

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INFORMÁTICA

**RED NEURONAL POLINOMIAL PARA EL
PRONÓSTICO DE CAPTURA MENSUAL DE
LA ANCHOVETA DE ALTA Y BAJA
FRECUENCIA EN LA ZONA NORTE DE
CHILE USANDO OPTIMIZACIÓN POR
ENJAMBRE DE PARTÍCULAS**

FELIPE GABRIEL MOLINA ROJAS

INFORME DE PROYECTO
PARA OPTAR AL TÍTULO PROFESIONAL DE
INGENIERO CIVIL INFORMÁTICA

ENERO 2014

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INFORMÁTICA

**RED NEURONAL POLINOMIAL PARA EL
PRONÓSTICO DE CAPTURA MENSUAL DE
LA ANCHOVETA DE ALTA Y BAJA
FRECUENCIA EN LA ZONA NORTE DE
CHILE USANDO OPTIMIZACIÓN POR
ENJAMBRE DE PARTÍCULAS**

FELIPE GABRIEL MOLINA ROJAS

Profesor Guía: **Nibaldo Rodríguez Agurto**
Profesor Co-Referente: **José Miguel Rubio León**

INFORME DE PROYECTO
PARA OPTAR AL TÍTULO PROFESIONAL DE
INGENIERO CIVIL INFORMÁTICA

ENERO 2014

Índice

1. Introducción	1
1.1. Objetivos del Proyecto	3
1.1.1. Objetivo General	3
1.1.2. Objetivos Específicos	3
1.2. Organización del texto	3
2. Pronóstico con Redes Neuronales	4
2.1. Pronóstico	4
2.2. Aplicaciones de Redes Neuronales como herramienta de Pronóstico	5
2.3. Red Neuronal Polinomial	10
3. Optimización por Enjambre de Partículas	10
3.1. Introducción	10
3.2. Fundamentos del Movimiento de Partículas	11
3.3. Algoritmo PSO	11
3.4. Parámetros PSO	13
3.5. Topologías de vecindad local y global	14
3.6. PSO con actualizaciones síncronas y asíncronas	14
3.7. Exploración y Explotación con PSO	15
3.8. Variaciones de Algoritmos de PSO	15
3.8.1. PSO con peso de inercia decreciente (DIWPSO)	15
3.9. Aplicaciones	16
4. Pronosticador Neuronal	16
4.1. Métricas de Evaluación	17
4.2. Modelo Propuesto	17
4.2.1. Descripción del Modelo	19
4.2.2. Selección de Topología y Parámetros de la Red	19
4.3. Baja Frecuencia	20
4.3.1. Neuronas de Entrada	20
4.3.2. Nodos Ocultos	21
4.3.3. Número de Partículas	21
4.3.4. Cantidad de Iteraciones	21
4.4. Alta Frecuencia	22
4.4.1. Configuración de la Red	22
4.4.2. Número de Partículas	22
4.4.3. Cantidad de Iteraciones	22
5. Análisis de Resultados	23
5.1. Prónostico para el mes siguiente	23
5.1.1. Modelo de Baja Frecuencia	23
5.1.2. Modelo de Alta Frecuencia	24
5.1.3. Modelo Propuesto LF + HF	25

5.2. Pr3n3stico para el cuarto mes	26
5.3. Investigaciones Anteriores	27
6. Conclusi3n	29

Lista de Figuras

1.	Topologías de la población Global,(b) Local con $N=2$, (c) Local con $N=4$	14
2.	Modelo de Alta y Baja Frecuencia.	18
3.	Diagrama de Bloque.	19
4.	Gráficos de resultados para baja frecuencia.	23
5.	Gráficos de resultados para alta frecuencia.	24
6.	Gráficos de resultados para el modelo sumado.	25
7.	Métricas MAPE y RMSE a través de los meses.	26
8.	Gráficos de resultados para la suma de la variable anual e interanual.	27

Lista de Tablas

1.	Resultados al variar el número de entradas	20
2.	Resultados al variar el número neuronas ocultas	21
3.	Resultados al varia el número de partículas	21
4.	Resultados al variar el número de iteraciones	21
5.	Resultados al varia el número de partículas	22
6.	Resultados al variar el número de iteraciones	22
7.	Comparación de resultados con trabajos previos	28

*Dedicado a mi familia y a todos los que me
acompañaron en este camino.*

Agradecimientos

A mi esposa, mi familia y profesor guía.

Resumen

Este proyecto presenta un modelo que permite pronosticar el nivel de captura de anchoveta en la zona norte de Chile. Esta especie pelágica se encuentra presente en todo nuestro territorio y su importancia radica en el gran aporte económico para la industria pesquera. Desde 1983, la extracción de este recurso se encuentra sujeta a regulaciones, debido a la caída de la población de anchoveta. Es bajo este escenario, que se hace importante proponer modelos que apoyen la toma de decisiones que permita mantener o perpetuar la fauna marina. El modelo propuesto consiste en la implementación de una red neuronal polinomial, con ajuste de sus parámetros mediante una variante del algoritmo evolutivo, llamado optimización por enjambres de partículas (PSO por sus siglas en inglés), utilizando como entrada del sistema los datos históricos de captura de la anchoveta separados en datos de alta y baja frecuencia. Finalmente la mejor configuración obtuvo un 98.6 % de la varianza explicada para la captura del mes siguiente y un 94 % de la varianza explicada para el cuarto mes.

Abstract

This project presents a model that allows to forecast the capture level of anchovies in the northern part of Chile. This pelagic species is found throughout the entire Chilean territory and its importance relies on the economic contribution the catches have for the fishing industry. Since 1983, the capture activity of this resource is found subject to regulations due to the downfall of the anchovy population. Under this scenario, it becomes important to present models that will help make decisions that will ensure sustainability of the stock and the reproductive process. The model consists of the implementation of a polynomial neural network with optimized parameters. These parameters were adjusted by an alternative of the evolutionary algorithm called, Particle Swarm Optimization, using historical capture levels data of the anchovies as inputs and separating them as high and low frequency data. Finally, the improved configuration obtained an explained variance of 98.6 % for next months catch and an explained variance of 94 % for the fourth month.

1. Introducción

Las costas de nuestro país se ven afectadas en gran parte de su extensión por corrientes marinas, especialmente por la corriente de Humboldt. Esta corriente transporta hacia el norte, masas de agua más frías, de baja salinidad y alto contenido de oxígeno disuelto. Estas características que presenta la corriente de Humboldt, crean las mejores condiciones para la existencia de vida marina, convirtiéndolo en el ecosistema más rico del mundo. Es en este ecosistema donde se presentan dos eventos principales que hacen que sus condiciones varíen en lo que podría considerarse como condiciones óptimas. Por un lado se presenta el fenómeno de El Niño, el cual consiste en una invasión de aguas cálidas que avanzan de norte a sur, lo que provoca variadas alteraciones en el ecosistema, una de ellas tiene relación con la sardina, la anchoveta y el plancton, los cuales se profundizan. El otro fenómeno natural es La Niña, que consiste en la fase fría, en que la temperatura de las aguas desciende [1].

Dentro de las especies pelágicas que se encuentran en la zona norte de Chile está la *Engraulis Rigens* o Anchoveta. A mediados del siglo pasado comienza a extraerse este recurso de manera masiva lo que llevó a que en el año 1983 se tuvieron que establecer mecanismos de regulación debido a la caída en la población de estos ejemplares, sobre todo durante los períodos de reproducción más importantes.

A la fecha se desconocen los procesos y mecanismos del ambiente marino que explican en forma directa la dinámica de las poblaciones de recursos pelágicos. La recopilación de información de las fluctuaciones de abundancia de la anchoveta en la zona norte tiene como fin, ser la base desde la cual se dicten las decisiones gubernamentales respecto de la caza del recurso pelágico, es decir tomar las medidas necesarias para asegurar la disponibilidad del recurso para el futuro.

El desembarque acumulado de Anchoveta a Diciembre del 2012 alcanzó las 902.000 toneladas, presentando una baja de 24,2% respecto a igual fecha del 2011, ésta cifra representó el 36% del desembarque extractivo en el país. Su aporte económico es de aproximadamente 3500 millones de dólares. La cota global anual de captura para el año 2013 corresponde a 805.000 toneladas [2].

Es aquí donde se vuelve importante la habilidad de predecir con cierto grado de exactitud los niveles de abundancia de los recursos pelágicos en la zona, con el fin de planificar de manera responsable la captura de éstos, para evitar que las poblaciones de especies pelágicas sigan disminuyendo.

Es tarea del gobierno, mediante las instituciones correspondientes como el instituto de fomento pesquero y la subsecretaría de pesca y acuicultura controlar las actividades de las pesqueras que desempeñan sus labores en la zona. Informándoles los niveles de captura permitidos, así estos podrán planificar su esfuerzo de pesca de acuerdo a estas medidas. Para establecer las cotas de captura, es necesario pronosticar el comportamiento de los niveles de población de los recursos pelágicos. Es ese el motivo de este proyecto: entregar una herramienta para la predicción de los niveles de captura de anchovetas en la zona norte.

En el presente proyecto se tomarán, como series temporales, los datos históricos relacionados con los volúmenes de captura mensual de anchovetas en el norte

de Chile que abarcan el período comprendido entre los años 1963 y 2007, estos datos estarán separados en Alta y Baja Frecuencia y mediante el uso de Redes Neuronales Polinomiales con aprendizaje basado en la técnica de Optimización por Enjambre de Partículas (PSO) [3], se realizará la predicción mensual para esta especie.

Las Redes Neuronales Artificiales son modelos que intentan reproducir el comportamiento del cerebro humano [4] y consisten en un conjunto de elementos de procesamiento, llamados neuronas, los cuales se conectan entre sí.

La organización y disposición de las neuronas dentro de una red neuronal se denomina topología, y viene dada por el número de capas, la cantidad de neuronas por capa, el grado de conectividad, y el tipo de conexión entre neuronas.

Una vez determinada la topología de la red neuronal es necesario entrenarla. En la etapa de entrenamiento la red es capaz de aprender relaciones complejas entre entradas y salidas mediante el ajuste de los pesos de las conexiones entre neuronas.

Existen distintas técnicas para entrenar una red neuronal, algunas implican un aprendizaje supervisado, mientras que otras implican uno no supervisado. En el presente caso de estudio, se realiza un aprendizaje supervisado de la red mediante la utilización del algoritmo Optimización por Enjambre de Partículas (PSO).

Optimización por Enjambre de Partículas es un método estocástico de cómputo evolutivo inspirado en el comportamiento social de organismos tales como las bandadas de aves o los bancos de peces, erigiéndose en alternativa eficiente y atractiva a otros métodos estocásticos mucho más extendidos, tales como los Algoritmos Genéticos (GA) [5].

La utilización de esta técnica para definir el estado interno de la red se debe a sus grandes cualidades de exploración y explotación, de manera que se pueda abarcar todo el espacio de búsqueda y así obtener las mejores regiones donde poder encontrar un óptimo local que se acerque lo más posible a una solución óptima de la red.

Cabe mencionar que existen informes y estudios anteriores respecto a la predicción de los niveles de captura de la anchoveta en la zona norte, en este proyecto se busca demostrar que la separación de los datos históricos en alta y baja frecuencia reduce drásticamente los errores de predicción, para probar esto se tomarán como referencia las memorias realizadas por compañeros en años anteriores [6, 7, 8, 9, 10]

1.1. Objetivos del Proyecto

1.1.1. Objetivo General

Desarrollar y evaluar una Red Neuronal Polinomial utilizando optimización por enjambre de partículas para pronosticar las capturas mensuales de anchovetas en el área Norte de Chile.

1.1.2. Objetivos Específicos

- Explicar el funcionamiento de las redes neuronales polinomiales con aprendizaje basado en la técnica de optimización por enjambre de partículas (PSO).
- Diseñar la estructura y estimar los parámetros del modelo de pronóstico neuroevolutivo mediante el uso de PSO.
- Implementar y evaluar el rendimiento del modelo de pronóstico propuesto.

1.2. Organización del texto

El proyecto continúa en el Capítulo 2 con la descripción del estado del arte relacionado a las Redes Neuronales utilizadas como herramienta de pronóstico, en este apartado se hace un recorrido donde se dan a conocer las áreas de la ciencia en las cuales las redes neuronales han sido un real aporte como elemento predictor. Luego el Capítulo 3 describe las técnicas evolutivas presentando a Optimización por Enjambre de Partículas, definiendo sus elementos básicos, algoritmos, variantes y aplicaciones. El Capítulo 4 consiste en la Redes Neuroevolutivas, donde se plantean los diferentes modelos elaborados para este proyecto, las métricas a utilizar y los resultados de los modelos sobre estas. Para continuar, el Capítulo 5 en el cual se discuten los resultados provenientes de la sección 4, analizando uno por uno cada modelo propuesto, además de una comparación de rendimiento entre el mejor modelo encontrado con los modelos propuestos en otros proyectos. Finalmente, el Capítulo 6 consiste en las conclusiones del proyecto.

2. Pronóstico con Redes Neuronales

2.1. Pronóstico

Predecir es hacer declaraciones sobre algo que va a suceder, a menudo se basan en información del pasado y del estado actual. Por ejemplo, cada persona hace diversos pronósticos en su diario vivir, ya sea sobre el clima, el tiempo se demora en llegar a un cierto destino, etcétera.

En el aspecto técnico, los parámetros predecibles de un sistema generalmente pueden ser expresados y evaluados por medio de ecuaciones, por lo tanto, la predicción es simplemente la evaluación o solución de una posible ecuación. Sin embargo, existen problemas muy complejos y quizás imposibles de solucionar. Además, la solución computacional podría ser muy complicada y el resultado podría ser obtenido una vez que ya sucedido el evento a predecir.

Las redes neuronales artificiales corresponden a un modelo de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso en los animales. Consiste en un sistema de interconexión de neuronas artificiales formando una red que colabora para producir un estímulo de salida. Estas neuronas corresponden a unidades de procesamiento que se comunican unas a otras mediante el envío de señales por medio de una serie de conexiones, donde cada una de estas conexiones tiene un peso asociado [11].

Según [12], el modelo de una neurona presenta 3 elementos básicos:

- Una serie de sinapsis, donde cada una está caracterizada por un peso el cual puede tomar valores negativos como positivos.
- Un sumador que suma las señales de entrada pesadas por la respectiva sinapsis de la neurona.
- Una función de activación para limitar la amplitud de la salida de la neurona. Cabe destacar que ésta corresponde a una función no lineal que normaliza el estado interno de la neurona a un intervalo cerrado $[0,1]$ o alternativamente $[-1,1]$. Ejemplos de función de activación serían: la función escalón, sigmoideal, gaussiana, polinomial, entre otras.

Una de las mayores aplicaciones de la RNAs es el pronóstico. Las RNAs proveen una herramienta alternativa para los investigadores ya que poseen una serie de características que las hacen valiosas para las tareas de predicción. En primer lugar y al contrario de los métodos tradicionales estadísticos, las RNAs son métodos adaptativos basado en datos, donde a veces se tiene muy poco conocimiento del problema bajo estudio. Éstas aprenden en base a ejemplos capturando relaciones funcionales sutiles entre los datos aún cuando estas relaciones sean desconocidas o difíciles de describir. Además son muy adecuadas para aquellos problemas cuyas soluciones requieren un conocimiento difícil de especificar, pero para los cuales existen suficientes datos u observaciones [13].

Otra característica importante es que las redes neuronales son capaces de generalizar, luego de la fase de entrenamiento pueden inferir el futuro de la población

aun cuando los datos contengan información ruidosa. Como el pronóstico se realiza vía la predicción del comportamiento futuro en base al comportamiento pasado, el pronóstico es una aplicación ideal para las redes.

Además, las RNAs son aproximadores funcionales universales. La experiencia de trabajos previos ha demostrado que una red puede aproximar cualquier función continua con una precisión deseada. Las RNAs tienen una estructura más flexible que los métodos estadísticos tradicionales. Cualquier modelo de pronóstico asume que existe una relación fundamental entre las variables de entrada (valores pasados de una serie de tiempo en este caso u otras variables relevantes) y las salidas (valores futuros). Frecuentemente los métodos estadísticos tradicionales tienen limitaciones a la hora de realizar las estimaciones de estas relaciones fundamentales debido a la complejidad de los problemas reales.

Por último las redes neuronales no son lineales, por mucho tiempo el pronóstico ha estado dentro del dominio de la estadística lineal, los cuales asumen que las series de tiempo bajo estudio son generadas por procesos lineales. Los procesos lineales tienen la ventaja de que pueden ser entendidos y analizados en gran detalle, lo que los hace fácil de explicar e implementar. Sin embargo pueden ser inapropiados si el proceso fundamental bajo estudio no es lineal. No es razonable a priori asumir que la realización de cierta serie de tiempo es generada por un proceso lineal. De hecho en la vida real los sistemas a menudo no son lineales.

Actualmente las aplicaciones de las redes neuronales se extienden ampliamente en los campos de los negocios, industrias y ciencias, llevando a cabo con éxito investigaciones en variadas áreas del conocimiento humano, a continuación se presentarán algunas aplicaciones recientes de las redes neuronales aplicadas al pronóstico.

2.2. Aplicaciones de Redes Neuronales como herramienta de Pronóstico

A través del tiempo las redes neuronales aplicadas a pronóstico han tenido éxito en una gran cantidad de áreas de la ciencia, en este trabajo se investigó acerca de los casos de éxito de las redes en las siguientes áreas; clima, minería, producción y tratamiento del agua, energía y producción eléctrica, radiación solar, seguridad, vibraciones de superficie y sismos, por último tránsito y duración de tiempos de viaje en grandes urbes. A continuación se discute la importancia de estas investigaciones y sus resultados.

Un área de gran importancia por su impacto directo en la economía como en el medioambiente es la generación de energía, hoy en día existe una preocupación a nivel global acerca de la forma de buscar nuevas fuentes limpias energéticas que permitan reducir los niveles de contaminación global y la emisión de los gases contaminantes, además se hace de vital importancia poder predecir el consumo energético de las ciudades que permitan elaborar mejores políticas energéticas, es por esto que en [14] se propuso una red neuronal wavelet difusa para calcular el consumo energético anual de las ciudades, se tomaron en cuenta como entradas del sistema la data histórica de consumo eléctrico así como también valores

climáticos, en sus simulaciones utilizaron a modo de comparación un modelo de gray, el MAPE obtenido para 5 años con la red neuronal wavelet anda en torno al 1.57 % a 2.26 % versus el modelo de gray el cual solo obtiene MAPE entre 15 % y 22 %, lo que les permite afirmar no solo que el pronosticador neuronal utilizado es mejor que un modelo de gray sino que además el MAPE obtenido es muy bajo cercano a 0. Otra investigación similar de pronóstico de consumo eléctrico pero enfocada a la creciente de demanda de producción energética de los países en vías de desarrollo [15] hace énfasis en la importancia de hacer pronósticos que ayuden a la toma de decisiones y desarrollo de nuevas políticas energéticas, en sus experimentos lograron obtener un MAPE de 3.8 % utilizando una red neuronal de retro-propagación lo que nos confirma el hecho que un pronosticador neuronal efectivamente es capaz de pronosticar el consumo eléctrico futuro de una ciudad o país.

El desarrollo de las energías renovables apuntan a solucionar el problema de la crisis energética debido a que su generación es eficiente y limpia, para el año 2020 se espera que la demanda mundial por energía eólica será del 12 % [14]. Sin embargo dadas las características de las fuentes de energías renovables tales como la inestabilidad e intermitencia hace que sea todo un desafío la administración de la matriz energética, es por esto que poder realizar pronósticos acertados a corto y largo plazo ayudara a la toma decisiones respecto a cuándo estas energías inyectaran electricidad a la matriz energética, a continuación se darán a conocer las investigaciones realizadas con redes neuronales para la generación eólica y solar. Para el pronóstico de generación de energía eólica generalmente se toman como variables de entrada la velocidad del viento y la potencia de la curva de éste, sin embargo el factor clave es la velocidad del viento, por lo cual se hace necesario predecirlo y dado que la velocidad del viento está dada por factores ambientales como; temperatura, presión ambiental, densidad del aire, topografía se hace muy complicado de realizar pronóstico ya que corresponden a procesos caóticos y no lineales. Generalmente el pronóstico de la generación de energía a través del tiempo se puede dividir en tres tipos; pronóstico a largo plazo (3-10 días), a mediano plazo (6-72 horas) y a corto plazo (0-6 horas), donde se considera que el pronóstico más útil es a corto plazo [15]. Debido a la importancia de la velocidad del viento en la predicción, existen distintas técnicas que permiten hacer pronóstico, no obstante acá veremos aquellas que usan redes neuronales en sus modelos.

En [16] se utilizó una red neuronal basada en un algoritmo de back-propagation para construir un modelo de predicción del viento. Por otro lado en [17] implementó una red neuronal con algoritmo back-propagation para los pesos de la red pero se utilizó un modelo diferencial para procesar los datos de entrada. [18] uso una mejora del algoritmo PSO con una red neuronal back-propagation para acelerar la rapidez de convergencia de la red, logrando aproximar las relaciones no lineales establecidas entre la data de entrada y de salida en la fase de entrenamiento, además su modelo de predicción tomaba en cuenta la temperatura y otros factores meteorológicos. [19] utilizó un algoritmo de aprendizaje adaptativo para entrenar la red neuronal BP, ya que las redes tradicionales pueden caer en óptimos locales lo que llegó a mejorar la precisión de su modelo de predicción de

la velocidad del viento.

La generación de energía fotovoltaica es un tipo de energía que usa el efecto fotovoltaico de una interfaz semiconductor para convertir energía solar directamente en energía eléctrica. Se estima que para el año 2030 el 30 % de la matriz energética estará dada por energías renovables y un 10 % estará dado por este tipo de energía a nivel mundial. Sin embargo a diferencia de las técnicas utilizadas para la predicción de la energía eólica a través de la velocidad del viento, la energía solar aún está en una fase inicial de investigación y donde los casos propuestos en la literatura son relativamente pocos. G. Capizzi and F. Bonanno [20] usaron una descomposición wavelet bi-ortogonal para diseñar redes neuronales y poner la radiación solar como una señal del espacio wavelet y una señal con coeficientes comprimidos que son usadas para modificar correctamente la estructura adaptativa del algoritmo de aprendizaje recurrente para un pronosticador neuronal. Por último [21] propuso cuatro redes neuronales; perceptrón multicapa, función base radial, red neuronal recurrente y un ensamble de red neuronal para hacer pronóstico de 24 horas hacia adelante en un sistema fotovoltaico de 0 a 20 KW.

En un área totalmente distinta pero de no menor importancia es la predicción del tiempo de viaje promedio dentro de una ciudad con altos niveles de tráfico, a medida que las urbes comienzan a expandirse, es normal que comience a aumentar el tráfico dentro de éstas, es por esto que las redes neuronales cumplen un factor determinante a la hora de poder realizar pronósticos certeros para la toma de decisiones, en [22] se propuso un modelo híbrido usando una red neuronal estado-espacio y gráficos de independencia condicional para establecer las variables más relevantes para realizar una predicción futura, éste modelo probó ser muy preciso obteniendo un MAPE en torno al 2 % y además venció a un modelo tradicional de red neuronal donde el coeficiente de determinación para el modelo propuesto fue de 99 % vs un 94 % para una red tradicional. En [23] usaron una red neuronal de base radial para predecir el tiempo de viaje en una carretera, se establecieron tres tramos independientes de carretera de 4km de extensión, se colocó un radar y sensores para realizar la toma de variables, donde se consideraron la cantidad de vehículos y su velocidad, los resultados obtenidos para los tres sectores van desde MAPE entre 11 % y 15 % y coeficientes de determinación entre 85 % y 88 %. Por último en un tópico que difiere de las dos investigaciones anteriores pero que de todas formas afecta el tiempo de viajes el propuesto en [24], donde se propuso un modelo para pronosticar la cantidad de taxis utilizando una red neuronal wavelet, donde se afirma que este tipo de redes es superior al resto por sus características, ya que las funciones wavelet tienen un buen tiempo-frecuencia, localización de características y un rápido decaimiento. Para mejorar la precisión del pronóstico del número de taxis se utilizó una red neuronal wavelet para aproximar la no linealidad de la cantidad de vehículos y su factor de influencia, las simulaciones realizadas en la investigación mostraron evidencia suficiente para determinar que el modelo propuesto es más preciso que las redes neuronales de back-propagation.

Por otro lado, las series de manchas solares han servido como punto de referencia y han sido bien estudiadas en la literatura estadística. Ya que los datos que son no lineales, no estacionarios y no gaussianos, a menudo se utilizan como

un criterio para evaluar y comparar los nuevos métodos de predicción . Algunos autores [25] se centran en el uso de redes neuronales para mejorar la precisión en la predicción de las actividades de las manchas solares a través de los métodos tradicionales, mientras que otros utilizan los datos para ilustrar un método. Una industria que ha sido el motor de desarrollo en nuestra humanidad en los últimos siglos es la minería, muchos son los minerales que son extraídos en todo el mundo, pero uno que atrae bastante la mirada por su relevancia para pronosticar es el carbón, el carbón fue utilizado ampliamente en el siglo XIX para la producción de motores a vapor y aun es usado para producir electricidad, debido a que es una forma sucia de energía que daña el medioambiente también corren grandes riesgos los trabajadores que laboran en estas minas por los gases que existen en las capas inferiores de la tierra, es por esto que en China donde el 95 % de las mineras de carbón están bajo tierra se hace indispensable poder predecir la cantidad de gas existente en las capas inferiores para poder prevenir desastres y proteger las salud de los trabajadores. En [26] implementaron una red neuronal back-propagation para pronosticar la cantidad de gas en las minas y lo compararon con una regresión múltiple, para cada punto medido de un total de 42 se midió el error relativo entre el valor medido y el calculado tanto por la red neuronal y la regresión múltiple donde en la mayoría de los casos el error del modelo propuesto era un 50 % menor a la regresión, probando que las redes neuronales son capaces de pronosticar con un pequeño error la cantidad de gas en las minas de carbón. Otra área de la ciencia explotada por las redes neuronales es la predicción de la cantidad de agua dulce que debe ser producida, un estudio [27] bastante interesante establece la importancia crítica que tiene la producción de agua en países desérticos y productores de petróleo donde la mayoría del agua debe pasar por un proceso de desalinización para hacerla potable, además para el caso de la ciudad de la Mecca debido a la explosión de turistas que recibe durante los meses de celebraciones sagradas se hace de vital importancia saber cuánta agua debe ser desalinizada, en el modelo propuesto se utilizaron redes neuronales, y se tomó como variables relevantes los registros históricos de agua producida y la distribución de los visitantes a la ciudad. Los resultados obtenidos muestran que el rendimiento de la red neuronal es mejor que el de un modelo econométrico utilizado anteriormente. Siguiendo en una línea similar a la investigación anterior, [28] propone un modelo de predicción de lluvias para una ciudad argelina donde dadas las características y ubicación geográficas se hace indispensable tener un buen manejo de las aguas lluvias, es por esto que el modelo utilizado es una red neuronal con desfase de tiempo (TDNN por sus siglas en inglés) con dos algoritmos de aprendizaje; uno de primer orden (gradiente recursivo con pesos constantes compartidos) y el otro de segundo orden (algoritmo de Lavenberg-Marquardt), la correlación observada entra la data deseada y la obtenida por la red TDNN, concluye que el segundo algoritmo describe y pronostica mejor la lluvia. En una rama mucho menos explorada varios investigadores han estado utilizando redes neuronales para poder predecir la magnitud de los terremotos, Alarifi, Alarifi y Al-Humidan [29] , proponen un modelo basado en redes neuronales recurrentes con capas ocultas múltiples, para pronosticar las magnitudes de futuros terremotos utilizando las

magnitudes de eventos telúricos pasados en el Norte del Mar Rojo, península del Sinai, el golfo de Aqaba y golfo de Suez. En sus resultados afirman que su red neuronal obtuvo una ganancia del 32% respecto a otros métodos estadísticos tradicionales. Por último para concluir con esta sección podemos decir que trabajo previo a éste ha sido aplicado a la captura de especies pelágicas en Chile, [30] propuso un modelo con redes neuronales recursivas multiplicativas con algoritmo de aprendizaje PSO e IPSO, obteniendo coeficientes de determinación entre 92% y 94%, como data de entrada se tomó la serie histórica de captura de anchoveta en la zona norte, más adelante se verá que el modelo propuesto en éste trabajo supera ampliamente estos resultados.

2.3. Red Neuronal Polinomial

En este proyecto se utilizara como función de transferencia una Red Neuronal Polinomial cuya estructura es de la siguiente forma:

$$y = \sum_{i=1}^H W_k H_k \quad (1)$$

$$H_k = \prod_{i=1}^n Z_i \quad (2)$$

$$Z = X_{ij}^{v_{kj}} \quad (3)$$

$$\text{Donde } X_i = [X_{i1}, X_{i2}, \dots, X_{in}], \quad i - \text{esima muestra}, i = 1, 2, \dots, N \quad (4)$$

$$V_k = [X_{k,1}, X_{k,2}, \dots, X_{k,n}], \text{ Pesos Ocultos} \quad (5)$$

$$W_k = \text{Pesos de Salida} \quad (6)$$

$$H = \text{Nodos Ocultos} \quad (7)$$

3. Optimización por Enjambre de Partículas

3.1. Introducción

La optimización por enjambre de partícula, más conocida en la literatura científica como *Particle Swarm Optimization* (PSO), nace, al igual que otras técnicas estocásticas del cálculo evolutivo, en un intento por imitar y mimetizar el comportamiento de los procesos naturales. Siendo uno de los métodos más utilizados en la inteligencia computacional, busca imitar comportamientos sociales de un colectivo a partir de la interacción entre sus individuos y de éstos con su entorno. Sus orígenes remontan de los estudio realizados por [3], quienes intentaron simular de manera gráfica el movimiento sincronizado e impredecible de bancos de peces o bandadas de aves, intrigados por la capacidad que tienen estos grupos para separarse, reagruparse o encontrar su alimento. Paralelamente con trabajos previos en el ámbito de la biología y la sociología, concluyen que el comportamiento, inteligencia y movimiento de estas agrupaciones está relacionado directamente con la capacidad para compartir información y aprovecha de la experiencia acumulada por cada uno de los congéneres. En [30] introducen el término **partícula**, el cual representa a cualquier tipo de individuo que muestre algún tipo de comportamiento social como grupo en forma de una colección de agentes o partículas que interactúan entre sí. De acuerdo a los fundamentos teóricos del método, el movimiento de cada una de estas partículas hacia la consecución de un objetivo en común depende de 2 factores: la **memoria autobiográfica** de la partícula o **nostalgia** y la **influencia social** de todo el enjambre. Esto se puede extender, dependiendo del problema bajo análisis, a nivel computacional de D-dimensiones, siendo N el número de incógnitas. Básicamente, el proceso evolutivo se reduce a mover cada partícula dentro del espacio de soluciones con una velocidad que

variará de acuerdo a su **velocidad actual**, a la **memoria de la partícula** y a la **información global** que comparte el resto del enjambre, utilizando una función de **fitness** para cuantificar la calidad de cada partícula en función de la posición que ésta ocupe. Existen diversos tipos de esquemas para la implementación de PSO. Dependiendo de cómo se actualicen las posiciones de las partículas surgen versiones sincronías y asíncronas del método. Además, se puede distinguir entre PSO local y global dependiendo de cómo fluya la información a través del enjambre. En el presente capítulo se comienza con la base de PSO relacionada con el **fundamento del movimiento de partículas**, para luego dar a conocer el **algoritmo básico**. Posteriormente se estudian los **parámetros** y las **topologías** más comunes.

3.2. Fundamentos del Movimiento de Partículas

Para aplicaciones relacionadas con el ámbito de la vida artificial se deben respetar cinco principios básicos de lo que se conoce como inteligencia de grupo, estos se conocen como: proximidad, calidad, diversidad de respuesta, estabilidad y adaptabilidad.

- **Proximidad.** la población debiera ser capaz de realizar cálculos sencillos de espacio y tiempo, lo cual en PSO se traduce a movimientos en D-dimensiones llevados a cabo durante un intervalo de tiempo que coinciden con movimientos de la población a una determinada velocidad.
- **Calidad.** Los factores de calidad en PSO se consiguen en base a la memoria de la partícula junto con la historia o conocimiento social que comparten todos los congéneres.
- **Diversidad de respuesta.** Representadas por las tendencias marcadas por la memoria personal de cada partícula y por la historia de la mejor posición visitada por el conjunto.
- **Estabilidad y Adaptabilidad.** Resaltan aspectos contrapuestos, dado que por un lado, la población solo cambia su comportamiento como grupo cuando se actualiza la mejor posición históricamente visitada por alguno de los miembros que lo integran (principio de estabilidad) y, por otro lado, la población debe ser capaz de modificar su comportamiento y movimiento cuando hay alguna señal que así lo recomienda. Esto último, desde el punto de vista de ahorro computacional o mejora de precisión. En PSO esto se consigue cuando alguna partícula alcanza una solución global que mejora el resultado, en ese instante la población cambia de rumbo.

3.3. Algoritmo PSO

Una partícula se define como un vector formado por la velocidad, posición y memoria de la partícula. PSO se inicializa con partículas aleatorias y optimiza sus

soluciones a medida que aumentan las iteraciones. El enjambre es representado por $X = (x_{i1}, x_{i2}, \dots, x_{id})$ donde n corresponde a la i -ésima partícula y corresponde a sus dimensiones. En cada iteración cada una de las partículas del enjambre se actualizada en base a dos valores importantes:

- Mejor posición encontrada por una **partícula**, conocida como P_{best} . El P_{best} de una i -ésima partícula se representa por $P_i = (p_{i1}, p_{i2}, \dots, p_{id})$.
- Mejor posición que ha encontrado el **enjambre**, conocida como P_{gbest} .

Ambos de estos valores se actualizan en cada iteración y se utilizan ajustar la velocidad con la que se mueve una determinada partícula en cada una de las dimensiones. La influencia que tiene la mejor posición personal sobre la velocidad de una partícula se conoce como el **factor de cognición** y la influencia de la mejor partícula del enjambre se conoce como **componente social**. En [31] se incorpora el **factor de inercia** w en la ecuación que determina la velocidad, de tal manera de que se balancee la búsqueda global y local.

$$x_{id}(t + 1) = w \cdot v_{id}(t) + c_1 \cdot r_1(p_{id}(t) - x_{id}(t)) + c_2 \cdot r_2(p_{gd}(t) - x_{id}(t)) \quad (8)$$

$$x_{id}(t + 1) = x_{id}(t) + v_{id}(t + 1) \quad (9)$$

Donde r_1 y r_2 corresponden a valores aleatorios independientes en el intervalo $[0,1]$, y c_1 y c_2 corresponden a las variables que controlan la influencia de las componentes cognitiva y social. Para prevenir que las velocidades de las partículas incrementen infinitamente, se incorpora un parámetro V_{max} que delimita el rango de velocidad $[-V_{max}, V_{max}]$ que puede tomar una partícula.

$$if(v_{ij} > V_{max}) \text{ or } (v_{ij} < -V_{max}) \\ v_{ij} = sign(v_{ij}) \cdot V_{max}$$

A continuación se presentan los seis pasos relevantes que constituyen el algoritmo de PSO Tradicional:

- **Paso 1:** Se debe definir el tamaño n del enjambre y se deben inicializar las posiciones de las velocidades de las partículas en un espacio D -dimensional.
- **Paso 2:** Evaluar fitness de cada partícula en base a algún criterio de optimización (función de costo, mínimos cuadrados, error cuadrático medio).
- **Paso 3:** Comparar el fitness de cada partícula obtenido en el Paso 2 con el fitness de su mejor posición personal. Si el fitness actual es mejor que el histórico, se debe actualizar esta mejor posición por la nueva obtenida.
- **Paso 4:** Comparar el fitness actual de cada partícula con el mejor fitness encontrado por el enjambre, si el actual es mejor, entonces se actualiza la mejor posición global con el valor de la posición actual.

- **Paso 5:** Ajustar la velocidad y posición de la partículas en base a las ecuaciones de velocidad y posición descritas en este capítulo. Se debe verificar que las partículas no sobrepasen la velocidad máxima definida .
- **Paso 6:** Si se cumple algún criterio de término término, en caso contrario se debe volver al Paso 2. Los criterios de términos más conocidos pueden ser:
 - Número máximo de iteraciones
 - Número de iteraciones sin mejoras
 - Error mínimo de la función objetivo

3.4. Parámetros PSO

Tal como se visualizó con los GA, el ajuste de los parámetros condicionará el rendimiento posterior del algoritmo de optimización. La selección está directamente ligada a la naturaleza del problema que se desea optimizar, siendo preciso lograr un balance óptimo entre la exploración y la convergencia. La definición de la función **fitness** cumple un papel no menor, dado que una función que no introduzca una métrica adecuada para pesar la bondad de cada partícula, hará emerger las carencias del PSO, las cuales erróneamente pueden asociarse con el algoritmo. Para **acotar la velocidad de la partícula** se define un valor máximo V_{max} , el cual restringe a la velocidad en cada dimensión del intervalo $[-V_{max}, V_{max}]$. Si el valor de V_{max} tiende a ser demasiado pequeño, las partículas exploraran el espacio de soluciones muy lento y podrán quedar atrapadas dentro de soluciones locales. Por otro lado, la no restricción de la velocidad llevaría a la no convergencia del enjambre en un punto, ocurriendo lo que se conoce como **explosión del PSO**, el cual corresponde a un comportamiento oscilatorio y creciente de la posición de las partículas provocando la ineficiencia del PSO como algoritmo de optimización. El **peso de inercia w** , regula la compensación entre la capacidad de exploración global y locales del enjambre. A mayor w , facilita la exploración global y en caso contrario facilita la exploración local. En un comienzo w se asigna como una constante, sin embargo, resultados experimentales recomiendan la asignación de mayores valores en un comienzo para que en una primera instancia se examine de manera global el espacio de búsqueda y gradualmente se disminuya para comenzar a conseguir soluciones más refinadas. Los parámetros c_1 y c_2 , luego de una buena sintonización, pueden acelerar la convergencia del algoritmo y aliviarlo de los mínimos locales. Típicamente a estos parámetros se les asigna 2 o 1.49. Trabajos recientes revelan que podrían ser un parámetro cognitivo c_1 más grande que el social c_2 , siempre y cuando se cumpla que $c_1 + c_2 \leq 4$ [31]. El tamaño de la población debe ser seleccionado de manera rigurosa, dado que valores muy grandes pueden explorar de manera minuciosa el espacio de búsqueda pero costo computación se eleva de manera considerable debido al aumento del número de evaluaciones de la función fitness. Generalmente se utilizan poblaciones ente 10 y 50 o 100 y 200 para problemas completos.

3.5. Topologías de vecindad local y global

En PSO los individuos mejoran sus fitness dado que en cada iteración imitan los comportamientos y tendencias que encuentran en los mejores congéneres de la población. Establecer cómo se define la vecindad de un individuo tiene una trascendencia vital en el rendimiento del algoritmo. Dependiendo de la topología que adquiera la población, la transmisión de la información entre individuos puede acelerarse o ralentizarse, lo cual está íntimamente relacionado con la velocidad de convergencia y con la capacidad del algoritmo para escapar de soluciones locales. La topología más extendida es la **topología de vecindad global**, en la cual todos los individuos están interrelacionados y tienen acceso inmediato a los hallazgos de sus congéneres. Sin embargo, esta estructura social es vulnerable a soluciones locales, dado que dependiendo de la distribución puntual de las partículas sobre el espacio de soluciones, una de ellas apuntando a una solución local puede llegar a dominar al resto.

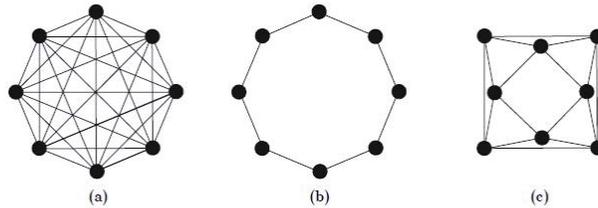


Figura 1: Topologías de la población Global,(b) Local con $N=2$,(c) Local con $N=4$

El modelo del **óptimo local** permite que cada individuo sea influenciado por un reducido número de miembros adyacentes. Estas partículas vecinas no tienen relación directa con el resto de las partículas que pertenecen a otra vecindad. Por lo general, como se visualiza en la Fig. 1, el óptimo local tienen 2 vecinos. Autores en [32] declaran que el modelo del óptimo global converge rápidamente hacia las soluciones del problema, pero tiene la debilidad de quedar atrapado en mínimos locales, mientras que el óptimo local tienen la desventaja de converger lentamente hacia la solución del problema.

3.6. PSO con actualizaciones síncronas y asíncronas

La diferencia entre una actualización asíncrona con una síncrona, radica en el instante en el cual se realiza la actualización de la memoria de cada partícula y el conocimiento social del grupo. En el modelo síncrono, todas las partículas se mueven en paralelo. El fitness de las partículas se evalúa en cada iteración, se actualiza su memoria p_{id} y el conocimiento social p_{gd} . Por otro lado, en el PSO asíncrono al momento de desplazarse, cada partícula aprovecha la información actualizada por sus inmediatos predecesores. Es decir, en cada iteración k , la i -ésima partícula se desplaza hacia un nuevo punto utilizando la información de los vectores p_{id} y p_{gd} , actualizados por las $i-1$ partículas previas. Posteriormente, la

partícula evalúa la calidad del nuevo punto y actualiza, si procede, las variables p_{id} y p_{gd} . Esta información se transmite a las restantes partículas. Al actualizar la información partícula a partícula, el modelo asíncrono acelera la optimización, aunque la naturaleza del modelo síncrono lo hace susceptible de ser ejecutado en paralelo, sobre múltiples procesadores.

3.7. Exploración y Explotación con PSO

Uno de los problemas de los algoritmos evolutivos, es que en ocasiones quedan atrapados en óptimos locales, lo cual en la literatura se conoce como **convergencia prematura** del algoritmo y corresponde a la incapacidad de salir la zona en la cual se encontró el mínimo local, impidiendo que se exploren otras regiones del espacio de búsqueda. Lo ideal es que las partículas representen soluciones de diferentes zonas del espacio de búsqueda. Esto se conoce como **diversificación** y en términos de los algoritmos corresponde a que las partículas sobrevuelen diferentes zonas del espacio de soluciones antes de que se comprometan con alguna en especial. Por otro lado, las partículas, una vez encontrada una zona comprometedoras deben ser capaces de mejorar la solución encontrada. Esto se conoce como **convergencia de la población**. Se define como **exploración** a la búsqueda de la solución del problema en el espacio de búsqueda de manera amplia, privilegiando la diversificación del enjambre. Por otro lado, la **explotación** consiste en la búsqueda de la solución del problema en un rango (sub-espacio) dentro del espacio de búsqueda donde el enjambre trate de tomar todos los valores posibles, privilegiando la convergencia del enjambre. Estudios realizados en [31] revelan que altos valores para el coeficiente de inercia, factores de cognición y V_{max} ($0,75 < w < 1$ y $2 < c_1, c_2 < 4$), favorecen el modo de exploración del PSO, mientras que valores bajos ($0,4 < w < 0,75$ y $0,1 < c_1, c_2 < 2$) favorecen el modo de explotación del PSO. La idea es ajustar los parámetros de tal manera que el algoritmo funcione en modo de exploración en las primeras etapas y una vez transcurridas unas cuantas iteraciones cambie a modo explotación.

3.8. Variaciones de Algoritmos de PSO

En la actualidad existen una serie de variaciones del algoritmo PSO tradicional, las cuales incorporan nuevos parámetros, nuevos métodos de actualización, e híbridos con otros algoritmos. La mayoría de estas mejoras características como la solución encontrada y la velocidad de convergencia. Pese a esto, algunas versiones agregan un mayor nivel de complejidad a la versión tradicional. A continuación se describe la variación utilizada en este proyecto.

3.8.1. PSO con peso de inercia decreciente (DIWPSO)

En lo referente al peso inercial, se utilizan comúnmente diferentes valores dependiendo del compromiso que se establezca entre exploración y rapidez de convergencia. Un peso inercial alto reduce la influencia de la nostalgia y de la coope-

ración sobre el movimiento de las partículas, favoreciendo la exploración. Por el contrario, un peso inercial muy bajo promueve la convergencia hacia la solución propuesta por la propia partícula P_i , y por la mejor solución de conjunto, G . En la literatura se sugiere un peso inercial constante de valor 0.8 como también se utiliza un valor constante para w de 0.4. Por otra parte, para una exploración y convergencia del espacio de soluciones, se suele utilizar una variante con un peso inercial que decrece linealmente entre w_{max} y w_{min} durante el transcurso de la simulación, con valores típicos de 0.9 y 0.4 para w_{max} y w_{min} , respectivamente.

$$w = w_{max} \frac{w_{max} - w_{min}}{K} * k \quad (10)$$

3.9. Aplicaciones

PSO ha recibido una gran atención en muchas aplicaciones y áreas de investigación durante los últimos años. Su carácter evolutivo junto con la gran cantidad de mejoras que presenta la literatura han dado pie para una gran cantidad de aplicaciones, como por ejemplo: en problemas electromagnéticos, en la síntesis de agrupaciones lineales de antenas y más aún en el aprendizaje de las Redes Neuronales. Este último es de gran importancia para motivos del presente caso de estudio. A continuación se presentan las aplicaciones listadas en el párrafo anterior y además cabe destacar que el uso de PSO en algunos de estos casos resulta como híbrido con GA.

- **Electromagnetismo.** [33] presenta un algoritmo híbrido evolutivo llamado GSO (Genetical Swarm Optimization) producto de la combinación de GA con PSO. En general, debido a la rápida búsqueda global en el espacio sin quedar atrapado en mínimos locales, GSO puede ser adoptado para resolver adecuadamente problemas de optimización en la síntesis de arrays lineales y planos, en materiales EBG o en el diseño de estructuras de protección.
- **Síntesis de Arrays.** (Correa y otros, 2005) compara PSO y GAS en la síntesis de arrays concluyendo que la aplicación de ambos aplicados a la síntesis de alimentaciones en agrupaciones lineales de antenas demuestra la mutua validez pero que la sencillez de implementación, ajuste, rapidez de convergencia hace que la aplicación de PSO prime sobre los GAS.
- **Entrenamiento de RNAs.** En [34] se ocupa un híbrido, formado por 2 variantes de PSO, para calcular los pesos de una RNA y definir la estructura, ambas entrenadas de manera simultánea demostrando una rápida convergencia.

4. Pronosticador Neuronal

Defínase a una Red Neuroevolutiva como toda aquella Red Neuronal Artificial cuya estimación de parámetros internos se realiza mediante la utilización de

Algoritmos Evolutivos. En este capítulo se presenta el modelo de Red Neuroevolutiva. El algoritmo utilizado para dicha estimación es DIWPSO. En una primera sección se presentan las principales métricas utilizadas para medir la calidad del modelo. Luego en las siguientes secciones se presentan los modelos desarrollados. Hay que señalar que el modelo genérico es una red Polinomial utilizando DIWPSO. Finalmente se presentan los resultados del estudio comparando los modelos presentados de manera que se determine cuál de ellos provee una mejor capacidad predictiva frente al pronóstico del nivel de captura de Anchovetas.

4.1. Métricas de Evaluación

En el presente estudio, para la selección de las topologías y, por consiguiente, de los modelos que presentan mejores resultados se utilizaron un conjunto de métricas de exactitud calculadas entre los datos observados (valores reales) y los datos pronosticados (entregados por la red). Estas métricas se nombran a continuación.

- Raíz Error Cuadrático Medio (RMSE). Consiste en la suma de las diferencias entre los datos observados y los datos proyectados por el modelo.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (d_i - y_i)^2}{N}} \quad (11)$$

- Coeficiente de Determinación. Mide la dependencia entre los datos reales y los pronosticados. El 0 muestra independencia y el 1 lo contrario.

$$R^2 = 1 - \frac{\sum_{i=1}^n (d_i - y_i)^2}{\sum_{i=1}^n (d_i - \bar{y}_i)^2} \quad (12)$$

- Porcentaje de Error Medio Absoluto (MAPE). Proporciona una indicación de que tan grandes son los errores de pronóstico comparados con los valores reales de la serie.

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{d_i - y_i}{d_i} \right|}{N} \times 100 \quad d_i \neq 0 \quad (13)$$

Se define para las formulas presentadas a al valor observado en el mes i , al valor pronosticado en el mes i , como la media de la data observada y como el número total de meses computados. Finalmente, cabe destacar que el equipo en el cual se realizarán todos los entrenamientos y pruebas corresponde a un Intel core i5 y 6gb Ram, los modelos están desarrollados en MATLAB.

4.2. Modelo Propuesto

Este modelo consiste en una Red Neuronal Polinomial cuyos parámetros serán estimados mediante el Algoritmo de Optimización por Enjambre de Partículas y

cuya topología, es decir, entradas y cantidad de nodos ocultos serán calculados en base al ensayo y error para la baja frecuencia, para la alta frecuencia o variable anual se usara una configuración de red neuronal basada en la característica de la data histórica. Los datos utilizados correspondan a la captura mensual de las anchovetas en la zona norte, estos datos están separados en datos de alta y baja frecuencia, se desarrollara un pronosticador para cada uno de estos tipos de datos y los resultados serán sumados al final para ser comparados en los datos esperados, como se describe en la siguiente figura:

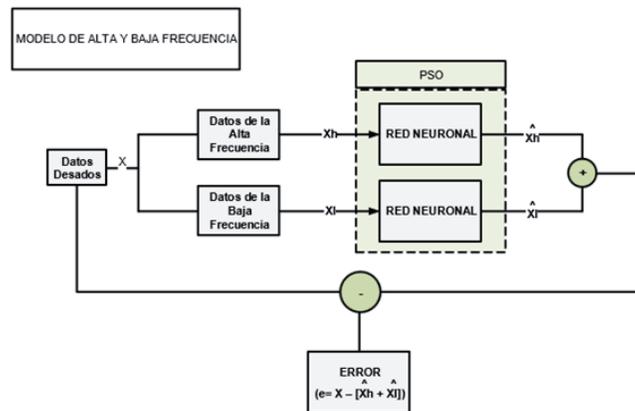


Figura 2: Modelo de Alta y Baja Frecuencia.

Una vista más cercana al modelo de la red neuronal nos permite ver con más detalle el modelo propuesto:

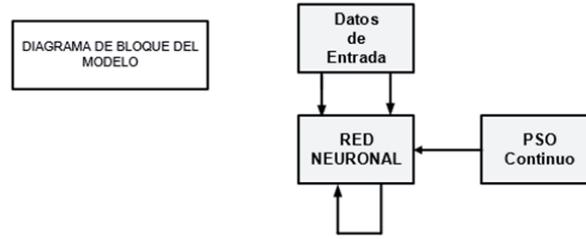


Figura 3: Diagrama de Bloque.

4.2.1. Descripción del Modelo

El proceso de entrenamiento comienza con la creación aleatoria de un enjambre inicial de N partículas, donde cada una de estas contendrá como posición pesos de la red (pesos no lineales y lineales), una velocidad generada aleatoriamente y una memoria personal (P_{best}) con el mismo valor que la posición. También definida la posición, velocidad y memoria de la partícula se procede al cálculo inicial del error. Finalmente se tendrá un enjambre inicial de N partículas definiendo como mejor posición global (P_{gbest}) al menor error encontrado. El proceso de actualización de las partículas se basa en el algoritmo descrito en la sección 3.3, específicamente aplicando las fórmulas (12) y (13) para la actualización de la velocidad y posición respectivamente. Luego se procede a la actualización del error de cada partícula, (P_{best}) y (P_{gbest}). Esto se hace reemplazando las nuevas posiciones obtenidas por las partículas en la red y cálculo del RMSE obtenido al procesar la data de entrenamiento y compararla con la data esperada (observada).

4.2.2. Selección de Topología y Parámetros de la Red

El método de selección de las configuraciones de prueba fueron hechas empleando la siguiente modalidad para cada uno de los modelos (alta y baja frecuencia):

- **Neuronas de Entrada:** Para obtener el número óptimo de entradas primero se fija el número de Neuronas ocultas en 10, el número de iteraciones en 4000 y finalmente el número de partículas en 20, luego se prueba el número de entradas desde un rango de 1 a 12 dejando la que presente el mejor MSE.
- **Neuronas Ocultas:** Para obtener el número óptimo de neuronas ocultas, se emplea el mismo mecanismo que se utilizó para la neuronas de entrada con la diferencia que estas se fijan con la mejor obtenida en el proceso anterior, luego se varían las neuronas ocultas en un rango de 1 a 12 dejando la que presente el mejor MSE.

- **Número de Partículas:** Finalmente al obtener el número óptimo de neuronas de entrada y ocultas, además del número óptimo de iteraciones se varía el número de Partículas en un rango de 10 a 30, dejando la que presente el mejor MSE.

En términos de representación, las topologías se denotarán de la forma **RM-AF-PSO (E, O, S)**, siendo **E** el número de neuronas de entrada, **O** el número de neuronas ocultas y **S** el número de neuronas de salida. Las tablas de resultados que se presentan a continuación, muestran los resultados al emplear dicho mecanismo detallando cada fase de este.

4.3. Baja Frecuencia

4.3.1. Neuronas de Entrada

Para la obtención de los nodos de entrada se fijaron los siguientes valores; diez nodos ocultos, cuatro mil iteraciones y veinte partículas.

A continuación los resultados obtenidos con estos parametros fijos y variando la cantidad de entradas.

NE	MSE	RMSE	MAE	MAPE	R	R cuadrado
1	0,00043547	0,02086799	0,01464844	0,05473172	0,97858113	0,95762103
2	0,00010861	0,01042163	0,00741267	0,02768622	0,99471022	0,98944842
3	6,30E-05	0,00793711	0,00556275	0,02075959	0,99693953	0,99388843
4	0,0001229	0,011086	0,00798229	0,02974724	0,99403356	0,98810271
5	7,80E-05	0,00883155	0,00632564	0,02356118	0,99621799	0,99245028
6	0,00026509	0,0162816	0,01259701	0,04658411	0,98764509	0,97544282
7	6,96E-05	0,00834265	0,00604656	0,0225132	0,99663745	0,9932862
8	0,00031838	0,01784329	0,0138547	0,0511936	0,98521682	0,97065217
9	0,00027699	0,01664315	0,01285626	0,04776254	0,98681457	0,97380299
10	0,0007129	0,02670021	0,02146024	0,07903109	0,96693713	0,93496741
11	0,00013253	0,01151196	0,00838691	0,03127713	0,99362842	0,98729743
12	0,00032837	0,018121	0,01433231	0,05313744	0,9848121	0,96985488

Tabla 1: Resultados al variar el número de entradas

4.3.2. Nodos Ocultos

Para la obtención de los nodos ocultos se fijaron los siguientes valores; cuatro nodos de entrada, cuatro mil iteraciones y veinte partículas.

NO	MSE	RMSE	MAE	MAPE	R	R cuadrado
1	0,00082788	0,02877299	0,02353883	0,08662309	0,96106857	0,9236528
2	0,00035791	0,01891859	0,01504446	0,05559704	0,98334234	0,96696216
3	0,00010188	0,01009362	0,00732167	0,02728724	0,99505492	0,9901343
4	0,00013928	0,01180169	0,00839064	0,03126448	0,99323566	0,98651708
5	0,00028177	0,01678602	0,01308675	0,04838529	0,98700929	0,97418734
6	7,61E-05	0,00872519	0,00603618	0,0225048	0,99630521	0,99262408
7	0,00023859	0,0154463	0,01194969	0,04424257	0,98889183	0,97790705
8	0,00015509	0,01245357	0,00923725	0,03433551	0,99258325	0,98522152
9	6,63E-05	0,00814161	0,00582228	0,02170421	0,99678317	0,99357669
10	0,00022752	0,01508366	0,01160568	0,04296651	0,98943269	0,97897704
11	7,13E-05	0,00844158	0,00584822	0,0218034	0,99654061	0,99309319
12	6,87E-05	0,00829099	0,00595521	0,02220007	0,99666365	0,99333843

Tabla 2: Resultados al variar el número neuronas ocultas

Se puede observar que el número de nodos ocultos que tiene las mejores métricas es 9 por ende este será el valor para los nodos ocultos de los datos de baja frecuencia.

4.3.3. Número de Partículas

Para la obtención del número de partículas se fijaron los siguientes valores; tres nodos de entrada, nueve nodos ocultos, doscientas iteraciones.

PARTICULAS	10	20	30
MSE	0,00035791	6,63E-05	7,13E-05

Tabla 3: Resultados al varia el número de partículas

4.3.4. Cantidad de Iteraciones

Para la obtención del número de iteraciones se fijaron los siguientes valores; tres nodos de entrada, seis nodos ocultos y veinte partículas.

ITERACIONES	500	1000	2000	3000	4000
MSE	0,0028	0,0009	0,0015	0,0019	0,00022752

Tabla 4: Resultados al variar el número de iteraciones

Finalmente se encontró que la topología **RM-BF-DIWPSO (3, 9,1)**, fue la que mejores resultados obtuvo, además de encontrar que el número óptimo de iteraciones y partículas para el modelo **RM-BF-DIWPSO (2, 9,1)** es de **4000** y **20** respectivamente.

4.4. Alta Frecuencia

4.4.1. Configuración de la Red

Al observar la data histórica de la anchoveta se puede apreciar que los periodos de mayor presión se dan cada doce meses por lo que para pronosticar la captura de alta frecuencia se utilizara un modelo que considera doce meses de la variable interanual con doce meses de la variable anual por lo tanto la cantidad de entradas será fijada en 24 y se utilizara la regla del logaritmo de la cantidad de muestras para la obtención de los nodos ocultos, es decir, en este caso serán seis.

4.4.2. Número de Partículas

Para la obtención del número de partículas se fijaron los siguientes valores; ocho nodos de entrada, nueve nodos ocultos, cuatro mil iteraciones.

PARTICULAS	10	20	30
MSE	0,08	0,0013	0,003

Tabla 5: Resultados al varia el número de partículas

4.4.3. Cantidad de Iteraciones

Para la obtención del número de iteraciones se fijaron los siguientes valores; ocho nodos de entrada, seis nodos ocultos y veinte partículas.

ITERACIONES	500	1000	2000	3000	4000
MSE	0,09	0,007	0,06	0,006	0,0013

Tabla 6: Resultados al variar el número de iteraciones

Finalmente se encontró que la topología **RM-AF-DIWPSO (24, 6, 1)**, fue la que mejores resultados obtuvo, además de encontrar que el número óptimo de iteraciones y partículas es de **4000** y **20** respectivamente.

5. Análisis de Resultados

En esta sección se busca contrastar los resultados obtenidos para los dos modelos definidos en el capítulo 4, los cuales se enfrentarán por separado a 5 ejecuciones respectivamente, luego los mejores pronósticos tanto para alta y baja frecuencia serán sumados y contrastados con los datos deseados para ver el rendimiento de este modelo. En todos los casos se utilizó la configuración obtenida en la sección anterior ya que presentaba los mejores resultados frente al problema. A continuación se presentan las tablas correspondientes a los resultados finales además de los gráficos de cada una de las configuraciones elegidas para cada modelo.

5.1. Pronóstico para el mes siguiente

5.1.1. Modelo de Baja Frecuencia

A continuación se presentan los resultados para el pronóstico de baja frecuencia para el primer mes.

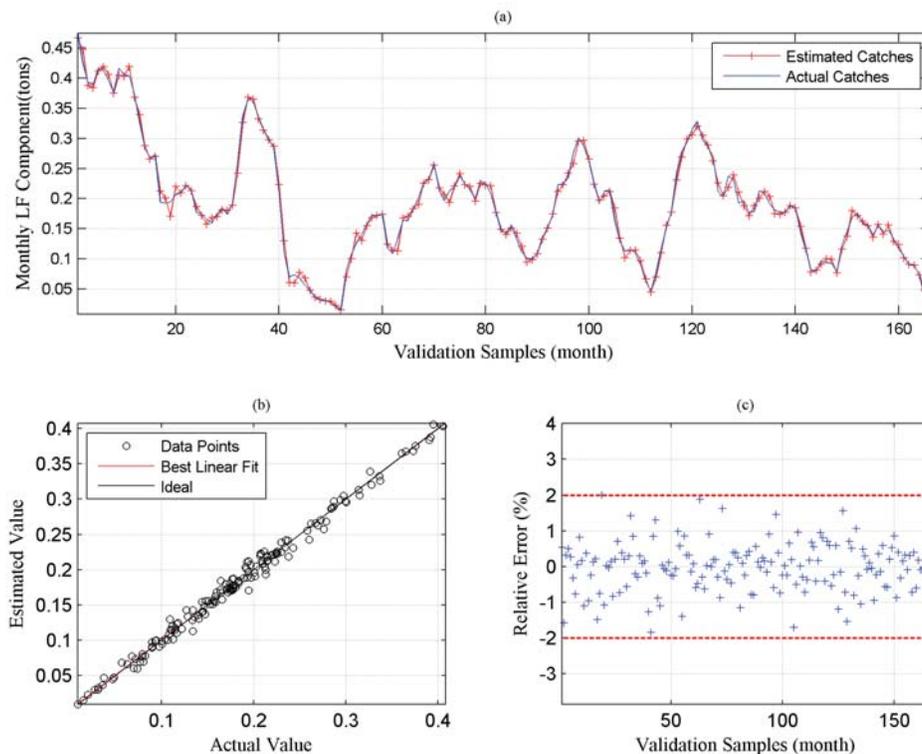


Figura 4: Gráficos de resultados para baja frecuencia.

En la figura 4 subfigura (a), podemos observar la recta de la data estimada es casi idéntica a la data histórica, es decir, los datos pronosticados por la red para el caso de baja frecuencia son ideales, por otro lado la subfigura(b), nos permite

afirmar que ambas rectas siguen la misma dirección con la misma pendiente, el valor del coeficiente de determinación de 99.26 % nos confirma este hecho, en la figura (c), el gráfico corresponde al error relativo, a partir de éste podemos inferir que para el 99.4 % de los datos de prueba el error de pronóstico será menor al 2 %.

5.1.2. Modelo de Alta Frecuencia

A continuación el mejor pronóstico para el siguiente mes de alta frecuencia:

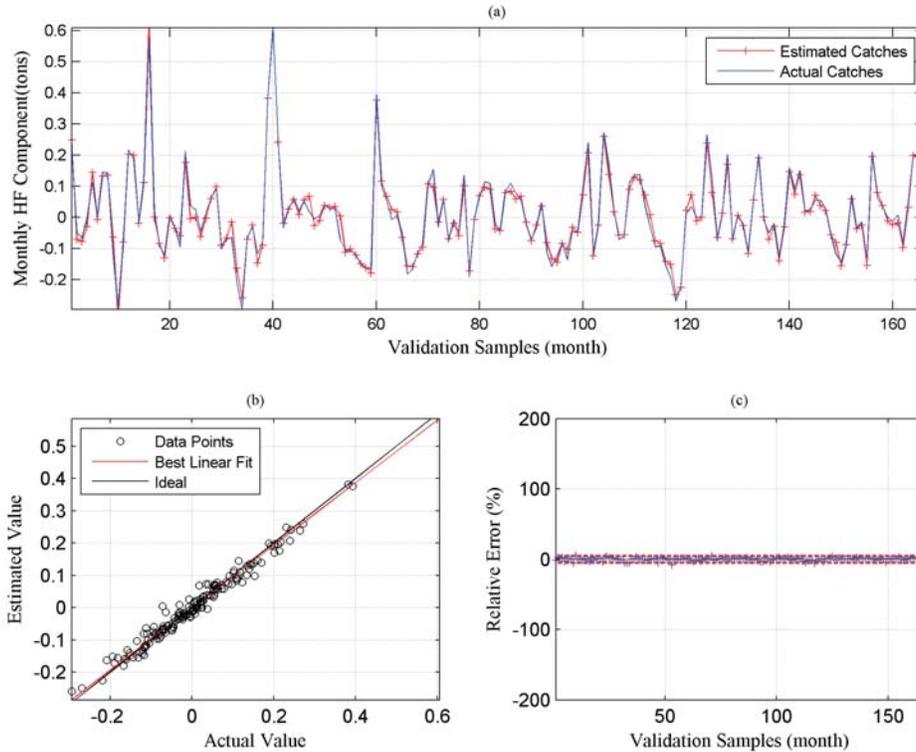


Figura 5: Gráficos de resultados para alta frecuencia.

En la imagen 5 subfigura (a), podemos afirmar que a pesar que la curva deseada oscila más que en el caso de baja frecuencia, el pronóstico realizado sigue la tendencia de forma casi perfecta, por otro lado la subfigura(b), el gráfico de dispersión nos muestra que nuevamente las rectas siguen la misma dirección con un coeficiente de determinación de 97.77 %, por último en la figura (c), podemos afirmar que el 96 % de los datos pronosticados tendrá un error relativo menor al 5 %

5.1.3. Modelo Propuesto LF + HF

A continuación se mostraran los resultados finales obtenidos en este proyecto y del modelo propuesto, para obtener estos resultados se sumaron la data obtenida de la red neuronal tanto de alta y baja frecuencia (LF + HF), una vez realizado esto se contrastaron con los datos sin separar, que representan los datos reales de captura de anchoveta en la zona norte:

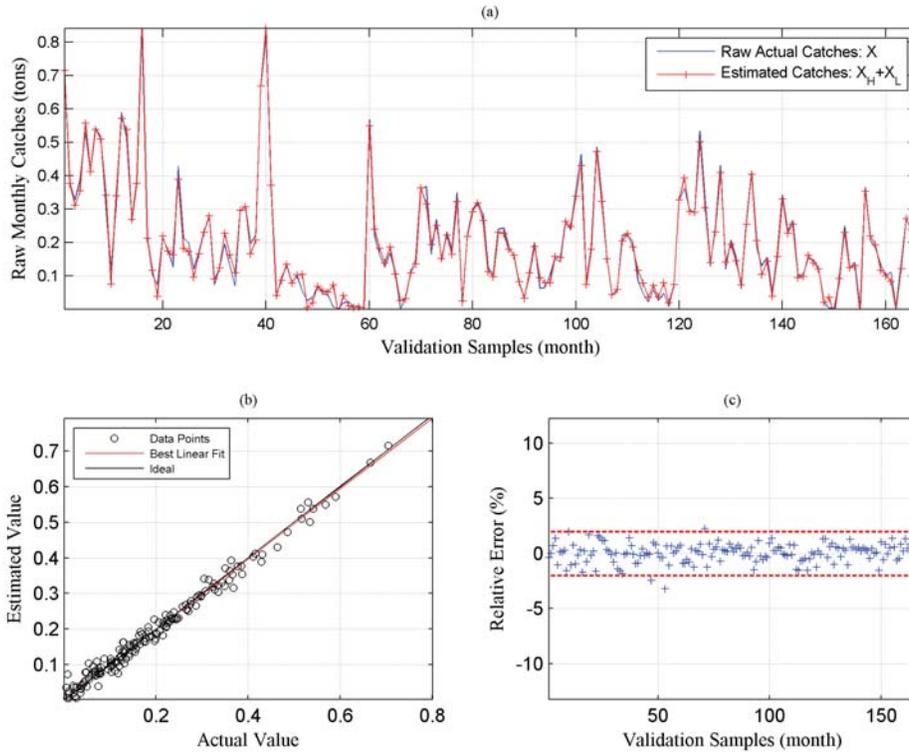


Figura 6: Gráficos de resultados para el modelo sumado.

Debido a los buenos resultados obtenidos por las variables separadas, se observa en los resultados obtenidos al sumar ambas variables que esta tendencia se mantiene, el MAPE obtenido es 0.66 % y el RMSE 0.01 %, en la figura 6 subfigura (a), podemos ver que el pronóstico gráficamente es preciso, la subfigura (b) nos indica que el coeficiente de determinación es de 98.63 %, es decir, la variabilidad explicada aumenta al sumar la variable anual con la interanual, el error relativo disminuye en comparación con la alta frecuencia y nos permite inferir que el 98.4 % de los datos pronosticados tiene un error menor al 2 %.

5.2. Pr3nóstico para el cuarto mes

Al afirmar que el modelo de predicci3n anda bien para el pron3stico del siguiente mes, se realiz3 el trabajo de extrapolar este modelo para realizar una estimaci3n hasta el cuarto mes, a continuaci3n una tabla que permite ver el comportamiento de las m3tricas de error, RMSE y MAPE a medida que extrapolamos el modelo desde el mes siguiente hasta el cuarto mes:

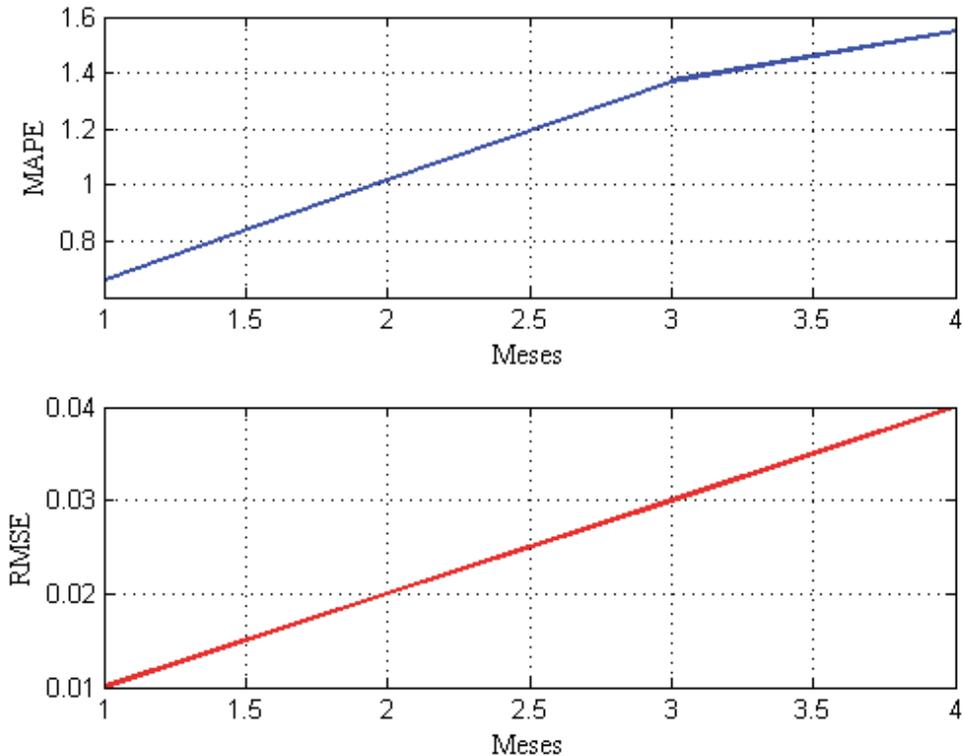


Figura 7: M3tricas MAPE y RMSE a trav3s de los meses.

Como se observa en la figura anterior, tanto los valores para el MAPE y el RMSE son bastante bajos, 1.55% y 0.042 respectivamente, lo que solamente representa una diferencia negativa de 0.9% para el caso del MAPE y 0.024 para el RMSE respecto al pron3stico del primer mes, es decir, aun siendo el cuarto mes podemos afirmar que el modelo es capaz de inferir el comportamiento de la captura de la anchoveta en la zona Norte con un gran grado de precisi3n.

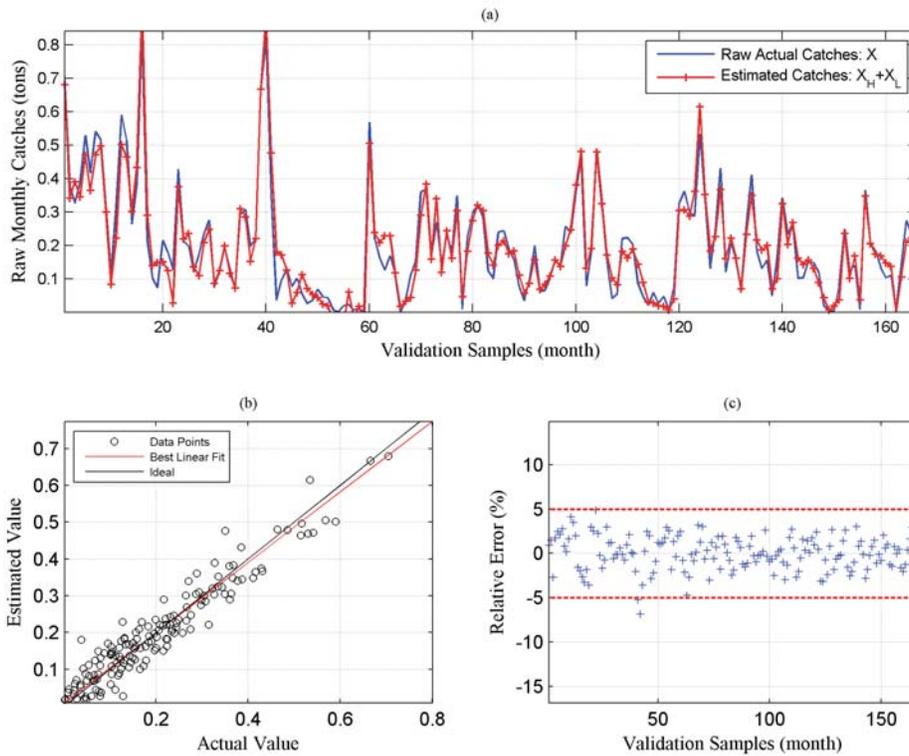


Figura 8: Gráficos de resultados para la suma de la variable anual e interanual.

Por último en la imagen 8 subfigura (a), podemos observar que a pesar de que estamos pronosticando cuatro meses en el futuro, la curva de pronóstico sigue muy cerca a la data real observada, en el gráfico (b) el modelo presenta un 93 % de la varianza explicada, el error relativo aumenta en comparación al primer mes, sin embargo tiene un valor muy pequeño que nos permite afirmar que el 98 % de la data pronosticada tendrá un error menor al 5 %.

Para poder comparar los resultados obtenidos en esta investigación se tomaron en consideración trabajos realizados en años anteriores con técnicas similares y utilizando también datos de anchoveta en la zona norte, para poder compararlos se utilizarán las métricas obtenidas para el pronóstico del siguiente mes de este proyecto contra la mejor técnica utilizada en las investigaciones previas, como muestra la figura a continuación.

5.3. Investigaciones Anteriores

Las investigaciones asociadas a la captura de la Anchoveta en la zona Norte se han venido realizando hace algunos años utilizando técnicas similares a las propuestas en este proyecto, en este apartado se presentan dichos trabajos y se compara con esta memoria. En [6] se presentaron tres modelos distintos con redes neuronales multiplicativas, RRM-IPSO(3,6,1) red multiplicativa con tres

nodos de entrada seis nodos ocultos y uno de salida con IPSO (variante del PSO tradicional), RRM-PSO(4,7,1) con cuatro nodos de entrada, siete nodos ocultos y uno de salida con PSO tradicional y RRM-DIWPSO(4,6,1) con cuatro nodos de entrada, seis nodos ocultos y uno de salida. Luego del análisis de los distintos modelos propuestos, se pudo concluir que el modelo RRM-IPSO(3,6,1) presenta un mejor desempeño en el pronóstico de las capturas de anchovetas por sobre los otros dos modelos. Estos resultados fueron obtenidos tanto en las métricas propuestas como en la velocidad de convergencia de cada modelo, obteniendo un 94,94 % de la varianza explicada y un 20 % de MAPE como el mejor modelo de esa investigación.

Por otro lado en [9] se utiliza un vector de soporte autoregresivo, con aprendizaje PSO y otras variantes. Para la selección del mejor modelo, se utilizaron diversas métricas de evaluación, a partir de ellas, se logró concluir que el modelo LS-SVM + LAPSO, era capaz de representar de mejor manera los datos del problema, arrojando mejores resultados por sobre las otras variantes (IWPSO y AMPPO). La ganancia del mejor modelo fue de un 0,25 % con respecto al mejor modelo IWPSO y de un 0,75 % para el caso del AMPPO, obteniendo un MAPE de un 60,41 % y un coeficiente de determinación de 94.63 %.

La tabla que se presenta a continuación nos permite comparar los resultados de estas investigaciones previas con el trabajo realizado en este proyecto:

Metricas	F. Villaroel [2012] (RRM-IPSO)	C. Cosming [2012] (LS-SVM + LAPSO)	P. Sepúlveda [2011] (RR-PSO)	F. Molina [2013] (RP-DIWPSO)
R2	94.9 %	94.63 %	85.66 %	98.6 %
RMSE	0,03	0,035	0.037	0.017
MAPE	20 %	61 %	24 %	1.17 %

Tabla 7: Comparación de resultados con trabajos previos

Como se observa en la tabla 13, los resultados obtenidos en este proyecto, muestran una ganancia aproximada del 4 % respecto a la varianza explicada de los proyectos anteriores, sin embargo la mayor disminución viene dada por el MAPE obtenido en esta investigación donde se obtuvo un valor de 0.66 % versus el 20 % obtenido por F. Villaroel, lo que nos permite inferir que el factor determinante a la hora de la obtención de mejores métricas es la separación de las variables de alta y baja frecuencia, ya que las técnicas utilizadas en trabajos anteriores son similares a las utilizadas en este proyecto.

6. Conclusión

En este Proyecto se presentó e implementó las técnicas que utilizadas para el desarrollo del modelo propuesto para predecir la captura de alta y baja frecuencia de anchovetas en la zona norte, mediante redes polinomiales con aprendizaje basado en PSO.

Debido a la necesidad de establecer las cotas de captura de la Anchoveta en la zona Norte de nuestro país, se hace imprescindible la realización de modelos que permitan predecir los volúmenes de captura de anchovetas para fijar las cotas de captura, de esta manera se puede evitar dañar el ecosistema marino y proteger la economía de la industria pesquera.

En base a la investigación acerca de las Redes Neuronales, el estudio de sus distintas topologías, y condiciones internas, junto con las aplicaciones encontradas en la literatura, han dejado claro que las RNAs corresponden a un modelo computacional capaz de resolver una gran cantidad de problemas, y dependiendo de lo que se desee resolver se debe configurar una red que satisfaga dichos requerimientos.

En los últimos quince años, las técnicas de computación evolutiva se han consolidado como un conjunto de meta heurísticas efectivas y robustas como consecuencia de su alta aplicabilidad a la resolución de problemas en múltiples áreas

Tomando en cuenta los numerosos y exitosos antecedentes que refieren a la utilización de técnicas evolutivas para la resolución de problemas de predicción, El algoritmo de optimización basado en enjambre de partículas (PSO), presenta un funcionamiento bastante definido y robusto, por lo que resulta ser una buena alternativa para su incorporación en la creación de un modelo en el cual los parámetros suelen ser complicados de hallar. Esta técnica busca dentro del espacio la solución más óptima. La existencia de varias soluciones al problema, hacen que la tarea de encontrar el óptimo global sea más difícil, porque existe la posibilidad de que el algoritmo converja hacia un óptimo local y no pueda salir de ese subespacio de soluciones. Para esto, es que se han definido algunas variaciones de su funcionamiento, con el afán de recorrer todo el espacio de solución en busca del óptimo global sin caer antes en uno local. Esto lo realizan a través de técnicas que buscan ejercer una buena compensación entre la exploración y la explotación del espacio de soluciones.

Luego del análisis de los distintos modelos propuestos, se puede concluir que el modelo RP-DIWPSO(3,9,1) para la baja frecuencia presenta un mejor desempeño en el pronóstico de las capturas de anchovetas, obteniendo un 99 % de la variabilidad explicada y un MAPE de 0.52 %, además podemos afirmar que el 99.4 % de los datos pronosticados tendrán un error relativo menor al 2 %. Por otro lado para el pronóstico de la alta frecuencia se obtuvo que la mejor configuración está dada por RP-DIWPSO(24,6,1), logrando un 97 % de coeficiente de determinación y un MAPE del 1.53 %, respecto al error relativo podemos afirmar que solo el 3.6 % tendrá un error por sobre el 5 %. Luego de obtener los resultados de ambas variables por separados, se sumó los datos pronosticados interanuales con los anuales obteniendo un 98.6 % de la variabilidad explicada y un MAPE de 0.66 %, además

el 1.8 % de los puntos estimados tendrá un error relativo mayor al 2 %, lo que nos confirma que este proyecto realizado propone un predictor de calidad.

Al comparar estos resultados con los obtenidos en proyectos de años anteriores, donde se utilizaron los mismos datos pero sin separar en variables de alta y baja frecuencia, se obtuvo en promedio una ganancia del 5 % aproximado en el coeficiente de determinación y una disminución de un MAPE de 20 % hasta 0.6 % obtenido en este proyecto, lo que representa una mejora substancial en las métricas de evaluación.

Debido a que el predictor neuronal obtenido ofrece resultados precisos, se extrapoló el modelo propuesto para pronosticar hasta el cuarto mes, a través del pronóstico directo, obteniendo un MAPE de 1.55 % y un coeficiente de determinación del 93 %, lo que nos confirma que este nuevo modelo que separa ambas variables, permite hacer buenos pronósticos no sólo del mes siguiente, sino que también hasta cuatro meses en el futuro.

Por último cabe destacar que esta investigación es parte del proyecto FONDECYT N° 1131105, se espera que los resultados obtenidos puedan ser utilizados por las autoridades correspondientes para elaborar mejores políticas de cotas de captura de los recursos pelágicos de nuestro país, para preservar la fauna marina y cuidar la industria pesquera nacional.

Referencias

- [1] Dirección meteorológica de Chile, Dirección General de Aeronáutica Civil.
- [2] Subsecretaría de Pesca y Acuicultura. Informe sectorial de pesca y acuicultura. Diciembre 2012.
- [3] R. Kennedy, J. y Eberhart. Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks*, 1995.
- [4] Hilera J. y Martpínez V. Redes neuronales artificiales. fundamentos, modelos y aplicaciones. 1995.
- [5] Back T. Self-adaptation in genetic algorithms. en towards a practice of autonomous systems. *Proceedings of the first European Conference on Artificial Life*, pages 263–271, 1992.
- [6] Villaroel F. Redes neuronales multiplicativas para pronóstico de captura de anchoveta en la zona norte de Chile. *PUCV, Chile*, Junio 2012.
- [7] Osorio J. Vector de soporte autoregresivo para pronosticar captura de anchoveta en la zona centro-sur de Chile. *PUCV, Chile*, Mayo 2012.
- [8] Caselli N. Pronóstico de captura de anchoveta con algoritmos genéticos y máquinas de vector de soporte. *PUCV, Chile*, Octubre, 2011.
- [9] Cosming C. Pronóstico de captura de anchoveta de la zona norte de Chile usando vector de soporte autoregresivo. *PUCV, Chile*, Abril, 2012.
- [10] Quezada P. Redes rrbf con aprendizaje híbrido para la predicción de captura de anchoveta. *PUCV, Chile*, 2009.
- [11] D. Matich. Redes neuronales conceptos básicos y aplicaciones. 1999.
- [12] S. Haykin. Neural networks, a comprehensive foundation. *Pearson Prentice Hall*, Second Edition, 1999.
- [13] Michael Y. Hu. Guoqiang Zhang, B. Eddy Patuwo. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14:pages 35–62, 1998.
- [14] D. Deligiorgi K. Philippopoulos. Stochastic modeling of hourly average wind speed sequences in national observatory of Athens[c]. *9th Conference on Environmental Science and Technology, Rhodes Island, Greece*, pages pp. 729–734, September 1st - 3rd, 2005.
- [15] Y. Z. Jafri L. Kamal. Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan[j]. *Solar Energy* 61 1, 1997.

- [16] S.D. Huang K. Y. Huang, D. Lang. Wind prediction based on improved bp artificial neural network in wind farm. *Electrical and Control Engineering (ICECE)*, 2010.
- [17] A. G. Tsikalakis et al Y. A. Katsigiannis. Improved wind power forecasting using a combined neuro-fuzzy and artificial neural network model. *Advances in Artificial Intelligence*, 2011.
- [18] S. Al-Humidan N. Al-Arifi. Local and regional earthquake magnitude calibration of tabuk analog sub-network, northwest of saudi arabia. *Journal of King Saud University - Science*, 2011.
- [19] C. Chengui J. Wang, Y. Xie. Solar radiation prediction based on phase space reconstruction of wavelet neural network. *Procedia Engineering*, volume 15:pages 4603–4607, 2011.
- [20] F. Bonanno G. Capizzi. A wavelet based prediction of wind and solar energy for long-term simulation of integrated generation systems. *roceedings of the 2010 International Conference on Modeling, Identification and Control*, 2010.
- [21] R. Ichikawa H. Hayashi A. Chaouachi, R. M. Kamel and K. Nagasaka. Neural network ensemble-based solar power generation short-term forecasting. *World Acad.Sci., Eng. Technol*, 2009.
- [22] Ajay K. Singh Ghassan Abu-Lebdeh. Modeling arterial travel time with limited traffic variables using conditional independence graphs & state-space neural networks. *Procedia - Social and Behavioral Sciences*, volume 16:pages 207–2016, 2011.
- [23] Hilmi Berk Celikoglua. Travel time measure specification by functional approximation: application of radial basis function neural networks. *Procedia - Social and Behavioral Sciences*, volume 20:pages 613–620, 2011.
- [24] Zhang Shaoyang Yang Yingjun Yang Yingjun, Hua Cui. A prediction model of the number of taxicabs based on wavelet neural network. *Procedia Environmental Sciences*, volume 12, part B:pages 1010–1016, 2012.
- [25] C. Chengui J. Wang, Y. Xie. Solar radiation prediction based on phase space reconstruction of wavelet neural network. *Procedia Engineering*, Volume 15:pages 4603–4607, 2011.
- [26] Zhang Jianqing. Study on the gas content of coal seam based on the bp neural network. *Procedia Engineering*, volume 26:pages 1554–1562, 2008.
- [27] E. Ali A. Ajbar. Prediction of municipal water production in touristic mecca city in saudi arabia using neural networks. *Journal of King Saud University - Engineering Sciences*, online:online, 2013.

- [28] M. Boulif K. Benmahdjoub, Z. Ameer. Forecasting of rainfall using time delay neural network in tizi-ouzou (algeria). *Energy Procedia*, Volume 36:pages 1138–1146, 2013.
- [29] Alarifi y Al-Humidan Alarifi. Earthquakes magnitude predication using artificial neural network in northern red sea area. *Journal of King Saud University - Science*, 2012.
- [30] Eberhart RC Shi Y. A modified particle swarms optimiser. *Proceedings of the IEEE international conference on evolutionary computation*, 2001.
- [31] Eberhart R. y Shi Y. Kennedy, J. Swarm intelligence. *The Morgan Kaufmann Series in Artificial Intelligence*, 2001.
- [32] Guo H. y Liu. H. Abraham, A. Swarm intelligence: Foundations, perspectives and applications, studies in computational intelligence (sci). *Springer-Verlag*, 2006.
- [33] F. Martinelli R. Zich E. Alfassio Grimaldi, M. Brenna. Em radiated field by a branched and tortuous cg discharge, even considering a direct stroke on an aircraft. *IEICE Transactions*, 88-B:3300–3306, 2005.
- [34] Jiao Lin Shen G. Yu. Shen, Jiang. Hybridized particle swarm algorithm for adaptative structure training of multilayer feed-forward neural network. *QSAR Studies of Bioactivity of Organic Compounds. Wiley InterScience*, 2004.