenidos por mes se discutirán más adelante.

En la Figura 6.10 se muestran los estadísticos de cada boya como función del mes. Esta figura muestra que todas las boyas tienen un comportamiento similar en sus resultados, exceptuando a las boyas que se encuentran el Caribe (42056 y 42057). Estas dos boyas muestran un sesgo y RMSE mucho mayor que las demás, así como valores de IOA menores. Sin embargo, llama la atención que sus valores de correlación no son particularmente bajos en comparación con el resto de las boyas. De hecho, las boyas 42040 y 42055 llegan a presentar valores de r menores en julio y septiembre, respectivamente.



Figura 6.10: Indices estadísticos para H_S como función del mes para todas las boyas analizadas.

De forma más general, no se aprecia una tendencia estacional clara en los resultados de la evaluación. La excepción son nuevamente las boyas del Caribe, para las cuales hay un peor desempeño entre mayo y septiembre. En el caso del coeficiente de correlación se observa que entre diciembre y marzo los valores de r se mantienen por encima de 0.8 y son muy similares entre todas las boyas. Además se observa que en abril y septiembre hay una baja generalizada de las correlaciones.



Figura 6.9: Diagramas anuales de dispersión para las ocho boyas analizadas para la altura significante. El color de los puntos representa el intervalo de horas de simulación al que pertenece el dato, la línea roja indica la recta que mejor se ajusta a los datos y la línea gris indica la función identidad (y = x).

De la Figura 6.10 también se aprecia que los valores de RMSE son casi iguales a los de RMSEb, mientras que los de EM son muy pequeños (sin tomar en cuenta las boyas del Caribe). Esto indica que casi todo el error del modelo está contenido en el error aleatorio y muy poco corresponde al error sistemático. Los valores de EM (que oscilan entre -0.2 y 1.1 m) representan entre el 0.2 y 32.4 % de la altura significante media en las primeras 24 horas de simulación y entre el 0.7 y 40.81 % en las últimas 24 horas de simulación.

En la Figura 6.11 se muestran los índices como función del intervalo de simulación, del cual presentan una dependencia clara. Los índices de error (EM, RMSE y RMSEb) aumentan con el tiempo de simulación, mientras que los índices de concordancia (r e IOA) disminuyen. Nuevamente se identifica que los resultados para las boyas del Caribe se separan del resto, con valores de EM mucho más altos y de IOA mucho más bajos (negativos a partir de las 48 horas para la boya 42057).



Figura 6.11: Índices estadísticos para H_S como función del intervalo de simulación.

Para todas las boyas, los valores de r se mantienen por encima de 0.8 en las primeras 24 horas y entre 0.6 y 0.8 para las 120 horas. El IOA es, por su lado, tiene valores entre 0.65 y 0.75 para las primeras 24 horas y entre 0.5 y 0.65 hacia las 120 horas (sin tomar en cuenta las boyas del Caribe).

Los valores de RMSEb tienen una pendiente muy pronunciada, que puede interpretarse como que el error numérico (no asociado a las parametrizaciones del modelo), que es aleatorio, aumenta con el tiempo de simulación. Esto es un comportamiento normal en los modelos numéricos.

Este comportamiento es el esperado y quedó mostrado en los diagramas de dispersión anuales, donde

se observaba que los puntos correspondientes a las primeras 24 horas de simulación presentaban una menor dispersión que para los siguientes periodos de simulación. Para resaltar esta dependencia temporal, en la Figura 6.12 se muestra un ejemplo de las series de tiempo y sus correspondientes diagramas de dispersión, para el periodo de simulación de 24 y 120 horas para la boya 42055 en el mes de mayo. En las series de tiempo se aprecia que el modelo reproduce muy bien la altura significante para ambos intervalos, aunque sí es notorio que se aleja un poco más de las observaciones para el intervalo de 120 horas. Esta diferencia en el desempeño se ve más claramente en las gráficas de correlación, donde la dispersión de la nube de puntos es mayor en el periodo de 120 horas y el coeficiente de correlación disminuye.



Figura 6.12: Serie de tiempo y diagrama de dispersión para HS en la boya 42055 para el periodo de simulación de 24 horas (arriba) y 120 horas (abajo).

Finalmente, cabe resaltar que las tendencias de sesgo para esta variable mostradas por los gráficos de dispersión anuales, son ratificadas en las figuras 6.10 y 6.11. Se aprecia que las boyas 42035 y 42040 tienen un sesgo (EM) negativo persistente y las boyas 42056 y 42057 tienen un sesgo positivo muy alto.

6.2.3. Periodo del oleaje

En la Figura 6.13 se muestran los diagramas de dispersión anuales para el periodo de oleaje. Se aprecia que esta variable en general no es bien representada, pues la dispersión en todas las boyas es alta, además de que no hay una tendencia tan clara en la distribución de los puntos. Esto queda evidenciado por los bajos valores de los coeficientes de correlación obtenidos (comparados con los obtenidos para HS). La boya que es mejor representada por el modelo es la 42055, con un r=0.75. A pesar de ser valores bajos, cinco de las ocho boyas tienen correlaciones mayores a 0.5. Para la boya 42057, la recta de ajuste indica que el modelo en general sobreestima los valores de periodo. Para el resto de las boyas, se observa una tendencia como la presentada en los diagramas de HS, donde los valores bajos son sobreestimados y los altos son subestimados.



Figura 6.13: Diagramas anuales de dispersión para las ocho boyas analizadas para el periodo de oleaje. El color de los puntos representa el intervalo de horas de simulación al que pertenece el dato, la línea roja indica la recta que mejor se ajusta a los datos y la línea gris indica la función identidad (y = x).

En la Figura 6.14 vemos, como en el caso de H_S , que aunque los índices para todas las boyas presentan cierta dispersión, en general los de las boyas del Caribe son lo que presentan los valores de EM y RMSE más altos y los valores de IOA más bajos. Los valores de EM llegan a representar hasta 32% de la media de la variables para las primeras 24 horas de simulación y hasta el 41% para las 120 horas. Al igual que con la altura significante, vemos que el coeficiente de correlación lineal no captura la magnitud de los errores en las boyas del Caribe, pues los coeficientes para estas dos boyas son muy similares a los de las otras seis. Se observa también que para esta variable hay coeficientes de correlación negativos.

Cabe resaltar que para la boya 42001, los errores incrementan considerablemente hacia agosto (cuando deja de presentar datos, ver Figura A.1), que bien podría ser un indicador del deterioro de los sensores.



Figura 6.14: Índices estadísticos para T como función del mes para todas las boyas analizadas.

En la Figura 6.15 se observa que los índices obtenidos para el periodo de oleaje también muestran una dependencia al tiempo de simulación, aunque no es tan marcada como la de los índices para la altura significante. En particular, el EM y el IOA presentan poca variación con respecto al tiempo de simulación, aunque existe. El RMSE, RMSEb y el coeficiente de correlación presentan una tendencia más clara.


Figura 6.15: Índices estadísticos para T como función del mes para todas las boyas analizadas.

A pesar de que en la Figura 6.15 no se aprecie la dependencia del periodo de oleaje al tiempo de simulación de manera muy clara como en el caso de HS, al observar por separado un ejemplo (Figura 6.16) de las series de tiempo y los diagramas de dispersión de los periodos de 24 y 120 horas para una boya seleccionada, se aprecia que sí existe una mejor representación de esta variable en las primeras 24 horas de simulación. Este ejemplo además indica que el periodo sí puede llegar a ser bien representado por el modelo.



Figura 6.16: Serie de tiempo y diagrama de dispersión para T en la boya 42055 para el periodo de simulación de 24 horas (arriba) y 120 horas (abajo).

6.2.4. Dirección del oleaje

Para esta variable se presentan las series de tiempo anuales (Figura 6.17) en vez de diagramas de dispersión, pues estos últimos se verían afectados por puntos de valores cercanos a $0^{\circ}/360^{\circ}$ que aparentarían dispersión mucho mayor a la real. Debido a que la boya 42001 no tiene datos de dirección, se presentan las gráficas para siete boyas. Es notorio que el modelo (azul) reproduce muy bien a las observaciones (rojo), lo cual se ve reflejado en los altos valores del coeficiente de correlación compleja que se reportan. En el caso de esta variable, las boyas del Caribe están muy bien representadas, siendo la boya 42057 la que tiene el r_{ang} más alto. Se puede ver también que en la temporada de nortes (entre octubre y febrero) hay más variabilidad en las direcciones de oleaje.

En la Figura 6.18 se aprecia que los valores de sesgo angular (EM_{ang}) son muy pequeños (entre 1.2° y 1.6°) y no presentan un patrón estacional claro. Por su lado, el coeficiente de correlación compleja muestra una ligera estacionalidad, al presentar sus valores más bajos en los meses de invierno (en particular en enero y febrero) y los valores más altos en los meses de primavera y verano (de abril a agosto).



Figura 6.17: Series anuales de dirección de oleaje.



Figura 6.18: Índices estadísticos para DIR como función del mes para todas las boyas analizadas.

La Figura 6.19 muestra que a diferencia del sesgo en las variables anteriores, el sesgo angular no presenta una dependencia clara del tiempo de simulación, salvo para las boyas 42020 y 42040. Aún así, la media del sesgo (línea negra) sí presenta una ligera tendencia a incrementar conforme aumenta la hora de simulación. El coeficiente de correlación compleja sí muestra una dependencia al tiempo de simulación, aunque los valores se mantienen altos durante todo el periodo, siendo el mínimo el de la boya 42002 a las 120 horas ($r_{ang} = 0.77$).



Figura 6.19: Índices estadísticos para DIR como función del intervalo de simulación.

Al igual que en el caso del periodo, al examinar la serie de tiempo y el coeficiente de correlación para un mes y boya en particular (Figura 6.20), se puede ver que la DIR también es mejor pronosticada para las primeras horas de simulación.



Figura 6.20: Serie de tiempo de DIR en la boya 42035 para el periodo de simulación de 24 horas (arriba) y 120 horas (abajo) en el mes de marzo.

6.2.5. Distribución direccional de la energía

La distribución direccional de energía del modelo se puede aproximar con rosas de oleaje, las cuales muestran la frecuencia de diferentes intervalos de altura significante en cada dirección. En la Figura 6.21 se muestran las rosas para las boyas con las que se hizo la validación y las rosas del modelo para la ubicación de la boya correspondiente. Se aprecia que en general el modelo aproxima bien la distribución presentada por las boyas, pues las rosas del modelo son muy similares a las de las obervaciones.

En general, las olas vienen de direcciones en los cuadrantes NE y SE. Para las boyas 42002 y 42020, dominan las olas viniendo de entre 120° y 150° (grados degT). Para las boyas 42035 y 42040 la dirección dominante es ligeramente más al sur, mientras que para las boyas 42055, 42056 y 42057 la dirección predominante es casi el este (entre 60° y 120°).

Pese a la gran similitud entre las distribuciones del modelo y de las boyas, se aprecian diferencias en la frecuencia de los intervalos de altura significante. Los casos más evidentes son las boyas del Caribe, donde el modelo muestra una mayor frecuencia en los intervalos de $1m \leq H_S < 2m$ y de $2m \leq H_S < 3m$ que las boyas. Este comportamiento es consistente con la sobreestimación generalizada de la altura significante que hace el modelo en la ubicación de estas boyas. Para las demás boyas no es clara esta relación con sus sesgos, pues para la mayoría se mostró que el modelo subestima las alturas más grandes, de modo que se esperaría que las rosas de boyas presenten una mayor frecuencia de las alturas significantes más grandes. Es posible que estas diferencias de frecuencia se deban a que las alturas mayores están distribuidas en direcciones distintas entre el modelo y la boya y, ya que la frecuencia general de H_S altas es poca, al quedar distribuidas en varias direcciones su frecuencia en cada una de ellas es aún menor.

6.3. Discusión de resultados

Con los resultados antes mostrados podemos ver que el desempeño del modelo es bueno para la altura significante y muy bueno para la dirección, sin embargo, no es tan bueno para el periodo.

Las diferencias entre el desempeño del modelo en la ubicación de cada boya mostradas por los diagramas de dispersión de altura significante muestran que hay un patrón espacial en el desempeño. La zona peor representada es la del Caribe, en la cual se presenta un sesgo positivo muy marcado en el pronóstico de la altura significante y el periodo de oleaje. Para el caso de la boya 42057, en la validación de los vientos forzantes usados se observó que esta boya era la peor representada por el modelo WRF, mostrando también un sesgo positivo muy marcado, de modo que es posible que el error presentado en el oleaje para este punto sea heredado del modelo atmosférico. Esto es consistente con la literatura revisada (Capítulo 3), en particular con Chawla et al. (2009), donde se encontró que un sesgo positivo persistente en los vientos GFS originó un sesgo positivo persistente en la HS de las simulaciones con WWIII. Sin embargo, la otra boya del Caribe, la 42056, que también fue mal simulada por el modelo de oleaje, parece ser bien representada por el modelo atmosférico. Un proceso que podría estar afectando las simulaciones para esta zona es que el oleaje esté interactuando fuertemente con la Corriente del Caribe y, considerando que nuestra simulación no incluye información de corrientes, estas interacciones no se resuelven adecuadamente.

Para las boyas 42035 y 42040 se observó un patrón de sesgo negativo persistente, en contraste con el resto de las boyas que muestran un sesgo positivo inicial (para alturas de menos de 2m) que luego se vuelve negativo (para alturas mayores a 2m). Como ya se mencionó, estas dos boyas se encuentran en



Figura 6.21: Rosas de oleaje anuales para las siete boyas con información de dirección.

zonas muy someras (de menos de 200 m de profundidad), por lo que puede ser que este comportamiento tenga que ver con que se ven sometidas a procesos de somerización a los cuales el modelo es sensible (como muestran los trabajos de Tolman et al. 2004 y Mentaschi et al. 2015). Sin embargo, la boya 42020 también se localiza en aguas muy someras (84.1 m) y no mostró el mismo patrón que estas boyas, de modo que no se puede atribuir el comportamiento únicamente a la profundidad. Para las demás boyas, el modelo mostró una tendencia a sobreestimar las alturas de ola más bajas y a subestimar las más altas. Esto es consistente con lo reportado por Tolman (2002) y Mentaschi et al. (2015).

Al comparar las evaluaciones de la altura significante y la magnitud del viento, se observa que esta variable del oleaje hereda los rasgos del forzante utilizado. Además del ejemplo ya comentado del comportamiento de la boya 42057, en los diagramas de dispersión anuales de ambas variables se puede ver que todas las boyas tienen en general un coeficiente de correlación alto y que la boya 42002 es la mejor representada. Por su lado, en las gráficas de estadísticos como función del mes se observó que para ambas variables hay dos meses con un desempeño particularmente malo: abril y septiembre.

En cuanto al periodo de oleaje, Bi et al. (2015) reportó que junto con la dirección de oleaje, eran los parámetros mejor representados, caso totalmente contrario a los resultados encontrados en este trabajo. La mala representación del periodo indica que en el dominio espectral, $F(f, \theta)$, la energía no se está distribuyendo adecuadamente entre las diferentes frecuencias. Esto debe ser originado por el esquema utilizado para representar las entradas por viento (S_{in}) y la disipación (S_{ds}) , que fue el ST2. Entre los autores que compararon varios esquemas (Bi et al. 2015,Mentaschi et al. 2015 y Lee 2015), ninguno recomendó el uso del paquete ST2. En particular, Rogers et al. (2012) recomienda usar el paquete ST4 para simulaciones con dominios anidados, que es el caso del sistema pronóstico analizado en este trabajo. Esto también tiene implicaciones para la tendencia de subestimar valores bajos de altura significante y sobreestimar los valores altos, pues dicha tendencia indica que la disipación y la trasnferencia de energía entre las componentes del espectro no se está dando adecuadamente, lo cual debe tener origen, nuevamente en la elección del paquete ST2.

Sobre la aparente inexistencia de una estacionalidad en la calidad de la simulación, Chawla et al. (2009) encontraron que al hacer mapas de error (con datos satelitales), sí se aprecia una estacionalidad. Sin embargo, para las métricas de error (*Scatter index*) que utilizaron no se aprecia ninguna estacionalidad. Los resultados presentados en este trabajo son consistentes con eso último. Pese a que no hay una estacionalidad notoria, cabe mencionar que los meses con mejor desempeño para la altura significante fueron los meses de invierno (con probable presencia de nortes). En este sentido, Chawla et al. (2009) menciona que observaciones de altura significante muy pequeñas (como en condiciones de calma) pueden llevar a estimaciones de error mayores. En el año analizado (2015) no hubo huracanes cuya trayectoria se desarrollara en el interior del dominio del Golfo de México (únicamente en los márgenes), por lo que los meses de verano presentarían condiciones de mayor calma.

También se observó que, de manera general, el error sistemático (EM) del modelo es pequeño y casi todo el error (RMSE) se debe al error aleatorio (RMSEb). Esto es consistente con los reportado por Rogers (2002). Sin embargo, las boyas del Caribe son una excepción a este comportamiento, pues el error sistemático asociado a ellas es comparable con el error total. Como ya se mencionó, para la boya 42057 este error parece ser heredado del modelo atmosférico. Sin embargo, no es así para la boya 42056. Esto refuerza la hipótesis de que el error con el que se pronostica esta zona tiene que ver con problemas de parametrización del modelo.

Finalmente, la disminución en la calidad del pronóstico con el tiempo de simulación era un comportamiento esperado. Los resultados muestran que las primeras 48 horas de pronóstico tienen un muy buen desempeño, al presentar los valores de error más bajos y las correlaciones más altas, consistente con los reportado por Chawla et al. (2009).

Capítulo 7 Conclusiones

Para este trabajo se realizó la implementación de un sistema de pronóstico de oleaje con el modelo numérico WWIII en el clúster Ometéotl del CCA de la UNAM. Este sistema de pronóstico ha estado trabajando de forma operativa desde hace dos años y diariamente genera un pronóstico para los siguientes cinco días para tres dominios: el océano mundial, el Pacífico mexicano y el Golfo de México. Dicha implementación se logró generando un conjunto de scripts con los cuales se automatizó el preprocesamiento de datos de entrada, la ejecución del modelo y el post-procesamiento de las salidas para ser convertidas a formato NetCDF y posteriormente graficadas. Los productos gráficos obtenidos de este sistema de pronóstico son publicados diariamente en la página web del Grupo IOA.

Se realizó también la validación del sistema de pronóstico para el dominio del Golfo de México. Para este fin, se simuló el año 2015 completo y las salidas del modelo fueron comparadas con datos de ocho boyas del NDBC distribuidas por todo el dominio. La comparación se hizo para tres parámetros integrales del oleaje, considerados los de mayor interés en el contexto del proyecto: la altura significante, el periodo de oleaje y la dirección de oleaje. Además, se hizo la validación de la magnitud del viento que se usó como forzante para el dominio del Golfo de México, el cual fue obtenido del modelo meteorológico WRF. Para la altura significante, el periodo de oleaje y la magnitud del viento se utilizaron cinco métricas para evaluar el desempeño del modelo: el sesgo (EM), la raíz del error cuadrático medio (RMSE), el error no sistemático (RMSEb), el coeficiente de correlación lineal (r) y el índice de concordancia (IOA); por su lado, para la dirección de oleaje se utilizó el coeficiente de correlación compleja (r_{ang}) y el sesgo angular (EM_{ang}) . Para realizar esta evaluación, también se generaron una serie de scripts que automatizaron el tratamiento de las observaciones, el preprocesamiento de las salidas del modelo, la comparación estadística de las series y el graficado de los resultados.

De forma general, se observó que el uso de varias métricas permitió entender y discutir a mayor detalle el desempeño del modelo. Gracias al análisis del EM, el RMSE y el RMSEb se pudo determinar que la mayor parte del error de los resultados corresponde a error aleatorio. Por su lado, el uso del IOApermitió ver que el coeficiente de correlación lineal debe ser interpretado con reserva: incluso en casos donde se presentaba un error medio muy grande, los valores de r eran altos, mientras que los valores de IOA eran considerablemente más bajos. En estos casos el modelo reprodujo la forma y tendencia de las observaciones, pero no la magnitud. Esto indica que valores altos de r indican que el modelo está reproduciendo la forma de la señal observada, pero el valor de IOA debe ser verificado para decir que también se está reproduciendo la magnitud. Aunque para el caso de la dirección de oleaje no se usaron métricas que pudieran compararse entre sí, los resultados fueron siempre consistentes: valores de correlación altos y valores de sesgo bajos. Además, al observar las gráficas de las series de tiempo y las rosas de oleaje se comprueba que el modelo reproduce muy bien esta variable.

Los resultados mostraron que el modelo de oleaje se desempeña muy bien al pronosticar la dirección de oleaje y la altura significante. Sin embargo, su desempeño para pronosticar el periodo de oleaje no es tan bueno. En el caso de la altura significante, se identificó un patrón espacial en la calidad del desempe- $\tilde{n}o$: para las boyas ubicadas en el Caribe (42056 y 42057) el modelo mostró un sesgo positivo muy alto; para las boyas ubicadas en las zonas más someras (42035 y 42040) se observó un sesgo negativo persistente; mientras que para el resto de las boyas (42001, 42002, 42020 y 42055) se observó una tendencia a sobreestimar las alturas significantes bajas (sesgo positivo) y a subestimar las alturas significantes altas (sesgo negativo). Destaca que la boya 42020 también se localiza en aguas muy someras, sin embargo no se comportó de la misma manera que las boyas 42035 y 42040. Al comparar estos resultados con la validación de la magnitud del viento forzante, se observa que, de forma similar al modelo de oleaje, el modelo meteorológico se desempeña bien para todas las boyas en general. Coinciden en que la boya mejor reproducida es la 42002 y que para la boya 42057 se presentan sesgos positivos muy altos. Aunque en la validación del viento no se observó un patrón espacial como en el de oleaje, al graficar las diferentes métricas estadísticas como serie de tiempo anual, tanto para el modelo de oleaje como para el meteorológico se identificaron dos meses con el desemepeño más pobre: abril y septiembre. De forma general, esto coincide con la literatura, la cual reporta que los modelos de oleaje (en particular para la variable de altura significante) son muy sensibles a la calidad de los vientos que reciben como forzantes, de modo que a menudo se observa que el modelo de oleaje reproduce los patrones de error observado en los vientos.

La variable de dirección de oleaje es la mejor pronosticada por el modelo, lo cual también coincide con lo reportado por trabajos previos. Por su lado, a diferencia de lo reportado en la literatura, el periodo de oleaje no es muy bien representado por el modelo de oleaje. Esto puede deberse al esquema usado para representar los procesos de entradas por viento y disipación. Sin embargo, las características de la validación realizada en este trabajo hace que los resultados obtenidos aquí no sean suficientes para comprobar que este sea el origen.

Se confirmó que el modelo pronostica bien las variables evaluadas durante las primeras 48 horas de simulación y la calidad va disminuyendo con el tiempo de simulación. Este comportamiento era el esperado para un modelo numérico.

7.1. Trabajo a futuro

Como trabajo a futuro, sería conveniente ampliar la validación, de modo que se pueda determinar con certeza el origen de los errores identificados en el modelo. Para esto, lo ideal sería realizar llevar a cabo lo siguiente:

- 1. Realizar varias versiones de la implementación con diferentes parametrizaciones para identificar la más adecuada para el dominio de interés.
- 2. Realizar la validación sobre las particiones del espectro para identificar si el problema se origina en la generación del oleaje local o al momento de transferir la energía a sistemas de *swell*.
- 3. Ampliar los campos forzantes utilizados.
- 4. Realizar la validación con observaciones satelitales para verificar a detalle la existencia de patrones espaciales.
- 5. Realizar la validación con un periodo de tiempo más largo de modo que sea representativo de las condiciones medias de oleaje del Golfo.

6. Realizar la validación también para los otros dominios, de modo que se identifiquen diferencias locales.

En el caso de los primeros dos puntos, estos deberían ser cubiertos considerando también el enfoque operativo de este sistema de pronóstico, es decir, que la relación entre la calidad del desempeño y el tiempo de cómputo debe ser tal que permita que el modelo se ejecute en un tiempo razonable y arroje resultados útiles para los analistas.

7.2. Observaciones finales

Finalmente, cabe destacar que el principal aporte de esta tesis es que se generó un sistema de pronóstico de oleaje del cual se conocen sus virtudes, pero también sus limitaciones. Esto lo hace muy útil para fines de protección civil, pero también para otras actividades, como, por ejemplo, la planeación de campañas oceanográficas. Pero además del aporte de tener el sistema de pronóstico en sí mismo, la arquitectura construida para realizar las simulaciones para la validación junto con todos los scripts desarrollados son fácilmente adaptables para casos de estudio específicos, de modo que pueden ser la base para realizar trabajos tan diversos como un análisis de potencial energético de oleaje, revisar y pronosticar los efectos de eventos extremos o caracterizar el oleaje en regiones de interés. Apéndice A

Registros de boyas utilizados en la validación de las variables de interés: rapidez del viento, altura significante, periodo del oleaje y dirección del oleaje



Figura A.1: Registros de la boya 42001 para el año 2015.







Figura A.3: Registros de la boya 42020 para el año 2015.



Figura A.4: Registros de la boya 42035 para el año 2015.



Figura A.5: Registros de la boya 42040 para el año 2015.



Figura A.6: Registros de la boya 42055 para el año 2015.



Figura A.7: Registros de la boya 42056 para el año 2015.



Figura A.8: Registros de la boya 42057 para el año 2015.

Bibliografía

- Affholder, M. and Valiron, F. (2001). Descriptive Physical Oceanography. Taylor & Francis.
- Alcorn, R. (2014). Chapter 17 wave energy. In Letcher, T. M., editor, *Future Energy (Second Edition)*, pages 357 382. Elsevier, Boston, second edition edition.
- Amante, C. and Eakins, B. W. (2009). Etopo1 arc-minute global relief model: procedures, data sources and analysis. Technical report, NOAA Technical Memorandum NESDIS NGDC-24.
- Appendini, C. M., Torres-Freyermuth, A., Oropeza, F., Salles, P., López, J., and Mendoza, E. T. (2013). Wave modeling performance in the Gulf of Mexico and western Caribbean: Wind reanalyses assessment. Applied Ocean Research, 39:20–30.
- Ardhuin, F. and Orfila, A. (2018). Wind waves. In Chassignet, E., Pascual, A., Tintoré, J., and Verron, J., editors, New Frontiers in Operational Oceanography, pages 393 – 422. GODAE OceanView, FRANCE, SPAIN.
- Arinaga, R. A. and Cheung, K. F. (2012). Atlas of global wave energy from 10 years of reanalysis and hindcast data. *Renewable Energy*, 39(1):49 - 64.
- Bi, F., Song, J., Wu, K., and Xu, Y. (2015). Evaluation of the simulation capability of the Wavewatch III model for Pacific Ocean wave. *Acta Oceanologica Sinica*, 34(9):43–57.
- Bidlot, J.-R., Holmes, D. J., Wittmann, P. A., Lalbeharry, R., and Chen, H. S. (2002). Intercomparison of the performance of operational ocean wave forecasting systems with buoy data. Weather and Forecasting, 17(2):287–310.
- Blanchette, C., O'Donnell, M., and Stewart, H. (2008). Waves as an ecological process. In Jørgensen, S. E. and Fath, B. D., editors, *Encyclopedia of Ecology*, pages 3764 – 3770. Academic Press, Oxford.
- Brumer, S. E. (2017). Wave breaking at high wind speeds and its effect on air-sea gas transfer. PhD thesis, Columbia University.
- Cavaleri, L., Fox-Kemper, B., and Hemer, M. (2012). Wind waves in the coupled climate system. Bulletin of the American Meteorological Society, 93(11):1651-1661.
- Cavaleri, L. and Rizzoli, P. M. (1981). Wind wave prediction in shallow water: Theory and applications. Journal of Geophysical Research: Oceans, 86(C11):10961–10973.
- Chawla, A., Spindler, D. M., and Tolman, H. L. (2013). Validation of a thirty year wave hindcast using the climate forecast system reanalysis winds. *Ocean Modelling*, 70:189–206.
- Chawla, A., Tolman, H. L., Hanson, J. L., and Gerald, V. M. (2009). Validation of a multi-grid wavewatch iii modeling system. *NOAA/MMAB Contribution*, 281.

- Chu, P. C., Qi, Y., Chen, Y., Shi, P., and Mao, Q. (2004). South China sea wind-wave characteristics. part I: validation of WAVEWATCH-III using TOPEX/Poseidon data. *Journal of atmospheric and oceanic technology*, 21(11):1718–1733.
- Folley, M. (2017). The wave energy resource. In Pecher, A. and Kofoed, J. P., editors, *Handbook of Ocean Wave Energy*, pages 43–79. Springer International Publishing, Cham.
- Guiberteau, K. L., Liu, Y., Lee, J., and Kozman, T. A. (2012). Investigation of developing wave energy technology in the gulf of mexico. *Distributed Generation & Alternative Energy Journal*, 27(4):36–52.
- Hanson, J. L. and Phillips, O. M. (2001). Automated analysis of ocean surface directional wave spectra. Journal of atmospheric and oceanic technology, 18(2):277–293.
- Hanson, J. L., Tracy, B. A., Tolman, H. L., and Scott, R. D. (2009). Pacific hindcast performance of three numerical wave models. *Journal of Atmospheric and Oceanic Technology*, 26(8):1614–1633.
- Hasselmann, K., Barnett, T., Bouws, E., Carlson, H., Cartwright, D., Enke, K., Ewing, J., Gienapp, H., Hasselmann, D., Kruseman, P., et al. (1973). Measurements of wind-wave growth and swell decay during the joint north sea wave project (jonswap). *Ergänzungsheft 8-12*.
- Hasselmann, S., Hasselmann, K., Allender, J., and Barnett, T. (1985). Computations and parameterizations of the nonlinear energy transfer in a gravity-wave spectrum. part ii: Parameterizations of the nonlinear energy transfer for application in wave models. *Journal of Physical Oceanography*, 15(11):1378–1391.
- Holmes, P. (2001). Phd,"a course in coastal defense systems i chapter 5 coastal processes: Waves,". CDCM Professional Training Programme.
- Holthuijsen, L. H. (2007). Waves in Oceanic and Coastal Waters. Cambridge University Press.
- Isotalo, J. (2001). Basics of statistics.
- Janssen, P. (2004). The Interaction of Ocean Waves and Wind. Cambridge University Press.
- Janssen, P. A. (2008). Progress in ocean wave forecasting. *Journal of Computational Physics*, 227(7):3572 3594. Predicting weather, climate and extreme events.
- Kinsman, B. (1984). Wind Waves: Their Generation and Propagation on the Ocean Surface. Dover Books on Chemistry and Earth Sciences. Dover Publications.
- Komen, G. J., Cavaleri, L., Donelan, M., Hasselmann, K., Hasselmann, S., and Janssen, P. A. E. M. (1994). Dynamics and Modelling of Ocean Waves. Cambridge University Press.
- Konda, M. (s.f.). Study of the impacts of wind waves and energy dissipation on air-sea momentum exchange. Consultado el 05 de diciembre 2019.
- Kubben, P., Dumontier, M., and Dekker, A. (2019). Fundamentals of Clinical Data Science. Springer.
- Kundu, P. K. (1976). Ekman veering observed near the ocean bottom. *Journal of Physical Oceanography*, 6(2):238–242.
- Lee, H. S. (2015). Evaluation of WAVEWATCH III performance with wind input and dissipation source terms using wave buoy measurements for october 2006 along the east korean coast in the East Sea. *Ocean Engineering*, 100:67–82.

- Lefèvre, J.-M. and Cotton, P. (2001). Chapter 7 ocean surface waves. In Fu, L.-L. and Cazenave, A., editors, *Satellite Altimetry and Earth Sciences*, volume 69 of *International Geophysics*, pages 305 xix. Academic Press.
- Li, X. (2010). Ocean surface wave measurements using SAR wave mode data. PhD thesis, Universitat Hamburg.
- Lighthill, M. (1962). Physical interpretation of the mathematical theory of wave generation by wind. Journal of Fluid Mechanics, 14(3):385–398.
- Lionello, P. (2012). The Climate of the Mediterranean Region: From the Past to the Future. Elsevier insights. Elsevier Science.
- Massel, S. (1996). Ocean Surface Waves: Their Physics and Prediction. Advanced series on ocean engineering. World Scientific.
- Mentaschi, L., Besio, G., Cassola, F., and Mazzino, A. (2015). Performance evaluation of Wavewatch III in the Mediterranean Sea. *Ocean Modelling*, 90:82–94.
- National Center for Atmospheric Research (s.f.). Weather Research and Forecasting Model. Consultado el 12 de noviembre 2019.
- National Centers for Environmental Prediction (s.f.). Global Forecast Syrtem (GFS). Consultado el 12 de noviembre 2019.
- National Data Buoy Center (2018a). Do ndbc's meteorological and oceanographic sensors measure data for the entire hour? Consultado el 29 de agosto 2019.
- National Data Buoy Center (2018b). How are spectral wave data derived from buoy motion measurements? Consultado el 29 de agosto 2019.
- Panchang, V., Jeong, C., and Demirbilek, Z. (2013). Analyses of extreme wave heights in the gulf of mexico for offshore engineering applications. *Journal of Offshore Mechanics and Arctic Engineering*, 135:031104.
- Park, B. and Lu, R. (2015). Hyperspectral Imaging Technology in Food and Agriculture. Food Engineering Series. Springer New York.
- Rew, R., Davis, G., Emmerson, S., Davies, H., and Hartnett, E. (1997). Netcdf user's guide. Unidata Program Center, June, 1:997.
- Rogers, W. E. (2002). The US Navy's Global wind-wave models: An investigation into sources of errors in low-frequency energy predictions. Technical report, NAVAL RESEARCH LAB STENNIS SPACE CENTER MS OCEANOGRAPHY DIV.
- Rogers, W. E., Dykes, J. D., Wang, D., Carroll, S. N., and Watson, K. (2012). Validation test report for WAVEWATCH III. Technical report, NAVAL RESEARCH LAB STENNIS DETACHMENT STENNIS SPACE CENTER MS.
- Rumsey, D. (2007). Intermediate Statistics For Dummies. For dummies. Wiley.
- Sawan, A. (2003). Theories of wind generated waves. Consultado el 01 de octubre 2019.
- Stewart, R. H. (2008). Introduction to physical oceanography. Texas A and M University College Station.

- Sullivan, P. P. and McWilliams, J. C. (2010). Dynamics of winds and currents coupled to surface waves. Annual Review of Fluid Mechanics, 42:19–42.
- The MathWorks Inc. (s.f.). Signal smoothing. Consultado el 29 de agosto 2019.
- The SWAMP Group (2013). Ocean Wave Modeling. Springer US.
- Tolman, H. (2002). Validation of WAVEWATCH III version 1.15 for a global domain. Technical report, NOAA/NWS/NCEP/OMB, Tech. Note 213.
- Tolman, H., Alves, J., and Chao, Y. (2004). A review of operational forecasting of wind generated waves by hurricane Isabel at NCEP. Technical report, NOAA/NWS/NCEP/MMAB, Tech. Note 235.
- Tolman, H. L. and Chalikov, D. (1996). Source terms in a third-generation wind wave model. *Journal* of *Physical Oceanography*, 26(11):2497–2518.
- Walther, B. A. and Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, 28(6):815–829.
- WAVEWATCH III Development Group (2016). User manual and system documentation of wavewatch iii version 5.16. Technical report, NOAA/NWS/NCEP/MMAB, Tech. Note 329.
- Wilks, D. (2005). *Statistical Methods in the Atmospheric Sciences*. International Geophysics. Elsevier Science.
- Willmott, C. J. (1982). Some comments on the evaluation of model performance. Bulletin of the American Meteorological Society, 63(11):1309–1313.
- Willmott, C. J., Robeson, S. M., and Matsuura, K. (2012). A refined index of model performance. International Journal of Climatology, 32(13):2088-2094.