



Previsión del Índice de Oscilación Peruano usando un modelo de red neuronal artificial no lineal autorregresivo

Peruvian Oscillation Index forecast using an autoregressive nonlinear neural network model

CARLOS QUISPE ^{1,*} & JORGE TAM ¹

¹ Centro de Investigaciones en Modelado Oceanográfico y Biológico Pesquero (CIMOBP), Instituto del Mar del Perú (IMARPE), Apdo. 22, Callao, Lima, Perú.

*cquispe@imarpe.pe

Resumen

Durante El Niño, el incremento de la temperatura superficial del mar (TSM), frente a las costas de Perú, está asociado a una disminución de nutrientes y producción primaria, que reduce la disponibilidad de la anchoveta; y por otro lado, el aumento de lluvias torrenciales provoca inundaciones que destruyen la infraestructura física y las áreas de cultivos, generando efectos negativos sobre los sectores productivos pesquero y agrícola. Para la toma de decisiones sobre inversiones económicas y medidas de prevención, es necesario contar con previsiones de la TSM local con varios meses de anticipación. En este trabajo se investiga el uso del modelo de redes neuronales artificiales (RNA) como herramienta para la previsión de la variación temporal de la TSM frente a las costas del Perú, representado por el Índice de Oscilación Peruano (IOP), que es un índice que integra la variabilidad de las TSM de varias estaciones costeras.

Se obtuvieron previsiones del IOP con tiempos de avance de 2, 5 y 8 meses, logrando prever el efecto de los eventos El Niño sobre las temperaturas costeras frente a Perú. Se obtuvo un coeficiente de correlación de 0.79 ($p < 0.05$) entre los valores observados y simulados para un horizonte de 2 meses, y un coeficiente de correlación de 0.55 ($p < 0.05$) para un horizonte de 5 meses.

Palabras claves: Red neuronal; Previsión; Temperatura superficial del mar; Perú; ENOS.

Abstract

During El Niño, sea surface temperature (SST) increase off the coast of Peru is associated to a decline in nutrients and primary production, which reduces anchovy availability, and the increase in heavy precipitations cause floods that destroy infrastructure and cultivated areas generating negative effects on the fishery and agriculture productive sectors. For investment decision-making and preventive measures, it is necessary to have forecasts of local SST with several months in advance. In this work, the use of artificial neural network models (ANN) is investigated as a tool to forecast the temporal variation of SST off the Peruvian coast, represented by the Peruvian Oscillation Index (POI), which is an index that integrates the variability of the SST from several coastal stations.

POI forecast was obtained for 2, 5 and 8 months in advance, simulating the effects of the main El Niño events based on the coastal temperatures off the coast of Peru. A correlation coefficient between observed and simulated data of 0.79 ($p < 0.05$) for a 2 months lead time, and a correlation coefficient of 0.55 ($p < 0.05$) for a 5 months lead time were obtained.

Keywords: Neural network; Forecast; Sea surface temperature; Peru; ENSO.

INTRODUCCIÓN

El área de estudio abarca la franja marina frente a la costa del Perú (entre los 7°S y 18°S) hasta las 30 mn mar adentro. Por su ubicación geográfica el mar peruano asociado a un afloramiento costero, está caracterizado por mantener la mayor productividad pesquera del mundo (Chávez et al., 2008). Los cambios en la TSM alteran esta productividad, por lo que es necesario contar con modelos de previsión de la TSM local con varios meses de anticipación, permitiendo la toma de decisiones sobre medidas de prevención e inversiones económicas.

En el Perú, existen pocos trabajos sobre modelos de previsión de la TSM local, resaltando el modelo autorregresivo de Quispe & Purca (2007) y el modelo estadístico empírico de Matellini et al. (2007).

El ciclo El Niño-Oscilación del Sur (ENOS) es una fluctuación entre dos eventos extremos en el Pacífico Tropical: El Niño (fase cálida de ENOS) y La Niña (fase fría de ENOS) (McPhaden et al., 2006). Aunque El Niño se origina en el Pacífico Ecuatorial, tiene múltiples efectos en diferentes regiones del mundo (Trenberth et al., 1998). En el mar peruano, los efectos de El Niño se manifiestan como incrementos de la temperatura superficial del mar, asociados a una disminución de nutrientes y de la producción primaria, con la consecuente reducción en la disponibilidad de anchoveta. A nivel internacional, continúan los esfuerzos para mejorar los modelos de previsión de ENOS en el Pacífico Tropical (Chen et al., 2004), así como para desarrollar teorías sobre su funcionamiento (Sheinbaum, 2003).

Para la predicción de la TSM se han usado diversos modelos basados en métodos matemáticos y estadísticos. El IRI (2010) muestra las previsiones de modelos matemáticos (e.g. modelo océano-atmósfera acoplado intermedio, modelo de circulación general acoplado) y modelos estadísticos (e.g. modelo de redes neuronales, modelo de Markov, regresión lineal múltiple, regresión polinomial múltiple, análisis de correlación canónica, modelo lineal inverso, modelo análogo construido).

Sin embargo, los modelos lineales han presentado limitaciones en la previsión de los efectos de ENOS en el Perú. Por ejemplo, Quispe & Purca (2007) presentaron un modelo autorregresivo ARIMA cuyas previsiones de temperaturas tuvieron una buena habilidad predictiva solo en el corto plazo (hasta 2 meses de anticipación), y Matellini et al. (2007) no pudieron reproducir bien la intensidad de las temperaturas frente al Perú durante El Niño 1982-83. Por tal motivo, es necesario desarrollar modelos no lineales, como las redes neuronales artificiales (Hsieh & Tang, 1998; Wu et al., 2006; Bishop, 1995) para la previsión de efectos de eventos como El Niño y La Niña frente al Perú.

La no linealidad de ENOS permite hipotetizar que un modelo de red neuronal no lineal tendrá una mejor habilidad predictiva que un modelo lineal.

Con el fin de contar con un índice que represente el efecto del ENOS sobre los cambios de la TSM frente a la costa peruana, Purca (2005) elaboró un Índice de Oscilación Peruano (IOP) que integra la variabilidad espacial conjunta de las anomalías de temperatura superficial del mar (ATSM) de las estaciones costeras peruanas.

Así, el objetivo del presente trabajo es prever la variabilidad temporal del Índice de Oscilación Peruano usando un modelo de red neuronal artificial autorregresivo.

DATOS Y MÉTODOS

Datos

Los datos fueron registros históricos mensuales de temperatura superficial del mar (TSM) (1950-2008) obtenidos de las estaciones costeras fijas del Instituto del Mar del Perú (IMARPE), distribuidas a lo largo del litoral peruano (Fig. 1) y registran manualmente la temperatura superficial del mar usando un termómetro de mercurio con una precisión de 0.1 °C, tres veces al día. El Índice de Oscilación Peruano (IOP) se calculó como la primera función ortogonal empírica (EOF) de las anomalías

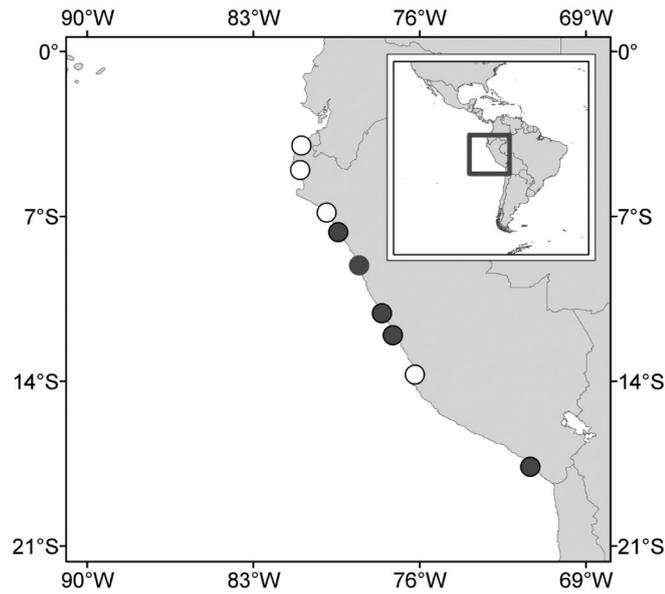


Fig. 1. Ubicación de las estaciones costeras del IMARPE (círculos claros) y de las estaciones usadas para calcular el Índice de Oscilación Peruano (IOP) (círculos oscuros). En las estaciones costeras se registra manualmente la temperatura superficial del mar 3 veces al día usando un termómetro de mercurio.

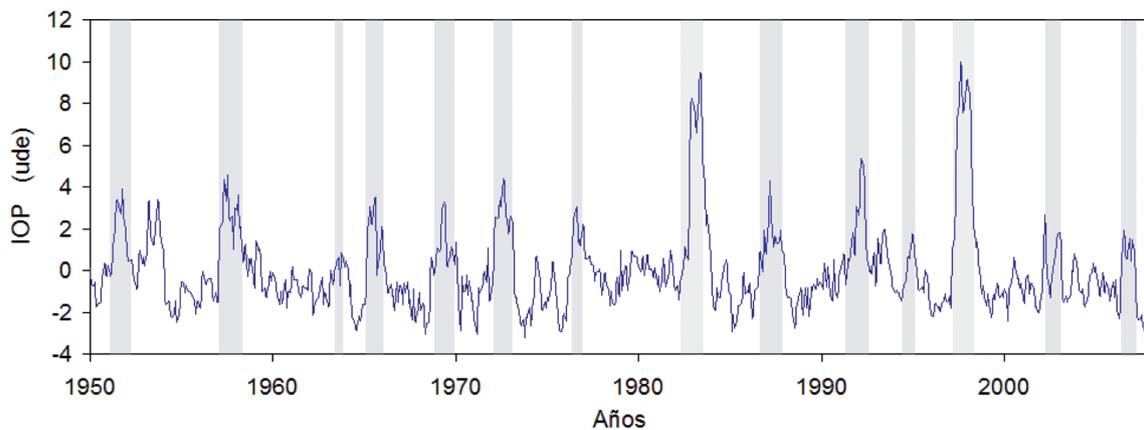


Fig. 2. Variación temporal del Índice de Oscilación Peruano (ude: unidades de desviación estándar) de 1950 al 2008. Las áreas sombreadas indican periodos de ocurrencia de eventos El Niño.

de TSM costeras (Purca, 2005), de modo que valores positivos (negativos) del IOP estarían asociados a eventos El Niño (La Niña). Las estaciones que contribuyeron al cálculo del IOP fueron Chicama ($07^{\circ}42'00''$ S, $79^{\circ}27'00''$ W), Chimbote ($09^{\circ}04'00''$ S, $78^{\circ}36'00''$ W), Huacho ($10^{\circ}16'00''$ S, $78^{\circ}36'00''$ W), Callao ($12^{\circ}04'00''$ S, $77^{\circ}09'00''$ W), Ilo ($17^{\circ}39'00''$ S, $71^{\circ}21'00''$ W) (Fig. 2).

Modelo

Se usaron redes neuronales artificiales (RNA) por tener la habilidad de modelar relaciones no lineales complejas sin la asunción a priori de la naturaleza de la relación (BuHamra et al., 2003; Warner & Misra, 1996). Se usó un modelo autorregresivo por su simplicidad y por tener la propiedad de realizar las predicciones

robustas en base a los valores previos de la propia serie de tiempo (Halide & Ridd, 2008). La Fig. 3 muestra la estructura de una RNA compuesta por capas dentro de las cuales hay nodos que operan en paralelo, y cada nodo se comunica con el otro a través de conexiones (pesos o parámetros del modelo) emitiendo información de una capa a otra. El presente trabajo utiliza una RNA con arquitectura perceptrón multicapas, que consiste en una capa de entrada, una oculta y una capa de salida con función de transferencia lineal. Las conexiones entre los nodos i y h son cuantificadas a través del peso w_{ih} . Cada nodo opera primero ponderando las entradas x_i con el peso w_{ih} , y luego aplicando una función de transferencia “ $f(\cdot)$ ” como muestra la ecuación (1).

$$z_h = f\left(\sum_i x_i w_{ih} + \beta_h\right) \quad (1)$$

donde $i = 1, 2, 3, \dots, I$ y $h = 1, 2, 3, \dots, H$, siendo I el número de entradas y H el número de nodos en la capa oculta, β_h es el umbral de activación del nodo h y puede ser escrito como una entrada más ($i = 0$) con valor $x_0 = +1$ y cuyo peso w_{i0} sería β_h . El término “tiempo de avance” se define como el tiempo entre el final del último periodo observado (predictor) y el principio del período predicho.

Para la previsión del IOP se usó el método directo (Hsieh & Tang, 1998), en el cual cada valor predicho Y fue calculado de manera autorregresiva en base a valores observados del IOP (Fig. 3), con tiempos de avance de 2, 5 y 8 meses.

Para este trabajo se usó una función de transferencia sigmoide tangente hiperbólica en los nodos de la capa oculta, que es una función no lineal y diferenciable asintótica en forma de S cuyo rango de valores va de -1 a $+1$ (Warner & Misra 1996, Bishop, 1995), que permite a las redes neuronales multicapas aprender de las relaciones no lineales y lineales entre los datos de entrada y salida (Demuth et al., 2008).

En la capa de salida se usó una función lineal debido a que permite extrapolar más allá del rango de los datos de entrenamiento, a

diferencia del uso de una función no lineal que en las salidas podría limitar el rango de valores entre -1 y 1 (Bishop, 1995; Demuth et al., 2008). Los parámetros de los modelos fueron determinados mediante un proceso de optimización denominado “entrenamiento de la red”. El entrenamiento de la red neuronal usó el algoritmo de retropropagación del error (Hilera & Martínez, 1995) en una topología multicapa hacia adelante con aprendizaje supervisado.

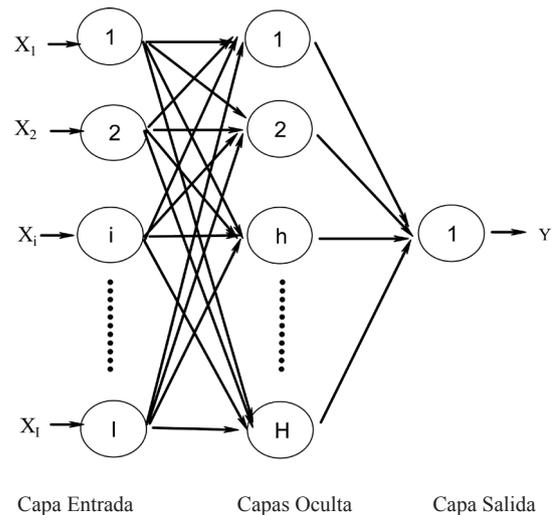


Fig. 3. Arquitectura de una Red Neuronal Artificial (RNA). Las flechas son las interconexiones entre las capas de entrada, oculta y salida. X son las variables predictoras, Y es la variable predictante y H es el número de nodos en la capa oculta.

El proceso de optimización consiste en minimizar la función del error, usando el enfoque regularización bayesiana implementado en el Neural Network Toolbox del software MATLAB R2008a (Demuth et al., 2008), que evita el problema de mínimos locales durante el entrenamiento de la red neuronal mediante la estimación de un parámetro óptimo de regularización que controla el compromiso entre el grado de suavidad de la solución y el grado de ajuste de los datos de entrenamiento (Poggio & Girosi, 1990).

La habilidad del modelo de RNA autorregresivo no lineal se evaluó calculando el coeficiente de correlación entre los valores simulados y observados del IOP,

y graficando la correlación en función del tiempo de avance y el mes previsto (Wu et al., 2006).

Adicionalmente, se comparó la habilidad de los modelos RNA autorregresivos lineal (sin capa oculta y con función lineal en el nodo de la capa de salida) y RNA no lineal (con función sigmoide tangente hiperbólica en los nodos de la capa oculta, y función lineal en el nodo de la capa de salida). La RNA lineal, al omitir la capa oculta y la función de transferencia en el nodo de la capa de salida es lineal, resulta ser equivalente a un modelo estadístico lineal autorregresivo (AR).

RESULTADOS Y DISCUSIONES

El proceso de optimización de los parámetros de la RNA lineal y no lineal resultó en una configuración óptima con un vector de entrada $[X1, X2, X3, X4, X5]$ igual al vector $[yt, yt-3, yt-6, yt-9, yt-12]$, donde X_i son los predictores, $yt-n$ es el IOP y t es el tiempo en meses.

La Fig.4 muestra la habilidad predictiva de los modelos RNA autorregresivos lineal y no lineal para el periodo 1952-2007. Las correlaciones de la validación cruzada

(Fig. 4) muestran que a un tiempo de avance de 2 meses ambos modelos presentaron habilidades predictivas similares, sin embargo la RNA no lineal presentó una mejor habilidad predictiva conforme se incrementó el tiempo de avance (5 y 8 meses), lo cual coincide con lo encontrado por Wu et al. (2006).

Por otro lado, el modelo de RNA no lineal presentó mejor habilidad predictiva que otros modelos que son lineales, por ejemplo, fue mejor que el modelo autorregresivo y de media móvil integrado (ARIMA), que tiene un horizonte de previsión del IOP de solo 3 meses (Quispe & Purca, 2007); y fue mejor que el modelo basado en el volumen de agua cálida (VAC), que subestimó las intensidades del IOP durante algunos eventos El Niño (Mate-llini et al., 2007).

El modelo de RNA no lineal simuló las intensidades y fechas del IOP con una aceptable habilidad predictiva entre los datos simulados y observados del IOP, para un tiempo de avance de 2 meses ($r = 0.79$ y $p < 0.05$) en los años 1957-58, 1968-69, 1982-83, 1986-87, 1992-93 y 1997-98 (Fig. 5a), y para un tiempo de avance de 5 meses ($r = 0.55$ y $p < 0.05$) en los años 1957-58, 1968-69, 1986-87 y 1997-98 (Fig. 5b) considerados periodos de eventos El Niño.

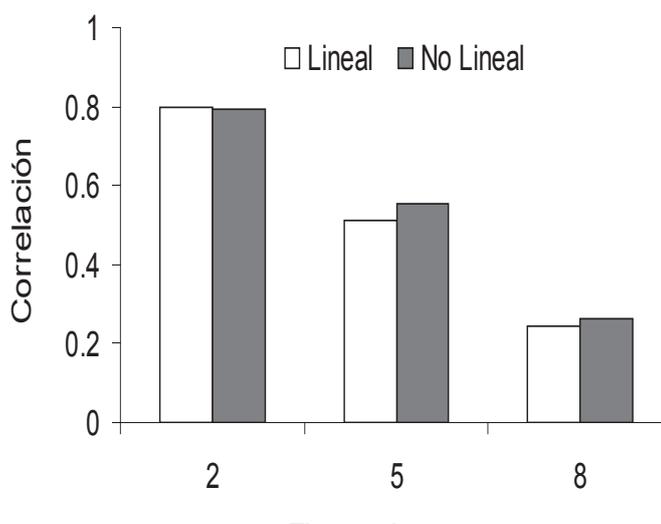


Fig. 4. Comparación de la habilidad predictiva de los modelos de RNA autorregresivos lineal y no lineal (1952-2007).

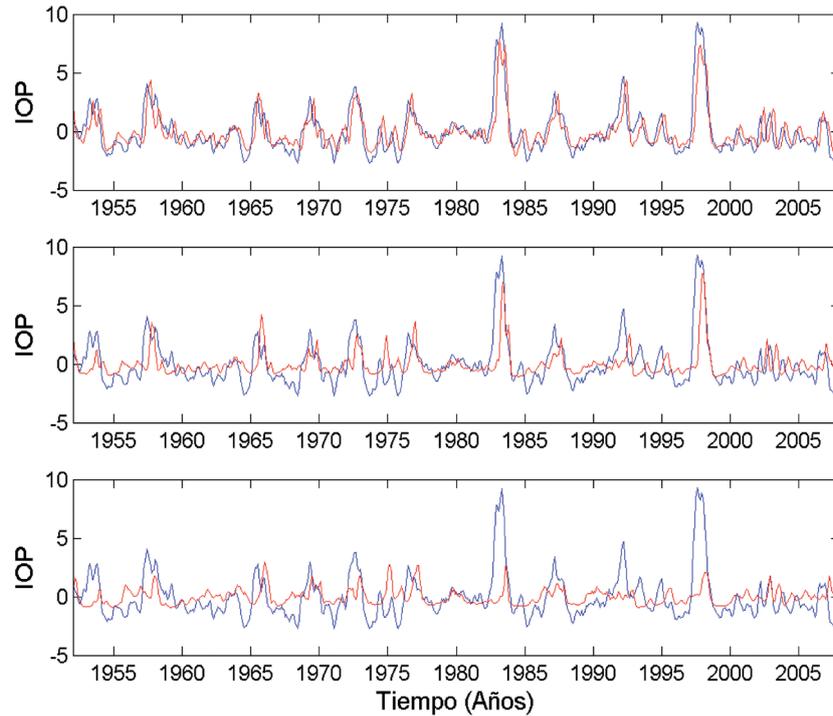


Fig. 5. Variación temporal del IOP observado (línea azul) y predicho (línea roja) por modelos de redes neuronales artificiales no lineales, para tiempos de avance de a) 2 meses, b) 5 meses y c) 8 meses.

La Fig. 6 muestra la habilidad predictiva del modelo de RNA no lineal, que fue mayor durante el verano austral (invierno boreal) con correlaciones máximas de 0.87 ($p < 0.05$), 0.65 ($p < 0.05$) y 0.36 ($p < 0.05$) en enero para tiempos de avance de 2, 5 y 8 meses respectivamente, y fue menor durante el otoño austral (primavera boreal) con correlaciones

mínimas de 0.72 ($p < 0.05$) en abril y 0.179 ($p > 0.05$) en mayo para tiempos de avance de 2 y 8 meses respectivamente, coincidente con la conocida barrera predictiva de otoño austral (primavera boreal) (Webster & Yang, 1992; Chen et al., 2004), además para un tiempo de avance de 5 meses tuvo un mínimo coeficiente de correlación de 0.46 ($p < 0.05$) en julio.

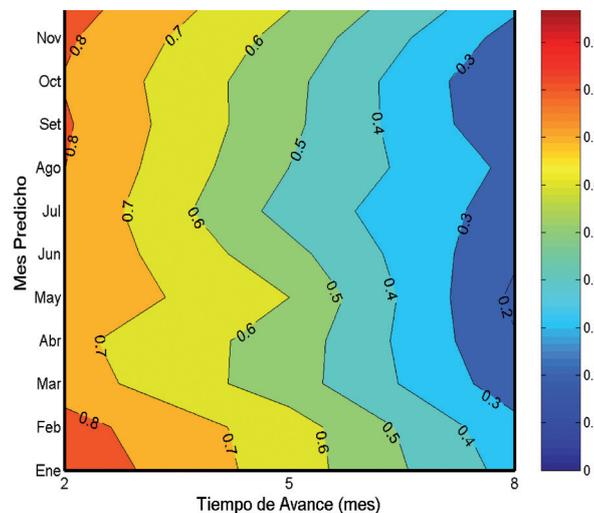


Fig. 6. Correlación entre el valor observado y el predicho por la RNA para cada tiempo de avance (mes) y mes predicho.

Considerando aceptables las correlaciones entre los datos observados y simulados estadísticamente significativas ($p < 0.05$) y mayores a 0.5 (Halide & Ridd, 2008), el modelo de RNA con un tiempo de avance de 5 meses representa un compromiso adecuado entre la habilidad predictiva y el tiempo de avance para la previsión del IOP.

La habilidad predictiva del presente modelo ($r = 0.55$, $p < 0.05$, tiempo de avance de 5 meses) explicó el 30 % de la varianza del IOP, sin embargo, fue comparable a la habilidad predictiva del modelo de redes neuronales artificiales usado por Wu et al. (2006), para predecir el índice de la región Niño 1+2 ($r = 0.635$, tiempo de avance de 6 meses).

CONCLUSIONES Y RECOMENDACIONES

La configuración óptima del modelo de RNA comprendió 5 predictores autorregresivos con desfases de 0, 3, 6, 9 y 12 meses, el modelo de RNA no lineal tuvo una mejor habilidad predictiva que el modelo de RNA lineal a tiempos de avance de 5 y 8 meses, apoyando la hipótesis planteada.

La predicción del IOP usando RNA no lineal presentó mejor habilidad predictiva durante el verano austral ($r = 0.65$, $p < 0.05$) en enero para un tiempo de avance de 5 meses.

El modelo de RNA no lineal permitió simular el Índice de Oscilación Peruano (IOP) con una aceptable habilidad predictiva ($r = 0.55$, $p < 0.05$) para un tiempo de avance de 5 meses.

El modelo de RNA no lineal logró simular la variación temporal del IOP durante los principales eventos El Niño (e.g. 1986-87 y 1997-98).

Se recomienda usar modelos de RNA multivariados que incluyan como predictores a variables tanto oceánicas como atmosféricas.

Agradecimientos

Se agradece a la Dirección de Investigaciones Oceanográficas (DIO-IMARPE) y a los laboratorios costeros por la información de TSM. Se agradece el apoyo permanente de todo el personal del CIMOBP durante la ejecución del presente trabajo, en especial a la Dra. Sara Purca por proporcionar la serie del IOP. Finalmente, se agradece al Dr. Matthieu Lengaigne por su valioso asesoramiento sobre redes neuronales artificiales.

BIBLIOGRAFÍA

- Bishop, C. M. (1995).** Neural Networks for Pattern Recognition, Clarendon Press. Oxford, U.K. 482 p.
- BuHamra, S., Smaoui, N., & Gabr, M. (2003).** The Box-Jenkins analysis and neural networks: prediction and time series modeling. Applied Mathematical Modeling, 27(10): 805-815.
- Chavez F.P., Bertrand A., Guevara-Carrasco R., Soler P. & Csirke J. (2008).** The northern Humboldt Current System: brief history, present status and a view towards the future. Progress in Oceanography 79:95-105.
- Chen, D., Cane, M. A., Kaplan, A., Zebiak, S. E. & Huang, D. (2004).** Predictability of El Niño over the past 148 years. Nature, 428: 733-736.
- Demuth, H., Beale, M. & Hagan, M. (2008).** Neural Network Toolbox version 6, for use with Matlab. The Mathworks Inc, Natick, MA 01760, USA, 906 p.
- Halide, H., & Ridd, P., (2008).** Complicated ENSO models do not significantly outperform very simple ENSO models. Int. J. Climatol. 28: 219-233.
- Hilera, J. & Martínez, V. (1995).** Redes Neuronales Artificiales, fundamentos modelos y aplicaciones. Editorial RA-MA, Madrid.
- Hsieh, W. W. & Tang, B. (1998).** Applying neural network models to prediction and data analysis in meteorology and oceanography. Bull. Am. Meteorol. Soc., 79(9): 1855-1870.
- IRI (International Research Institute) (2010).** Summary of ENSO Model Forecasts. Additional information on Models. En

línea: http://iri.columbia.edu/climate/ENSO/currentinfo/SST_table.html.

Matellini, B., Tam, J. & Quispe, C. (2007). Modelo empírico para previsión de la temperatura superficial del mar peruano. *Rev. Perú. biol.*, 14(1): 101-108.

McPhaden, M.J., Zebiak, S.E. & Glantz, M.H. (2006). ENSO as an integrating concept in Earth science, *Science* 314: 1740–1745.

Poggio, T. & F. Girosi. (1990). Regularization algorithms for learning that are equivalent to multilayer networks. *Science*. 247:978-979.

Purca, S. (2005). Variabilidad temporal de baja frecuencia en el Ecosistema de la Corriente Humboldt frente a Perú. Tesis de Doctorado, Universidad de Concepción, Chile.

Quispe, C. & Purca, S. (2007). Previsión de la temperatura superficial del mar frente a la costa peruana mediante un modelo autorregresivo integrado de media móvil. *Rev. Perú. biol.*, 14 (1): 109- 115.

Sheinbaum, J. (2003). Current theories on El Niño-Southern Oscillation: A review *Geofísica Internacional*. 42:291-305.

Trenberth, K. E., Branstator, G. W., Karoly, D., Kumar, A., Lau, N.-Ch. & Ropelewski, Ch. (1998). Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.*, 103(C7): 14291- 14324.

Warner, B. & Misra, M. (1996). Understanding Neural Networks as Statistical Tools. *American Statistician*, 50(4): 284-293.

Webster, P. J. & Yang, S. (1992). Monsoon and ENSO: Selectively Interactive Systems. *Quart. J. Roy. Meteor. Soc.*, 118(507): 877-926.

Wu, A, Hsieh, W. W. & Tang, B. (2006). Neural network forecasts of the tropical Pacific sea surface temperatures. *Neural Networks*, 19(2): 145-154.