



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA  
INGENIERÍA CIVIL – HIDRÁULICA

**IMPLEMENTACIÓN DE REDES NEURONALES PARA GENERAR OLEAJE UTILIZANDO  
DATOS DE VIENTO EN LA COSTA DE SISAL, YUCATÁN.**

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
MAESTRO EN INGENIERÍA

PRESENTA:  
LUIS MAURICIO BRAVO CORTÉS

TUTOR  
JOSÉ LÓPEZ GONZÁLEZ, INSTITUTO DE INGENIERÍA

SISAL YUCATÁN, MÉXICO. ENERO 2016

**JURADO ASIGNADO:**

Presidente: Dra. Enriquez Ortiz Cecilia E.

Secretario: Dr. Torrez Freyermuth Alec

Vocal: Dr. López González José

1<sup>er.</sup> Suplente: Dr. Figueroa Espinoza Bernardo

2<sup>do.</sup> Suplente: M. en C. Appendini A. Christian M.

Lugar o lugares donde se realizó la tesis: LABORATORIO DE INGENIERÍA Y PROCESOS COSTEROS. UNIDAD ACADÉMICA SISAL, YUCATÁN. MÉXICO.

**TUTOR DE TESIS:**

JOSÉ LÓPEZ GONZÁLEZ

-----  
**FIRMA**

(Segunda hoja)

## AGRADECIMIENTOS

*A los investigadores y académicos del LIPC (Laboratorio de Ingeniería y Procesos Costeros), que todos me han ayudado a ser mejor en estos menesteres de los procesos costeros.*

*A mis compañeros que sin ellos y su convivencia no hubiera podido concluir con este trabajo, gracias a todos.*

*A José López González por su paciencia y ayuda para hacer este trabajo de tesis; así como todo su aporte y conocimiento que me brindo durante mi estancia en el LIPC.*

*A Paulo Salles por darme luz verde y la oportunidad de ser parte del LIPC.*

*A Tonatiuh Mendoza y familia, por brindarme su confianza durante mis primeros días en Itzimná, Mérida.*

*A todos los integrantes de comité de evaluación en mi defensa de tesis que me ayudaron a revisar el documento y aportaron en su conocimiento, para que cumpliera con la calidad necesaria y obtener el grado de maestría.*

*Al proyecto internacional del II 4314 “Sistema operacional de predicción de oleaje para las costas y mares mexicanos”.*

*A toda persona que labora o laboro en el LIPC, que no por no mencionarlos no signifique que he aprendido bastante de ellos, gracias.*

*A mi familia por estar siempre presente en mi vida día a día y todo el apoyo que me brindan.*

*A mi madre Marysol Cortés Rodríguez por todo su apoyo y ánimos que día a día me da.*

*A mis abuelos Luis Cortés y Silvia Rodríguez, por todos sus consejos que siempre me brindan.*

## CONTENIDO

Agradecimientos .....	i
Índice de Figuras.....	iv
Índice de Tablas.....	vii
Resumen .....	viii
1. Introducción .....	1
1.1. Motivación .....	2
1.2. Antecedentes.....	3
1.3. Objetivo .....	4
1.4. Objetivos Particulares.....	4
2. Oleaje .....	6
3. Máquinas de Aprendizaje .....	9
3.1. Analogía entre red neuronal biológica y RNA .....	9
3.2. Historia de la RNA .....	12
4. Descripción de redes Implementadas y algoritmos de entrenamiento .....	16
4.1. Red Neuronal Lineal (RNL).....	16
4.2. Entrenamiento RNL .....	18
4.3. Red Neuronal Bayesiana - RNB .....	23
5. Área de Estudio .....	29
6. Descripción de los datos.....	31
6.1. Oleaje.....	31
6.2. Viento .....	34
6.3. Hindcast .....	36
7. Implementación y casos de aplicación de las redes neuronales .....	38
7.1. Implementación de la RNA .....	38
7.2. Determinación del tiempo y número de datos para entrenamiento. ....	39

7.3.	Validación de la RNA.....	45
7.4.	Selección de las horas de generación de oleaje.....	48
7.4.1.	Fuera de la costa.....	48
7.4.2.	Cerca de la costa.....	49
7.5.	Validando la RNA con datos medidos del 2015.....	50
7.6.	Comparando modelos neuronales (RNA) contra modelos numéricos (MT).....	54
7.6.1.	Comparación con el SWAN.....	56
7.7.	Casos de Aplicación RNA.....	57
7.7.1.	Propagación del oleaje.....	57
7.7.1.1.	Dirección de oleaje propagado.....	60
7.7.1.2.	Periodo pico (Tp) del oleaje propagado.....	61
7.7.2.	Aplicación de RNA para generar un Hindcast de oleaje.....	63
7.7.2.1.	Momentos estadísticos.....	65
7.7.2.2.	Primer caso hindcast.....	66
7.7.2.3.	Estadísticos primer caso de aplicación.....	69
7.7.2.4.	Segundo caso Hindcast.....	72
7.7.2.5.	Estadísticos segundo caso.....	74
7.7.2.6.	Tercer caso Hindcast.....	77
7.7.2.7.	Conclusiones del Hindcast.....	79
7.7.3.	Oleaje extremal.....	81
8.	Discusiones.....	83
9.	Conclusiones.....	84
9.1.	Futuras líneas de investigación.....	86
10.	Referencias.....	87

## Índice de Figuras

Figura 1. Representación de una interconexión de dos neuronas biológicas. ....	10
Figura 2. Ejemplificación de la transferencia de información entre dos neuronas biológicas. ....	11
Figura 3. Analogía de la función de una neurona biológica contra una neurona artificial (tipo perceptrón). ....	11
Figura 4. Representación del modelo perceptrón, capa de entrada y capa de salida, conectados directamente. ....	14
Figura 5. Modelo multicapa. ....	15
Figura 6. Modelo con función de transferencia lineal. ....	17
Figura 7. Esquema donde se ejemplifica el problema de sobreajuste. ....	20
Figura 8. Comportamiento del error, donde se muestra el punto "stopped training". ....	20
Figura 9. Representación general de un tipo de red neuronal multicapa (MLP). ....	24
Figura 10 Imagen tomada de Google Earth, localización del Puerto de Sisal dentro del territorio mexicano. ....	29
Figura 11. Localización de los equipos de medición ADCP, con respecto a la línea de costa de Sisal. ....	32
Figura 12. Ejemplo del instrumento de medición de oleaje ADCP. ....	33
Figura 13. Rosas de dirección de oleaje. Datos medidos por el ADCP-10m-Monitor (a) y ADCP-4.5m-Sentinel (b) respectivamente. ....	34
Figura 14. Localización de la torre meteorológica donde se encuentra el anemómetro de donde se adquirieron los datos de viento. ....	34
Figura 15. Rosa dirección de viento para el periodo de medición en Sisal. ....	36
Figura 16. Localización del nodo 12918 del hindcast con respecto a la costa de Sisal y el punto de medición. ....	37
Figura 17. Generación de Hs usando RNA en el punto cerca de la costa, comparando con datos medidos. ....	40
Figura 18. Generación de Hs usando RNA en el punto fuera de la costa, comparando con datos medidos. ....	41
Figura 19. Error rms de los datos del ADCP-10m-Monitor. ....	43

Figura 20. Grafica 3D donde se representan los errores obtenidos con las diferentes combinaciones de entrenamiento, para el punto fuera de la costa.....	44
Figura 21. Error rms de los datos del ADCP-4.5m- Sentinel. ....	44
Figura 22. Grafica 3D donde se representan los errores obtenidos con las diferentes combinaciones de entrenamiento, para el punto cerca de la costa. ....	45
Figura 23. Salidas de los modelos neuronales contra datos medidos a) ADCP-Monitor y b) ADCP-Sentinel.....	46
Figura 24. Correlación de la distribución de los datos a) para el punto cercano de la costa ADCP-Sentinel-4.5m y b) para el punto fuera de la costa ADCP-Monitor-10m. ....	47
Figure 25. Comparación entre los mejores resultados obtenidos por newlind + stepwise (línea negra), RNB + stepwise usando 1N y 4N en la capa oculta (línea roja) y Bootstrap + RNB + stepwise, con 1N and 4 N (líneas azul). Punto fuera de la costa ADCP-Monitor. .	49
Figura 26. Comparación entre los mejores resultados obtenidos por newlind + stepwise (línea negra), RNB + stepwise usando 1N y 4N en la capa oculta (línea roja) y Bootstrap + RNB + stepwise, con 1N and 4 N (líneas azul). Punto cercano a la costa (ADCP-Sentinel). .....	50
Figura 27. Comparación entre el oleaje medido ADCP y la salida del modelo neuronal para el año 2015.....	52
Figura 28. Comparación de la distribución de los dos series de datos de Hs del ADCP y de la salida del modelo neuronal. ....	53
Figura 29. Primera grafica corresponde a la Hs registradas por los ADCP (línea continua-Monitor y línea segmentada-Sentinel). Segunda grafica corresponde a velocidades de viento registradas por el anemómetro para el mismo periodo de tiempo.....	54
Figura 30. Localización del ADCP entre Yucalpeten y Chuburná, así como el nodo 12972 de la malla del MIKE21. Figura tomada del trabajo de Medellín (2015). ....	55
Figura 31. Comparación de las Hs medidas (ADCP) y la salida del MIKE21. ....	55
Figura 32. Comparación de la RNA y un modelo numérico tomado del trabajo de Lira, 2015. .....	56
Figura 33.Comparación entre los mejores resultados obtenidos por newlind + stepwise (línea negra), RNB + stepwise + Bootstrap usando 2N, 4N y 8N en la capa oculta (líneas rojas y azul). Propagación del oleaje al punto cercano a la costa (ADCP-Sentinel). ....	60
Figura 34. Comparación de las rosas de dirección de oleaje de los datos medidos (a) y la dirección pronosticada por el modelo neuronal (b).....	61

Figura 35. Comparación del pronóstico de periodo pico (Tp) generado con el modelo neuronal.....	62
Figura 36. Correlación del Tp entre los datos medidos y la salida del RNA. ....	62
Figura 37. Localización del Nodo 12918 del hindcast y el ADCP fuera de la costa de Sisal, Yucatán. ....	64
Figura 38. Resultado obtenido de la etapa de evaluación usando la RNB con 10N en la capa intermedia.....	68
Figura 39. Comparación de la distribución de los dos series de datos de Hs, entrenando con 2 años la RNB.....	69
Figura 40. Comparación de las medias anuales del MIKE21 vs. RNA, entrenando con 2 años.....	71
Figura 41. Comparación de la desviación estándar anuales del MIKE21 vs. RNA, entrenando con 2 años. ....	71
Figura 42. Comparación de la Hs anuales del MIKE21 vs. RNA, entrenando con 2 años.	72
Figura 43. Comparación de los hindcast con MIKE21 y con RNA entrenada solo con 1 año. ....	73
Figura 44. Comparación de la distribución de los dos series de datos de Hs, entrenando con 1 año la RNB. ....	74
Figura 45. Comparación de las medias anuales del MIKE21 vs. RNA, entrenando con 1 año. ....	76
Figura 46. Comparación de las desviaciones estándar anuales del MIKE21 vs. RNA, entrenando con 1 año. ....	77
Figura 47. Comparación de las Hs anuales del MIKE21 vs. RNA, entrenando con 1 año. ....	77
Figura 48. Comparación de la distribución de los dos series de datos de Hs, entrenando con 1 (1992-1993) año la RNB. ....	79
Figura 49. Grafica del periodo de retoro generado a 100 años con los datos del MIKE21. ....	82
Figura 50. Grafica del periodo de retoro generado a 100 años con los datos de la RNA. ....	82

## Índice de Tablas

Tabla 1. Tipos de anemómetros localizados en la torre meteorológica de la UAS. ....	35
Tabla 2. Error medio cuadrático con diferentes porcentajes de entrenamiento y tiempo de generación.....	42
Tabla 3. Representación de la RNA utilizada para validar la metodología. ....	51
Tabla 4. Tabla con diferentes errores obtenidos con diferente número de neuronas. ....	51
Tabla 5. Tabla de comparación RNA y modelo numérico. ....	57
Tabla 6. Arreglo de los datos para entrenar y evaluar las RNA, en la propagación de oleaje (del 26 de Marzo al 22 de Mayo, del 2014). ....	59
Tabla 7. Errores obtenidos con diferente número de neuronas en la capa oculta, RNB entrenada con 2 años. ....	67
Tabla 8. Tabla de comparación de estadísticos anuales, entrenando con dos años de datos. ....	69
Tabla 9. Tabla de errores obtenidos con diferente número de neuronas en la capa oculta para la RNB entrenada con un año 2006-2007. ....	73
Tabla 10. Tabla de comparación de estadísticos anuales, entrenando con un año de datos ....	75
Tabla 11. Tabla de errores obtenidos con diferente número de neuronas en la capa oculta para la RNB entrenada con un año 1992-1993. ....	78
Tabla 12. Comparación de los errores obtenidos, entrenando con diferentes años. ....	80
Tabla 13. Alturas de olas calculadas para cada periodo de retorno correspondiente, usando las dos series de datos. ....	81

## RESUMEN

Las redes neuronales artificiales (RNA), de manera general, son máquinas de aprendizaje que imitan el funcionamiento que tiene una neurona biológica. Las RNA permiten analizar gran cantidad de datos con la finalidad de obtener información futura basada en la experiencia, mediante un proceso de aprendizaje y entrenamiento. En este trabajo, se presenta la aplicación de las RNA para generar oleaje, a partir de datos de viento en la costa de Sisal, Yucatán. Para cumplir con el objetivo, nos hemos enfocado en aplicar modelos neuronales con la ayuda de métodos probabilísticos para obtener los parámetros de oleaje (altura de ola significativa, periodo pico y dirección media), debido a que nos proporcionan la información necesaria para planear y diseñar actividades marítimo-terrestre.

En este estudio se han empleado dos tipos de redes neuronales supervisadas, la primera corresponde a una red neuronal lineal – RNL, la segunda es una red neuronal bayesiana – RNB en la cual se variaron el número de neuronas en la capa intermedia. A los modelos neuronales dependiendo el caso de estudio se le agregaron métodos probabilísticos para optimizar el entrenamiento. Los resultados obtenidos con los modelos neuronales fueron validados con datos reales medidos en el área de estudio; lo cual nos permitió obtener 3 tipos de redes neuronales listas para usarse en distintos casos de implementación y aplicación: la primera corresponde a una RNB para generar de manera puntual oleaje a partir de datos de viento, la segunda corresponde a una RNL utilizada para propagar el oleaje de un punto en aguas intermedias a un punto en aguas someras y la tercera es una RNB con la cual se pueden generar bases de datos históricos a partir de una base de datos histórica de viento. Se concluye que el emplear las RNA para obtener oleaje a partir de datos de viento medidos es una herramienta alternativa, eficiente y confiable.

*Palabras clave: red neuronal artificial, entrenamiento, oleaje, métodos probabilísticos.*

## 1. INTRODUCCIÓN

Las redes neuronales artificiales (RNA) son modelos computacionales basados en las matemáticas y algoritmos de aprendizaje; son una imitación del funcionamiento del sistema neuronal biológico y en específico del cerebro humano. El cerebro es un órgano muy poderoso que contiene aproximadamente  $10^{11}$  interconexiones de neuronas lo cual lo hace ser más robusto y complejo de lo que podría llegar a ser una imitación hecha por una computadora y una RNA. Sin embargo las RNA contienen ciertas características que permiten resolver problemas planteados, ya que son sistemas dinámicos auto-adaptativos, es decir que tiene la capacidad de autoajuste, debido a que son capaces de cambiar constantemente para adaptarse a las nuevas condiciones.

Por lo general una de las mayores prioridades en los modelos computacionales es realizar procesos con datos de forma rápida y eficiente. Las redes neuronales se adaptan bien a esta necesidad, debido a la implementación paralela. La mayoría de las redes puedan operar en un entorno de tiempo real, la necesidad de cambio en los distintos parámetros de entrenamiento es mínima (Matich, 2001), más adelante se explicara con mayor detalle cada una de estos parámetros y su funcionamiento. A partir de la década de los 80's las RNA se han utilizado para resolver problemas relacionados con las ciencias ambientales, meteorología y oceanografía por mencionar algunas aplicaciones. En este trabajo nos enfocaremos en aplicar modelos neuronales con la ayuda de métodos probabilísticos en la parte oceanográfica, específicamente en el tema de oleaje, por lo que a partir de datos de viento y oleaje se entrenarán dos diferentes tipos de redes neuronales para posteriormente con datos de viento generar alturas de ola, para un sitio en específico.

La información de altura de ola en un sitio, es un dato que es considerado fundamental para el diseño de obras portuarias y/o estructuras costeras, así como en operaciones y/o actividades portuarias. De igual forma, durante la operación de las obras de construcción es necesario considerar un pronóstico de altura de ola de un par de horas o días, antes y después de las actividades (Deo & Naidu, 1999).

En la actualidad, el desarrollo de modelos de tercera generación (MT- Model tree) como: WAM, SWAN y WAVEWATCH (Komen 1996), en combinación con modelos atmosféricos como el Nort American Mesoscale (Janjić, 2003), el Weather Research and Forecasting

Model (Skamarock, et al. 2005) y el Resolution Atmosphere Model (Zhao, et al. 2009), por mencionar algunos, ha permitido la implementación de estos modelos numéricos para la predicción y/o generación de oleaje a nivel global y regional (Appendini, 2012). Estos modelos requieren un esfuerzo computacional muy alto, así como una gran dependencia de datos de alta resolución temporal y espacial (Ej.: batimetría, vientos). Sin embargo, para puntos específicos como es el área de estudio de este trabajo, se puede optar por usar sistemas computacionales inteligentes más simples de generación de oleaje, como las Redes Neuronales Artificiales – RNA (Jain, Deo, Latha, & Rajendran, 2011), las cuales se han comprobado que funcionan satisfactoriamente.

### *1.1. Motivación*

Considerando los altos costos de instalación y mantenimiento para los instrumentos de medición, así como, las dificultades que se presentan durante largos periodos de tiempo en las campañas de medición para mantener redes de boyas de oleaje operando (ADCP, sensores de presión) y obtener datos; se han propuesto metodologías alternas para de esta manera conseguir información de otras fuentes que se encuentran más accesible como, estaciones meteorológicas costeras y el uso en conjunto de metodologías con herramientas computacionales (Jain et al., 2011).

Debido a lo anterior, este trabajo se centra en proporcionar una herramienta y metodología alternativa usando métodos de las máquinas de aprendizaje (conocidas comúnmente en Ingles como Machine Learning,), tomando como base las redes neuronales y modelos probabilísticos, para obtener parámetros de oleaje (altura de ola significativa, periodo pico y dirección media), para sitios puntuales o un lugar en específico; de manera más accesible en cuanto a recursos de cómputo y personal especializado, ya que es una herramienta de libre uso (open source) de muy bajo costo en comparación con una licencia de un modelo numérico tercera generación MT y relativamente sencilla de operar y manipular.

Cabe mencionar que esta herramienta la han implementado en muchos trabajos (ej. Deo & Naidu, 1999; Tsai, Lin, & Shen, 2002; Makarynskyy et al., 2004; entre otros), donde realizaron predicciones de elevación del nivel del mar (mareas), predicción de marea de tormenta y oleaje. Sin embargo en México existen muy pocos trabajos en el tema de la oceanografía costera o ingeniería de costas, razón por la cual se plantea en este trabajo obtener parámetros de oleaje a partir del aprendizaje de un modelo neuronal entrenado

con datos de viento y oleaje. Una vez que la red neuronal ha sido entrenada, esta será capaz de obtener a partir de datos de viento parámetros del oleaje.

### *1.2. Antecedentes*

- (Hsieh & Tang, 1998), describen las técnicas de superar las dificultades que se presentan comúnmente aplicando modelos neuronales para predicción y análisis de datos meteorológicos y oceanográficos.
- El trabajo de Deo y Naidu (1999), realiza predicción de altura de significativo ( $H_s$ ), para un sitio en la costa este de India dentro de la Bahía de Bengala. Las redes usadas fueron entrenadas a partir de datos de oleaje medido, los cuales corresponden a 16 meses de Mayo 1983 a Agosto de 1984, usan el 80% de los datos para entrenar la red y el 20% restante para evaluar los resultados obtenidos por el modelo neuronal. Las predicciones que se hicieron corresponden a 3, 12 y 24 horas.
- Tsai, Lin, & Shen(2002), aplican un modelo neuronal de propagación hacia atrás (back-propagation) para predecir altura de ola significativa ( $H_s$ ) y periodo pico ( $T_p$ ) para un punto en específico, haciendo una interconexión de pesos de los datos de oleaje entre redes neuronales correspondientes a otras estaciones de medición. A partir de las salidas de una red neuronal, se alimenta la siguiente red neuronal.
- Jain et al( 2011), realizan una comparación de predicciones de altura de ola significativa y periodo pico, implementando redes neuronales RNA y un modelo tercera generación MT; a partir de series de datos históricos de viento en la costa este de India, dentro de la Bahía de Bengala. Las predicciones de los parámetros de oleaje ( $H_s$  y  $T_p$ ) son de 24 horas en intervalos de tiempo 0, 6, 12 y 24 horas. Los resultados tanto del modelo neuronal como del modelo numérico fueron comparados entre si y resultaron ser muy similares.
- Lee (2006) realizó una predicción de marea de tormenta usando redes neuronales. De acuerdo a sus resultados, obtuvo una alta correlación (0.9864), lo cual implica una buena concordancia entre los datos observados y los predichos. Este estudio concluye en que la herramienta puede ser eficiente para predecir la marea de tormenta, usando como datos de entrada cuatro factores: velocidad de viento, dirección del viento, presión y un análisis armónico del nivel de marea.

- Mandal y Prabakaran(2006), describieron una red neuronal artificial usando un algoritmo actual (Rprop), el cual es un sistema de aprendizaje eficiente que realiza una adaptación directa en la etapa de ponderación, basada en el gradiente de información local. Finalmente realizaron una comparación con el trabajo de Deo y Naidu, (1999), Subba R.,y Mandal S,(2001), presentando los valores más altos de coeficientes de correlación, concluyendo que se puede realizar una predicción más eficiente con este tipo de algoritmo.

### *1.3. Objetivo*

La finalidad de este trabajo es implementar una red neuronal artificial como una herramienta alternativa, para generar oleaje a partir de datos de viento, en la costa de Sisal, Yucatán. Para cumplir con este objetivo principal es necesario definir los siguientes objetivos particulares.

### *1.4. Objetivos Particulares*

- Determinar la mejor arquitectura de red neuronal, evaluando diferentes modelos y algoritmos de entrenamiento
- Validar la red neuronal con datos medidos de viento y oleaje en la zona de Sisal, Yucatán.
- Propagar el oleaje que incide a la costa de Sisal, utilizando para esto datos de dos equipos oceanográficos colocados a 10 y 4 metros de profundidad.
- Sacar parámetros de clima marítimo con los datos de oleaje obtenidos con el modelo neuronal y comparar con los parámetros de clima marítimo generados con el modelo numérico MIKE21.
- Evaluar los resultados del modelo neuronal contra los de modelos numéricos tercera generación MT (MIKE21 y SWAN).

El presente trabajo estará estructurado de la siguiente manera e incluirá algunos capítulos introductorios e importantes para el entendimiento del trabajo realizado:

- ✓ Capítulo 1.- Se da una introducción al trabajo que se realiza en esta tesis, planteando la motivación, antecedentes más importantes del tema de tesis y se plantean los alcances de trabajo.
- ✓ Capítulo 2.- Introducción del oleaje, donde se describen los tipos de oleaje que existen, así como los parámetros estadísticos que ayudan a describir y entender mejor el comportamiento del oleaje en el mar.
- ✓ Capítulo 3.- Definición de las máquinas de aprendizaje y de las Redes Neuronales Artificiales (RNA). Así como una breve historia de las mismas.
- ✓ Capítulo 4.- Se presentan las principales diferencias entre las redes neuronales utilizadas: red neuronal lineal (RNL), la cual tiene la característica de presentar dos capas simplemente con una función de transferencia lineal y, la red neuronal bayesiana (RNB), que es una aproximación del modelo neuronal multicapa usando el algoritmo de optimización Levenberg-Marquard. Así como la descripción de la etapa de entrenamiento y la etapa de evaluación. También se abordan algunos problemas típicos que se presentan, al implementar las RNA en la oceanografía.
- ✓ Capítulo 5.- Descripción del área de estudio y su localización.
- ✓ Capítulo 6.- Descripción de los datos para entrenar y validar los modelos neuronales usados en el trabajo; así como la localización de los puntos donde fueron registrados los datos y una breve descripción de cómo funcionan los instrumentos de medición.
- ✓ Capítulo 7.- En este capítulo se presentan algunos casos de implementación y aplicación de las RNA para validar la metodología usada como: propagación oleaje, generación de una serie de datos históricos de oleaje, cálculo de momentos estadísticos con base a las Hs generadas con los modelos neuronales, así como resultados de un análisis de oleaje extremal y finalmente algunas comparaciones de los resultados obtenidos con un modelo tercera generación (MT).
- ✓ Capítulo 8.- Se presentan las discusiones de los resultados en general.
- ✓ Capítulo 9.- Conclusiones, así como las futuras líneas de investigación del presente trabajo.

## 2. OLEAJE

A través del tiempo, el conocimiento de las características del oleaje ha representado una necesidad, debido a que nos proporciona la información necesaria para planear actividades marítimo – terrestre, diseñar estructuras costeras así como planear y diseñar actividades de recreación (pesquerías; construcciones costeras y de mar abierto; deportes acuáticos y recreativos; transporte y/o navegación; entre otros). En la actualidad la importancia de estudiar las olas no puede ser subestimada, sino por el contrario, debido al aumento e importancia de las actividades humanas en la zona costera y en mar abierto. Al considerar el gran crecimiento demográfico en las zonas costeras del mundo, el consecuente desarrollo de infraestructura, así como la explotación de recursos marinos (biomasa, petróleo, energía oceánica, por citar algunos) y la intensificación del transporte marítimo, el conocimiento de los procesos oceánicos y costeros se hace cada día más relevante (Appendini, 2012).

Durante el siglo pasado el interés en la predicción del oleaje creció durante la segunda guerra mundial, una de las razones fue; obtener el conocimiento del estado de mar durante las operaciones de aterrizaje. La primera predicción de operación se basó en el trabajo de Sverdrup y Munk (1947), quienes introdujeron una descripción paramétrica del estado de mar y utilizaron las leyes empíricas de vientos marinos y swell (Komen 1996).

En el mar, se genera un fenómeno llamado oleaje, que se forma debido a un intercambio de momentum entre la atmosfera y el océano a través del viento. El viento es la fuerza generadora del oleaje, cuando este se encuentra actuando sobre la superficie del mar, forma ciertas perturbaciones; además, si es continuo e intenso, genera ondas (olas). La generación de las olas depende de la velocidad del viento y su duración (duración de la tormenta), durante este evento se forman olas de diversos periodos y tamaños (altura) a este tipo de oleaje se le conoce como **sea** (*oleaje de viento o local*). El área de la superficie del mar que está bajo la influencia del viento y donde se genera el oleaje se le conoce como “*Fetch*” Komar, 1998. Cuando el oleaje abandona el área de generación se dispersa en todas direcciones, debido a que la velocidad está en función directa del periodo, las olas más largas viajan más rápido que las más cortas; de esta manera el oleaje al propagarse hacia la costa lo hace en grupos o paquetes de olas con mismas características de periodo y altura, a este tipo de oleaje se le conoce como **swell**. Así el oleaje que se aleja del área

de generación es más uniforme, forma frentes de cresta muy grandes y las diferencias entre los periodos y longitudes de onda son mínimas, surge una periodicidad, así como direcciones predominantes (Silva, 2005).

El oleaje a medida que se aproxima a la costa, sufre procesos de transformación y atenuación. Estos procesos se pueden observar cuando el oleaje pasa de aguas profundas (profundidad donde la ola el fondo marino no siente el fondo marino) a aguas transicionales (profundidad en la cual el efecto del oleaje influencia un cierto movimiento en el suelo marino). Debido a la disminución de la profundidad conforme nos acercamos a la línea de costa la ola se ve afectada reduciendo su longitud de onda y celeridad, así como la modificación en la dirección de las crestas de ola (refracción), la altura de la ola aumenta (asomeramiento) ya que siente el efecto del fondo, esa fricción de la ola con el fondo da como resultado una disipación de la energía y finalmente una rotura de ola, debido a que llega a su límite de crecimiento.

Como se mencionó, el oleaje tipo **sea**, presenta oleaje de diferentes direcciones y está compuesto de olas con alturas y periodos diferentes. Por tal motivo, es necesario tratar los parámetros de oleaje en términos estadísticos lo cual complica el análisis, pero describe la superficie del mar de manera más aproximada a la realidad. Existen parámetros estadísticos de las olas que son útiles para caracterizar y saber las condiciones del oleaje, los más comunes son: altura de ola, periodo y la dirección del oleaje, los cuales son valores representativos de las series temporales registradas, con los instrumentos de medición.

La altura de ola significativa ( $H_s$  o  $H_{1/3}$ ), fue presentado por primera vez por Sverdrup and Munk (1947), y surgió de la necesidad de establecer un parámetro estadístico que relacionara las alturas de ola obtenidas en el registro instrumental del oleaje y las establecidas a través de observaciones visuales de un estado de mar. A partir de un registro de oleaje (serie de tiempo del desplazamiento vertical de la superficie libre), la altura de ola significativa se define, como la media aritmética del tercio de olas de mayor altura del registro:

$$H_s = \frac{3}{N} \sum_{i=1}^{\frac{N}{3}} H_i \quad \text{Ec. 1}$$

Donde  $H_i$  es la serie de alturas de ola individuales del registro, ordenada de mayor a menor y N es el número total de olas individuales del registro.

Un segundo parámetro es el periodo pico ( $T_p$ ), está asociado al periodo de los distintos oleajes con mayor contenido energético, este lo podemos observar en el espectro de energía como un pico máximo en las gráficas de las frecuencias. El espectro o espectro de energías es una forma de representar como está distribuida la energía del oleaje en función de las frecuencias que integran una señal de oleaje en particular.

Otro parámetro que es de suma importancia es la dirección de oleaje ya que este nos dice el ángulo con el cual incide el oleaje a la costa o algún sitio de interés. Una representación práctica del oleaje asociando la intensidad y frecuencia de aparición de la altura de ola respecto a su dirección son las rosas de oleaje. En estos gráficos podremos discernir la dirección predominante del oleaje que presenta mayor altura y la dirección del oleaje que posiblemente no presente la mayor altura pero si el mayor porcentaje de ocurrencia.

De esta manera, los parámetros más comunes para describir el estado de mar son: la altura de ola, el periodo y la dirección del oleaje, que son valores representativos de las series de tiempo registradas con los instrumentos de medición. El conocer las características principales de los parámetros de oleaje mencionados, es de gran importancia para planear construcciones de protección costeras y estructuras marítimas mar adentro (ej., plataformas petroleras). De igual manera, estas características son relevantes para el estudio, desarrollo e implementación de energías renovables, dado que son una ayuda para estimar la capacidad energética que podemos obtener de este recurso natural (Makarynsky, et al., 2004).

Es relevante conocer, a partir de los registros históricos de datos cuál es la distribución del oleaje en condiciones medias (año medido) o bien determinar las características de oleaje en condiciones extremas. En el primer caso hablamos de una caracterización del oleaje a mediano plazo, régimen del oleaje o clima medio, mientras que en el segundo caso hablamos de régimen de temporales o clima extremal; en ambos casos el estudio se realiza de una forma estadística, más adelante en el capítulo 7 se tratara a detalle el estudio y las metodologías usadas en este trabajo para obtener la caracterización del oleaje.

### 3. MÁQUINAS DE APRENDIZAJE

Las máquinas de aprendizaje, son una rama importante de la inteligencia artificial entendiendo como tal a la ciencia e ingenio de hacer maquinas inteligentes, que ayuden a resolver problemas mediante procesos inductivos del conocimiento utilizando como paradigma la inteligencia humana y métodos estadísticos. Como parte de las máquinas de aprendizaje tenemos a las redes neuronales artificiales (RNA), que como su nombre lo dice es un intento de imitación del funcionamiento de como procesan la información las redes neuronales del cerebro humano para resolver problemas complejos reales. Además de las redes Neuronales, otro tipo de máquinas de aprendizaje son:

- Árboles de decisiones.- Esta es una técnica que analiza **decisiones secuenciales** basadas en el uso de resultados y probabilidades asociadas, arroja valores lógicos y los resultados generalmente son binarios.
- Algoritmos genéticos.- Estos son métodos adaptativos que se usan para resolver problemas de **búsqueda y optimización**, basados en el procesos genético de los organismos vivos. Este tipo de algoritmos son usados por ejemplo para resolver problemas de genética en las razas humanas.
- Máquinas de vectores de soporte.- Estas a diferencia de las RNA consiste en un conjunto de algoritmos de aprendizaje supervisados que tiene como objetivo resolver **problemas de clasificación y regresión**.

Los enfoques presentados anteriormente son algunos tipos diferentes de máquinas de aprendizaje, sin embargo en este trabajo, nos enfocaremos en uso de las RNA, la cuales han demostrado un buen desempeño en los problemas planteados.

#### *3.1. Analogía entre red neuronal biológica y RNA*

Si bien el cerebro es un órgano asombroso que presenta  $10^{11}$  interconexiones de células neuronales llamadas “neuronas”. Estas neuronas poseen tres principales componentes, las dendritas, el cuerpo de la célula o soma y el axón. Las dendritas, son arboles receptores de la red, son fibras nerviosas que cargan de señales eléctricas el cuerpo de la célula. El cuerpo de la célula, realiza la suma de esas señales de entrada. El axón, es una fibra larga

que lleva la señal desde el cuerpo de la célula hacia otras neuronas. El punto de contacto entre un axón de una célula y una dendrita de otra célula es llamado sinapsis, la longitud de la sinapsis es determinada por la complejidad del proceso químico que estabiliza la función de la red neuronal (Figura 1).

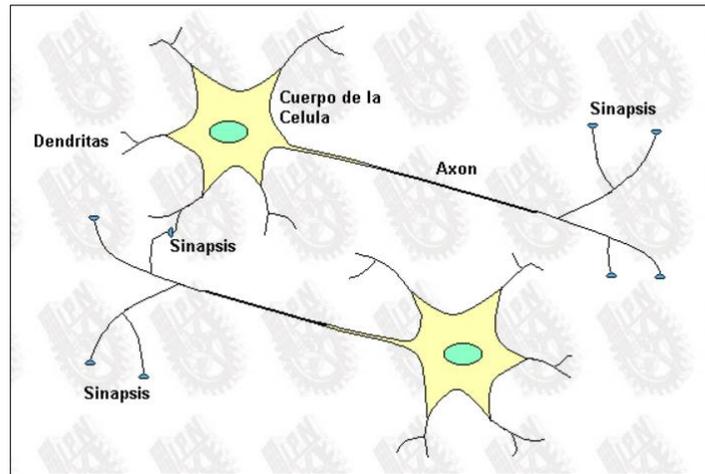


Figura 1. Representación de una interconexión de dos neuronas biológicas.

Se tiene identificada la analogía entre las RNA y una red neuronal biológica, esto lo podemos entender cuando se genera la **sinapsis** por la conexión entre el axón y las dendritas (red neuronal biológica) y por otro lado la conexión que hay entre los datos de entrada y los pesos (RNA). La sinapsis es un proceso importante ya que aquí es donde se transmite la información a otra neurona con la que se encuentra interconectada, información que servirá para la etapa de aprendizaje. En la Figura 2, se ilustra la manera en que se transfiere una señal entre dos neuronas a través de la sinapsis. El axón recibe señales o estímulos del cuerpo de la célula, estos estímulos son acarreados hacia la sinapsis, que realizará conectando el axón de una célula emisora (neurona transmisora) con las dendritas de una neurona vecina (neurona receptora). Al llegar la señal a la membrana presináptica los neurotransmisores se liberan de las vesículas en cantidades proporcionales a la intensidad de la señal entrante. Posteriormente los neurotransmisores se depositan en el espacio sináptico y esos son los encargados de excitar la membrana postsináptica, y finalmente transferirse a las dendritas de la neurona vecina (Basheer y Hajmeer, 2000).

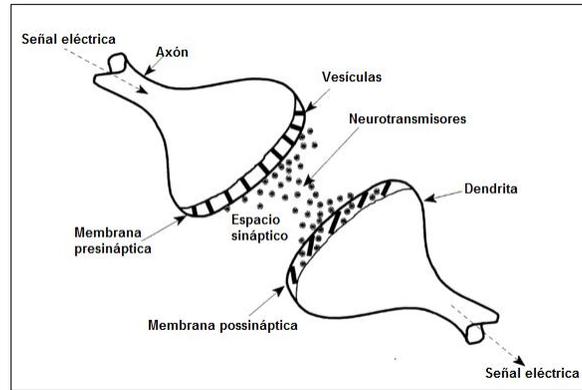


Figura 2. Ejemplificación de la transferencia de información entre dos neuronas biológicas.

Otra similitud básica es la actividad que se genera en el soma la cual está representada por el procesamiento que se genera en la función de transferencia. En la Figura 3, podemos observar que los valores de  $x_1, \dots, x_n$ , representan las señales o valores de entrada los cuales generan una sinapsis con la matriz  $\mathbf{W}$  de pesos, esa información es enviada al núcleo de la célula y en su contraparte a la función de transferencia, para procesar la información y generar una salida o estímulo, el cual puede ser enviado a otra neurona o al ambiente, generando una red de “ $\mathbf{N}$ ” neuronas, dependiendo el arreglo deseado.

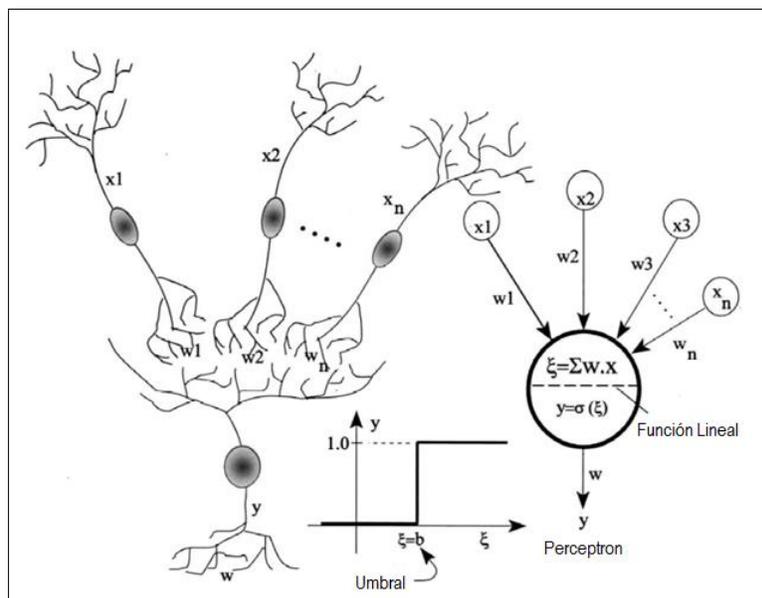


Figura 3. Analogía de la función de una neurona biológica contra una neurona artificial (tipo perceptrón).

La realidad es que las RNA no alcanzan la complejidad del cerebro, sin embargo hay aspectos similares entre redes biológicas y artificiales: una es que presentan interconexiones entre si y que esas interconexiones determinan la función de la red; resuelven problemas de no-linealidad lo cual permite un mejor ajuste a los datos; el funcionamiento y procesado de información es en paralelo lo cual implica una eficiencia computacional tolerante a fallos; tiene la capacidad de aprender de la experiencia y son adaptativas , es decir, contiene un conjunto de datos numéricos (pesos) que son ajustados y actualizados en respuesta a cambios del entorno (Basheer & Hajmeer, 2000). El principal objetivo de las RNA es desarrollar algoritmos matemáticos para aprender a procesar la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano “la neurona”. Las RNA son modelos artificiales simplificados del funcionamiento del cerebro humano, son el ejemplo perfecto de sistemas que son capaces de adquirir conocimiento a través de la experiencia.

### *3.2. Historia de la RNA*

Las redes neuronales fueron las precursoras en el tema de las máquinas de aprendizaje a finales de los 80's, mientras que los métodos de las redes neuronales combinadas con métodos probabilísticos (“kernel”), llegaron como segunda oleada a mitad de los 90's, misma década en la que las máquinas de aprendizaje comenzaron a implementarse en las ciencias ambientales, en el proceso de datos satelitales, modelos generales de circulación (de tipo climático océano-atmosfera), así como en pos-procesamiento de los datos de los modelos físicos, en la generación de pronósticos del clima, calidad del aire, predicciones hidrológicas y oceanográficas.

Los investigadores en medicina fueron los primeros que se vieron interesados en el funcionamiento de las redes neuronales artificiales, sin embargo investigadores y científicos de otras disciplinas (la ingeniería, la economía y las ciencias ambientales, por mencionar algunos) también se vieron interesados en desarrollar redes neuronales artificiales , como técnicas computacionales e inteligencia artificial.

El primer modelo moderno de redes neuronales y el cual ha servido de inspiración para el desarrollo de otros modelos corresponde al modelo de McCulloch y Pitts (1943), a partir de un enfoque neurobiológico describen que las neuronas reciben estímulos (señales) de otras

neuronas vecinas y si el total de esos estímulos excede un umbral, las neuronas se convierten en una neurona activa la cual genera una señal de salida. Este modelo lo que realiza es una sumatoria ponderada de los valores de entrada y los pesos de otras unidades y da salida con valores lógicos ya sea 1 o 0, dependiendo si la suma excede el umbral. El valor 1 correspondiendo a una neurona activa y el valor 0 corresponde a una neurona que permanece en reposo. Para una neurona, si  $X_i$  corresponde a las señales de entrada de la  $i$ -ésima neurona vecina, el cual es multiplicado por parámetro de pesos  $W_i$ , la salida de la neurona está dada por:

$$Y = H\left(\sum_i W_i X_i + b\right) \quad \text{Ec. 2}$$

Donde  $b$  es llamado bias o sesgo,  $H$  es el paso de una función:

$$H(z) = \begin{cases} 1 & \text{si } Z \geq 0 \\ 0 & \text{si } Z \leq 0 \end{cases}$$

El modelo de McCulloch y Pitts es muy similar a las puertas lógicas, lo cual carece de un algoritmo para poder encontrar los parámetros de pesos y sesgos apropiados que ayude a resolver problemas que no dependan solamente de una respuesta lógica.

Para resolver el problema mencionado anteriormente uno de los mejores avances en el tema de la RNA es el modelo de Rosenblatt (1958, 1962), así como un trabajo similar realizado por Widrow y Hoff (1960). El modelo **perceptrón** consiste en una capa de entrada conectada directamente con una capa de salida (Figura 4); lo novedoso de este modelo es la introducción de una función de activación, el cual encuentra el valor de los pesos y sesgos para resolver un problema en particular. Sin embargo la función de activación presenta limitaciones, al resolver problemas en los que no pueden ser linealmente separables.

En este caso la salida de una neurona está dada por:

$$y_j = f(\sum_i W_{ji} X_i + b_j) \quad \text{Ec. 3}$$

Donde  $X_i$  nuevamente corresponde a los valores de entrada,  $f$  es una especie de función de transferencia conocida también como función de activación,  $W_{ji}$  corresponde a los pesos conectados con la  $i$ -ésima neurona para salida de la  $j$ -ésima neurona y  $b_j$  es un sesgo o bias de la  $j$ -ésima neurona de salida.

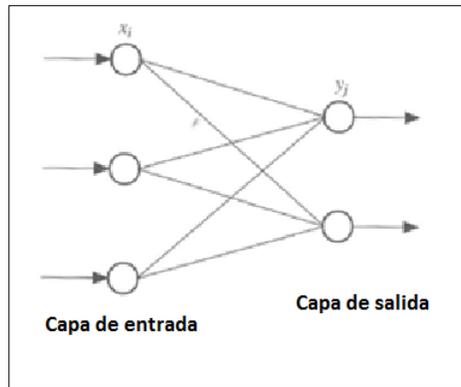


Figura 4. Representación del modelo perceptrón, capa de entrada y capa de salida, conectados directamente.

Afortunadamente Rumelhart et al. (1986) y Hertz et al. (1991) redescubren el algoritmo de propagación hacia atrás (backpropagation) el cual resuelve exitosamente los problemas de encontrar los pesos en un modelo multicapa perceptrón (**MLP**), el cual incluye una capa oculta entre la capa de entrada y la capa de salida (Figura 5). En general una optimización no lineal es difícil y la convergencia puede ser muy tardada o numéricamente poco precisa, razón por la cual uno necesita hallar un buen algoritmo de optimización.

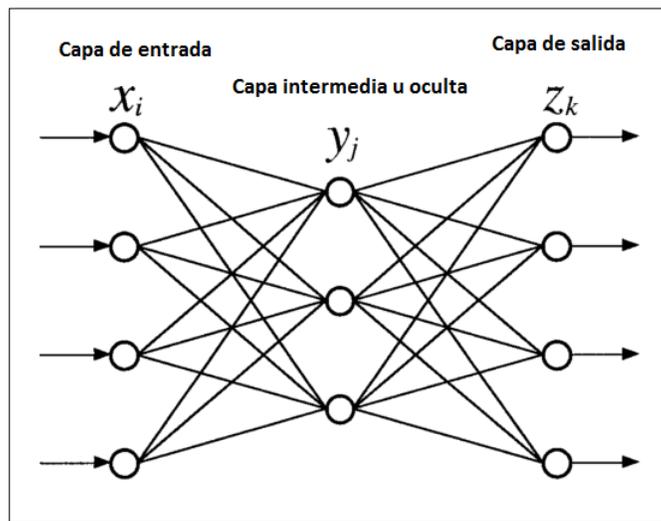


Figura 5. Modelo multicapa.

## 4. DESCRIPCION DE REDES IMPLEMENTADAS Y ALGORITMOS DE ENTRENAMIENTO

En el tema de redes neuronales, a la minimización del error le han aplicado el concepto de **optimización** y sirve para mejorar los resultados del modelo que se está usando. En este trabajo usaremos dos tipos de optimización.

- El primero utiliza un algoritmo creado por Widrow y Hoff (1960), conocido como el de mínimos cuadrados LMS (por sus siglas en ingles *Least Mean Square*), el cual realiza una aproximación para hallar el error mínimo.
- El segundo tipo de optimización que se utilizará es el Bayesiano, llamado así debido a que toma como base el Teorema de Bayes, Mackay (1992 b, a) nombra a este tipo de red como; red neuronal bayesiana (RNB) la cual utiliza el algoritmo Levenberg-Maquard, del que se obtiene una medida del número de pesos que están siendo usados de manera efectiva en la red neuronal para reducir el valor de la función de error.

Las RNA pueden consistir de una o más interconexiones de capas y una capa puede estar compuesta de varias neuronas, que a su vez están conectadas con otras capas. Los patrones de interconexión entre capas y neuronas se le conocen como la arquitectura de la red (Makarynskyy, Makarynska, Kuhn, & Featherstone, 2004). Por lo tanto, en este trabajo, se determinara el número de neuronas y/o capas de la red neuronal (arquitectura). Una notación comúnmente utilizada y práctica es la presentada en el diagrama de la Figura 5, donde tenemos una red 4-3-4, lo cual significa que en la primer capa de los datos de entrada tiene “4” entradas, el siguiente número corresponde al número de neuronas que contiene la capa oculta “3” y finalmente el número “4” corresponde al número de salidas que hay en esta última capa (Hsieh, 2009).

### 4.1. Red Neuronal Lineal (RNL)

El primer tipo de red neuronal que se estará implementando en este trabajo corresponde a una red neuronal lineal – RNL, utiliza una función de transferencia lineal (Figura 6), la cual contiene el algoritmo LMS (*Least Mean Square*), este es un algoritmo de filtrado lineal adaptativo que consiste en el cómputo de la salida de un filtro lineal en respuesta a una señal de entrada y por otro lado la generación de una estimación del error mediante la

comparación de esta salida con la señal deseada. Este es un proceso adaptativo, que involucra el ajuste automático de los parámetros de acuerdo al error calculado entre la salida del modelo neuronal y el valor esperado.

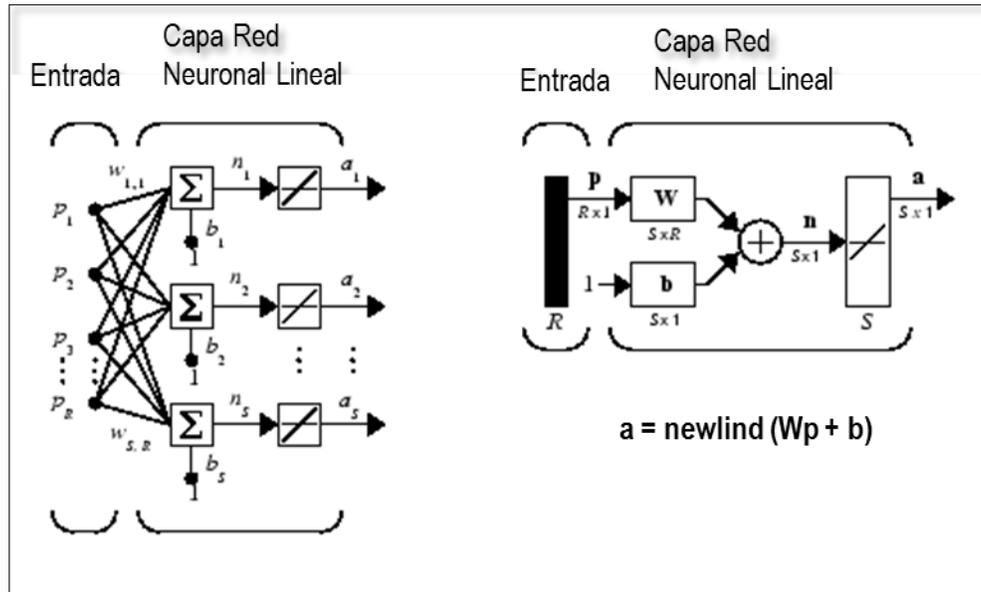


Figura 6. Modelo con función de transferencia lineal.

Los datos de entrada están representados por el vector “**P**” de **R** número de elementos, este es multiplicado por una matriz de pesos “**w**” y el producto es uno de los términos de la sumatoria, el otro término de entrada es una desviación inicial o sesgo “**b**” (bias), el resultado de esta sumatoria ( $\mathbf{n} = \mathbf{wp} + \mathbf{b}$ ), es enviado a la función de transferencia lineal, generando valores de salida (**a**), **S** se refiere al número de neuronas en la capa. Los valores de salida estarán dentro del rango  $(-1/a \text{ y } 1/a)$ , por encima o de bajo de esta zona se fija la salida en 1 o -1, respectivamente; cuando  $a = 1$  la salida es igual a la entrada. Posteriormente se calcula el error restando el valor “**a**”, generando por la RNA menos el valor esperado “**T**” (ecuación  $\mathbf{Error} = \mathbf{a} - \mathbf{T}$ ). Si el error es mayor a una tolerancia dada el valor de los pesos y el valor del sesgo se tiene que modificar de manera que el error cada vez sea menor, de tal manera que el proceso se realiza iterativamente hasta que se minimice el error, a estos se le llama entrenamiento o aprendizaje.

En la literatura para las RNA, existen principalmente dos tipos de aprendizaje, el supervisado y el no supervisado. En este caso se trabajará con el tipo de aprendizaje supervisado, en el cual durante la etapa de entrenamiento del modelo neuronal se proporcionan datos de entrada  $x_1, \dots, x_n$  así como datos esperados  $y_1, \dots, y_n$  los cuales corresponden a las respuestas correctas y a partir de los datos de entrada el modelo genera datos de salida  $y'_1, \dots, y'_n$  durante la etapa de evaluación del modelo neuronal (Hsieh, 2009).

Se le llama aprendizaje supervisado en el sentido de que la salida del modelo ( $y'_1, \dots, y'_n$ ), es guiado hacia los datos esperados, con la finalidad de minimizar la función objetivo. En este caso, la salida de la red neuronal serán datos de oleaje a partir de datos de viento, los cuales serán comparados con los datos esperados que corresponden a mediciones de oleaje.

#### 4.2. Entrenamiento RNL

El aprendizaje de una red neuronal consiste en encontrar o ajustar el valor de los pesos con la finalidad de minimizar el error entre los datos generados por la RNA y los datos objetivo (conjunto de datos esperados “targetdata”), para lo cual, es necesario contar con una función objetivo y minimizarla “ $J$ ” (también conocida como función costo, función de error o función de pérdida), definida como el cálculo del error medio cuadrático *EMC*

$$J = \frac{1}{N} \sum_{n=1}^N \left[ \frac{1}{2} \sum_k [Y_k^{(n)} - Y_{dk}^{(n)}]^2 \right] \quad \text{Ec. 4}$$

Donde  $Y_{dk}$  corresponde a los datos esperados (targetdata),  $Y_k$  es la salida del modelo neuronal y para fines de esta trabajo  $n=1, \dots, N$  observaciones o mediciones de altura de ola significativa. Esta forma de calcular el error es el más utilizado y comúnmente reportado en trabajos similares: (Deo & Naidu, 1999); (Jain et al., 2011); (Tsai et al., 2002). El *EMC* (*Error medio cuadrático*) es la forma más común para representar la función objetivo en problemas de regresión no lineal, como minimizar el EMC es equivalente a maximizar la probabilidad de la función asumiendo que el error tiene una distribución Gausiana.

En general, al inicio del proceso de optimización los valores de los pesos son elegidos aleatoriamente, las entradas son llevadas hacia la red neuronal, la información es procesada para obtener una salida, posteriormente se calcula el error entre la salida de la red neuronal y el valor real, este resultado es propagado hacia atrás con la finalidad de ajustar el valor de los pesos. Este proceso se realiza iterativamente hasta que la función

objetivo satisfaga algún criterio de convergencia. Por lo que cada uno de los valores de datos de entrada se ingresa a la red neuronal y cada uno de sus pesos se actualizan, generando un “epoch o ciclo de entrenamiento”. Y una vez encontrados los parámetros óptimos, el entrenamiento es suspendido y la red neuronal esta lista para realizar una evaluación con nuevos datos de entrada.

Normalmente el registro de datos de entrenamiento se divide en dos partes, con la primera parte se realiza el entrenamiento de la red y la segunda parte es usada para la evaluación de la misma.

Debido al gran número de parámetros con los cuales puedes entrenar y la gran flexibilidad que tienen las redes neuronales, la salida del modelo neuronal se puede ajustar muy bien en la etapa de entrenamiento pero durante la etapa de evaluación es posible que se genera un ajuste muy pobre, resultado de un “**sobreajuste**”. Es muy común que los datos usados para entrenar los modelos neuronales contengan ruido, el ruido lo consideramos como todo aquello en una señal o información en general, que no es de interés, la degrada o distorsiona, impide y/o limita un estudio o un uso de tal información (Márquez, 2012). Durante la etapa de entrenamiento, los modelos neuronales tienen la capacidad de aprender a partir de los datos, incluyendo aquellos que presentan ruido. Por lo que el error obtenido en la etapa de evaluación primero decrece y posteriormente incrementa, debido a que el modelo está aprendiendo de la señal con ruido (presente en las mediciones) durante el periodo de entrenamiento. En la Figura 7, se representa un esquema que ejemplifica el sobreajuste, la curva segmentada representa un buen ajuste hacia los datos con ruido (indicados con cuadrados), por otro lado, la curva continua representa el sobreajuste, donde vemos que el ajuste es bueno en la etapa de entrenamiento, pero es pobre en la etapa de evaluación (círculos).

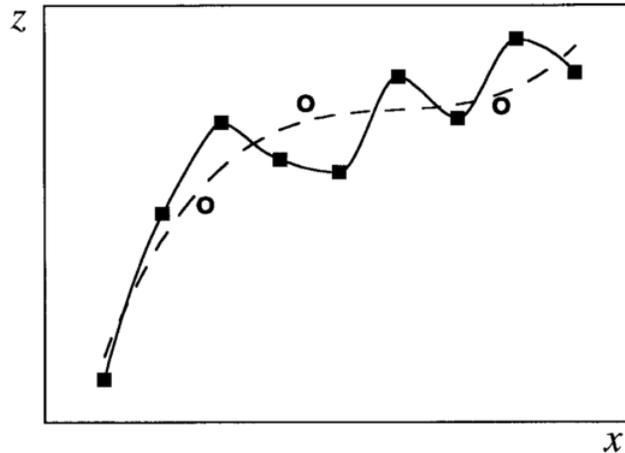


Figura 7. Esquema donde se ejemplifica el problema de sobreajuste.

El sobreajuste es uno de los problemas que a menudo se presentan al implementar los modelos neuronales, para evitarlo es necesario determinar la etapa en que la red neuronal ya no necesita más entrenamiento, es esto se le denomina "early stop o stopped training", el cual lo podemos observar en la Figura 8, donde, conforme aumenta el número de ciclos de entrenamiento (epochs), la función objetivo (" $J$ ") disminuye en la etapa de entrenamiento (línea segmentada), sin embargo para la etapa de evaluación podemos observar que al inicio, la función " $J$ " decrece, pero eventualmente se incrementa (curva continua), punto en el cual representa que el entrenamiento debe ser suspendido, ya que puede iniciar una sobreajuste (Hsieh y Tang, 1998).

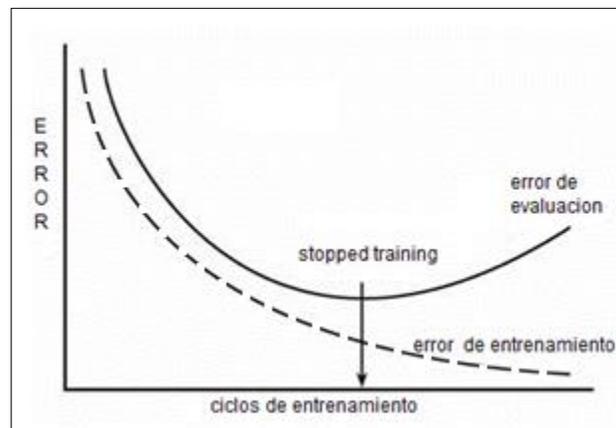


Figura 8. Comportamiento del error, donde se muestra el punto "stopped training".

Para evitar el problema de sobreajuste, uno de los métodos comúnmente usados es el llamado método estadístico “**stepwise**” el cual agrega y elimina términos basado en un modelo de regresión múltiple, a cada paso se calcula una regresión lineal y de ella se obtiene los estadísticos, prueba F de Fisher, que se aplican para determinar la relación y grado de asociación entre la variable dependiente y sus supuestos valores de la variable explicativa. De esta manera vamos a evitar usar datos redundantes y que no son necesarios para obtener un buen resultado y de esa manera hacer más eficiente el tiempo de cómputo.

Este tipo de regresiones va paso a paso calculando estadísticos, pruebas y análisis a medida que vamos introduciendo y/o sacando variables independientes en el modelo; mientras que el análisis de regresión lineal simple, resolvía el proceso en un paso al contar solamente con la relación de un par de variables. El objetivo es encontrar entre todos los valores de la variable explicativa (viento) aquellos que mejor expliquen la variable dependiente (oleaje) sin que ninguna de ellas sea combinación lineal de las restantes.

Esto implica de manera general;

1. En cada paso se introducen los datos que cumplen con los criterios de entrada lo cual consiste en, verificación de los criterios de probabilidad de entrada, el p-valor asociado al estadístico F (Fisher) o probabilidad de entrada, nos indica si la información proporcionada por cada una de las variables es redundante, lo cual se verifica, si este es menor que un valor crítico de la probabilidad de inicio (digamos 0.05)
2. En este paso se valora si alguna de los datos cumple los criterios de salida, en este caso, si el p-valor asociado al estadístico F, o probabilidad de salida, es mayor que un determinado valor crítico (digamos 0.10), entonces la variable será eliminada.
3. En este paso, se valora la bondad del ajuste de los datos con el modelo de regresión lineal y se calculan los parámetros del modelo verificado en dicho paso. El proceso se inicia sin ningún valor y se van introduciendo valores paso a paso y el proceso concluye cuando no queda ningún valor fuera de la ecuación que satisfaga el criterio de selección (garantiza que las variables seleccionadas son significativas) y/o el criterio de eliminación (garantiza que una variable seleccionada no es redundante).

Si algún término en el modelo no se encuentra, se debe a que tiene un valor de probabilidad menor que la tolerancia de entrada

Si algún término en el modelo tiene un p-valor mayor que la tolerancia de salida (es decir que la hipótesis nula puede ser rechazada) y se elimina el que tenga un valor mayor de “p” y se regresa al paso anterior; si sucede lo contrario termina el cálculo. El método termina cuando no hay más mejoras del modelo.

En la etapa de entrenamiento o aprendizaje se enseña al modelo neuronal a relacionar la variable viento usando componentes  $u$  y  $v$ , así como la magnitud del viento; con la altura de ola. Este tipo de regresión múltiple calcula estadísticos, pruebas y coeficientes de correlación entre la altura de ola y cada uno de los componentes y velocidad de viento, para cada paso de tiempo. De tal manera que el método identifica y elige las componentes que son más representativas para generar una determinada altura de ola.

El modelo estadístico “stepwise” nos arroja un matriz del mismo tamaño que la matriz de entrada (donde se encuentran las componentes de viento y la velocidad), con valores lógicos que nos indican la posición del valor de la variable que ha sido elegida por el modelo estadístico, por lo que esa matriz lógica es usada para entrenar la red neuronal y posteriormente también evaluarla.

Hsieh et al. (1998) identificaron el problema de contar con un registró de datos muy corta y que no sea suficiente para generar una buena optimización de la función objetivo. Presentándose inestabilidades no-lineales, esto ocurre, cuando las series de datos que tenemos para entrenar, no son suficientemente largas en relación con la escala de tiempo del fenómeno que uno quiere analizar o estudiar. En el área de oceanografía y meteorología es un problema que muy a menudo se presenta.

Para resolver las inestabilidades no lineales, en este trabajo se utiliza un método estadístico de re-muestreo llamado originalmente “bagging” (abreviación en ingles de **bootstrap aggregating**), debido a que no se cuenta con el suficiente registro de datos para la escala de los fenómenos que deseamos analizar con las RNA. Este método consiste esencialmente, en la extracción de un gran número de muestras repetidas de los propios datos, y sobre esta base se realizan posteriormente descripciones e inferencias estadísticas.

Debido a lo anterior, la técnica estadística de bootstrapping (Efron, 1979; Efron y Tibshirani, 1993) o llamado (resampling) remuestreo de datos, se basa en la teoría de la probabilidad y la inferencia estadística la cual a partir de los datos observados genera nuevas submuestras simuladas de igual tamaño que la muestra original a la que llamamos submuestras. Este método nos ayuda aproximar el sesgo y la varianza del análisis estadístico, así como para construir intervalos de confianza. Debido a que con este método generaremos más datos, es necesario evitar un sobreajuste usando el método de “early stop” ya descrito previamente y de esa manera no caer en un sobreajuste usando datos redundantes. En este trabajo, se está usando la función “*bootstrp*” de Matlab que ya tiene por default implícito el método de “**early stop**”.

Revisando los trabajos citados en la parte de antecedentes, se puede observar que los autores usan como mínimo un año de datos para realizar un entrenamiento de la red neuronal y obtener buenos resultados; por tal motivo y con la finalidad de validar la metodología seguida con las RNA para generar oleaje en el área de estudio, se trabajara con una base de datos de oleaje histórica de 30 años. En tal caso, se cuenta con más de un año de registro de datos que serán usados para entrenar los modelos neuronales, por lo que el uso de la técnica de re-muestreo no será necesario usarla.

El número de datos ideales para entrenar los modelos neuronales, depende de cada caso de estudio, sin embargo el objetivo para lograr un buen entrenamiento es optimizar la función objetivo, y eso depende también del fenómeno ambiental del cual estamos estudiando así como la escala del fenómeno. Tangang et al. (1998<sup>a</sup>).

#### *4.3. Red Neuronal Bayesiana - RNB*

Este tipo de red corresponde a una regularización bayesiana de tipo multicapa MLP (multilayer perceptrón model), el cual a diferencia del modelo neuronal lineal presenta tres capas; la primera; es la que corresponde a los datos de entrada, la segunda es la capa intermedia o capa oculta, donde se encuentra una o más neuronas con función de transferencia no lineal y finalmente la capa de salida donde se generan los resultados, en esta capa se encuentra una función de transferencia lineal (Figura 9). Como se mencionó anteriormente en el apartado de las redes multicapa, los parámetros que pueden variar en

este tipo de redes neuronales son: la cantidad de datos de entrada, número de neuronas en la capa intermedia, número de neuronas en la capa de salida y el número de ciclos de entrenamiento.

En este caso, el tipo de algoritmo que se está usando en la capa intermedia de acuerdo a Foresee y Hagan (1997) donde aplican una aproximación del modelo neuronal multicapa usando el algoritmo de optimización Levenberg-Marquard; introduciendo un tipo de red Bayesiana, llamada así debido a que toma como base el Teorema de Bayes, MacKay (1992b,a).

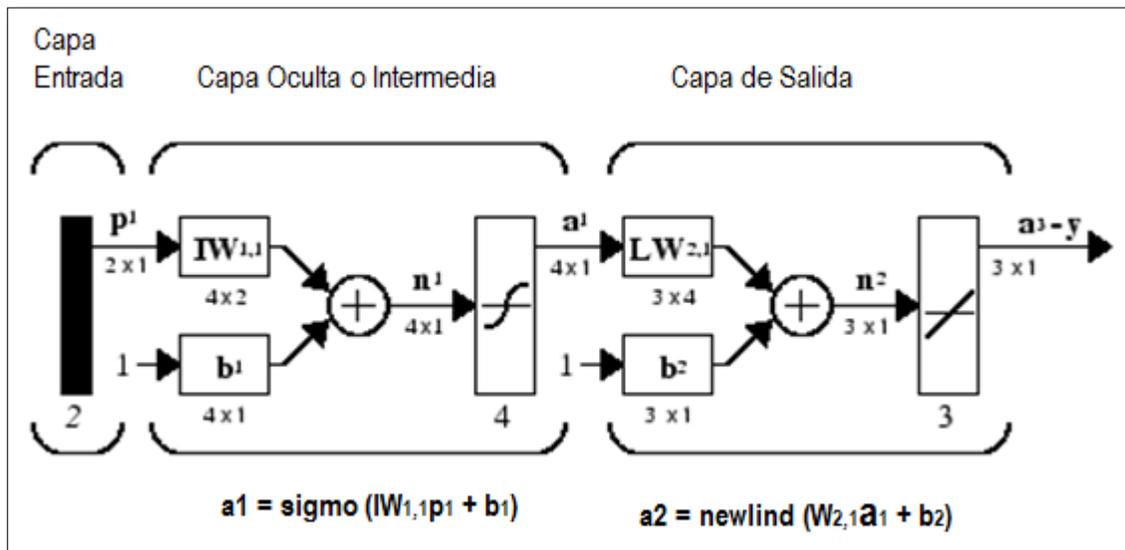


Figura 9. Representación general de un tipo de red neuronal multicapa (MLP).

En la figura 8 se puede observar que los datos de entrada están representados por un vector  $P$  que es multiplicado por una matriz de pesos  $W$  y a ese producto se les suma el sesgo " $b$ " ( $n = Pw+b$ ), este producto es enviado a la función de transferencias la cual actualiza los pesos y los sesgos de acuerdo al algoritmo Levenberg-Marquard,

El algoritmo Levenberg-Maquard (L-M) es una optimización que provee soluciones numéricas al problema de minimizar funciones no lineales, el algoritmo interpola usando como base el método de Gauss-Newton (G-N) cuando se encuentra cerca de la solución y un descenso de gradiente cuando se encuentra lejos de la solución, por lo que el algoritmo

L-M es más robusto en comparación con el método de G-N. El secreto del L-M es alterar la diagonal de la matriz cuadrada del producto de Jacobianos en las ecuaciones normales del G-N. Recordando que el método de G-N, es el método más usado para resolver problemas lineales por mínimos cuadrados.

A partir del resultado obtenido por el algoritmo L-M, genera un resultado que está dado por **(a1 = sigmo (IW1,1p1 + b1))**, posteriormente esta salida es la entrada de la siguiente neurona que se encuentra en la capa de salida por lo que ahora el resultado generado por el modelo neuronal es **a2 = newlind (W2,1a1 + b2)**, la función de transferencia en esta capa es lineal, la cual contiene el algoritmo LMS (*Least Mean Square*), para minimizar en la etapa final el error entre la salida del modelo neuronal y el dato esperado.

Como se ha mencionado anteriormente, el principal objetivo en la etapa de entrenamiento es minimizar el error obtenido por el modelo neuronal y los datos conocidos (Ec. 5), y al mismo tiempo este tipo de optimización trata de minimizar la magnitud de los parámetros de los pesos (Ec. 6)

$$Ed = \frac{1}{2} \sum_{i=1}^N (Y_i - Y_{di})^2 \quad \text{Ec. 5}$$

$$Ew = \frac{1}{2} \sum_{j=1}^{Nw} W_j^2 \quad \text{Ec. 6}$$

Donde w contiene todos los pesos y sesgos de los parámetros de los modelos de redes neuronales MLP, con un total de **Nw** parámetros, por lo que la función objetivo queda de la siguiente manera:

$$J(w) = \beta Ed + \alpha Ew \quad \text{Ec. 7}$$

Donde  $\alpha$  corresponde a los hiperparámetros de los pesos. Anteriormente  $\beta$  se simplificaba a 1, pero en la regularización Bayesiana, el valor de  $\beta$ , es usado para describir el comportamiento Gaussiano de los datos esperados. La optimización Bayesiana será automáticamente determinado por dos valores llamados **hiperparámetros  $\alpha$  y  $\beta$** . Si  $\alpha \ll \beta$  entonces el modelo se ajustará a un número cercano; pero si pasa lo contrario donde  $\alpha \gg \beta$  entonces el modelo se forzará para obtener un ajuste más cercano, haciendo un ajuste suavizado de los datos.

Cuando se entrenan redes neuronales usando este método, los valores de los pesos de la red neuronal se abordan desde un punto de vista de probabilidad. Aplicando la ley de Bayes para la probabilidad posterior de los pesos pero considerando una medida del error que garantice obtener una red neuronal con buena regularización, se tiene la expresión mostrada en la Ec. 8:

$$P(\mathbf{w}/\mathbf{D}, \alpha, \beta, \mu) = \frac{P(\mathbf{D}/\mathbf{w}, \alpha, \beta, \mu)P(\mathbf{w}/\alpha, \beta)}{P(\mathbf{D}/\alpha, \beta, \mu)} \quad \text{Ec. 8}$$

Donde:

$P(\mathbf{D}/\mathbf{w}, \alpha, \beta, \mu)$  : Densidad de la probabilidad a posteriori de los pesos

$P(\mathbf{w}/\alpha, \mu)$  : Densidad de la probabilidad a priori de los pesos

$P(\mathbf{D}/\mathbf{w}, \beta, \mu)$  : Densidad de la probabilidad condicional de los pesos

$P(\mathbf{D}/\alpha, \beta, \mu)$  : Factor de normalización.

Nota.- En todo momento se considera que nuestro caso de estudio presenta una distribución de tipo Gaussiana.

La expresión para la probabilidad condicional de los pesos se muestra en la Ec. 9, en la cual se observa que esta probabilidad depende del parámetro  $\beta$ . Lo anterior significa que los pesos que tenga la red influyen la salida de la misma y de esta manera afecta el error  $E_d$ .

$$P(\mathbf{D}/\mathbf{w}, \beta, \mu) = \frac{1}{Z_d(\beta)} \exp(-\beta E_d) \quad \text{Ec. 9}$$

Donde  $Z_d(\beta) = (\pi/\beta)^{p/2}$

La expresión para la probabilidad a priori de los pesos se muestra en la Ec. 10. Se observa que esta probabilidad depende del parámetro  $\alpha$ . Lo cual significa que los pesos afectan directamente la medida del error  $E_w$  por medio de dicho parámetro.

$$P(\mathbf{w}/\alpha, \mu) = \frac{1}{Z_w(\alpha)} \exp(-\alpha E_w) \quad \text{Ec. 10}$$

Donde:  $Zw(\alpha) = (\pi/\alpha)^{\frac{N}{2}}$

Por lo que anterior, la expresión para la probabilidad a posteriori de los pesos queda de la siguiente manera:

$$P(\mathbf{w}/D, \alpha, \beta, \mu) = \frac{1}{ZM(\alpha)} \exp(-M(\mathbf{w})) \quad \text{Ec. 11}$$

Al remplazar en las (Ec. 9 y Ec. 10) las expresiones para las diferentes densidades de probabilidad se tiene el resultado mostrado en la Ec. 12 y aplicando el teorema de Bayes se llega finalmente a la Ec. 13:

$$P(\mathbf{w}/D, \alpha, \beta, \mu) = \frac{\frac{1}{ZM(\alpha)} \frac{1}{ZM(\alpha)} \exp(-(\beta Ed + \alpha Ew))}{\text{Factor de normalizacion}} \quad \text{Ec. 12}$$

$$P(\mathbf{w}/D, \alpha, \beta, \mu) = \frac{\frac{1}{ZM(\alpha)} \frac{1}{ZM(\alpha)} \exp(-J)}{Z(\alpha, \beta)} \quad \text{Ec. 13}$$

Donde “**J**” corresponde a la función objetivo y  $Z(\alpha, \beta)$  corresponde al factor de normalización. Hasta aquí el método encuentra el óptimo  $W$ , que se necesita para maximizar la densidad de la probabilidad posteriori, el siguiente paso consiste en encontrar los valores óptimos de los hiperparámetros  $\alpha$  y  $\beta$ ; aplicando nuevamente el método de Bayes.

Los pasos a seguir en resumen para aplicar la optimización bayesiana es la siguiente:

1. Inicializar  $\alpha$  y  $\beta$  y los pesos de la red neuronal ( $w$ ). Una opción de los valores que se les puede asignar es:  $\alpha=0$ ,  $\beta=2$  de esta manera al inicio del entrenamiento se garantiza que la red trate de minimizar el error **Ed** sin penalizar la magnitud de los pesos.

2. Minimizar la función  $J(\mathbf{w})$  con el método de Levenberg-Maquardt, ya que en el mismo método realizan la estimación de la matriz Hessiana  $H$  que será usado en los siguientes pasos del algoritmo. Siendo la  $H$  la matriz cuadrada de las segundas derivadas parciales.
3. Con el valor del cálculo de la  $H$ , se calcula el número efectivo de parámetros:

$$\gamma = Nw - \alpha MPtr((H^{MP})^{-1})$$

4. Donde el método de aproximación de Gauss-Newton se aplica en la matriz Hessiana implícitamente en el algoritmo Levenberg-Maquardt, el cual está dado de la siguiente manera:

$$H \approx \beta J^T J + \alpha I$$

Donde  $J$  corresponde a la matriz Jacobiana e  $I$  a la matriz identidad.

El Jacobiano se refiere a la matriz formada por las derivadas parciales de primer orden de una función (en este caso de la función objetivo " $J$ "), y se usa aquí para hacer aproximaciones lineales.

5. El paso 2 al 4 se realiza iterativamente hasta que el entrenamiento converge.

Una vez que conocemos los hiperparámetros  $\alpha$  y  $\beta$  la función objetivo puede ser minimizada, generando los pesos óptimos  $\mathbf{W}_{MP}$  para el modelo neuronal. Por lo que encontrar estos valores se resume a encontrar el valor que minimiza la expresión del error.

## 5. ÁREA DE ESTUDIO

El estado de Yucatán se encuentra sobre la península de Yucatán (Figura 10), la cual presenta una plataforma continental extensa conformada mayormente por rocas sedimentarias depositadas sobre un basamento Paleozoico (Lugo- Hubp *et al.*1992), la plataforma continental presenta una pendiente de 1:1000.

El puerto de Sisal se localiza a 53 km al Noroeste de la ciudad de Mérida capital del estado de Yucatán, se encuentra limitado geográficamente al este con el Puerto de Chuburná, al oeste con la reserva de Celestún y al Sur con el municipio de Hunucmá (Figura 10). Históricamente fue el principal puerto del estado desde el siglo XVI hasta inicios del siglo XX, fundado para servir a la industria henequenera, así como las transacciones marítimas de importación y exportación de productos de otros estados, países y continentes.



Figura 10 Imagen tomada de Google Earth, localización del Puerto de Sisal dentro del territorio mexicano.

La costa de Sisal se puede clasificar geomorfológicamente de tipo isla de barrera (Carranza *et al.* 1975), está compuesta de una playa de barrera en la parte oceánica, en la que se dan los procesos de transporte de material sedimentario, así como los procesos físicos debidos a la energía del oleaje, corrientes, mare y viento; haciendo dicha zona muy dinámica. En la

parte interior hacia tierra se puede observar humedales, conocidos localmente como ciénaga.

En el área de estudio se presenta oleaje de baja energía con una altura media anual de 60 cm, con dirección reinante del noreste. Sin embargo en temporadas de “Nortes” se presenta oleaje dominante del norte y noroeste con alturas máximas de entre 1.60 y 1.80 m asociadas a periodos de 2 y 8 segundos. La costa yucateca se caracteriza por estar dominada por vientos locales y un sistema de brisas marinas las cuales juegan un papel muy importante sobre los parámetros de oleaje (Appendini et al. 2012).

## 6. DESCRIPCIÓN DE LOS DATOS

En este trabajo se usaron datos medidos in-situ de viento y de oleaje de los años 2013, 2014 y 2015. Así como datos generados por modelos numéricos (oleaje) a partir de datos de reanálisis de 30 años (viento); con estos datos se entrenaron y validaron los modelos neuronales usados en los diferentes casos de aplicación.

### *6.1. Oleaje*

El oleaje fue registrado con dos equipos ADCP de la marca RD instruments (por sus siglas en inglés - Acoustic Doppler Current Profile). Estos instrumentos como su nombre lo dice son perfiladores acústicos que miden la velocidad de las corrientes por medio del efecto Doppler. Esto consiste en que el instrumento emite pulsos de sonido a frecuencias constantes en el agua, a medida que estas ondas de sonido viajan en el medio marino, estas serán eventualmente reflejadas en las partículas (sedimento o plantcon) que se encuentran suspendidas en el medio marino y las cuales viajan a la misma velocidad del agua. El sonido se transmite en cuatro rayos inclinados en ángulo recto el uno del otro, el corrimiento en la frecuencia del sonido reflejado respecto al transmitido, permite conocer la velocidad de la partícula en la dirección del rayo emitido (velocidad radial) y al conocer la frecuencia con que el aparato emite la señal y el tiempo que tarda en regresar, podemos calcular la distancia. Los diferentes tiempos de llegada indican el sonido reflejado desde diferentes distancias respecto a los transductores, así que un ADCP proporciona la información no sólo sobre la rapidez de la corriente y su dirección en un punto del océano, sino en todo un rango de profundidad es decir produce un perfil de corriente contra profundidad.

Los equipos ADCP utilizados, fueron programados para que midieran oleaje cada hora durante 20 minutos a una frecuencia de 2 Hz, de los datos recabados de la superficie libre se obtuvo altura de ola significativa, periodo pico, dirección espectral, corrientes, profundidad y temperatura. Estas mediciones son almacenadas en una tarjeta de memoria interna, una vez que se recupera simplemente se descarga la información a una computadora para ser procesada, con un software del mismo distribuidor del equipo llamado Waves Mon.

Uno de los ADCP, se colocó aproximadamente a 11 km de la costa y a una profundidad de 10 metros; en las coordenadas  $21.276^\circ$  latitud norte y  $90.0357^\circ$  longitud oeste, el cual lo etiquetaremos de aquí en adelante como ADCP-10m o ADCP-Monitor. Las mediciones se registraron del 26 de marzo del 2014 al 29 de mayo del 2014. El otro, se colocó a 1 km de la costa aproximadamente y a una profundidad de 4.5 metros; en las coordenadas  $21.1732^\circ$  latitud norte y  $90.0442^\circ$  longitud oeste ADCP-4.5m o ADCP-Sentinel (Figura 11) y las mediciones de este equipo van del 25 de Marzo del 2014 al 22 de Mayo del 2014.

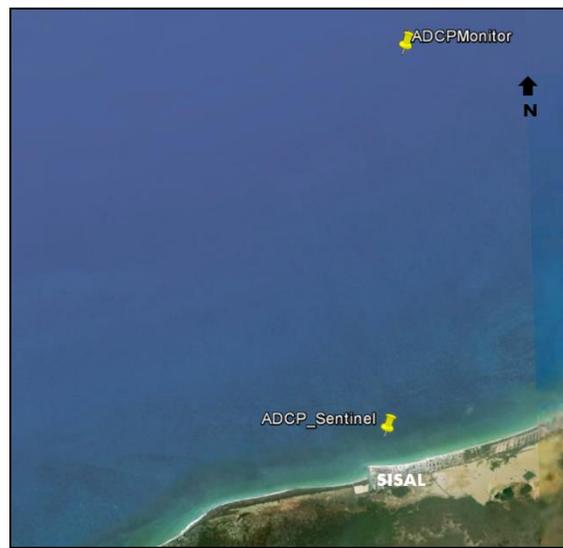


Figura 11. Localización de los equipos de medición ADCP, con respecto a la línea de costa de Sisal. Los instrumentos de medición se colocaron en el fondo del mar, cada uno sobre una estructura tipo trípode en la cual se monta el ADCP (Figura 12), en cada una de las patas tiene peso, con la finalidad de fijarlo al fondo y anclarlo lo suficiente para que las corrientes y el oleaje no lo muevan. La instalación se hizo a bordo de una embarcación del LIPC; con la ayuda de dos buzos y otras dos personas en la embarcación para cumplir con la maniobra de colocación del instrumento.



Figura 12. Ejemplo del instrumento de medición de oleaje ADCP.

En la Figura 13, se puede observar los datos de oleaje del aparato que se encuentra más alejado de la costa (ADCP-10m-Monitor) y del que se encuentra cerca de la costa (ADCP-4.5m-Sentinel). En la Figura 13, se observa que el oleaje que incide con mayor frecuencia proviene de la dirección noreste y por otro lado el que tiene menor frecuencia, pero en mayor intensidad es el oleaje que proviene de dirección noroeste y norte, teniendo direcciones de oleaje de tres cuadrantes. De igual manera podemos observar que la dirección de oleaje que registró el ADCP-4.5m-Sentinel, ha sufrido modificación debido a los fenómenos de propagación y transformación del oleaje (ej. refracción). Por tal motivo, las direcciones de oleaje con mayor frecuencia son del noreste (Figura 13, b).

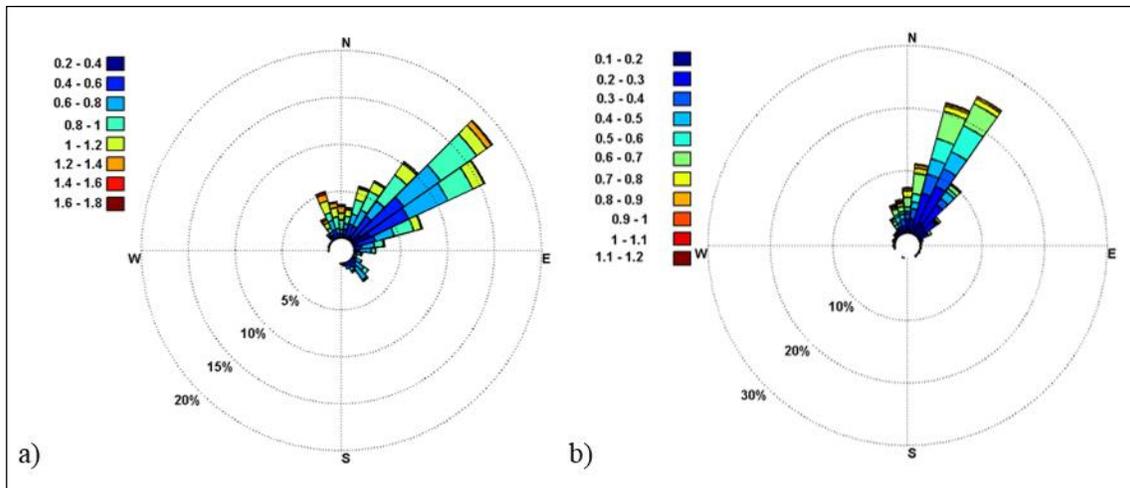


Figura 13. Rosas de dirección de oleaje. Datos medidos por el ADCP-10m-Monitor (a) y ADCP-4.5m-Sentinel (b) respectivamente.

## 6.2. Viento

Los datos de viento utilizados para entrenar y evaluar la red neuronal fueron colectados con un anemómetro ultrasónico, el cual se encuentra ubicado sobre una torre meteorológica de la Unidad Académica Sisal – UAS, localiza en los coordenadas: latitud:21.165° norte y longitud: 90.047° oeste (Figura 14).

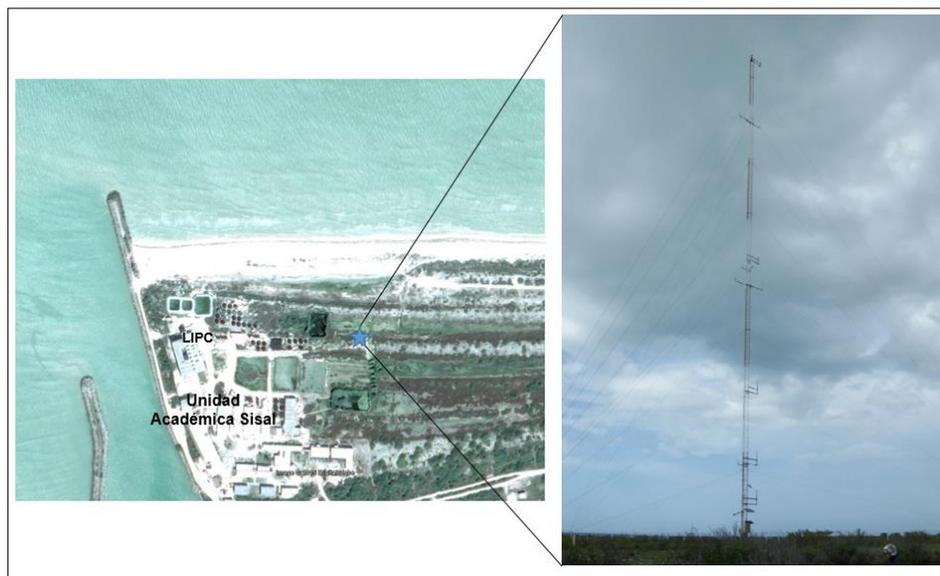


Figura 14. Localización de la torre meteorológica donde se encuentra el anemómetro de donde se adquirieron los datos de viento.

Sobre la torre se encuentra instalados 5 anemómetros a diferentes alturas (Tabla 1), la columna uno es la forma en que se identifican los anemómetros instalados en la torre, la segunda columna hace referencia a 2D si mide dos componentes de velocidad de viento (**u y v en los ejes x, y**) o 3D si registra en las tres componentes de velocidad (u, v y w en los ejes x, y, z respectivamente); y finalmente la tercera columna corresponde a la altura sobre el nivel del piso donde se encuentran instalados.

Tabla 1. Tipos de anemómetros localizados en la torre meteorológica de la UAS.

ID	tipo	h(m)
0	2D	3
2	2D	6
3	3D	12.5
<b>1</b>	<b>2D</b>	<b>25</b>
4	3D	51

Para fines de este trabajo, se usaron los datos correspondientes a 25 m sobre el nivel del piso, por lo que solo se obtuvieron velocidades para las componentes **u** y **v** (m/s). La frecuencia de adquisición de los datos con el anemómetro fue de 10Hz, dichos datos se guardan automáticamente en archivos en formato ASCII y son tratados con la técnica de covarianza de remolinos (Figuroa-Espinoza y Salles, 2014) para estudiar las correlaciones de escala que caracterizan el ambiente costero particular. Los datos que fueron usados específicamente para este trabajo son promedios de las componentes medidas (u y v) para cada hora.

En la Figura 15 podemos observar que los vientos reinantes provienen del sureste y este, estos son los que presentan mayor frecuencia de ocurrencia pero no son tan intensos como los vientos dominantes los cuales, proviene del noreste, noroeste y norte, estos son más intensos como ya se mencionó anteriormente y con mayor frecuencia durante la temporada invernal.

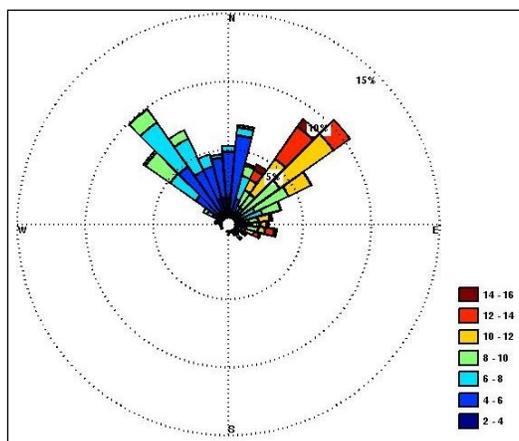


Figura 15. Rosa dirección de viento para el periodo de medición en Sisal.

### 6.3. Hindcast

Además de los datos de oleaje recolectados directamente por el laboratorio, también se utilizaron datos de una base histórica (hindcast) de oleaje generada en el LIPC (Laboratorio de Ingeniería y Procesos Costeros) para el Golfo de México y Mar Caribe, la cual incluye el área de estudio de la costa de Sisal, Yucatán. Los datos fueron generados a partir de campos de viento del North American Reanalysis utilizando el modelo MIKE 21 SW. La base de datos histórica, tiene una cobertura temporal de enero de 1979 al 31 de diciembre de 2008 a las. Los parámetros de oleaje que contiene esta base de datos son: altura significativa ( $H_s$ ), periodo pico ( $T_p$ ), periodo medio ( $T_m$ ) y dirección del viento así como velocidad y dirección del viento del NARR interpolados en la malla de cálculo del modelo de oleaje (Appendini et al., 2013). Los datos se encuentran promediados cada 3 horas del total de los 30 años (<http://observatoriocostero.iingen.unam.mx/web/>).

De esta serie de datos, se utilizaron los datos del nodo 12918 (Latitud:  $21.286329^\circ$  y Longitud:  $-90.070053^\circ$ ) que se encuentra a 3km de distancia del equipo de medición de oleaje (ADCP-10m-Monitor) Figura 16. Los datos de oleaje se usaron para entrenar la red neuronal y para validar los resultados arrojados por los modelos neuronales. Y los datos de viento del NAAR también fueron utilizados para entrenar y generar un hindcast, usando los modelos neuronales, para el punto específico de donde se extrajo las series de datos.

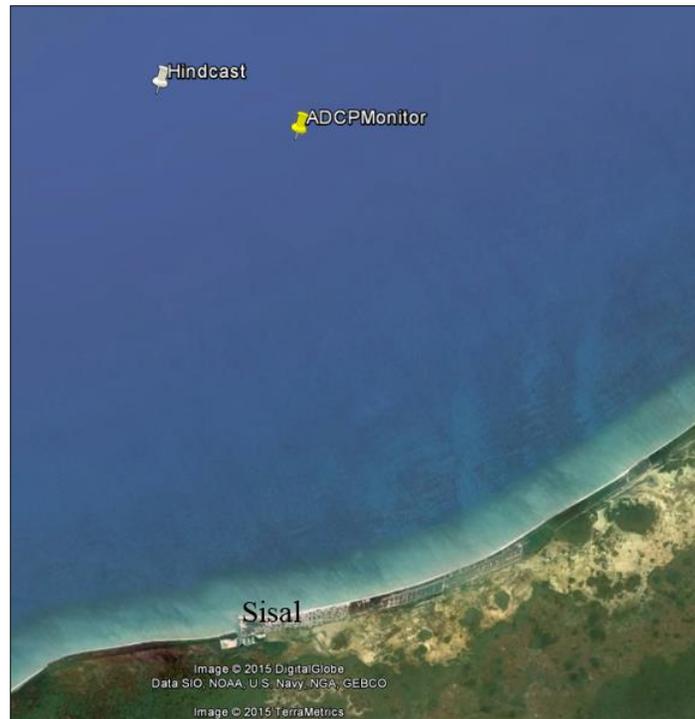


Figura 16. Localización del nodo 12918 del hindcast con respecto a la costa de Sisal y el punto de medición.

## 7. IMPLEMENTACIÓN Y CASOS DE APLICACIÓN DE LAS REDES NEURONALES

La utilización de RNA la podemos dividir en dos partes, la primera corresponde a la implementación donde se determina la arquitectura de la RNA determinando la mejor configuración que corresponde a números de capas, números de neuronas, método de entrenamiento y tipo de función de transferencia; así como tiempo y número de datos para entrenamiento con la finalidad de optimizar la etapa de aprendizaje y contar una RNA que genere resultados confiables.

### *7.1. Implementación de la RNA*

En esta sección del trabajo se están usando los datos que se colectaron del 26 de marzo al 29 de mayo del 2014. Los datos con los que se entrenó corresponden a datos promedio de cada hora de: viento (magnitud y dirección) y datos de oleaje (*Hs*) para los puntos donde se ubicaron los diferentes ADCP (Figura 11), recordemos que del 100% del registro de datos que se usan en estas metodologías un porcentaje está destinado para entrenar la red neuronal y el otro para evaluar. Una vez que se generan las salidas de los modelos neuronales se calcula el error con los datos esperados. Los pasos a seguir fueron los siguientes:

- Se descargaron los datos de los equipos de medición de oleaje (ADCP's), se obtuvo el promedio de cada hora, de altura de ola significativa, periodo pico y dirección.
- Se descargaron los datos de viento del portal del observatorio costero del LIPC (<http://132.247.45.41/vientosisal/concat.php>), que corresponden a las fechas de medición de los ADCP's.
- Posteriormente se inició usando una red neuronal lineal-RNL (newlind), realizando combinaciones con diferentes porcentaje de datos para entrenar (de 40% al 80%) y tiempo de generación de oleaje de 1 a 5 días con la finalidad de obtener la mejor combinación.
- Se evaluó la RNA comparando el resultado obtenido contra los datos correspondientes (valores esperados – *Hs* registrada de los ADCP).

- Se calculó el error medio cuadrático (*ECM*), así como correlaciones de los datos obtenidos por la RNA contra los medidos, para saber qué tan precisa y eficiente es la metodología, para generar altura de ola significativa, para los sitios de interés.
- El error fue calculado de la siguiente manera:

$$ECM = \frac{(\frac{1}{N} \sum (H_{med} - H_{RNA})^2)^{\frac{1}{2}}}{H_{rms}}$$

Donde:

$N$  = Número total de  $H_s$

$H_{rms}$  = Altura de ola media cuadrática

$H_{med}$  =  $H_s$  registrada por el ADCP

$H_{RNA}$  =  $H_s$  obtenida por la RNA

Nota.- La precisión del modelo neuronal mejora al momento en que el error medio cuadrático se aproxima a cero.

- Una vez definida la combinación de porcentaje de datos y horas de predicción que arroja el menor error, se usa el otro tipo de red neuronal RNB y ahora de manera sistemática se le cambian el número de neuronas en la capa intermedia, y se agrega los modelos probabilísticos “stepwise” para evitar un sobreajuste, lograr una optimización en la etapa de aprendizaje y por consiguiente minimizar el error.
- Posteriormente se comparan gráficamente las salidas de los modelos neuronales RNL y RNB contra los datos medidos, para saber qué modelo genera mejores resultados o se acerca más a los datos de oleaje real (datos medidos por los ADCP’s) en cada uno de los puntos de interés.

### *7.2. Determinación del tiempo y número de datos para entrenamiento.*

Como ejercicio inicial, la RNA fue entrenada usando el 50 % de datos de viento y oleaje (magnitud, dirección y  $H_s$ ) para obtener alturas de ola cada 24, 48, 72, 96 y 120 horas, estos resultados fueron comparados con los datos medidos del mismo periodo de tiempo (Figura 17). Como se puede observar en esta figura, el oleaje generado con la RNA es muy similar a la línea que corresponde a los datos medidos, sin embargo hay algunos momentos

en que los resultados sobreestiman la altura de ola y en otras los subestima. Con la finalidad de obtener una medida cualitativa del desempeño de la RNA se calculó el error medio cuadrático que hay entre los datos medidos y los generados con la RNA, los resultados para cada uno de los periodos se muestran en la Tabla 2.

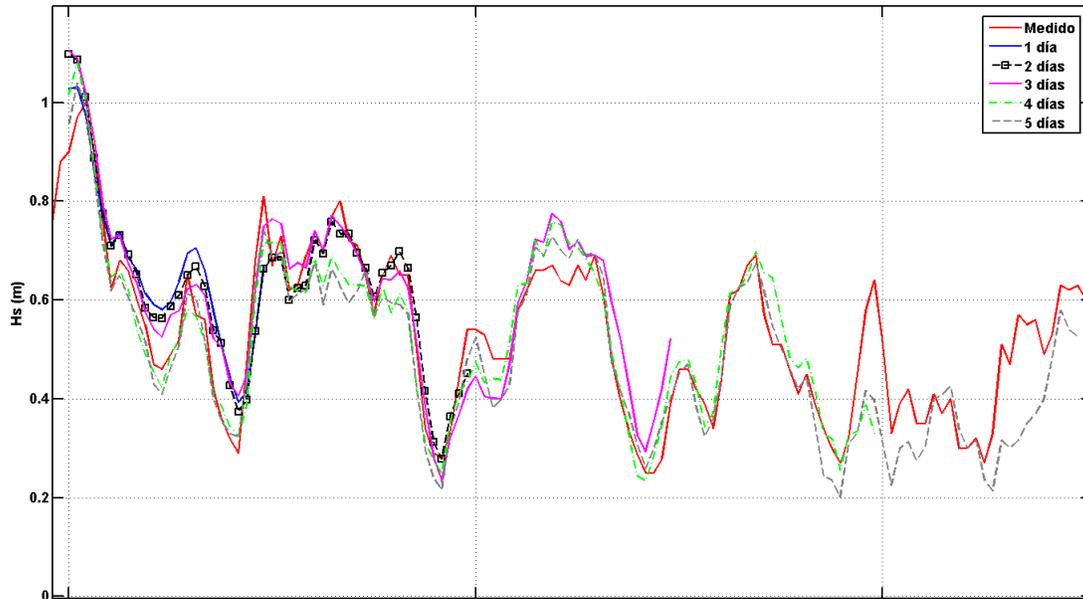


Figura 17. Generación de Hs usando RNA en el punto cerca de la costa, comparando con datos medidos.

Utilizando los datos del equipo ADCP-4.5-Sentinel que se encontraba cerca de la costa, se obtuvieron los resultados que se presentan en forma gráfica en la Figura 18 donde se puede observar un buen ajuste entre la salida de la RNA en comparación con los datos medidos para diferentes periodos de tiempo, la línea de color rojo corresponde a los datos esperados (datos medidos por el ADCP) y las demás líneas como lo indica la leyenda representan diferentes horas de generación de oleaje.

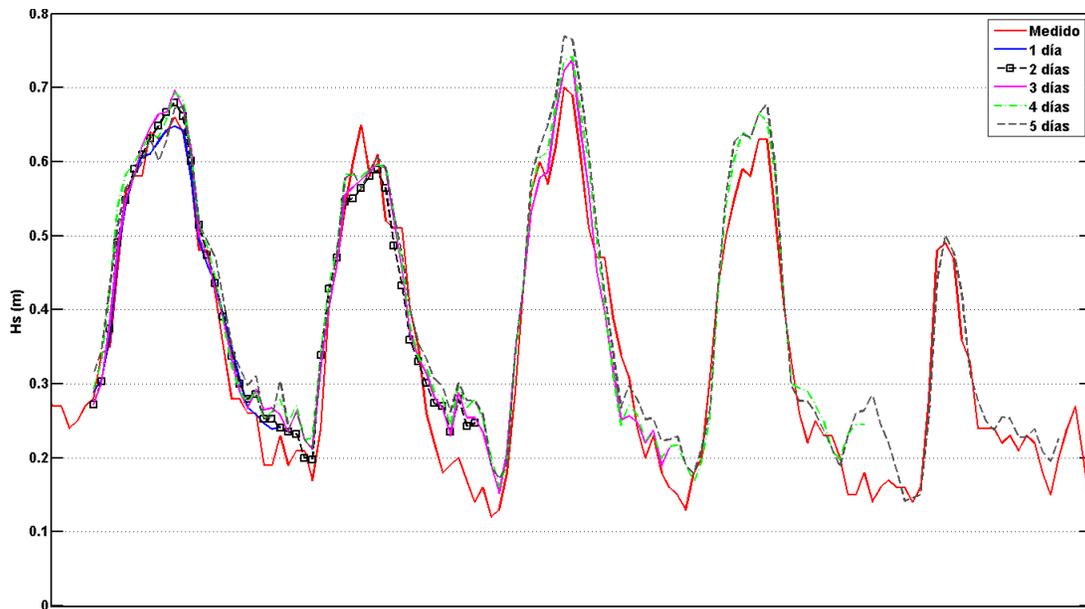


Figura 18. Generación de Hs usando RNA en el punto fuera de la costa, comparando con datos medidos.

Los ejercicios anteriores (Figura 17 y Figura 18) mostraron que, variando las horas de entrenamiento, se obtienen diferentes resultados. Con la finalidad de encontrar la mejor arquitectura como siguiente paso se realizaron variaciones usando el porcentaje de datos de entrenamiento del 40% al 80% del total de los datos disponibles y se varió el número de datos de generación de oleaje de 1 a 5 días (24 – 120 horas), con lo cual se obtuvieron los resultados mostrados en

En la Figura 19 se muestran las curvas del error medio cuadrático para los datos del ADCP-Monitor, la línea que corresponde al 60% presenta los mínimos errores (0.145 y 0.152) para 24 y 48 horas respectivamente.

Tabla 2; en la cual, podemos observar las combinaciones que presentan menor error para entrenar y generar oleaje. Para el ADCP-Monitor con 60 % de datos de entrenamiento y con 1 o 2 días de generación de oleaje se obtuvo los mínimos errores ( $EMC=0.145$  y  $0.153$

correspondientemente). Para el caso de del ADCP-Sentinel, las combinaciones que presentan menor error es con 60% de datos con 24 horas de generación de oleaje ( $EMC=0.217$ ) y 80% de datos con 4 días de generación de oleaje ( $EMC=0.191$ ).

En la Figura 19 se muestran las curvas del error medio cuadrático para los datos del ADCP-Monitor, la línea que corresponde al 60% presenta los mínimos errores (0.145 y 0.152) para 24 y 48 horas respectivamente.

Tabla 2. Error medio cuadrático con diferentes porcentajes de entrenamiento y tiempo de generación

Datos de ADCP Monitor					
	Tiempo de generación de oleaje (Días)				
% Entrenamiento	1	2	3	4	5
40	0.202	0.203	0.226	0.228	0.237
50	0.156	0.180	0.213	0.208	0.200
<b>60</b>	<b>0.145</b>	0.153	0.159	0.157	0.161
70	0.152	0.161	0.164	0.178	0.166
80	0.178	0.167	0.150	0.159	0.146
Datos del ADCP Sentinel					
	Días de generación de oleaje				
% Entrenamiento	1	2	3	4	5
40	0.227	0.251	0.254	0.260	0.254
50	0.228	0.254	0.264	0.273	0.252
<b>60</b>	<b>0.217</b>	0.248	0.257	0.236	0.239
70	0.221	0.219	0.220	0.248	0.224
80	0.237	0.221	0.222	<b>0.191</b>	0.222

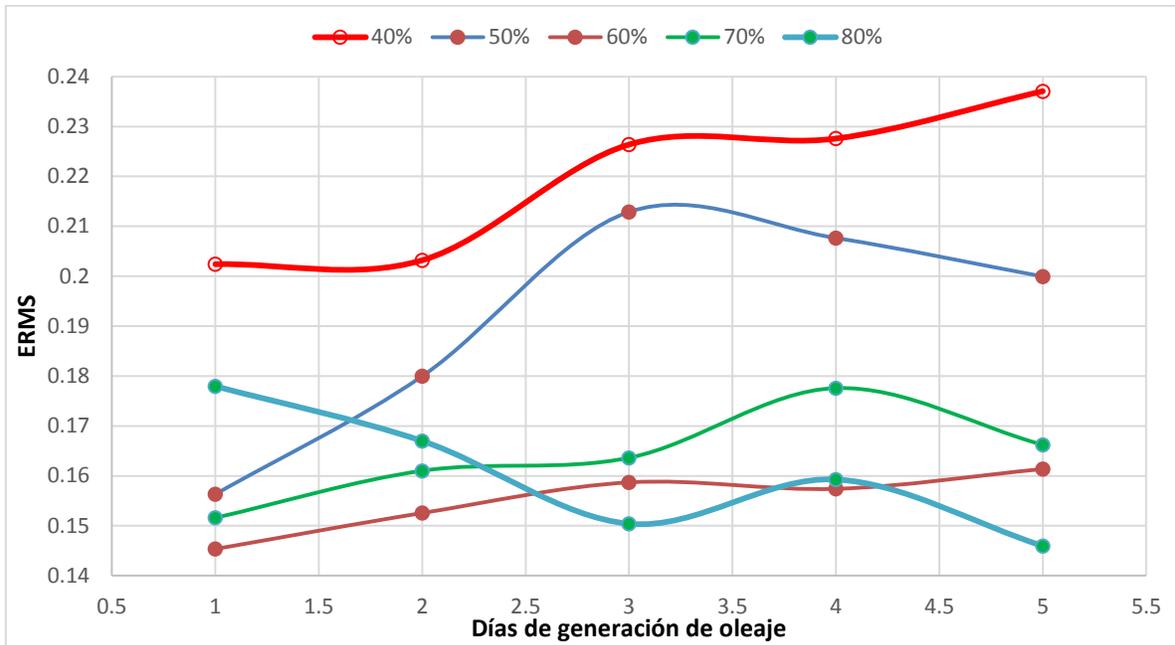


Figura 19. Error rms de los datos del ADCP-10m-Monitor.

La Figura 20, muestra una forma alternativa de mostrar el error mediante un gráfico de superficie, en la cual se grafica en 3D el porcentaje de datos a entrenar, las horas de generación de oleaje y el error obtenido, los colores hacia los rojos representan valores altos de error y por lo contrario los colores hacia el azul intenso, representa errores menores o bajos.

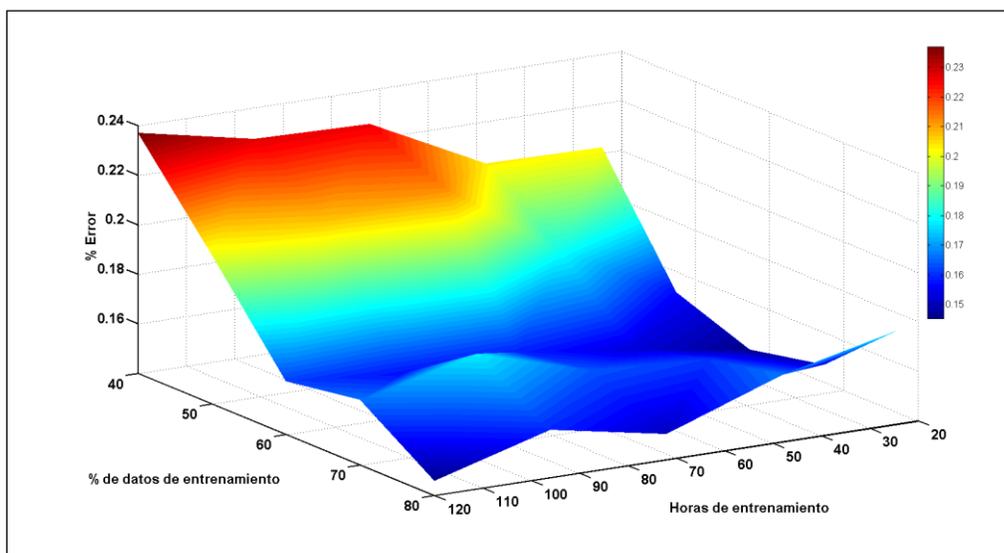


Figura 20. Grafica 3D donde se representan los errores obtenidos con las diferentes combinaciones de entrenamiento, para el punto fuera de la costa.

En la Figura 21, se observa las curvas del error medio cuadrático correspondientes al punto ADCP-Sentinel; la línea que corresponde al 60% presenta un error mínimo de 0.217 usando 24 horas de datos de entrada para entrenar la RNA. En la gráfica se muestra un mínimo error exagerado que corresponde a la línea de 80% de datos y 96 horas de datos para entrenar, el mínimo error es de 0.19, sin embargo la realidad es que la diferencia del error obtenido con la pareja de datos de entrenamiento (60% y 24 horas) menos el error obtenido con la pareja de datos de entrenamiento (80% y 96 horas) es del orden de centésimas.

Así mismo, en la Figura 22 se presenta un gráfico de superficie de error, en la cual se grafica en 3D el porcentaje de datos a entrenar, las horas de generación de oleaje y el error obtenido, los colores hacia los rojos representan valores altos de error y por lo contrario los colores hacia el azul intenso, representa errores menores o bajos

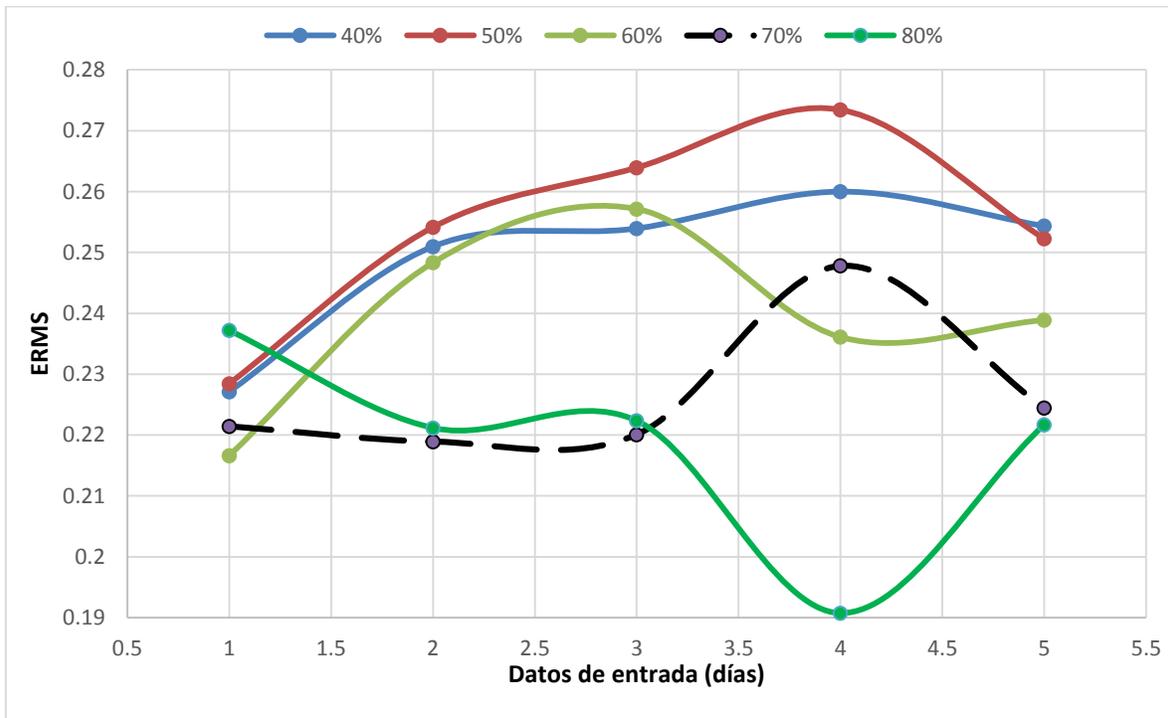


Figura 21. Error rms de los datos del ADCP-4.5m- Sentinel.

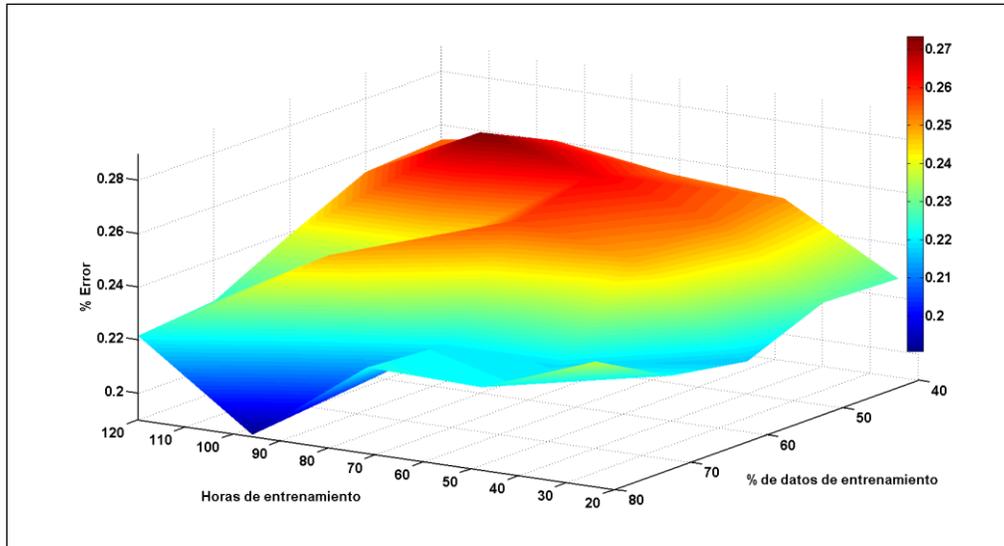


Figura 22. Grafica 3D donde se representan los errores obtenidos con las diferentes combinaciones de entrenamiento, para el punto cerca de la costa.

De acuerdo a los resultados obtenidos para el sitio cercano a la costa (ADCP-4.5m-Sentinel), podemos observar que el mínimo de error (0.19) lo obtenemos usando 80% del total de los datos y con 96 horas de entrenamiento (Figura 21), sin embargo usando 60% del total de datos y con 24 horas de entrenamiento obtuvimos un error de 0.21. Para ser consistentes en el tipo de RNA que usaremos tanto para el punto fuera de la costa así como para el punto cercano a la costa, se decidió que la mejor pareja de entrenamiento de % de datos y hora a entrenar para los dos puntos de interés es usando, 60% de los datos y 24 horas de entrenamiento. Los errores obtenidos para el punto fuera de la costa indican que usando el 60% y 24 horas de entrenamiento, genera resultados aceptables.

Otra razón por la que se realizará el entrenamiento con esa pareja de (60 % y 24 horas) es debido a que usar un 80% del total de datos, solo nos permitiría evaluar el aprendizaje con el 20% de los datos restantes, con lo cual, podríamos incurrir en un sobrentrenamiento.

### *7.3. Validación de la RNA*

El entrenamiento para obtener oleaje en los dos sitios de interés (ADCP-Monitor-10m y ADCP-Sentinel-4.5m) se realizó utilizando el 60% de los datos medidos de oleaje y viento. Posteriormente en la etapa de evaluación la red neuronal se alimentó con el 40% restante de los datos de viento, para obtener alturas de oleaje (línea de color azul) y esos datos posteriormente fueron comparados con el otro 40% restante de los datos medidos de oleaje

(datos esperados o reales). En la Figura 23 a, tenemos los resultados obtenidos para el punto que se encuentra fuera de la costa ADCP-Monitor y la Figura 23 b, corresponde a los resultados obtenidos para el punto que se encuentra cerca de la costa ADCP-Sentinel. Podemos observar a simple vista que los resultados de las RNA son semejantes a los datos de oleaje medidos.

Cabe mencionar y recordar que la series de datos de oleaje medidos, que corresponde al punto fuera de la costa, ADCP-Sentinel va del 25 de Marzo del 2014 al 22 de Mayo del 2014; y para el ADCP-Monitor las mediciones se registraron del 26 de marzo del 2014 al 29 de mayo del 2014.

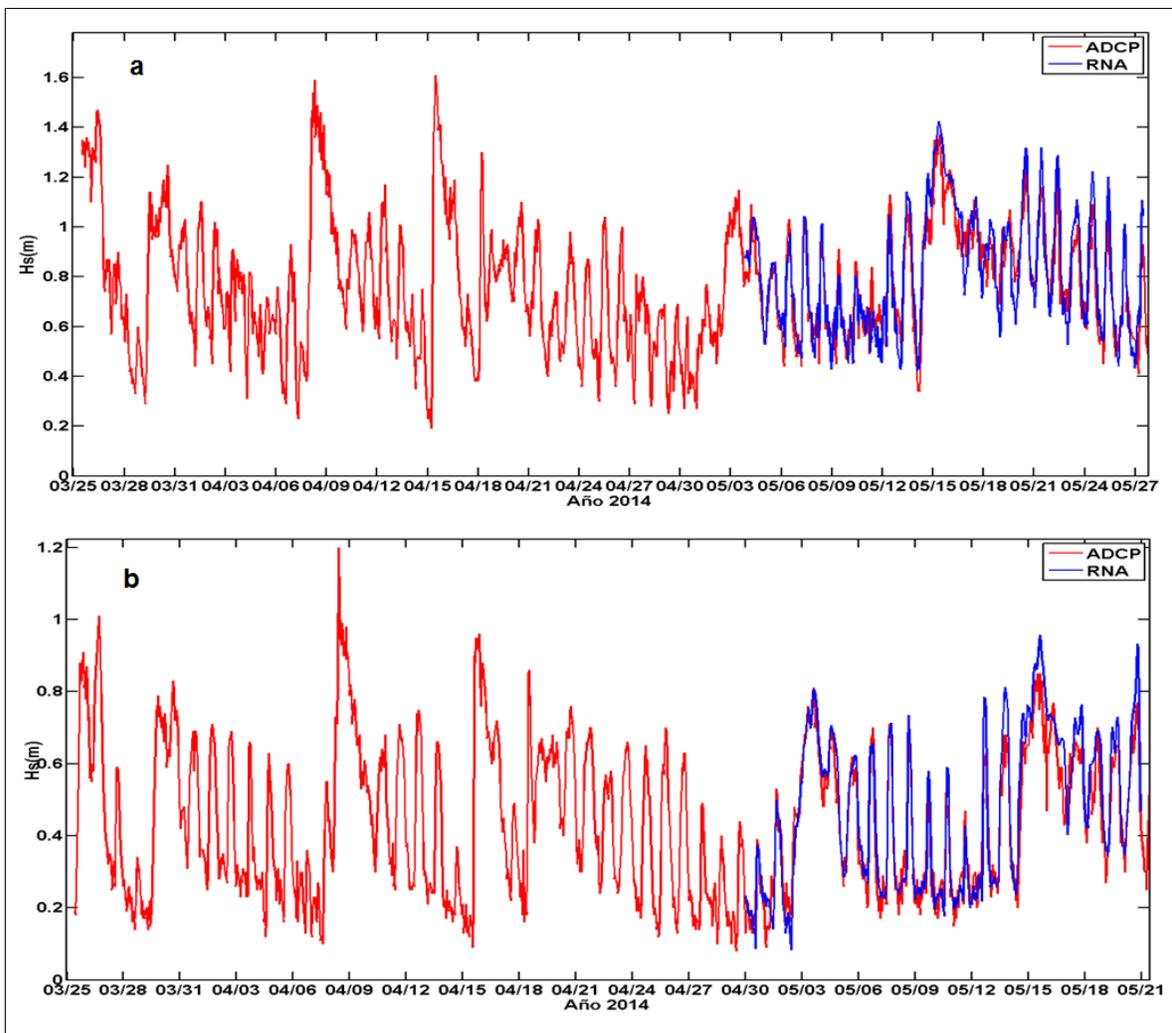


Figura 23. Salidas de los modelos neuronales contra datos medidos a) ADCP-Monitor y b) ADCP-Sentinel.

En la Figura 24, podemos observar dos graficas a) corresponde a los datos del punto que se encuentra cercano a la costa ADCP-Sentinel y b) corresponde a los datos del punto que se encuentra fuera de la costa ADCP-Monitor. Con la finalidad de determinar el grado de semejanza y asociación entre las Hs generadas con el modelo neuronal y las Hs medidas por el ADCP (datos reales). Se calculó el coeficiente de correlación a partir de la distribución de los datos, para los dos sitios donde se colocaron los ADCP's. Las líneas de color azul indican que tanto, el comportamiento de los datos generados por el modelo neuronal, se ajustan a la línea de color negra la cual corresponde a los datos esperados (mediciones del ADCP), si los puntos se encuentran sobre la línea de color negro, significa un buen ajuste y alta correlación. El coeficiente de correlación que se calculó para cada uno de los sitios de interés se encuentra sobre cada una de las gráficas, como se muestra en la Figura 24, el CC para el punto ADCP-Sentinel cercano a la costa es de 0.958, mejor que el que obtuvimos para el punto ADCP-Monitor fuera de la costa, que fue de 0.930. Lo cual indica que el modelo neuronal arroja Hs con mayor precisión en el punto cercano a la costa, debido a que el comportamiento del oleaje en ese sitio, se encuentra mayormente dominado por los vientos locales, más de lo que podría estar el punto fuera de la costa; sin embargo las dos correlaciones son aceptables y altas.

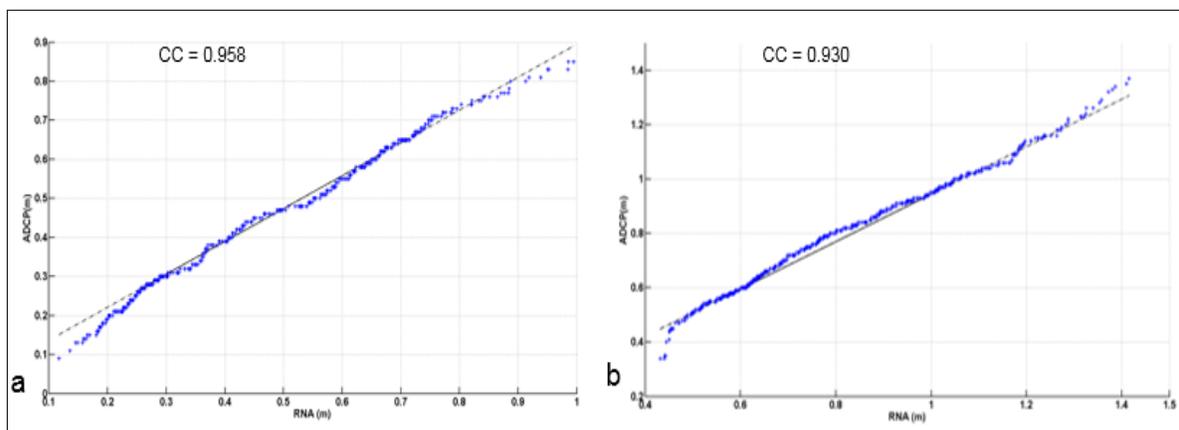


Figura 24. Correlación de la distribución de los datos a) para el punto cercano de la costa ADCP-Sentinel-4.5m y b) para el punto fuera de la costa ADCP-Monitor-10m.

#### *7.4. Selección de las horas de generación de oleaje.*

En esta sección del trabajo se usaron diferentes modelos neuronales y se evaluaron, obteniendo los errores para cada uno de los modelos usados. Los errores que se obtuvieron son el promedio de 24 datos de oleaje generados en cada paso de tiempo (cada una hora), menos los datos esperados (mediciones del ADCP). Se generaron las gráficas de errores para cada uno de los puntos de interés donde se colocaron los instrumentos de medición. Estos modelos neuronales fueron entrenados nuevamente con el 60 % del total de los datos de viento y oleaje, y evaluados con el 40% restante de viento del total de los datos, para obtener Hs.

##### *7.4.1. Fuera de la costa*

En esta sección se minimizó el error generado entre la salida del modelo neuronal – los datos medidos, así mismo se resolvieron los problemas de inestabilidad no lineal, usando algunos métodos estadísticos descritos en la sección 0. En la gráfica de la Figure 25, se muestran los errores obtenidos con cada una de las redes neuronales usadas adicionando los modelos probabilísticos (stepwise y bootstrap). Para discernir cual es el modelo neuronal que genera mejores datos y arquitectura que genera una Hs, vamos a comparar los resultados obtenidos con (RNB + stepwise + bootstrap) con los resultados que se obtuvo usando (RNB + stepwise) y con el método lineal (newlind + stepwise). El modelo neuronal que minimiza mejor el error (RNB + stepwise + bootstrap) con 4N en la capa oculta (línea discontinua de color azul), obteniendo errores por debajo del 0.10m (Figure 25).

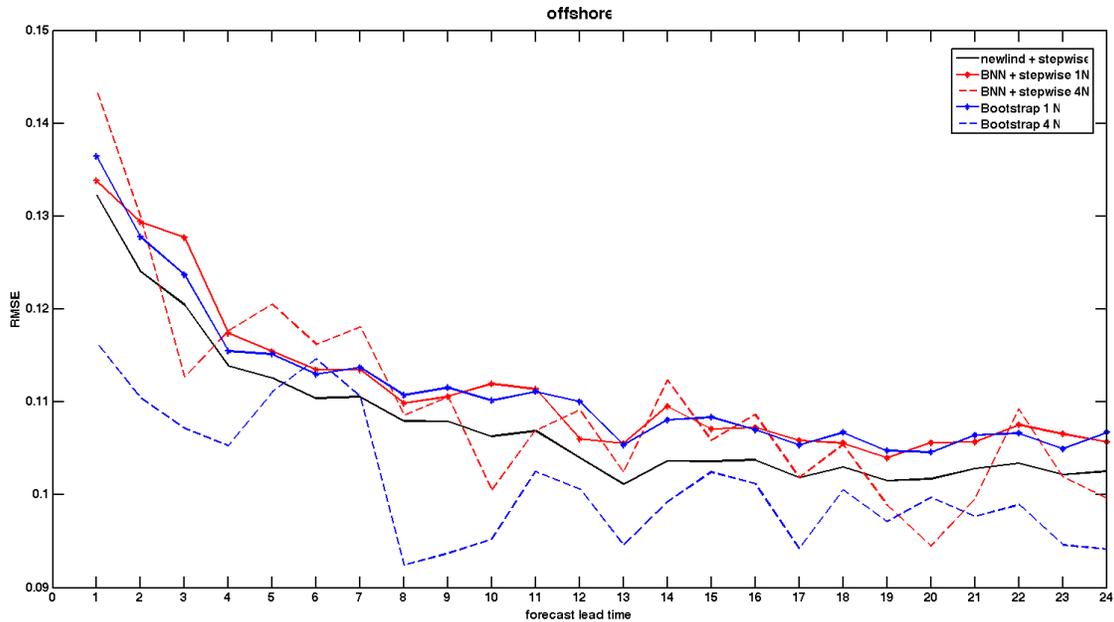


Figure 25. Comparación entre los mejores resultados obtenidos por newlind + stepwise (línea negra), RNB + stepwise usando 1N y 4N en la capa oculta (línea roja) y Bootstrap + RNB + stepwise, con 1N and 4 N (líneas azul). Punto fuera de la costa ADCP-Monitor.

#### 7.4.2. Cerca de la costa

De la misma manera a lo realizado en el punto fuera de la costa ADCP-Monitor lo aplicamos para el punto (ADCP-Sentinel) cercano a la costa; podemos observar que las líneas de error tienen diferente comportamiento. El primer tipo de comportamiento corresponde a las líneas del modelo neuronal lineal (línea continua de color negra, línea con puntos de color rojo y línea con puntos de color azul). Estas tienen la característica de presentar errores altos en las primeras dos horas de predicción, posteriormente disminuye hasta llegar a un mínimo de error alrededor de 0.07m, posterior a las 9 horas de predicción el error incrementa hasta finalizar con un error mayor a 0.075m (Figure 26).

El segundo tipo de comportamiento corresponde a las líneas segmentadas de color rojo y azul, estas presentan errores grandes en las primeras 6 horas de predicción (principalmente la línea de color rojo), posterior a estas horas de predicción el error decrece hasta finalizar con errores menores a 0.075m; los menores errores los obtenemos usando (RNB +stepwise+bootstrap), con 4N en la capa oculta, semejante al punto fuera de la costa (Figure 26).

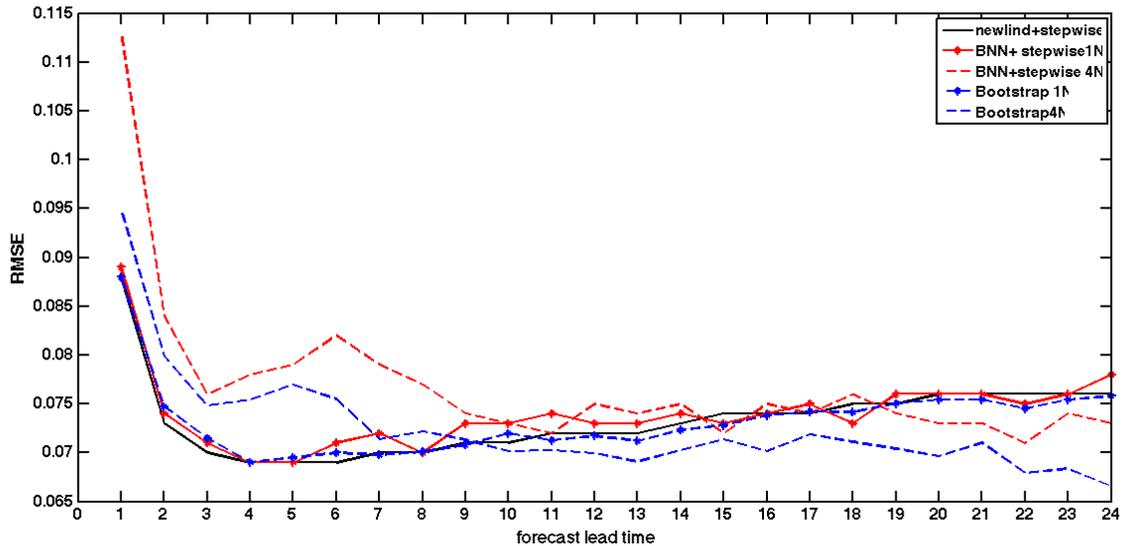


Figure 26. Comparación entre los mejores resultados obtenidos por newlind + stepwise (línea negra), RNB + stepwise usando 1N y 4N en la capa oculta (línea roja) y Bootstrap + RNB + stepwise, con 1N and 4 N (líneas azul). Punto cercano a la costa (ADCP-Sentinel).

Cabe remarcar que el tiempo de computo es considerablemente mayor debido a que al aplicar el Bootstrap aumenta el número de muestras para entrenar (tardando 60 min en generar oleaje), sin embargo, aumenta la precisión en la generación de Hs. Esta combinación de red neuronal con métodos probabilísticos resulta ser la mejor para la implementación de RNA y generar Hs en cada uno de los puntos donde se ubicaron los ADCP's.

### 7.5. Validando la RNA con datos medidos del 2015

En la primera parte de la metodología se comprobó que la RNB, generó buenos resultados entrenando con el 60% del total de los datos de viento y oleaje correspondientes al periodo

de medición del 2014 (26 de Marzo al 29 de Mayo del 2014) y evaluando con el 40% restante. En esta sección, se valida la RNB entrenando con el 100% de los datos medidos de oleaje y viento que corresponden al periodo de medición del 26 de Marzo al 29 de Mayo del 2014; y una vez entrenada la red neuronal, se evaluará con datos de viento medidos con el anemómetro ultrasónico de la misma temporada estacional marzo, abril y mayo del **2015**; generando “Hs” y comparar la salida del modelo neuronal contra datos medidos de las mismas fechas (Tabla 3).

Tabla 3. Representación de la RNA utilizada para validar la metodología.

	Etapas				Salida
	Entrenamiento		Evaluación		
Red Neuronal	Año	Parámetro	Año	Parámetro	
RNB	2014  (Marzo, Abril, Mayo)	Viento: Velocidad, Direc°	2015  (Marzo, Abril, Mayo)	Viento: Velocidad, Direc°	Hs del 2015  (Marzo, Abril y Mayo)
		Oleaje: Hs			

La RNB fue entrenada nuevamente con diferente número de neuronas en la capa oculta para minimizar el error y evitar el sobreajuste, ya que a diferencia de las RNB entrenadas anteriormente, ahora se están utilizando más datos para entrenar, lo cual hace que los pesos varíen y se tenga que encontrar un nuevo ajuste para este ejercicio, si observamos en la Tabla 4, ahora con 6N en la capa oculta se obtiene el error mínimo.

Tabla 4. Tabla con diferentes errores obtenidos con diferente número de neuronas.

RNB	% Error
2N	0.138
4N	0.136
<b>6N</b>	<b>0.123</b>
8N	0.130
10N	0.134

La salida del modelo neuronal obtenida por la RNB con 6N en la capa oculta fue comparada contra datos medidos por el ADCP-Monitor que se reinstaló en el mismo sitio (Figura 11), para continuar con las mediciones durante el año 2015; de esa manera se pudo cuantificar

el error entre la salida del modelo – los datos medidos. En la gráfica de la Figura 27, la línea de color rojo corresponde a los datos medidos por el ADCP y de color azul corresponde a la salida generada por el modelo neuronal, podemos observar que hay un periodo que va del 2 de mayo del 2015 al 10 de mayo del 2015, en el cual no hay oleaje generado por la RNA, debido a que en esas fechas no hubo mediciones de viento, por fallas técnicas con el anemómetro.

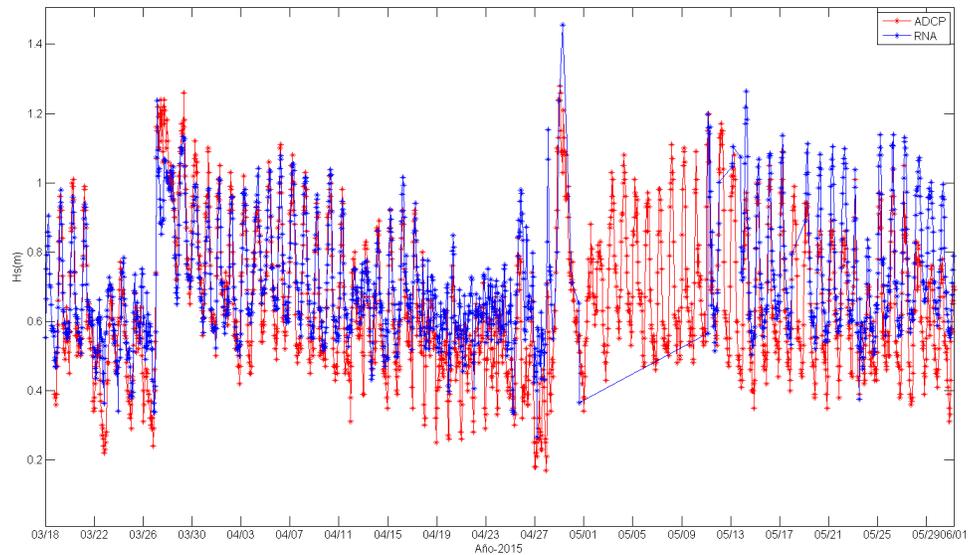


Figura 27. Comparación entre el oleaje medido ADCP y la salida del modelo neuronal para el año 2015.

En la Figura 28, se muestra la correlación correspondiente a la distribución de los datos medidos (ADCP) de los meses de marzo, abril y mayo del 2015, contra los datos generados por el modelo neuronal a partir de los datos de viento del 2015. Las Hs generadas con el modelo neuronal (línea discontinua de color azul) presentan un buen ajuste a la línea de los datos esperados obteniendo un coeficiente de correlación aceptable (CC=0.85).

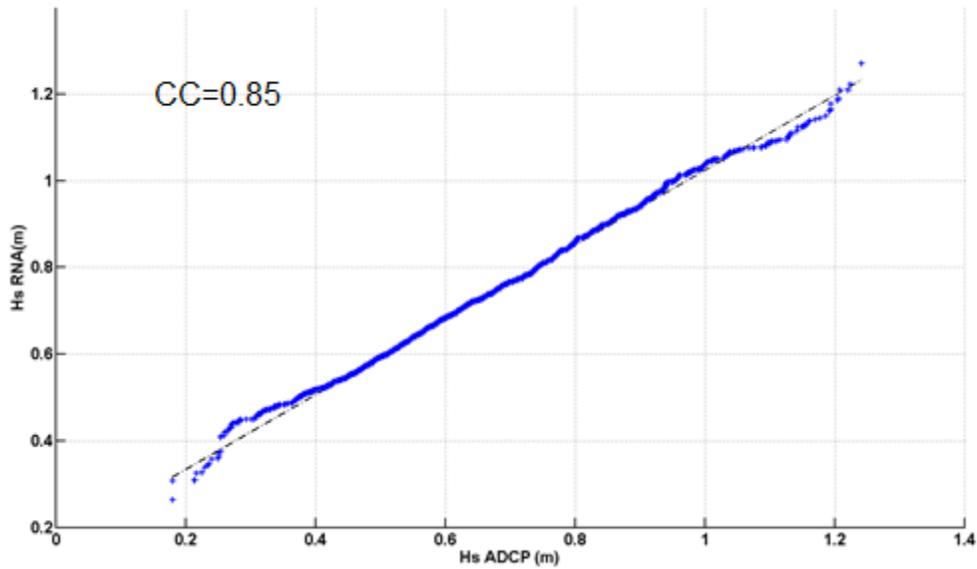


Figura 28. Comparación de la distribución de los dos series de datos de Hs del ADCP y de la salida del modelo neuronal.

En la Figura 29, se presentan los datos de los equipos ADCP (Hs) y magnitud de viento, se puede observar que el oleaje está directamente relacionado con la intensidad y dirección del viento, lo que hace que la implementación y entrenamiento de la red neuronal sea encaminada a encontrar esta relación. Ya que en las partes donde se presentan máximos de altura de ola en los dos sitios, coinciden con los máximos de viento, de igual manera cuando hay una disminución en la velocidad de viento la altura de ola disminuye proporcionalmente en aproximadamente de 1 a 10.

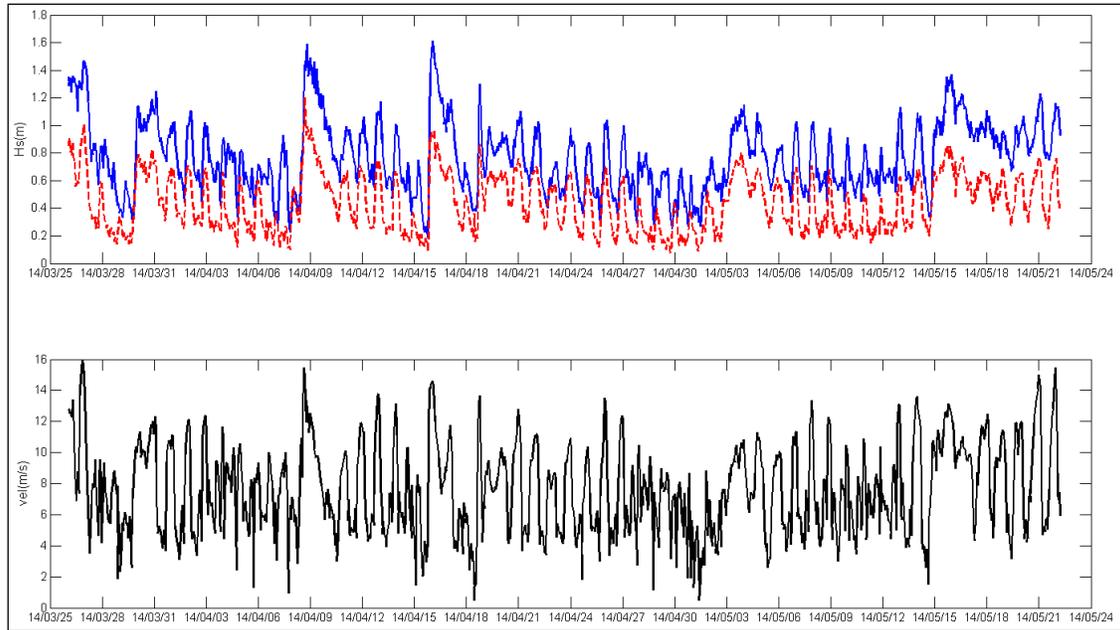


Figura 29. Primera grafica corresponde a la Hs registradas por los ADCP (línea continua-Monitor y línea segmentada-Sentinel). Segunda grafica corresponde a velocidades de viento registradas por el anemómetro para el mismo periodo de tiempo.

### *7.6.Comparando modelos neuronales (RNA) contra modelos numéricos (MT)*

Con la finalidad de evaluar el desempeño del modelo neuronal contra el de un MT, se presentan los resultados del MIKE21 obtenidos del trabajo realizado por Medellín (2015), donde realiza un comparación de datos medidos in-situ (ADCP) contra la salida del modelo MIKE21, para una localidad cercana a Sisal en la costa de Yucatán (Nodo 12972) donde presentan condiciones de oleaje muy semejantes al sitio de estudio de este trabajo. La profundidad donde se colocó el ADCP es de 8.5m y está localizado entre las localidades de Yucalpeten y Chuburná (Figura 30). Los resultados obtenidos de la correlación de Hs medidos (ADCP) y los generados por el modelo MIKE21 para  $H_s \leq 1.2m$  muestra un buen ajuste para valores de Hs entre 1.2m y 1.7m subestima las Hs y para valores de  $H_s \geq 1.7m$  el modelo sobrestima la Hs, toda las comparaciones se realizaron con respecto a los datos medidos por el ADCP.

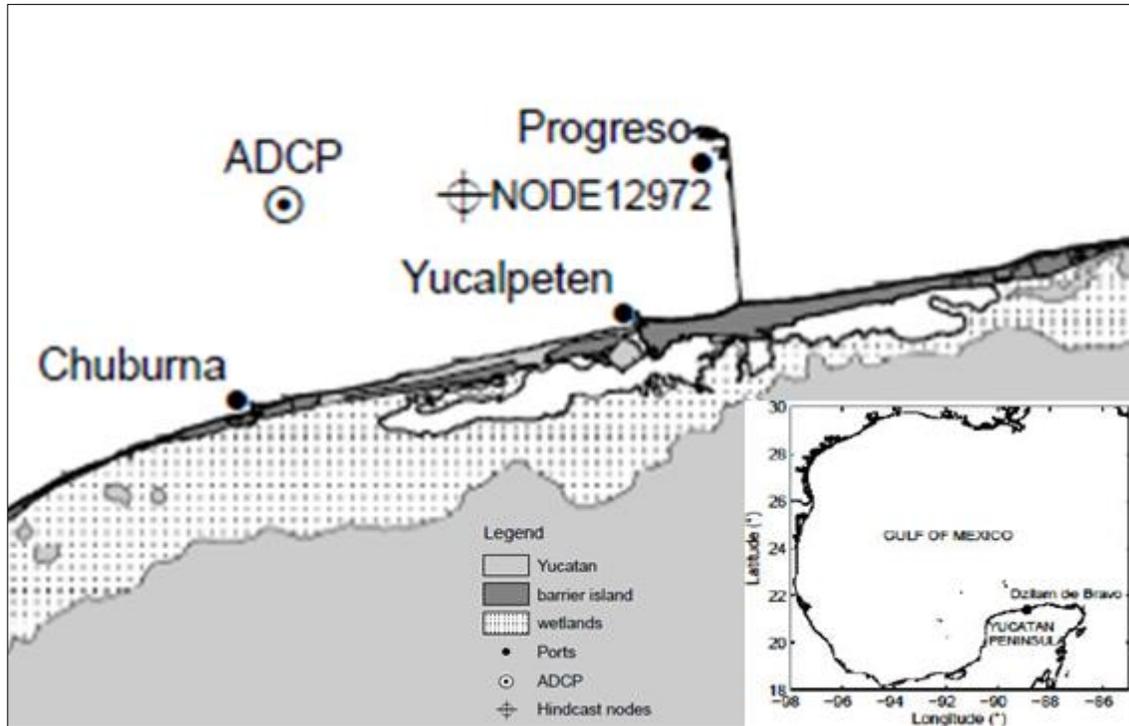


Figura 30. Localización del ADCP entre Yucalpeten y Chuburná, así como el nodo 12972 de la malla del MIKE21. Figura tomada del trabajo de Medellín (2015).

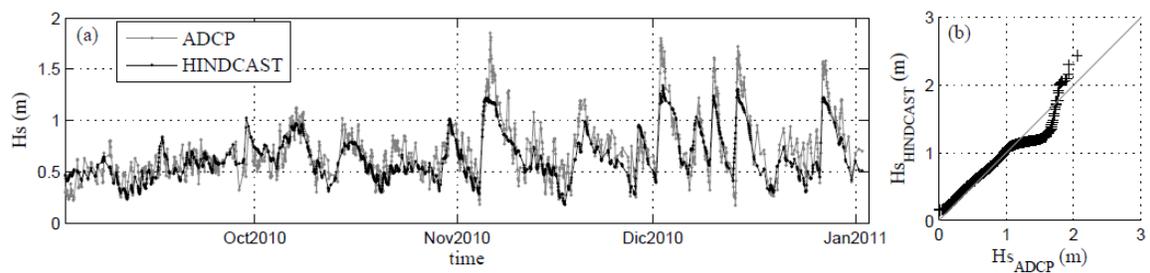


Figura 31. Comparación de las  $H_s$  medidas (ADCP) y la salida del MIKE21.

A simple vista, se puede decir que los resultados obtenidos por Medellín (2015) y los generados por este trabajo, presentan un buen ajuste a los datos medidos, por lo tanto la implementación de las redes neuronales para generación de oleaje a partir de datos de viento es equiparable a los resultados generados con un modelo numérico.

### 7.6.1. Comparación con el SWAN

Se compararon con los resultados del trabajo de Lira (2015); donde realizó una predicción de oleaje forzando un modelo numérico (SWAN – Simulating WAVes Nearshore) usando datos de pronóstico GFS (Global Forecast System) para el mismo punto donde se colocó el ADCP-Monitor-10m en el año 2014. En el panel superior de la Figura 32, tenemos el oleaje generado con el modelo neuronal y en el panel inferior de la Figura 32, tenemos lo generado por el modelo numérico SWAN. Podemos observar que las dos líneas de color azul que corresponde a los modelos se ajustan bien a la señal medida (líneas de color rojo), para cuantificar que tanto se parecen a los datos medidos para ese periodo de medición (marzo, abril y mayo del 2014), se calculó el coeficiente de correlación.

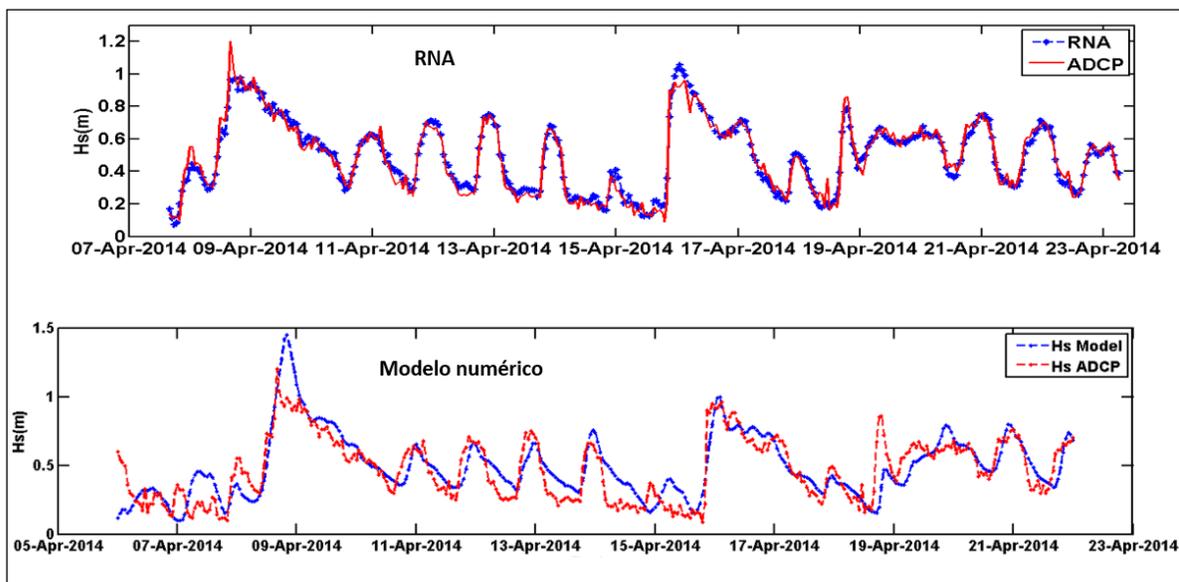


Figura 32. Comparación de la RNA y un modelo numérico tomado del trabajo de Lira, 2015.

Como se puede observar tanto el modelo numérico como la RNA, arrojan Hs muy similares a las registradas por el ADCP, por lo cual genera buenos resultados, el error calculado también es aceptable y la R muestra una buena correlación (Tabla 5) lo cual quiere decir que las dos metodologías son válidas. Sin embargo existe una gran diferencia entre estos dos métodos; 1) si no se requieren los datos de oleaje en todo el dominio computacional, la RNA es una buena opción ya que el tiempo de computo es significativamente más corto 2) La RNA solo da valores de oleaje en un punto en específico.

Tabla 5. Tabla de comparación RNA y modelo numérico.

Metodología	Error(RMS)	R
RNB 4N	0.14	0.89
Modelo Numérico	0.17	0.81

### *7.7. Casos de Aplicación RNA*

En esta parte del trabajo una vez que la red neuronal está entrenada se decidió realizar algunos casos de aplicación. La RNA, se alimenta con nuevos datos de viento y la salida de la RNA se compara con los datos del ADCP para su validación. Los casos de aplicación con los que se va a trabajar: a) propagación de oleaje y b) generación de un Hindcast de oleaje usando las RNA. Al inicio de cada apartado de caso de aplicación se explicará con que datos y de qué manera, fueron entrenadas y validadas los tipos de RNA utilizadas.

#### *7.7.1. Propagación del oleaje*

Si bien, sabemos que la batimetría es un factor importante en la propagación del oleaje, ya que la ola al sentir fondo se va modificando tanto la dirección como la altura del oleaje. En este trabajo no consideramos la batimetría del área de estudio de manera directa. El oleaje sufre ciertas atenuaciones y la dirección del oleaje se ve modificada conforme se va aproximando a la costa, debido a la interacción con el fondo marino. El ángulo de incidencia es de suma importancia ya que puede ocasionar desbalance de sedimento acrecentando o erosionando las playas; así como modificando el transporte de material sedimentario dentro de una celda litoral.

En esta sección se entrenaron diferentes modelos neuronales para que aprendan sobre los procesos de transformación de oleaje, refracción y asomeramiento. Los datos de oleaje y viento que se usaron corresponden al periodo de medición del 26 de Marzo al 22 de Mayo del 2014, con la finalidad de hacer coincidir los datos para el ADCP-Monitor, ADCP-Sentinel

y los datos de viento del anemómetro. Para realizar el entrenamiento se utilizó nuevamente el 60% del total de datos de oleaje y para la etapa de evaluación se usó el 40% restante de los datos de viento para generar oleaje. En esta aplicación de la metodología vamos a obtener el oleaje correspondiente a sus parámetros de Hs, Tp y Dir°, de tal manera que tendríamos 3 redes neuronales con una salida cada una de las redes.

Los modelos neuronales que fueron usados son: RNL (newlind + stepwise), RNB + stepwise con 2N, 4N y 8N, en la capa intermedia para este tipo de red neuronal bayesiana. Para estos diferentes modelos se calculó el RMSE entre la salida de cada uno de los modelos menos los datos esperados (mediciones del ADCP-Sentinel), este es el promedio de 24 Hs generadas a partir de 24 datos de viento cada hora o paso de tiempo de cálculo.

Un ejemplo de la manera en que se entrenaron, así como los parámetros y porcentajes de datos que se usaron para entrenar los modelos neuronales se puede observar en la Tabla 6. Se usó el 60% de los parámetros oleaje fuera del punto fuera de la costa (Hs, Tp, Dir°); el 60% de los datos de viento (Velocidad y dirección) y 60% del parámetro de oleaje a generar en el punto cercano a la costa (Hs, Tp o Dir°). Posteriormente en la etapa de evaluación se usó el 40 % restante de los datos de oleaje fuera de la costa (Hs, Tp, Dir°) y viento (Velocidad y dirección), para generar oleaje en el punto cercano a la costa correspondiente al 40 % restante del total de la serie de datos con la que se trabaja.

Tabla 6. Arreglo de los datos para entrenar y evaluar las RNA, en la propagación de oleaje (del 26 de Marzo al 22 de Mayo, del 2014).

RNL	Etapas				Predicción	
	Entrenamiento		Evaluación			Salida
	Porcentaje	Parámetro	Porcentaje	Parámetro		
1	60%	Oleaje-Monitor: Hs, Direc°, Tp	40%	Viento: Velocidad, Direc°; Oleaje-Monitor: Hs, Direc°, Tp	40 % restante de Hs	
	60%	Viento: Velocidad, Direc°				
	60%	Oleaje-Sentinel: Hs				
2	60%	Oleaje-Monitor: Direc°, Hs, Tp	40%	Viento: Velocidad, Direc°; Oleaje-Monitor: Hs, Direc°, Tp	40 % restante de Direc°	
	60%	Viento: Velocidad, Direc°				
	60%	Oleaje-Sentinel: Direc°				
3	60%	Oleaje-Monitor: Hs, Tp, Direc°	40%	Viento: Velocidad, Direc°; Oleaje-Monitor: Hs, Direc°, Tp	40 % restante de Tp	
	60%	Viento: Velocidad, Direc°				
	60%	Oleaje-Sentinel: Tp				

Como se muestra en la Figura 33 la RNL (newlind + stepwise) genera mejores resultados que la RNB. El error medio cuadrático se calculó usando la salida generada por los modelos neuronales menos los datos esperados (oleaje en el ADCP Sentinel); de acuerdo a lo descrito anteriormente Capítulo 7.1.

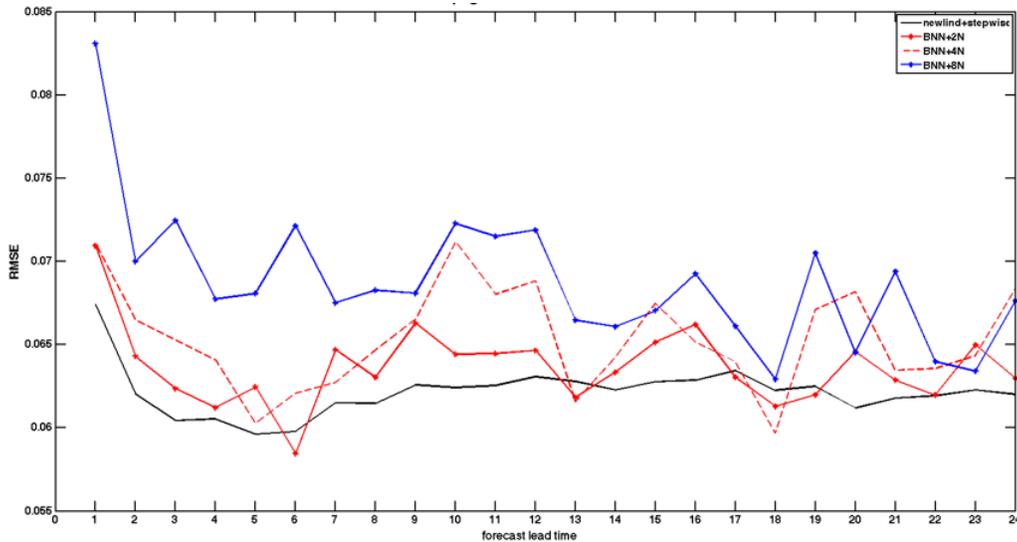


Figura 33. Comparación entre los mejores resultados obtenidos por newlind + stepwise (línea negra), RNB + stepwise + Bootstrap usando 2N, 4N y 8N en la capa oculta (líneas rojas y azul). Propagación del oleaje al punto cercano a la costa (ADCP-Sentinel).

Al inicio mencionamos que en estos modelo neuronal la batimetría no se considera de manera directa, esto significa que durante el entrenamiento no se ingresan valores de batimetría; sin embargo en este ejercicio se puede apreciar la transformación del oleaje de un punto fuera de la costa (ADCP-10m) a un punto cercano a la costa (ADCP-4m), por lo cual los modelos neuronales entrenados consideran de manera implícita el efecto del fondo y de manera indirecta la batimetría del área de estudio, ya que ese es el factor principal del cambio de dirección y altura del oleaje lo cual el modelo neuronal resuelve correctamente.

Una vez que determinamos que tipo de RNL + Stepwise es la mejor tipo de red neuronal para obtener buenos resultados en este caso de aplicación propagando el oleaje, debido que, los errores obtenidos son menores que los errores obtenidos con los otros tipos de redes neuronales, se elige para usarse y obtener los otros parámetros de oleaje (dirección-Dir° y periodo pico-Tp), que a continuación se presentan.

#### 7.7.1.1. Dirección de oleaje propagado

La dirección del oleaje propagado para el punto cercano a la costa (ADCP-Sentinel) la obtuvimos entrenando una red neuronal con los parametros de oleaje del punto fuera de la costa, y las componentes de dirección de viento; asociado a las componentes de dirección de oleaje del punto cercano a la costa (ADCP-Sentinel), posteriormente en la etapa de

evaluación se usan como datos de entrada al modelo neuronal los parámetros de oleaje y los datos de viento del 40% restante del total de los datos del periodo de medición.

En la Figura 34 **a** podemos observar que la componente de dirección de oleaje predominante es del noreste y la dirección de oleaje dominante es de la dirección norte. Con respecto a la rosa de dirección de oleaje generada con el modelo neuronal en la Figura 34 **b**, se observa mayor variabilidad en la dirección de oleaje, el comportamiento es similar ya que el oleaje conserva la dirección predominante del noreste y que las alturas de oleaje más altas las tenemos de la dirección norte.

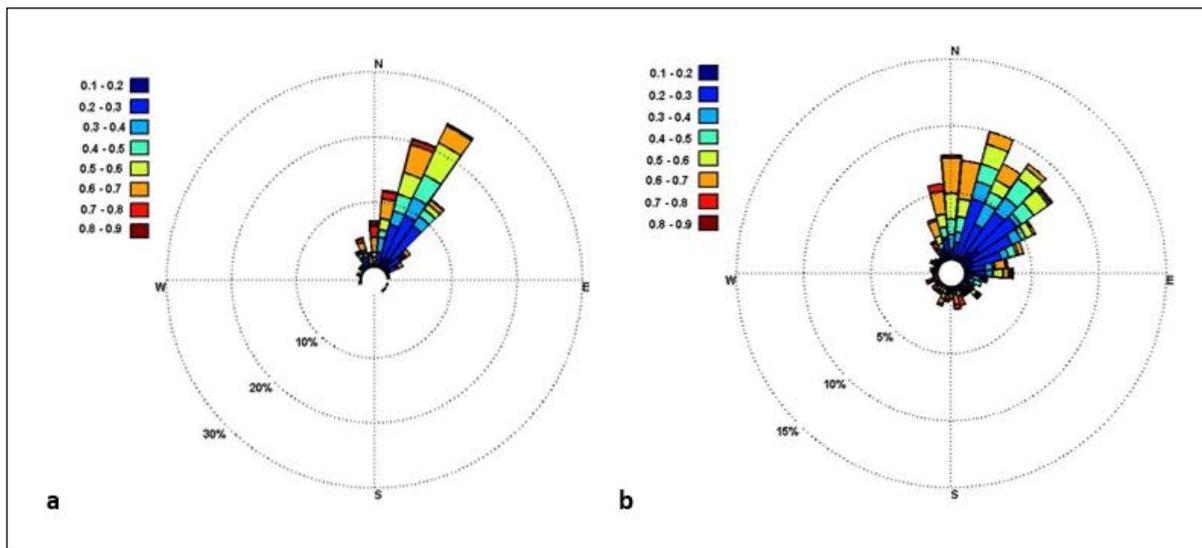


Figura 34. Comparación de las rosas de dirección de oleaje de los datos medidos (a) y la dirección pronosticada por el modelo neuronal (b).

#### 7.7.1.2. *Periodo pico ( $T_p$ ) del oleaje propagado*

Otro parámetro que fue generado con el modelo neuronal es el periodo pico ( $T_p$ ), el cual lo obtuvimos entrenando una red neuronal con los parámetros de oleaje del punto fuera de la costa, y las componentes de dirección de viento que hemos estado utilizando en los casos anteriores; asociado a los periodos de oleaje registrados con el ADCP-Sentinel en el punto cercano a la costa.

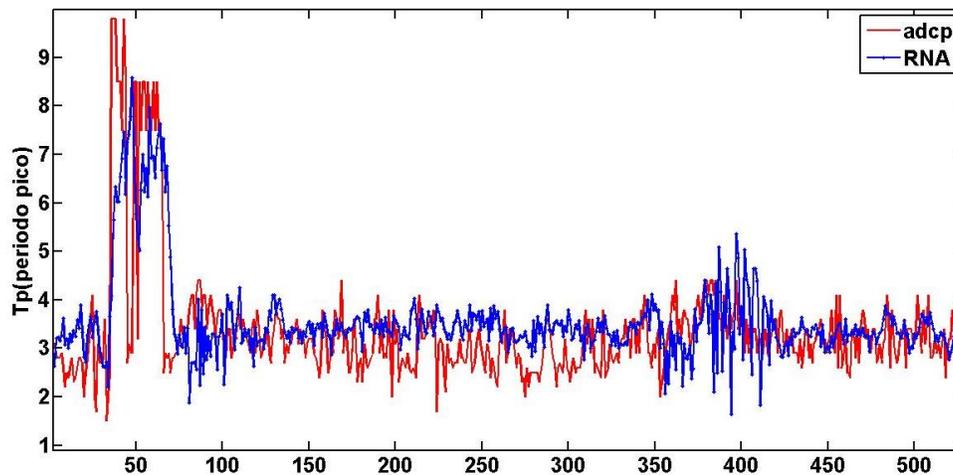


Figura 35. Comparación del pronóstico de periodo pico ( $T_p$ ) generado con el modelo neuronal.

Los resultados de  $T_p$  generados con el modelo neuronal tienen un comportamiento muy parecido a los datos medidos por el ADCP, y esto lo podemos observar en el coeficiente de correlación que se obtuvo relacionando los  $T_p$  obtenidos con el modelo neuronal contra los  $T_p$  registrados con el ADCP (Figura 36). El coeficiente de correlación del modelo neuronal es de 0.62, se encuentra por arriba de la correlación que obtuvo Lira (2014) de 0.46, a partir de la salida de un modelo numérico contra datos medidos.

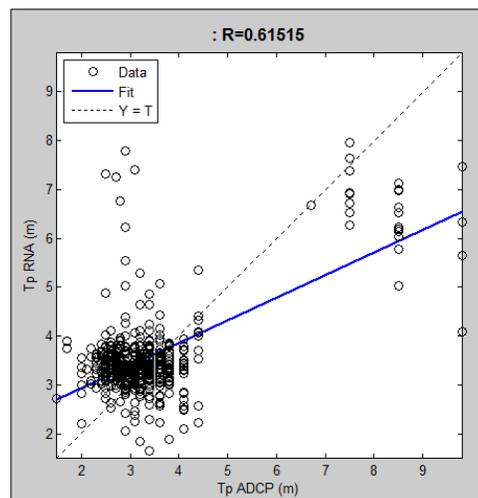


Figura 36. Correlación del  $T_p$  entre los datos medidos y la salida del RNA.

Para este caso de aplicación donde el proceso es la transformación del oleaje del punto fuera de la costa al punto cercano a la costa (ADCP-Sentinel), depende directamente de las condiciones de oleaje que se presentan en el punto fuera de la costa y la configuración topográfica, es suficiente usar un modelo neuronal lineal (RNL). Con los resultados obtenidos anteriormente implementando la RNL, nos permite hacer uso de las RNA para representar los procesos de transformación de oleaje, provocado por el cambio de la batimetría. Así mismo, se observa una relación en la transformación del oleaje de aguas intermedias a aguas someras debido a los procesos de transformación del oleaje por efectos del fondo, esta relación también puede ser relativamente bien ajustada por la red neuronal.

### *7.7.2. Aplicación de RNA para generar un Hindcast de oleaje*

En esta sección, se plantea una aplicación de la metodología usada en la cual se utilizó la red neuronal usando una base de datos histórica (hindcast) de oleaje para el área de estudio de la costa de Sisal, Yucatán. Los datos de oleaje que se usaran para entrenar la red neuronal fueron generados a partir de campos de viento del North American Reanalysis (NARR) utilizando el modelo MIKE 21 SW. La base de datos tiene una cobertura de enero 1979 a las 00:00 horas hasta el 31 de diciembre de 2008 a las 21:00 horas. Los datos que se presentan del hindcast hecho con el MT corresponden a los parámetros de oleaje: altura significativa (Hs), periodo pico (Tp), periodo medio (Tm); así como dirección y velocidad de viento del NARR interpolados en la malla de cálculo del modelo de oleaje (Appendini et al., 2013) los datos que tenemos son cada 3 horas y para fines de nuestro objetivo solo trabajaremos con Hs.

Los datos que se utilizaron para esta aplicación de la RNA, corresponden al nodo 12918 de la malla computacional generada con el MIKE21 por (Appendini et al., 2013) el cual se localiza a 14 km frente a la costa de Sisal Yucatán con las coordenadas latitud norte 21.286329° y longitud -90.070053°; a 3 km de distancia del punto donde se colocó el ADCP-10m-Monitor (Figura 37).

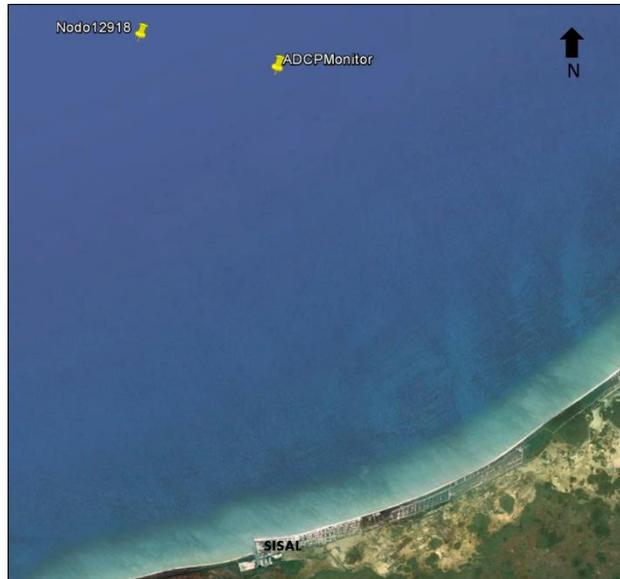


Figura 37. Localización del Nodo 12918 del hindcast y el ADCP fuera de la costa de Sisal, Yucatán.

La forma en que se trabajó con los 30 años del Hindcast es de la siguiente manera, cubrir toda la época de Nortes que de acuerdo con algunos autores (López-Méndez, 2009; Appendini et al., 2013) corren de septiembre a abril; por tal motivo en este trabajo un año estacional se contara de junio a junio (es decir, de Junio de 1979 a Junio de 1980) , con la finalidad de incluir la mayor cantidad de eventos que se pueden presentar en un año estacional; por consiguiente los 30 años históricos se reducirán a 29 para fines de la implementación de las RNA.

De acuerdo a los implementado y los casos de aplicación de las RNA se decidió usar para este caso solo la RNB + Stepwise y variar el número de neuronas en la capa intermedia con la finalidad de encontrar la configuración ideal para minimizar el error y evitar el sobreajuste.

A partir de la serie de datos históricos de oleaje generados con el MIKE 21 y a partir de los datos del NAR, se usaron diferentes años para entrenar el modelo neuronal.

- Primer caso que se realizó fue: entrenar con dos años (2005 al 2007) de datos de oleaje (MIKE21) y viento (NAR); para posteriormente en la etapa de evaluación usar los 27 años restantes de datos de viento del NAR para generar oleaje .

- Segundo caso que se realizó fue: entrenar con un año (2005 - 2006) de datos de oleaje (MIKE21) y viento (NAR); para posteriormente en la etapa de evaluación usar los 28 años restantes de datos de viento del NAR para generar oleaje.
- Tercer caso que se realizó fue: entrenar con un año (1992 - 1993) de datos de oleaje (MIKE21) y viento (NAR); para posteriormente en la etapa de evaluación usar los 28 años restantes de datos de viento del NAR para generar oleaje.

### 7.7.2.1. Momentos estadísticos

Con la finalidad de evaluar la RNA de manera cualitativa, se calculan los momentos estadísticos de las alturas de ola significativa  $H_s$  generadas con el modelo neuronal y se comparan con los resultados del Mike21

El primer momento que se calculó fue la media anual la cual corresponde al promedio aritmético representado con la siguiente ecuación:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad \text{Ec. 7-1}$$

Donde  $X_i$  es el dato individual de la muestra por cada año.

$N$  es el total de datos de la muestra.

El segundo momento estadístico corresponde a la varianza la cual está dada por la siguiente ecuación (este no es presentado en la tabla de resultados directamente, solo se presentará la desviación estándar):

$$S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2 \quad \text{Ec. 7-2}$$

El tercer momento de la muestra corresponde a la desviación estándar el cual corresponde a la raíz cuadrada de la varianza:

$$S_x = \sqrt{S_x^2} \quad \text{Ec. 7-3}$$

Por otro lado también se calcula la  $H_{1/3}$  de cada año, calculando la media aritmética del tercio de olas más altas del conjunto de  $N$  olas ( $H_s$ ) de los hindcast.

$$H_{1/3} = \frac{3}{N} \sum_{i=1}^{N/3} H_s \quad \text{Ec. 7-4}$$

Donde  $H_s$ , es la serie en este caso de los promedios de altura de ola significativa obtenidos por las series de datos de los hindcast, ordenadas de mayor a menor ( $H_i$  es la altura de ola máxima y  $H_N$  es la altura de ola mínima) y  $N$  es el número de olas individuales del registro.

Los momentos estadísticos anteriormente descritos fueron calculados tanto para los datos del hindcast ( $H_s$  del MIKE21), así como para los datos de  $H_s$  generados con los modelos neuronales.

#### *7.7.2.2. Primer caso hindcast*

Para generar el Hindcast con RNA, en un principio se decidió usar dos años (Junio del 2005 a Junio del 2007) debido a que, durante el 2005 y 2007 los registros de oleaje corresponden a años anómalos donde los eventos de ciclones dominaron el clima en la región del Golfo y Mar Caribe (Appendini et al., 2013). Una vez entrenada la red neuronal con estos datos, viento y oleaje, se generara un Hindcast con datos de viento de los años restantes y se evaluara el desempeño de la red neuronal, comparando los resultados con los obtenidos con el modelo numérico.

El tipo de red neuronal que se utiliza para esta aplicación es la tipo bayesiana RNB con diferente número de neuronas en la capa intermedia 2N, 4N, 8N y 10N, para encontrar el mínimo error y evitar una sobrestimación. En todos estos casos se calculó el error del entrenamiento, error en la parte de evaluación, coeficiente de correlación y tiempo de cómputo. Finalmente se eligió el caso que se ajusta mejor a los resultados del MIKE21 (Tabla 7). Una vez definida la mejor arquitectura de la red neuronal, con la cual obtenemos un error mínimo y no se tiene sobreajuste, se decidió usar los resultados ( $H_s$ ) generados por el modelo neuronal, para calcular los estadísticos de oleaje (sección 7.7.2.5).

Tabla 7. Errores obtenidos con diferente número de neuronas en la capa oculta, RNB entrenada con 2 años.

Red Neuronal	Entrenamiento	Evaluación	CC	Tiempo (seg)
RNB 2N	0.182	0.188	0.88	34.941
RNB4N	0.169	0.174	0.898	87.744
<b>RNB 6N</b>	<b>0.157</b>	<b>0.169</b>	<b>0.905</b>	<b>197.099</b>
RNB 8N	0.155	0.172	0.901	243.673
RNB 10N	0.151	0.172	0.901	227.213

Como se puede observar en la Tabla 7, usando la RNB con 6 neuronas en la capa oculta logramos obtener buenos resultados, el oleaje generado con el modelo neuronal se ajusta más al oleaje generado con el modelo MIKE21. En la Figura 38 tenemos el resultado de la evaluación usando el mayor porcentaje de los datos de viento del NARR (dejando fuera el año 2005 y 2006, con los que se entrenó el modelo neuronal) de color rojo las alturas de ola generados con el modelo MIKE21 y de color azul las alturas de ola generadas con la RNA. Podemos observar que el comportamiento de la línea azul es muy parecido al de la línea de color rojo, sin embargo la señal de color azul no presenta mínimos de altura de ola tan bajos como lo hace el MIKE21, y en algunos picos de altura de ola máxima, el modelo neuronal no los alcanza y en otras ocasiones la altura de ola obtenida con el modelo neuronal sobrestima la Hs.

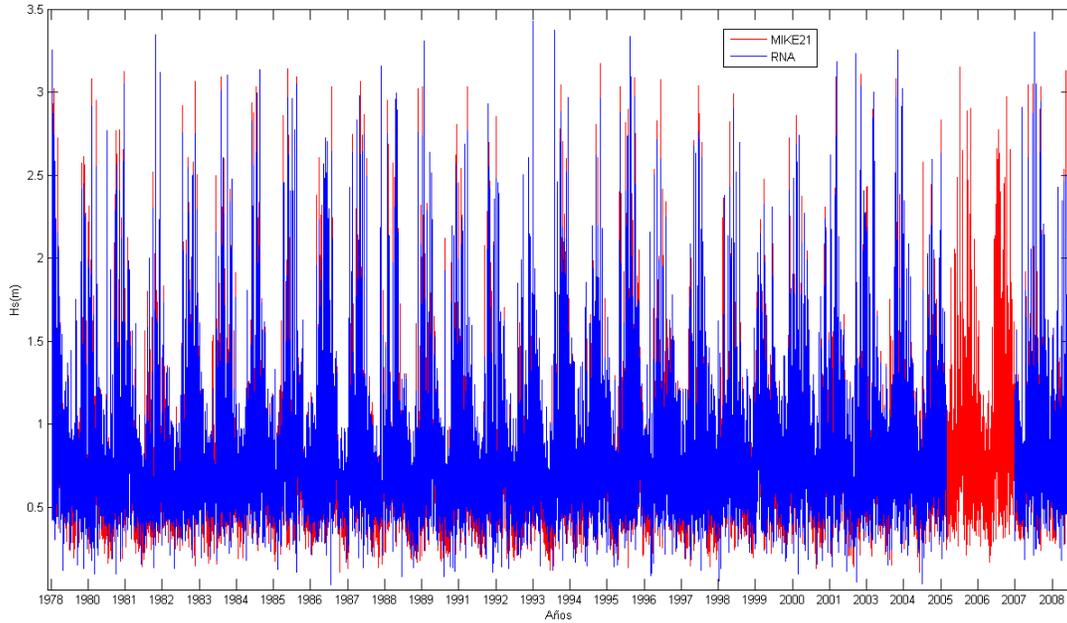


Figura 38. Resultado obtenido de la etapa de evaluación usando la RNB con 10N en la capa intermedia.

Para discernir que  $H_s$  del modelo neuronal se aproxima más a la  $H_s$  del Mike21, se decidió utilizar un gráfico cuantil-cuantil (función “**qqplot**” – Matlab), donde se calcula la correlación que existe entre las distribuciones de las dos series de datos de oleaje (Figura 39). La línea recta segmentada indica un ajuste perfecto entre las dos señales de oleaje que se están analizando, los puntos discretos de color azul corresponden al comportamiento de la señal de oleaje generado con el modelo neuronal, de manera general se observa un buen ajuste con una  $CC=0.905$ , sin embargo, se observa que para  $H_s > 1.5m$  las alturas son menores a las  $H_s$  que fueron generadas por el MIKE21.

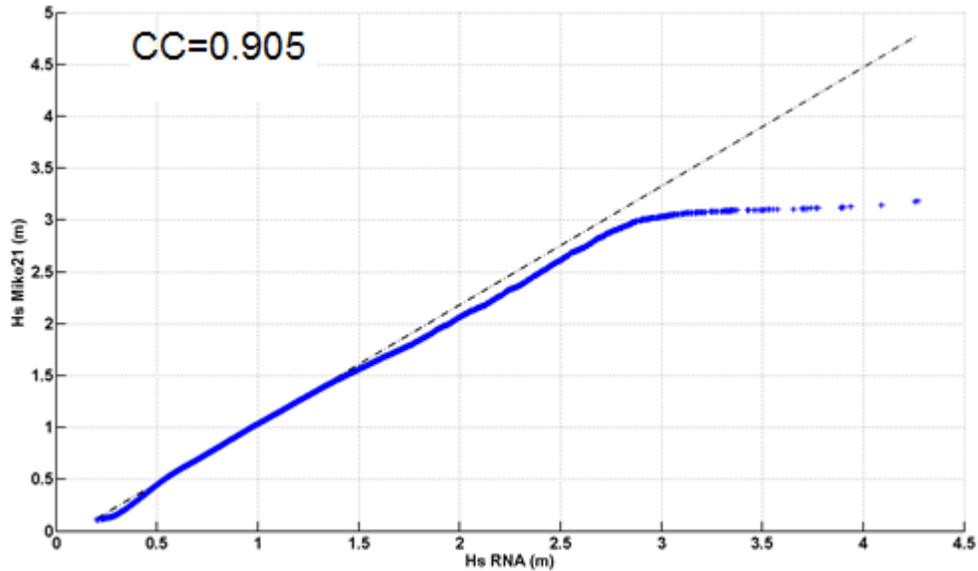


Figura 39. Comparación de la distribución de los dos series de datos de Hs, entrenando con 2 años la RNB.

### 7.7.2.3. Estadísticos primer caso de aplicación

Se calcularon los momentos estadísticos de los datos de oleaje de acuerdo a lo descrito en el apartado 7.7.2.1, para comparar la semejanza entre los datos del MT y los datos obtenidos con el modelo neuronal. En la Tabla 8, se presentan los estadísticos anuales calculados, con la serie de datos del MIKE21 y la serie de datos generada por la RNA. En las columnas que corresponden a los momentos estadísticos calculados con los datos de la RNA, en el renglón correspondiente al año 28 tiene la palabra “Train” lo cual significa que ese año se usó para entrenar el modelo neuronal y por consiguiente no hay momento estadístico evaluado con la RNA.

Tabla 8. Tabla de comparación de estadísticos anuales, entrenando con dos años de datos.

No. Año	Fechas	Media Mike	Media RNA	desv.std Mike21	desv.std RNA	Hs anual Mike	Hs anual RNA
1	79-80	0.768	0.772	0.38	0.333	1.159	1.113
2	80-81	0.794	0.816	0.419	0.372	1.239	1.213
3	81-82	0.723	0.72	0.347	0.289	1.096	1.015
4	82-83	0.79	0.795	0.389	0.357	1.21	1.168

5	83-84	0.768	0.77	0.422	0.349	1.198	1.115
6	84-85	0.78	0.779	0.383	0.332	1.161	1.102
7	85-86	0.809	0.819	0.433	0.393	1.246	1.216
8	86-87	0.783	0.807	0.379	0.357	1.186	1.178
9	87-88	0.796	0.808	0.424	0.377	1.254	1.198
10	88-89	0.759	0.799	0.4	0.374	1.161	1.164
11	89-90	0.742	0.742	0.401	0.341	1.153	1.079
12	90-91	0.732	0.757	0.384	0.359	1.128	1.12
13	91-92	0.745	0.763	0.375	0.34	1.139	1.114
14	92-93	0.721	0.741	0.354	0.328	1.092	1.074
15	93-94	0.766	0.78	0.406	0.358	1.199	1.144
16	94-95	0.771	0.797	0.329	0.305	1.116	1.118
17	95-96	0.834	0.821	0.467	0.413	1.329	1.248
18	96-97	0.801	0.798	0.39	0.342	1.227	1.16
19	97-98	0.763	0.793	0.398	0.358	1.189	1.174
20	98-99	0.771	0.784	0.371	0.338	1.151	1.125
21	99-00	0.782	0.793	0.315	0.276	1.131	1.101
22	00-01	0.825	0.847	0.371	0.352	1.229	1.225
23	01-02	0.829	0.832	0.395	0.357	1.255	1.215
24	02-03	0.786	0.823	0.439	0.401	1.237	1.239
25	03-04	0.852	0.861	0.393	0.378	1.274	1.256
26	04-05	0.824	0.831	0.362	0.329	1.216	1.191
27	05-06	0.865	Train	0.431	Train	1.308	Train
28	06-07	0.885	Train	0.435	Train	1.364	Train
29	07-08	0.886	0.88	0.436	0.391	1.342	1.284

En la Figura 40 se muestran las medias de Hs anuales correspondiente a los datos del hindcast y a los datos generados con el modelo neuronal. La línea de color azul corresponde a las medias del modelo neuronal y la línea de color rojo corresponde a las medias obtenidas con el modelo numérico MIKE21. La misma comparación se realizó con las desviaciones estándar de las dos diferentes bases de datos, aquí podemos observar que las desviaciones anuales del modelo neuronal son menores que las desviaciones anuales obtenidas con los datos del MIKE21 (Figura 41). De igual manera se presenta la  $H_{1/3}$  por año, de las Hs promedio (Figura 42). En las figuras: Figura 40, Figura 41 y Figura 42 las líneas de color azul que corresponden al modelo neuronal, no tiene datos en el número de año 27 (2005-2006) y 28 (2006-2007), debido a que esos años se usaron en la etapa de aprendizaje para entrenar a la red neuronal.

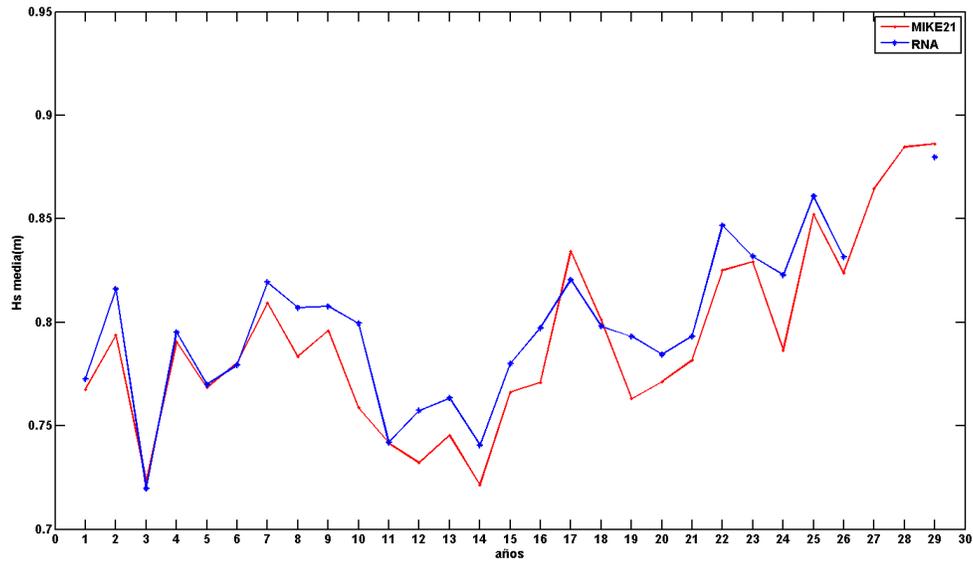


Figura 40. Comparación de las medias anuales del MIKE21 vs. RNA, entrenando con 2 años.

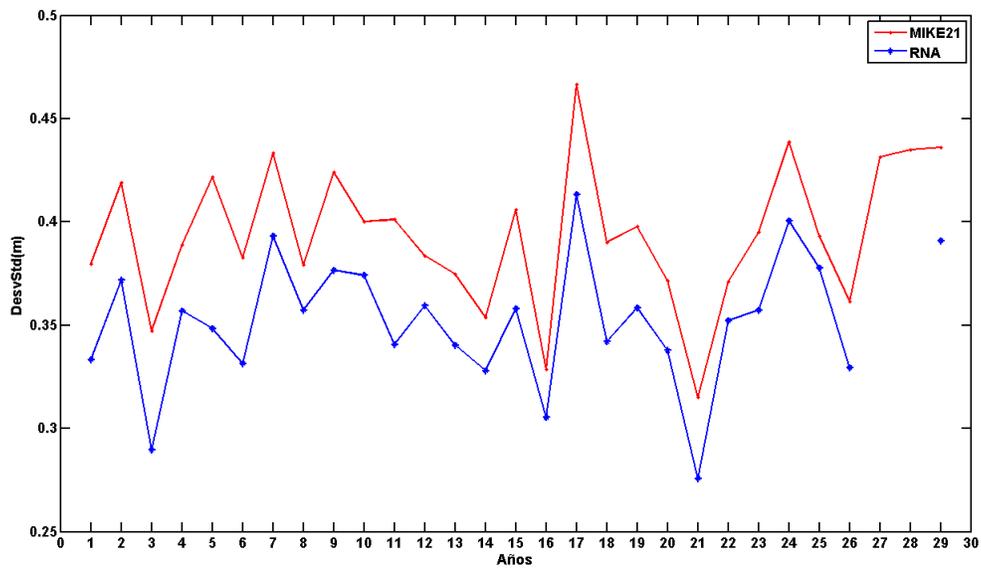


Figura 41. Comparación de la desviación estándar anuales del MIKE21 vs. RNA, entrenando con 2 años.

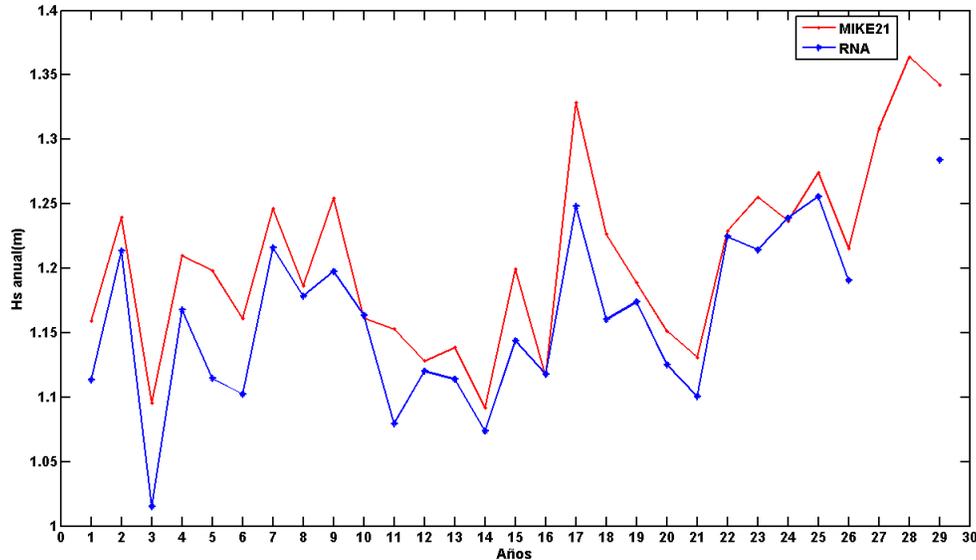


Figura 42. Comparación de la Hs anuales del MIKE21 vs. RNA, entrenando con 2 años.

#### 7.7.2.4. Segundo caso Hindcast

En este segundo caso de aplicación se utilizan menos datos para entrenar la red neuronal con la finalidad de usar solo los datos necesarios para generar buenos resultados, por tal motivo se considera que entrenar con un año es suficiente para tomar en cuenta los diferentes eventos estacionales o temporales representativos del área de estudio, que generan un clima marítimo de mayor intensidad como son los eventos de nortes; y por otro lado eventos de menor intensidad como las brisas locales. Por lo anterior se decidió entrenar con el año 2006 - 2007 que es donde se registraron 21 eventos de Nortes, durante toda la serie de datos históricos (Appendini et al., 2013).

Así también, se varió el número de neuronas en la capa intermedia para generar los mejores resultados de Hs, y evitar la sobrestimación. En la Tabla 9, podemos observar los errores que se obtuvieron tanto en la etapa de entrenamiento como en la etapa de evaluación del modelo neuronal, el coeficiente de correlación que existe entre la señal obtenida con el modelo neuronal contra la señal de oleaje del hindcast del Mike21, así como el tiempo de computo.

Tabla 9. Tabla de errores obtenidos con diferente número de neuronas en la capa oculta para la RNB entrenada con un año 2006-2007.

	Entrenamiento	Evaluación	CC	Tiempo (seg)
BNN 2N	0.168	0.193	0.875	41.976
<b>BNN 4N</b>	<b>0.151</b>	<b>0.177</b>	<b>0.896</b>	<b>59.323</b>
BNN 6N	0.145	0.177	0.896	92.908
BNN 8N	0.140	0.183	0.889	110.722
BNN 10N	0.134	0.185	0.894	161.342

Con los resultados obtenidos (Tabla 9) se observa que el mejor resultado se obtiene usando 4N en la capa intermedia, el cual genera un mínimo error, con esta arquitectura se genera el oleaje y los principales momentos estadísticos de igual manera a como se realizó en el apartado anterior usando 2 años de entrenamiento. De esta manera los resultados de la Figura 43 corresponden a una red neuronal de tipo bayesiana - RNB con 4N en la capa intermedia, entrenada con datos de viento del NARR y Hs modelado del MIKE21 y la evaluación se realiza alimentando la RNB con datos de viento de los 28 años restantes (de 1979 al 2006 y posteriormente 2007 - 2008). En la Figura 43, observamos la comparación entre los datos generados por el modelo neuronal (línea de color azul) contra los datos del MIKE21 (línea de color rojo), en la cual podemos observar un comportamiento muy similar sin embargo en algunos años se observan sobrestimación de alturas de ola.

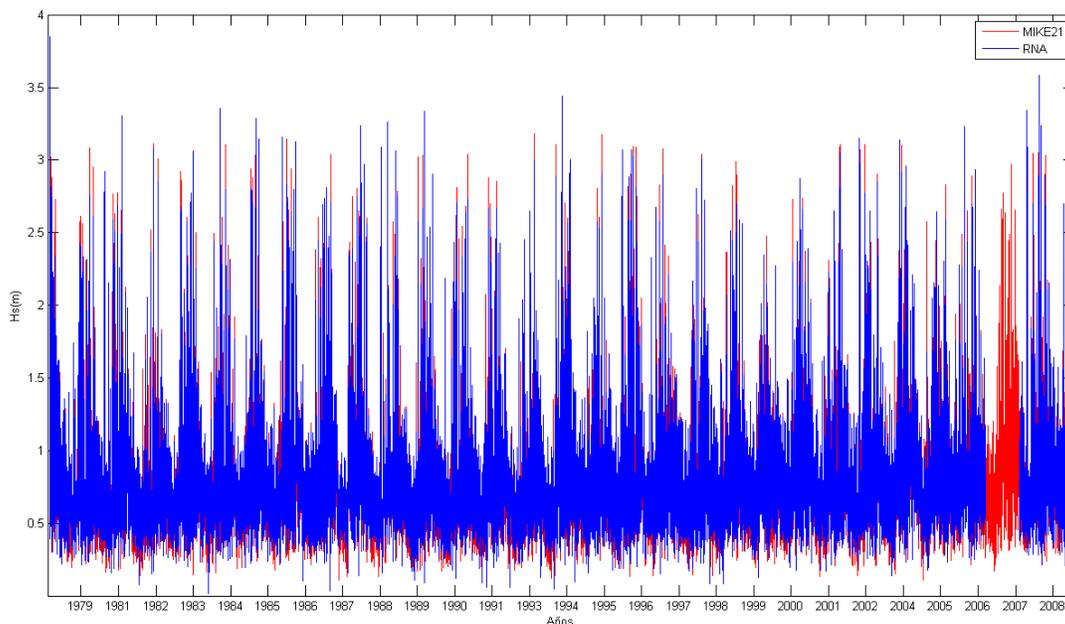


Figura 43. Comparación de los hindcast con MIKE21 y con RNA entrenada solo con 1 año.

De igual manera que en el caso 1 (entrenando con 2 años), se realiza una correlación de la distribución de las series de datos de Hs, para poder discernir las alturas de ola del modelo neuronal que se ajusten mejor a las alturas de ola del MIKE21 (Figura 44). En este caso, podemos observar que las Hs  $\leq 0.5\text{m}$  presentan una subestimación con respecto a la señal del MIKE21, posteriormente el modelo neuronal se ajusta perfectamente en valores de Hs entre 0.5m a 2m y finalmente para Hs  $>2.5\text{m}$  nuevamente las alturas son subestimadas.

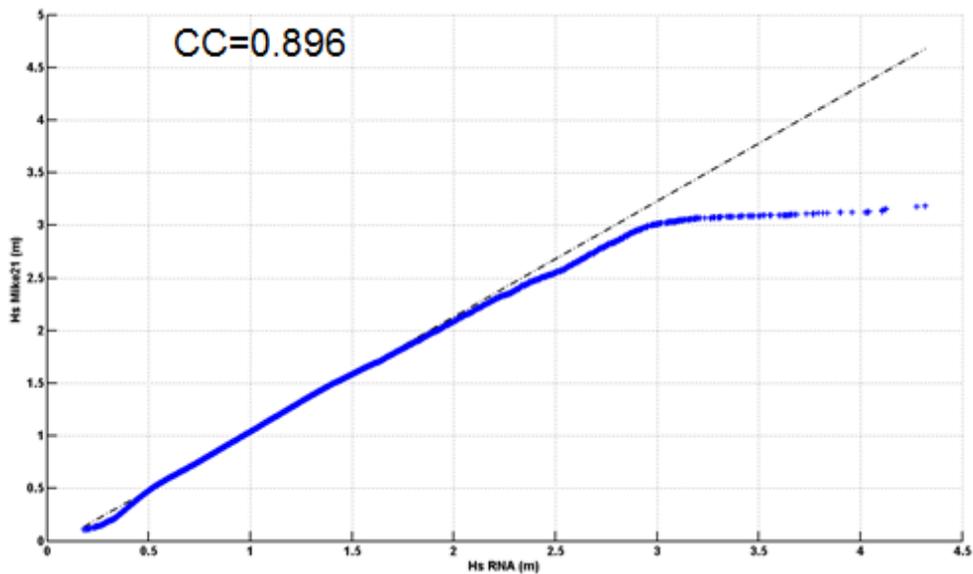


Figura 44. Comparación de la distribución de los dos series de datos de Hs, entrenando con 1 año la RNB.

#### 7.7.2.5. Estadísticos segundo caso

Se calcularon los momentos estadísticos de las señales de oleaje como se realizó en el apartado anterior para poder validar los resultados obtenidos con el modelo neuronal. En la Tabla 10, se presentan los estadísticos anuales calculados, con la serie de datos del MIKE21 y la serie de datos generada por la RNA. En las columnas correspondientes a los momentos estadísticos obtenidos para los datos tanto de la RNA y del MIKE21, en el año número 28 (correspondiente al 2006-2007) tiene la palabra “Train” lo cual significa que ese año se usó para entrenar el modelo neuronal y por consiguiente no hay momento estadístico en la etapa de la evaluado con la RNA.

Tabla 10. Tabla de comparación de estadísticos anuales, entrenando con un año de datos

No. Año	Fechas	Media Mike	Media RNA	desv.std Mike21	desv.std RNA	Hs anual Mike	Hs anual RNA
1	79-80	0.768	0.738	0.38	0.33	1.159	1.075
2	80-81	0.794	0.791	0.419	0.385	1.239	1.199
3	81-82	0.723	0.686	0.347	0.301	1.096	1.005
4	82-83	0.79	0.759	0.389	0.371	1.21	1.152
5	83-84	0.768	0.736	0.422	0.367	1.198	1.093
6	84-85	0.78	0.757	0.383	0.343	1.161	1.089
7	85-86	0.809	0.791	0.433	0.409	1.246	1.204
8	86-87	0.783	0.779	0.379	0.354	1.186	1.151
9	87-88	0.796	0.778	0.424	0.385	1.254	1.179
10	88-89	0.759	0.767	0.4	0.375	1.161	1.14
11	89-90	0.742	0.713	0.401	0.353	1.153	1.063
12	90-91	0.732	0.728	0.384	0.365	1.128	1.103
13	91-92	0.745	0.736	0.375	0.337	1.139	1.092
14	92-93	0.721	0.716	0.354	0.334	1.092	1.063
15	93-94	0.766	0.759	0.406	0.371	1.199	1.137
16	94-95	0.771	0.787	0.329	0.302	1.116	1.106
17	95-96	0.834	0.799	0.467	0.435	1.329	1.243
18	96-97	0.801	0.779	0.39	0.337	1.227	1.138
19	97-98	0.763	0.769	0.398	0.381	1.189	1.174
20	98-99	0.771	0.757	0.371	0.339	1.151	1.102
21	99-00	0.782	0.782	0.315	0.276	1.131	1.092
22	00-01	0.825	0.836	0.371	0.35	1.229	1.214
23	01-02	0.829	0.817	0.395	0.356	1.255	1.201
24	02-03	0.786	0.804	0.439	0.454	1.237	1.242
25	03-04	0.852	0.847	0.393	0.386	1.274	1.249
26	04-05	0.824	0.82	0.362	0.32	1.216	1.175
27	05-06	0.865	0.839	0.431	0.422	1.308	1.252
28	06-07	0.885	Train	0.435	Train	1.364	Train
29	07-08	0.886	0.886	0.436	0.443	1.342	1.314

El comportamiento de las medias anuales de  $H_s$  que fueron calculadas para los 27 años de evaluación con el modelo neuronal, presenta un comportamiento muy parecido a las medias obtenidas con los datos anuales del Hindcast (Figura 45). La línea de color azul corresponde a las medias del modelo neuronal y la línea de color rojo corresponde a las medias obtenidas con el modelo numérico MIKE21. La misma comparación se realizó con la desviación estándar de las dos diferentes bases de datos, aquí podemos observar que las desviaciones anuales del modelo neuronal son menores que las desviaciones anuales obtenidas con los datos del MIKE21 (Figura 46). De igual manera se decidió calcular la  $H_{1/3}$  por año, de las  $H_s$  promedio (Figura 47).

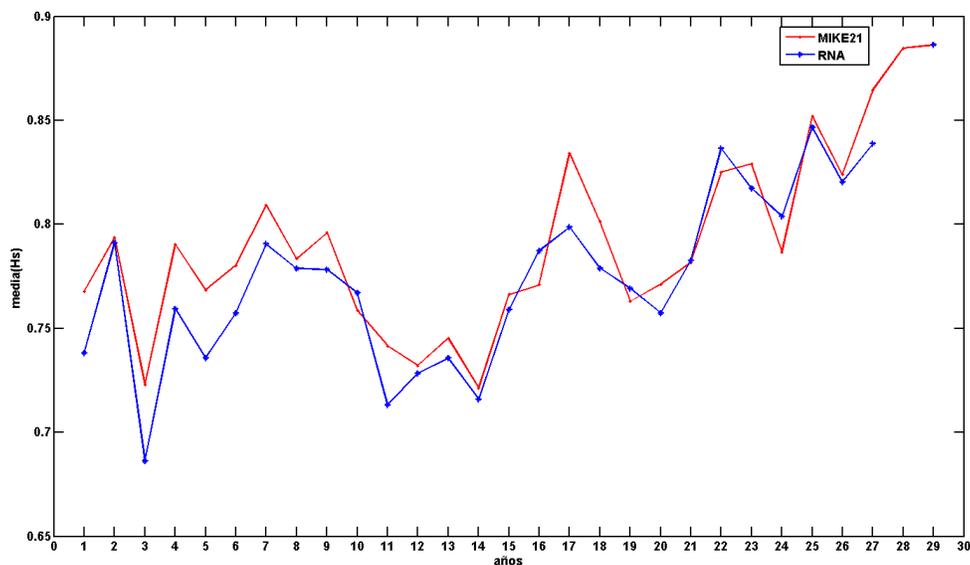


Figura 45. Comparación de las medias anuales del MIKE21 vs. RNA, entrenando con 1 año.

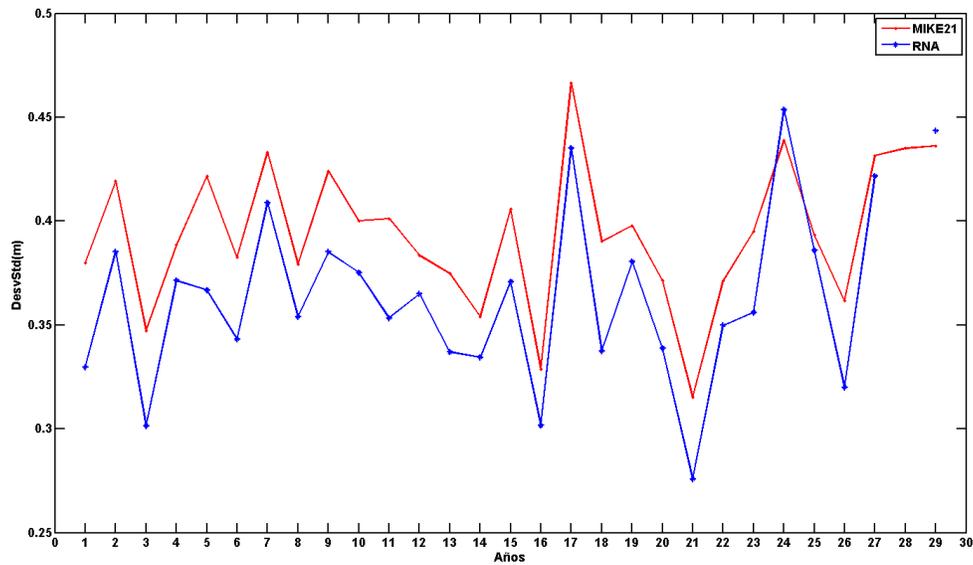


Figura 46. Comparación de las desviaciones estándar anuales del MIKE21 vs. RNA, entrenando con 1 año.

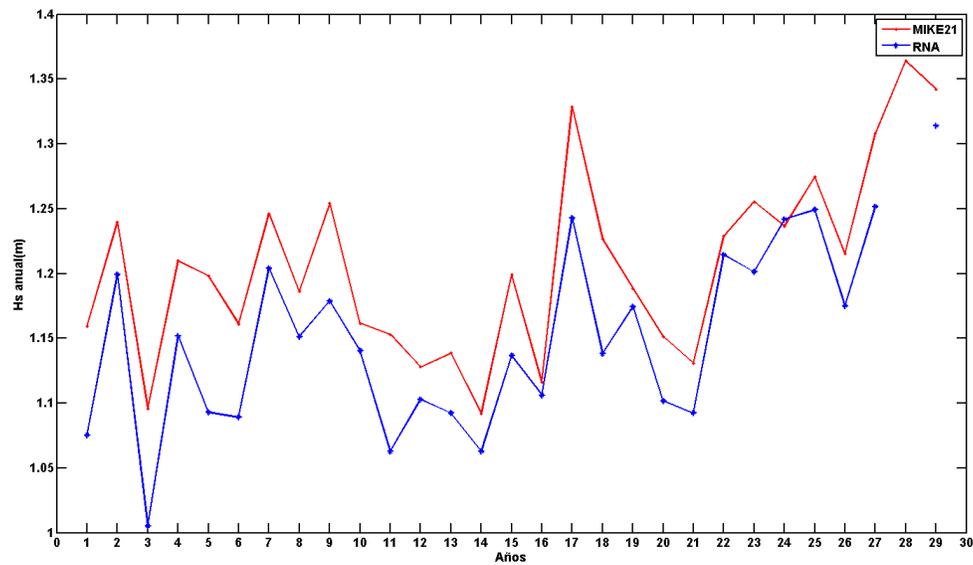


Figura 47. Comparación de las Hs anuales del MIKE21 vs. RNA, entrenando con 1 año.

### 7.7.2.6. Tercer caso Hindcast

Al observar que con un año de entrenamiento la RNA tiene un buen desempeño, se decidió utilizar los datos de otro año estacional donde las alturas de ola no corresponden a registros de oleaje extremo. En este tercer caso se utilizara el año 1992-1993 donde la altura de ola media es de 0.72m la cual es menor en comparación con la altura de ola para el año donde

se registraron más nortes en el segundo caso (año 2006-2007) la cual es de 0.88m. Esto con la finalidad de investigar si los resultados dependen o no del periodo de entrenamiento usando un año con mucha intensidad de nortes teniendo oleaje extremal, contra un año con poca intensidad sin oleaje extremo.

Debido a que los datos de entrenamiento corresponden a los años 1992-1993, se decidió nuevamente entrenar la red neural bayesiana usando diferentes número de neuronas en la capa oculta para generar los mejores resultados de Hs, y evitar la sobrestimación. En la Tabla 11, podemos observar los errores que se obtuvieron tanto en la etapa de entrenamiento como en la etapa de evaluación del modelo neuronal, el coeficiente de correlación que existe entre la señal obtenida con el modelo neuronal contra la señal de oleaje del hindcast del Mike21, así como el tiempo de computo.

Tabla 11. Tabla de errores obtenidos con diferente número de neuronas en la capa oculta para la RNB entrenada con un año 1992-1993.

	Error Entrenamiento	Error Evaluación	CC	Tiempo (seg)
BNN 2N	0.1690	0.193	0.876	37.863
<b>BNN 4N</b>	<b>0.142</b>	<b>0.178</b>	<b>0.896</b>	<b>57.861</b>
BNN 6N	0.152	0.187	0.884	92.276
BNN 8N	0.138	0.180	0.893	135.953
BNN 10N	0.133	0.185	0.888	172.806

Se realiza una correlación de la distribución de las series de datos de Hs y se observa que los resultados son muy similares por lo cual los resultados de la red neuronal no dependen del año de entrenamiento (Figura 48).

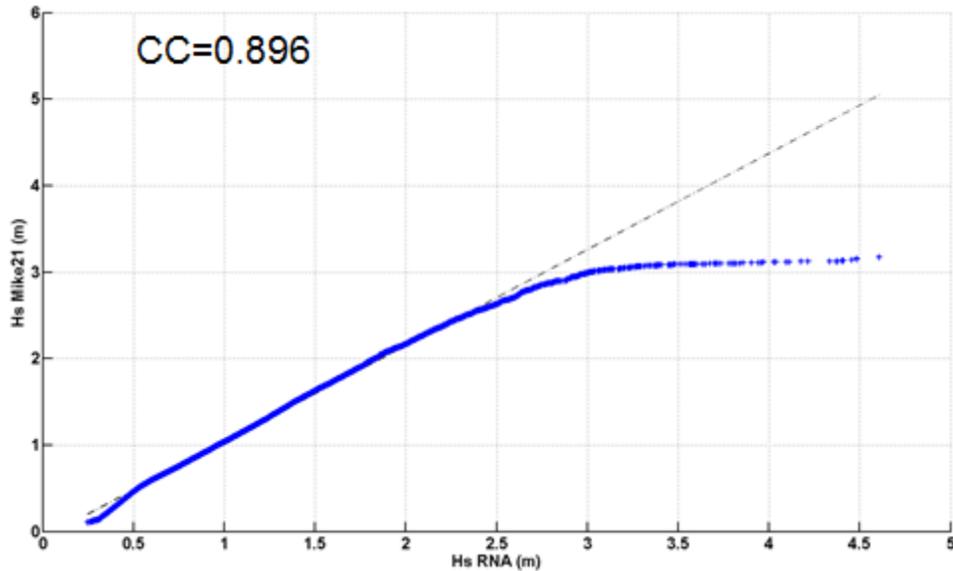


Figura 48. Comparación de la distribución de los dos series de datos de Hs, entrenando con 1 (1992-1993) año la RNB.

#### 7.7.2.7. Conclusiones del Hindcast

La diferencia entre la RNA entrenada con dos y con un año es mínima, entrenando con 2 años (2005 al 2007) el error fue de 0.169m; entrenando con un año (1992-1993) el error fue de 0.178m, por lo cual, si solo tenemos datos de un año para entrenar y querer generar una base de datos históricas es suficiente para obtener datos de Hs confiables a partir de datos de viento.

Los resultados de la RNA no dependen del clima de oleaje seleccionado por lo que, no es necesario entrenar con un año que haya estado expuesto a eventos de clima de oleaje extremo para que la RNB logre una buena optimización y genere resultados confiables. Ya que el modelo neuronal en la etapa de entrenamiento aprende a asociar las alturas de ola con la velocidad del viento, ya sea de mayor o menor escala. Lo anterior lo podemos observar en la Tabla 12, donde se muestra que el error de la etapa de evaluación que se calculó entrenando con 2006-2007 o 1993-1993, prácticamente es el mismo y los coeficientes de correlación son iguales.

Tabla 12. Comparación de los errores obtenidos, entrenando con diferentes años.

Años	RNA	Error Entrenamiento	Error Evaluación	CC	Tiempo (seg)
2006-2007	BNN 4N	0.151	0.177	0.896	59.323
1992-1993	BNN 4N	0.142	0.178	0.896	57.861

En este caso de aplicación, las alturas de ola que se ajustaron mejor de manera general, están directamente relacionadas con los vientos locales (oleaje tipo sea). Justamente los vientos que se consideraron para el entrenamiento son vientos puntuales (nodo 12918) y locales por lo que la red neuronal aprende a generar olas (de 0 a 2m); a diferencia de las alturas de ola mayores de 2m, las cuales son provocadas por el oleaje distante (oleaje tipo swell).

Otra posible razón de que las alturas mayores a 2m, no presenten un buen ajuste a los datos (graficas de correlación de la distribución de los datos Figura 39, Figura 44 y Figura 48), puede ser que el modelo MKE21 este sobrestimando las alturas de ola en comparación con las alturas reales, esto como resultado de la observación del ajuste de los datos del Mike 21 contra datos medidos del ADCP (sección 7.6) donde se puede observar en la Figura 31, que las  $H_s > 2m$  generadas con el MIKE21 están sobrestimando a la altura medida por el ADCP.

Finalmente podemos optar por usar el tipo de RNB-4N (con cuatro neuronas en la capa oculta) + Stepwise, ya que entrenando con un año como se hizo en el segundo caso de aplicación y en el tercer caso de aplicación, obtenemos los mínimos errores (Tabla 9 y Tabla 11) y las correlaciones más altas. Por otro lado entrenando con dos años los mínimos se obtuvieron usando RNB-6N + Stepwise (Tabla 7), sin embargo la diferencia usando 4N y 6N en la capa oculta es mínima (centésimas), por lo que si usamos 4N en lugar de 6N para el entrenamiento los resultados que obtenemos son aceptables. De esa manera tendríamos solo un tipo y arquitectura de red neuronal como producto para generar series de datos históricos de  $H_s$  a partir de datos de viento.

### 7.7.3. Oleaje extremal

En este apartado, con el  $H_{1/3}$  que se calculó en la sección de los momentos estadísticos se calculará el periodo de retorno y el oleaje extremal. El oleaje extremal es de gran importancia en la ingeniería costera ya que muchos criterios de diseño están basados en tal concepto, es decir, para construir una obra marítima es necesario conocer la altura de ola de diseño.

En esta parte del trabajo se realizó el cálculo del periodo de retorno a partir de las series de datos históricos del MIKE21 y los generados con el modelo neuronal, con la finalidad de comparar la altura de ola extremal obtenida a partir de estas dos fuentes y ver si son equiparables.

De los resultados obtenidos, se puede observar que las  $H_s$ , que corresponden a los datos del hindcast (MIKE21) son ligeramente mayores que las alturas generadas con el hindcast del modelo neuronal (Tabla 13). La diferencia que existe entre el promedio del oleaje extremal obtenido a partir del MIKE21 menos el promedio del oleaje extremal obtenido con los datos de la RNA resulta de 0.15m, esto quiere decir que los valores de oleaje extremal obtenido con los datos de la RNA son equiparables con los obtenidos con un MT.

Tabla 13. Alturas de olas calculadas para cada periodo de retorno correspondiente, usando las dos series de datos.

Años	$H_s$ (m) MIKE21	$H_s(m)_{RNA}$
1.5	1.17	1
2	1.19	1.03
5	1.26	1.1
10	1.3	1.14
15	1.32	1.17
20	1.34	1.19
30	1.36	1.21
50	1.39	1.24
75	1.42	1.27
100	1.43	1.28

En las Figura 49 y Figura 50, podemos observar el periodo de retorno obtenido a partir del ajuste de los datos con el método de Gumbel para diferentes periodos de retorno. Los puntos corresponden a las alturas obtenidas con los modelos correspondientes, la línea segmentada corresponde al ajuste del periodo de retorno y las líneas continuas que se encuentran por arriba y debajo de los datos corresponden a las desviaciones estándar y el área que se encuentra entre esas dos líneas corresponde al intervalo de confianza de los datos, el cual es el 95%.

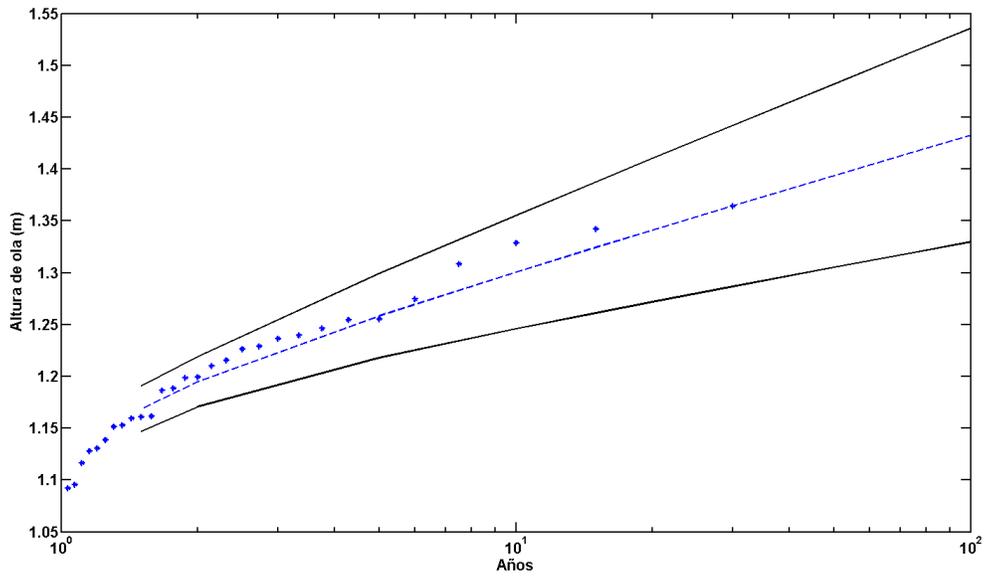


Figura 49. Grafica del periodo de retorno generado a 100 años con los datos del MIKE21.

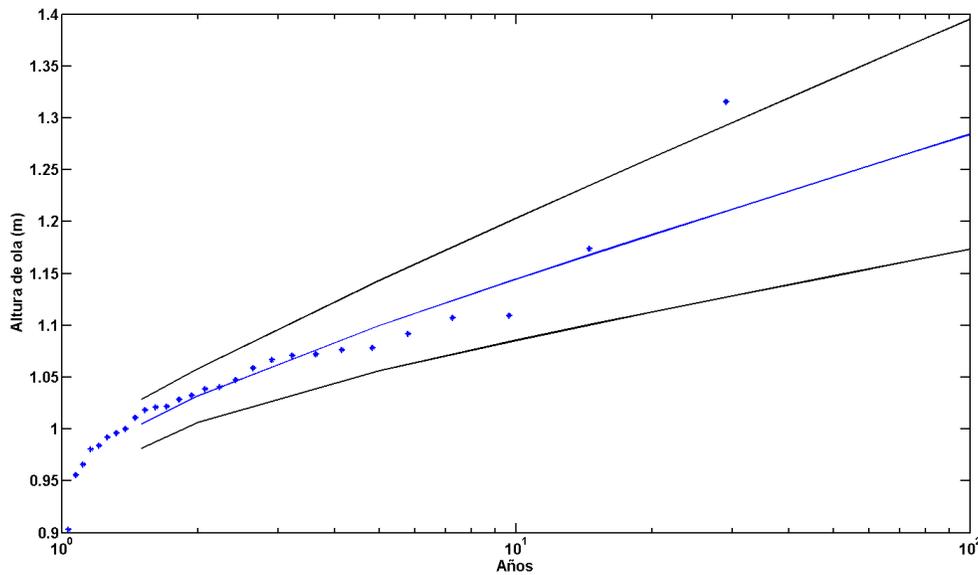


Figura 50. Grafica del periodo de retorno generado a 100 años con los datos de la RNA.

## 8. DISCUSIONES

Con el método de stepwise se hizo más eficiente los modelos neuronales ya que el modelo estadístico lo que nos arroja es una matriz lógica que indica la posición donde se encuentra las componentes de velocidad más representativas de la variable de predicción (viento) para poder explicar el comportamiento de la Hs y mejorar la minimización del error. En cuanto a tiempo de computo también se pudo observar una mejora ya que el modelo tarda un 80% menos en arrojar una salida de Hs usando el método estadístico; tiene sentido ya que al meter menos valores en el entrenamiento el cálculo es más rápido i.e. si con un modelo sin el stepwise para generar 24 horas de Hs, usas 24 datos de componente de viento u, 24 datos de componente de viento v y 24 datos de oleaje; en total usábamos 72 datos, para generar 24 datos de Hs. Ahora aplicando el método de stepwise solo elige de esos 72 valores, 13 o 15 variables siendo las más representativas estadísticamente para el modelo estadístico. Lo cual indica que las componentes de la variable de viento elegidas son las más importantes para obtener y explicar la generación del oleaje y los demás datos son redundantes e innecesarios para optimizar la red neuronal.

## 9. CONCLUSIONES

- La minimización del error se logró usando el método de regresión múltiple “Stepwise”. Una ventaja de usar este método radica en el tiempo de computo ya que hace más eficiente el tiempo de entrenamiento, otra ventaja importante es que evitamos tener un sobreajuste (overfitting), ya que para entrenar solo se utiliza el número necesarios de los valores de la variable predictiva.
- El trabajo realizado dejó como resultado tres tipos de redes neuronales listas para usarse dependiendo del caso de aplicación y objetivo.
  - La primera es una RNB-4N + Stepwise + Bootstrap, donde obtenemos los mejores resultados para obtener oleaje en cada uno de los puntos donde se colocaron los ADCP (donde se registraron 3 meses de medición), ya que los errores son mínimos entre los datos del salida del modelo y los datos esperados (mediciones ADCP), para el punto fuera de la costa como en el punto cerca de la costa para generar oleaje en tiempo real, si se disponen de mediciones de viento en tiempo real.
  - La segunda RNA corresponde al caso de aplicación de propagación de oleaje, la cual es una RNL + Stepwise, alimentando con datos de oleaje de fuera de la costa y viento local.
  - La tercera corresponde al caso de aplicación para generar un hindcast, esta RNA es un tipo RNB-4N + Stepwise, alimentada con datos de viento del nodo 12918 (Sisal, Yucatán).
- Para fines de este trabajo el usar el método de re-muestreo resulto ser una buena técnica solo en la parte de implementación de la red neuronal para cada sitio donde se colocaron los ADCP, debido que el modelo neuronal se entrena con menos de 2 meses de datos. Para el caso de aplicación (sección 7.7.2), no es necesario usar el re-muestreo, ya que el entrenamiento se optimiza usando como mínimo un año de datos.
- Es claro que la implementación de las RNA generan buenos resultados de Hs a partir de los datos de viento con los que se entrenan, por lo que se puede decir que el viento medido el cual corresponde a las brisas locales es el principal agente que

modula el comportamiento del oleaje (tipo sea) para el área de estudio, lo cual se mencionado en trabajos previos como los realizados por Appendini et al., ( 2013).

- Se considera que la metodología usada en este trabajo, así como el tipo de RNB puede implementarse en otro sitio de la costa de Yucatán, debido a que las condiciones geomorfológicas y de clima marítimo son similares a lo largo de la costa yucateca.
- La tercer RNA producto de este trabajo está lista para generar series de datos históricos (hindcast) del área de estudio a partir de series de datos históricas de viento local.
- La primer RNA producto de este trabajo puede usarse para rellenar huecos en series de tiempo incompletas debido a interrupciones en las mediciones con el ADCP, alimentando la red solamente con datos de viento.
- Usando los modelos neuronales se puede encontrar un tipo de red y arquitectura que a partir de datos de predicción sea capaz de aprender a generar pronósticos de Hs.
- Si se requiere conocer los parámetros de oleaje en un sitio en específico, no es necesario modelar todo un dominio, por lo que implementar esta metodología usando redes neuronales es conveniente, ya que requiere de menos esfuerzo computacional (el tiempo de cálculo es menor).
- De acuerdo a los resultados obtenidos con el modelo neuronal generando un Hindcast, se concluye que no es necesario usar un año extremal o dos años extremales para generar un buen entrenamiento de la RNB y generar Hs confiables, por lo que es suficiente entrenar solo con un año de medición, para poder generar a partir de datos de viento una serie histórica de oleaje.

### *9.1. Futuras líneas de investigación*

- Una validación que quedo pendiente, fue realizar la comparación de los resultados del modelo neuronal contra la salida de un modelo numérico para el mismo periodo de tiempo. Para esto se necesitaría tener una serie de datos medidos de oleaje y viento, del área de estudio durante los 30 años del Hindcast, de tal manera que el modelo neuronal se entrenaría con esos datos medidos y posteriormente se validaría solo con los datos de viento como se realizó en este trabajo. De tal manera que podríamos comparar las salidas del modelo numérico con datos reales (mediciones del ADCP) y las salidas del modelo neuronal contra datos reales (mediciones del ADCP), de esa manera podríamos concluir que resultado se ajusta mejor a los datos reales.
- Se recomienda para futuras investigaciones o trabajos del tema, considerar en el entrenamiento de los modelos neuronales, viento que se encuentra en aguas profundas para enseñar a la red neuronal sobre la influencia del oleaje tipo swell a zonas costeras, de esta manera la señal de oleaje generada con el modelo neuronal se ajustaría más a los datos reales y podríamos tener mejores resultados.
- Otra manera de generar un modelo más robusto en caso que se tenga interés por conocer los parámetros de oleaje en un dominio o área, es realizar mallas computacionales y en cada nodo tener una neurona la cual este interconectada a su vez con otra neurona y otro nodo, de tal manera que haya una interconexión de neuronas para aprender una de otras. En este sentido se deberá buscar nuevamente el mejor tipo de red neuronal a usar, algoritmos de optimización y arquitectura, para lograr un entrenamiento que cumpla con los objetivos planteados.

## 10. REFERENCIAS

- Appendini, C. M., Torres-Freyermuth, A., Oropeza, F., Salles, P., López, J., & Mendoza, E. T. (2013). Wave modeling performance in the Gulf of Mexico and Western Caribbean: Wind reanalyses assessment. *Applied Ocean Research*, 39, 20-30. doi:10.1016/j.apor.2012.09.004
- Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks : fundamentals , computing , design , and application, 43, 3-31.
- Deo, M., & Naidu, C. S. (1999). Real time wave forecasting using neural networks. *Ocean Engineering*, 26(3), 191-203. Recuperado a partir de <http://portal.acm.org/citation.cfm?doid=190468.190290>
- Figueroa-Espinoza, B., & Salles, P. (2014). On the wind power potential in the northwest of the Yucatan Peninsula in Mexico. *Atmósfera*, 27(1), 77-89. doi:10.1016/S0187-6236(14)71102-6
- Hsieh, W. W., & Tang, B. (1998). Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography. *Bulletin of the American Meteorological Society*, 79(9), 1855-1870. doi:10.1175/1520-0477(1998)079<1855:ANNMTP>2.0.CO;2
- Jain, P., Deo, M. C., Latha, G., & Rajendran, V. (2011). Real time wave forecasting using wind time history and numerical model. *Ocean Modelling*, 36(1-2), 26-39. doi:10.1016/j.ocemod.2010.07.006
- Makarynsky, O., Makarynska, D., Kuhn, M., & Featherstone, W. E. (2004). Predicting sea level variations with artificial neural networks at Hillarys Boat Harbour, Western Australia. *Estuarine, Coastal and Shelf Science*, 61(2), 351-360. doi:10.1016/j.ecss.2004.06.004
- Mandal, S., & Prabakaran, N. (2006). Ocean wave forecasting using recurrent neural networks. *Ocean Engineering*, 33(10), 1401-1410. doi:10.1016/j.oceaneng.2005.08.007
- Subba R., Mandal S., P. N. (2001). Rao(2001).pdf. *Proceedings of the International Conference in Ocean Engineering*, 11-14.
- Tsai, C.-P., Lin, C., & Shen, J.-N. (2002). Neural network for wave forecasting among multi-stations. *Ocean Engineering*, 29(13), 1683-1695. doi:10.1016/S0029-8018(01)00112-3

Appendinni C.M. Sistema operacional de predicción de oleaje para las Costa y Mares Mexicanos. Sisal: UNAM. Noviembre 2012.

Castillo E. 1988. Extreme Value theory in Engineering. Academic Press. Universidad de Cantabria. Santander, Spain. 389pp.

Hsieh W.W. 2009. Machine Learning Methods in the Environmental Sciences. Neural Networks and Kernels. University of British Columbia, Vancouver. 349pp.

Komar P. D. 1998. Beach Processes and Sedimentation. 2<sup>nd</sup> edition, Prentice Hall Oregon State University, 429pp.

Komen G.J. L. Cavaleri, M. Donelan, K. Hasselmann and P. A. E. M. Janssen, 1994. Dynamics and modelling of ocean waves. Cambridge University Press, 532 pp.

Lira-Pantoja M.A. 2015. Implementación y evaluación de un sistema operacional de predicción de oleaje para las costas y mares mexicanos. Tesis de Maestría. Universidad Nacional Autónoma de México.

López, G.J. 2005. Implementación de Redes Neuronales para predicción de mareas. Tesis de Maestría. Universidad Nacional Autónoma de México.

López-Mendez L.V. 2009. Análisis del evento meteorológico del 2007 relacionado con la inundación de Tabasco. Universidad Nacional Autónoma de México.

Matich D.J. (2001). Redes Neuronales: Conceptos Básicos y Aplicaciones. Universidad Tecnológica Nacional – Facultad Regional Rosario. Disponible en: <ftp://decsai.ugr.es>.

MacKay, Neural Computation, Vol. 4, No. 3, 1992, pp. 415–447.

Silva Rodolfo. 2005. Análisis y descripción estadística del oleaje. Serie de docencia. Instituto de Ingeniería. Universidad Nacional Autónoma de México.

"Acoustic Doppler Current Profilers" (PDF). Hydro International. Retrieved 2014-05-22.

"Acoustic Doppler Current Profiler (ADCP)". Woods Hole Oceanographic Institution. Retrieved 2014-05-22.