

UNIVERSIDAD NACIONAL DEL ALTIPLANO

ESCUELA DE POSGRADO

MAESTRÍA EN CIENCIAS: INGENIERIA QUÍMICA



TESIS

**FORMULACIÓN DE UN MODELO DE RED NEURONAL ARTIFICIAL PARA
EL PRONÓSTICO DE CONCENTRACIÓN DE OXÍGENO DISUELTO Y
CLOROFILA EN LA BAHÍA INTERIOR DE PUNO**

PRESENTADA POR:

ALFREDO QUISPE LUJANO

PARA OPTAR EL GRADO ACADÉMICO DE:

**MAGÍSTER SCIENTIAE EN INGENIERIA QUÍMICA
MENCIÓN EN SEGURIDAD INDUSTRIAL Y AMBIENTAL
PUNO, PERÚ**

2018

UNIVERSIDAD NACIONAL DEL ALTIPLANO

ESCUELA DE POSGRADO



MAESTRÍA EN CIENCIAS: INGENIERIA QUÍMICA

TESIS

FORMULACIÓN DE UN MODELO DE RED NEURONAL ARTIFICIAL PARA
EL PRONÓSTICO DE CONCENTRACIÓN DE OXÍGENO DISUELTO Y
CLOROFILA EN LA BAHÍA INTERIOR DE PUNO

PRESENTADA POR:

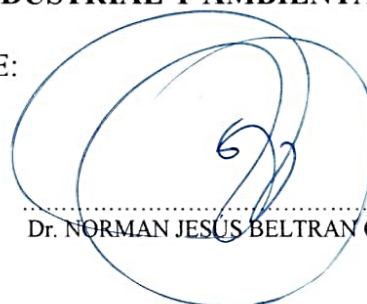
ALFREDO QUISPE LUJANO

PARA OPTAR EL GRADO ACADÉMICO DE:

MAGÍSTER SCIENTIAE EN INGENIERIA QUÍMICA
MENCION EN SEGURIDAD INDUSTRIAL Y AMBIENTAL

APROBADA POR EL JURADO SIGUIENTE:

PRESIDENTE:



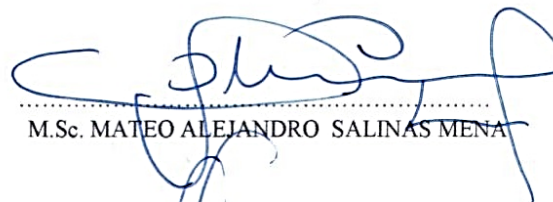
Dr. NORMAN JESÚS BELTRAN CASTAÑO

PRIMER MIEMBRO:




M.Sc. LIDIA ENSUEÑO ROMERO IRURI

SEGUNDO MIEMBRO:



M.Sc. MATEO ALEJANDRO SALINAS MENA

ASESOR DE TESIS:



M.Sc. GERMAN QUILLE CALISAYA

ÁREA: Investigación

TEMA: Ingeniería Química

LÍNEA: Seguridad Industrial y Ambiental

Puno, 08 de enero del 2018.

DEDICATORIA

A Dios por ser el guía de mi vida y darme la oportunidad de culminar mis estudios brindándome vida, salud y fortaleza cada día de mi vida. A mis padres que fueron siempre un ejemplo de cambio, perseverancia humildad y mucho sacrificio.

A mi familia: Presente Carmen, Marianela, Juan y Aymar Ghasem, que siempre están en mi entorno los cuales me han demostrado su esfuerzo, confianza, amistad, cariño y amor.

AGRADECIMIENTOS

- A mis padres Román y María quienes gestaron la idea de traer un vástago y que con el espacio tiempo transcurrido se concluya en un pequeño objetivo que es este documento.
- A mis asesores y amigos: M.Sc. Ing. German Guille Calisaya, Henry Pizarro Viveros quienes en todo este tiempo me han brindado su apoyo y conocimiento para poder culminar la presente tesis.
- A la Universidad Nacional del Altiplano de Puno a la Facultad de Ingeniería Química y más específicamente a los docentes de la Maestría en Seguridad Industrial y Ambiental por sus dosis de conocimiento, para la realización de este trabajo de investigación.
- Al Instituto del Mar del Perú IMARPE, con sede en Puno, por su gran apoyo en la data de variables fisicoquímicas.
- En general, quisiera agradecer a todas aquellas personas que colaboraron en esta labor, dándome fortaleza sugerencias sobre la presente investigación y la incansable búsqueda de desarrollar este objetivo, para llegar a una conclusión relevante.

ÍNDICE GENERAL

	Pág.
DEDICATORIA	iii
AGRADECIMIENTOS	ii
ÍNDICE GENERAL	iii
ÍNDICE DE TABLAS	vii
ÍNDICE DE FIGURAS	viii
ÍNDICE DE ANEXOS	viii
RESUMEN	ix
ABSTRACT	x
INTRODUCCIÓN	1
CAPÍTULO I	3
REVISIÓN DE LITERATURA	3
1.1 Marco Teórico	3
1.1.1 Parámetros físicos	7
1.1.1.1 Temperatura	7
1.1.1.2 pH	7
1.1.1.3 Transparencia	8
1.1.1.4 Conductividad	8
1.1.1.5 Solidos totales disueltos	8
1.1.2 Parámetros Químicos	9
1.1.2.1 Alcalinidad	9
1.1.2.2 Cloruros	9
1.1.2.3 Oxígeno Disuelto (OD)	10
1.1.2.4 Nitratos	10
1.1.2.5 Fosfatos	10
1.1.2.6 Clorofila-a	11
1.2 Redes neuronales biológicas	12
1.3 redes neuronales artificiales (RNA)	13
1.3.1 La Neurona Artificial	13
1.3.2 Análisis Del Perceptrón Simple	16
1.3.3 El Algoritmo Backpropagation	17
1.4 Perceptrón multicapa	18
1.5 Aplicaciones de las redes neuronales artificiales a la calidad del agua	19
1.5.1 Topología de red neuronal artificial para la clorofila alfa	21
1.5.1 Función de activación	23
1.5.1.1 Función tangente hiperbólica	23
	iii

1.5.1.2	Función sigmoïdal logarítmica	24
1.5.1.3	Función lineal o purelin	24
1.5.1.4	Algoritmo de Levenberg Marquardt	25
1.6	Antecedentes	26
CAPÍTULO II		30
PLANTEAMIENTO DEL PROBLEMA		30
2.1	Definición del problema	30
2.1.1	Problema General	30
2.1.1.1	Problemas específicos	31
2.2	Justificación	31
2.3	Objetivos	33
2.3.1	Objetivo General	33
2.3.2	Objetivos Específicos	33
2.4	Hipótesis	33
2.4.1	Hipótesis General	33
2.4.2	Hipótesis Específico	33
CAPÍTULO III		34
MATERIALES Y MÉTODOS		34
3.1	Lugar de estudio	34
3.2	Población	35
3.3	Muestra	35
3.4	Método	35
3.5	Modelamiento con redes neuronales artificiales	35
3.5.1	Formulación de modelos de redes neuronales artificiales	36
3.5.1.1	Determinación de la topología y parámetros de los modelos de redes neuronales artificiales	36
3.5.1.2	Fase de entrenamiento de los modelos de redes neuronales artificiales	36
3.5.1.3	Modelo de Levenberg–Marquardt	40
3.5.1.4	Aprendizaje en redes neuronales usando Regularización Bayesiana	43
3.5.1.5	Algoritmos basados en método de Gradiente Conjugado Escalado (SCG)	44
3.5.1.6	Selección de los mejores modelos de redes neuronales artificiales	46
3.5.1.7	Construcción de los modelos de red neuronal artificial	46
3.6	Resultados del rendimiento de parámetros de determinación	47
3.7	Evaluación del modelo	55
CAPÍTULO IV		57
RESULTADOS Y DISCUSIÓN		57
4.1	Presentación de resultados	57
4.1.1	Evaluación de resultados	57
4.1.2	Resultados de la prueba estadística	61
4.1.3	Resultado pronóstico y validación de los modelos entrenados con data	63
CONCLUSIONES		66



RECOMENDACIONES	68
BIBLIOGRAFÍA	69
ANEXOS	78

ÍNDICE DE TABLAS

	Pág.
1. Índice de Carlson	6
2. Variables suministradas para el modelo de la Bahía interior Puno.	35
3. Parámetros de determinación del rendimiento para Oxígeno disuelto (OD)	47
4. Parámetros de determinación del rendimiento para Clorofila – a	48
5. Modelo de RNA seleccionados para OD y Clorofila a para la Bahía interior Puno...	58
6. Análisis de Varianza para Oxígeno Disuelto	61
7. Análisis de Varianza para Clorofila Alfa.....	62

ÍNDICE DE FIGURAS

	Pág.
1. Estructura básica de una Neurona.....	12
2. Neurona biológica y neurona artificial.	14
3. Proceso en una neurona artificial.....	14
4. Perceptrón de una neurona con dos entradas.	16
5. Gráfica X_2 vrs. X_1	17
13. Mapa de localización geográfica del punto de monitoreo IMARPE.	34
14. Gráficos de regresión durante el desempeño del entrenamiento para OD con 4 neuronas y una capa oculta.	49
15. Gráficos de regresión durante el desempeño del entrenamiento para OD con 7 neuronas y una capa oculta.	50
16. Gráficos de regresión durante el desempeño del entrenamiento para OD con 10 neuronas y una capa oculta.	51
17. Gráficos de regresión durante el desempeño del entrenamiento para clorofila alfa con 4 neuronas y una capa oculta.	52
18. Gráficos de regresión durante el desempeño del entrenamiento para clorofila alfa con 7 neuronas y una capa oculta.	53
19. Gráficos de regresión durante el desempeño del entrenamiento para clorofila alfa con 10 neuronas y una capa oculta.	54
20. Evolución del error de aprendizaje y del error de generalización. En la figura 20 (a) se presenta una Situación idealizada, y en la figura 20 (b) una situación real.....	55
21. Coeficiente de regresión para el Oxígeno disuelto.	58
22. Coeficiente de regresión para la clorofila alfa	59
23. Entrenamiento y prueba de la RNA para OD con 4 neuronas en capa oculta y MSE como función de desempeño para la Bahía interior Puno.	60
24. Entrenamiento y prueba de la RNA para Clorofila a con 10 neuronas en capa oculta y MSE como función de desempeño para la Bahía interior Puno.	60
25. Gráfico de validación con data de 1 año del IMARPE.....	63
26. Coeficiente de regresión para los datos de OD.....	64
27. Gráfico pronóstico de Clorofila alfa con data de 1 año del IMARPE.	64
28. Coeficiente de regresión para los datos de Clorofila alfa	65

ÍNDICE DE ANEXOS

	Pág.
1. Pasos secuenciales de entrenamiento de neuronas	79
2. Pesos óptimos: con 4 variables de entrada 10 neuronas una capa oculta	82
3. Como ajusta los pesos entre capas ocultas.....	85
4. Como se deriva las formulas.....	86

RESUMEN

Se realizaron dos modelos de Redes Neuronales Artificiales (RNA) para la predicción de las concentraciones de clorofila presentes en un cuerpo de agua, ambos modelos fueron basados en una estructura de RNA de retro propagación; utilizando como herramienta para su desarrollo el programa de MATLAB. Los modelos pronósticos fueron para la estación de monitoreo del IMARPE, ubicada en el puerto muelle Puno. Se emplearon datos históricos de parámetros que determinan la calidad del agua para un periodo de 2011 - 2016. Objetivos: Formular un modelo pronóstico, utilizando RNA que permita analizar la calidad de un cuerpo de agua enfocado principalmente en oxígeno disuelto y clorofila alfa, cuyos objetivos específicos fueron probadas desarrollando el modelo de red neuronal artificial dando como resultado el algoritmo de entrenamiento gradiente a escala conjugada, la metodología empleada para el entrenamiento de red neuronal fueron los modelos de Levenverg Marquardt, Regularización Bayesiana y gradiente a escala conjugada, siendo esta última el algoritmo óptimo de entrenamiento. Resultados: Los nodos que los modelos presentaron en la capa oculta se variaron entre 4 y 10 para oxígeno disuelto y clorofila respectivamente ambos casos. Las funciones de transferencia de la capa oculta fueron tansig, y en la capa de salida fue lineal o purelin, para contrastar fue realizado un análisis estadístico para la fase de entrenamiento y prueba que permitió asegurar al 95% de nivel de confianza, comprobándose el desempeño de los modelos obtenidos. Finalmente se realizaron las pruebas de validación con los modelos entrenados dando un performance de 0.97% para efectos de oxígeno disuelto y 0.98% para la clorofila alfa.

Palabras clave: algoritmo de entrenamiento, Clorofila alfa, Gradiente a escala conjugada, Modelo de red neuronal, Oxígeno disuelto, Perceptrón multicapa.

ABSTRACT

Two models of Artificial Neural Networks (RNA) were made for the prediction of chlorophyll concentrations present in a body of water, both models were based on a retropropagation RNA structure; using the MATLAB program as a tool for its development. The forecast models were for the IMARPE monitoring station, located in the Puno dock port. Historical data of parameters that determine water quality for a period of 2011 - 2016 were used. Objectives: To formulate a prognostic model, using RNA to analyze the quality of a water body focused mainly on dissolved oxygen and chlorophyll alfa, whose objectives were developed by developing the artificial neural network model resulting in the conjugate-scale gradient training algorithm. The methodology used for the training of neural network were Levenverg Marquardt models, Bayesian regularization and conjugate scale gradient, the latter being the optimal training algorithm. Results: The nodes that the models presented in the hidden layer were varied between 4 and 10 for dissolved oxygen and chlorophyll, respectively, both cases. The functions of transfer of the hidden layer were tansig, and in the output layer it was linear or purelin, to contrast a statistical analysis was made for the training and testing phase that allowed to assure a 95% confidence level, verifying the performance of the obtained models. Finally validation tests were carried out with the trained models giving a performance of 0.97% for dissolved oxygen effects and 0.98% for chlorophyll alfa.

Keywords: training algorithm, Chlorophyll alfa, Conjugated scale gradient, neural network model, Dissolved oxygen, Multilayer perceptron.

INTRODUCCIÓN

Debido a la constitución y fundamentos de las RNA, presentan un gran número de características semejantes a las del cerebro (Abbas y Al-Bastaki, 2005). Por ejemplo, son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos, de separar características esenciales a partir de entradas que representan información irrelevante, etc.; esto hace que ofrezcan numerosas ventajas como son: (Maren *et. al.*, 1990)

Aprendizaje adaptativo. Es la capacidad de aprender a realizar tareas basadas en un entrenamiento o una experiencia inicial.

Las RNA son sistemas dinámicos auto adaptativos. Son adaptables debido a la capacidad de auto ajustarse de los elementos procesales (neuronas) que componen el sistema y son dinámicos, pues son capaces de estar constantemente cambiando para adaptarse a las nuevas condiciones.

Auto organización. Las RNA usan su capacidad de aprendizaje adaptativo para auto organizar la información que reciben durante el aprendizaje y/o la operación. Mientras que el aprendizaje es la modificación de cada elemento procesal, la auto organización consiste en la modificación de la RNA completa para llevar a cabo un objetivo específico.

La auto organización provoca la generalización; esto es, la facultad de las RNA de responder apropiadamente cuando se les presentan datos o situaciones a las que no habían sido expuestas anteriormente.

Tolerancia a fallos. Las RNA no pierden su funcionalidad en cuanto sufren un pequeño error de memoria, si se produce un fallo en un pequeño número de neuronas, aunque el comportamiento del sistema se ve influenciado, no sufre una caída repentina.

La razón por la que las redes neuronales son tolerantes a los fallos es que tienen su información distribuida en las conexiones entre neuronas, existiendo cierto grado de redundancia en este tipo de almacenamiento.

- **Operación en tiempo real.** Las RNA se adaptan bien a la necesidad de realizar grandes procesos con datos de forma muy rápida debido a su implementación paralela. Para que

la mayoría de las redes pueden operar en un entorno de tiempo real, la necesidad de cambio en los pesos de las conexiones o entrenamiento es mínima.

- Fácil inserción dentro de la tecnología existente. Debido a que una red puede ser rápidamente entrenada, comprobada, verificada y trasladada a una implementación hardware de bajo costo, es relativamente fácil insertar redes neuronales para aplicaciones específicas dentro de sistemas existentes. De esta manera, las redes neuronales se pueden utilizar para mejorar sistemas.

- No linealidad. Una neurona es básicamente un dispositivo no lineal, por consiguiente, la constitución de las interconexiones de neuronas en una RNA es no lineal. Es una propiedad altamente importante; particularmente si los mecanismos físicos responsables de la generación de los datos de entrada son no lineales; como es el caso de las ciencias ambientales. (Haykin, 1999; Martín y Sanz, 2007)

El presente trabajo se encuentra dividido en 4 capítulos; el primero titulado Introducción, el cual aborda los antecedentes y justificación del estudio; así como los objetivos planteados y la hipótesis; el segundo capítulo llamado Generalidades teóricas, donde se aborda la forma en la que se puede evaluar la calidad del agua, el análisis de los principales parámetros para su evaluación; de igual forma, se abordan los fundamentos de las RNA, así como algunas aplicaciones de estas en las disciplinas medioambientales; en el tercer capítulo Desarrollo del Modelo de RNA, donde se describe la metodología de la construcción, el entrenamiento y la prueba de los modelos de RNA para cada uno de los cuerpos de agua estudiados; en el cuarto capítulo Análisis de los modelos obtenidos, se analizan los resultados obtenidos, así como una evaluación estadística de los modelos de RNA y la obtención del índice de eutroficación; por último se presentan las conclusiones y recomendaciones a las que se llegaron en este trabajo de tesis. Además de los capítulos antes mencionados, se tiene un apartado titulado Anexos, donde se incluye información complementaria respecto de las generalidades teóricas; así como los valores de los vectores y pesos con los que trabajan los modelos obtenidos.

CAPÍTULO I

REVISIÓN DE LITERATURA

1.1 Marco Teórico

La calidad de las aguas superficiales es uno de los temas más importantes hoy en día, debido a sus efectos sobre los sistemas acuáticos eco-salud humana y. Con el aumento de la población, hay una presión de uso - sobre los recursos hídricos. la calidad del agua de superficie en una región está determinada en gran medida tanto por procesos naturales, incluyendo la litología de la cuenca, entradas atmosféricas, y las condiciones climáticas, y por aportes antropogénicos tal como descarga de aguas residuales municipales e industriales. Por otro lado, los ríos desempeñan un papel importante en la asimilación o el transporte de las aguas residuales y aguas de escorrentía municipal e industrial de las tierras agrícolas. (Singh et al , 2004)

Las aguas residuales municipales e industriales constituye una fuente de contaminación constante, mientras que la escorrentía superficial es un fenómeno estacional, en gran medida afectada por el clima dentro de la cuenca (Singh et al., 2004).

Hoy en día, el oxígeno disuelto (OD) para efectos de control de concentración es el método más ampliamente utilizado. El oxígeno es un sustrato clave; El oxígeno disuelto (OD) es uno de los parámetros importantes de calidad de agua de un ecosistema acuático y un indicador de estado significativo para los ecosistemas acuáticos. Las fuentes de OD en un cuerpo de agua incluyen aireación captada de la atmósfera, la producción de oxígeno fotosintético. Los sumideros incluyen oxidación de material carbonoso y nitrogenado, la demanda de oxígeno en sedimentos, y la respiración por plantas acuáticas (Kuo, et al, 2007)

Por lo tanto, el seguimiento de su consumo es uno de los métodos más eficaces de seguimiento del proceso de tratamiento (Shen et al, 2008)

La concentración de OD durante el proceso de lodo activado debería ser suficiente para suministrar oxígeno para los microorganismos en el lodo y por lo tanto degradar las sustancias orgánicas y convertir de amonio en nitrato alta. Sin embargo, un nivel excesivamente alto de la OD, requiere un alto nivel de velocidad de flujo, dará lugar a un alto nivel de consumo de energía y puede dañar la calidad de lodos. La investigación muestra que un alto nivel de OD en agua recirculada internamente también disminuye la eficiencia del proceso de Nitrificación (Shen et al, 2008)

La investigación significativa ha estudiado métodos de controlar DO como un medio para mejorar el proceso. Los métodos tradicionales que se hace referencia como el Proporcional-Integral (PI) (Carlsson et al, 1994) Y proporcional-integral-derivativo (PID) (Carlsson et al, 2002).

Una carga excesiva de DBO daña la calidad del agua del río. Esto causa baja OD (oxígeno disuelto) concentración y las condiciones de vida inadecuadas para la flora y la fauna en el río. Al mismo tiempo, las relaciones BOD-OD incluyen intercambio con el lecho del río y la nitrificación y desnitrificación (Radwan et al, 2003)

Varios modelos de calidad del agua, tales como los enfoques mecanicistas tradicionales se han desarrollado con el fin de manejar las mejores prácticas para la conservación de la calidad del agua. La mayoría de estos modelos necesitan varios datos de entrada diferentes que no son de fácil acceso y hacen que sea un proceso que consume muy caro y el tiempo (Suen et al, 2003)

El uso de las aguas de superficie de forma segura y adecuada requiere la determinación de la calidad del agua. La determinación de la calidad del agua se basa tradicionalmente en la clasificación teniendo en cuenta factores físicoquímicos cal o biológicamente de acuerdo con el rango de uso del agua en el contexto de las normas nacionales o internacionales (EPA 1994; SAFE 1995; WEP 1996, Boletín Oficial 2004).

Según estos estándares, en el caso de la calidad del agua, la supervisión implica el análisis de un gran número de variables tales como la temperatura, el oxígeno disuelto (OD), pH, conductividad eléctrica (EC), la demanda bioquímica de oxígeno, cloruro, fosfato total de, nitritos, nitratos, amoníaco, iones, metales pesados, de concentración total de sal, de coliformes fecales, y así sucesivamente. Por ejemplo, la concentración de OD es importante para el buen funcionamiento saludable de los ecosistemas acuáticos y es un indicador importante del estado de los ecosistemas acuáticos. Concentración de oxígeno es un parámetro frecuentemente utilizado para evaluar la calidad del agua en diferentes depósitos y cuencas. También, la concentración está fuertemente influenciada por una combinación de las propiedades físicas, químicas y características biológicas de las sustancias que necesita oxígeno de las corrientes, incluyendo la biomasa de algas,; (Cox 2003; Kalff 2002; Mulholland et al, 2005; Quinn et al, 2005; Spanou, et al, 1993; Williams et al, 2000)

El desarrollo y el progreso actual en la integración de las diversas técnicas de inteligencia artificial (sistema basado en el conocimiento, algoritmos genéticos, redes neuronales artificiales, y del sistema de inferencia difusa) en, por ejemplo, modelización de la calidad del agua, los sedimentos de transporte, la evaporación, la concentración de OD, estimación profunda e integral de la OD ha sido estudiado por muchos investigadores. Por ejemplo, las redes neuronales artificiales (RNA) se han utilizado con éxito en varios campos de la estimación y los métodos de previsión; Haykin 1994; Lerner et al. 1994; Haykin 1998; kisi 2005; Cigizoglu y Kisi 2006; kisi 2007; Hanbay et al, 2008; Dogan et al, 2009; kisi 2009;

Se identificaron modelos de RNA de alimentación hacia adelante, validados, y se ensayaron para el cálculo de OD (Brodnjak-Vonina et al. 2002; Sengorur et al. 2006; Schmid y Koskiaho 2006; Dogan et al. 2009; Akkoyunlu et al. 2011) y hacer DBO₅ (demanda bioquímica de oxígeno) (Singh et al. 2009).

Carlson en el año de 1977 diseñó un índice que es hasta hoy uno de los más utilizados para lagos templados, puede variar entre 0 (oligotrófico) y 100 (hipereutrófico) (Carlson, 1977). La fórmula que figura a continuación es utilizada para determinación del índice utilizando concentraciones de clorofila. La tabla 1.1 muestra la relación de

las concentraciones de clorofila con el índice trófico obtenido (y los grados de eutroficación correspondientes. TSI_{Cl}).

$$TSI_{Cl} = 10 \times \left(6 - \frac{2.04 - 0.68 \ln C1}{\ln 2} \right)$$

Donde:

Cl: se refiere a las concentraciones de clorofila expresada en mg/m^3

Tabla 1

Índice de Carlson

ESTADO TRÓFICO		TSI	CLOROFILA EN SUPERFICIE (mg/m^3)
OLIGOTRÓFICO (> 30)		0	0,04
		10	0,12
		20	0,34
		30	0,94
MESOTRÓFICO (> 30 - 60)		40	2,6
		50	6,4
		60	20
EUTRÓFICO (> 60 - 90)		70	56
		80	154
		90	427
HIPEREUTRÓFICO (> 90)		100	1183

Fuente: Brown, 1973

En el Perú, se ha realizado la evaluación del Índice de Calidad del Agua (ICA), con técnicas multiplicativas y ponderadas con la asignación de pesos específicos a parámetros de control de calidad del agua, obteniéndose a partir de una media geométrica que se debe a la influencia de los trabajos realizados por Brown y Mc Cleland (Brown y Mc Cleland, 1973).

Estos índices de estado trófico y la forma de estimarlos proporcionan información relevante para ser aplicada en la obtención de un índice de eutroficación a partir de los datos obtenidos por los modelos que serán construidos; principalmente acerca de las ventajas y desventajas de sus usos.

1.1.1 Parámetros físicos

Los parámetros físicos tienen directa incidencia sobre las condiciones estéticas y de aceptabilidad del agua; estas características son consideradas en el control de la calidad del agua.

1.1.1.1 Temperatura

Es uno de los parámetros físicos más importantes en el agua, pues por lo general influye en el retardo o aceleración de la actividad biológica, la absorción de oxígeno, la precipitación de compuestos, la formación de depósitos, la desinfección y los procesos de mezclado, floculación, sedimentación y filtración.

Múltiples factores, principalmente ambientales, pueden hacer que la temperatura del agua varíe continuamente, una variación de la temperatura afecta a la solubilidad de las sales ocasionando con ello cambios de la conductividad y el pH; además de afectar la tasa de crecimiento fitoplanctónico, ya que algunos organismos son sumamente sensibles a esas variaciones. (Ji, 2008)

1.1.1.2 pH

El pH expresa la intensidad de la condición ácida o alcalina de una solución cuyo factor más importante es habitualmente la concentración de anhídrido carbónico debido a la mineralización del agua. El pH presente en un cuerpo de agua se debe a la composición de los terrenos atravesados por la corriente de agua; un pH alcalino indica que los suelos son calizos, mientras que un pH ácido indica que estos son silíceos. La mayoría de los ecosistemas acuáticos tienen un pH que oscila entre 5.0 y 9.0, muy pocas especies pueden crecer a pH inferior a 2.0 o superiores a 10.

Vertimientos ácidos, $\text{pH} < 6$ en corrientes de agua con baja alcalinidad ocasionan disminuciones del pH del agua natural por debajo de los valores de tolerancia de las especies acuáticas (pH entre 5 y 9), lo mismo sucede con vertimientos alcalinos $\text{pH} > 9$. Los vertidos de pH ácido, disuelven los metales pesados y el pH alcalino los precipitan, teniendo con esto formación de sedimentos. Este parámetro puede ser considerado como químico o físico, por sus características. (Castro, 1987).

1.1.1.3 Transparencia

La luz, que es esencial para el crecimiento de las plantas, viaja más lejos en las aguas claras que en cualquier agua turbia que contiene sólidos en suspensión o agua con color. Normalmente se utilizan dos métodos para medir la transparencia o el grado de penetración de la luz en el agua: el disco Secchi y el tubo de turbiedad. La luz del sol proporciona la energía para la fotosíntesis, La transparencia decrece a medida que el color, los sedimentos en suspensión o la abundancia de las algas aumentan. El agua toma el color por la presencia y la acción de algunas bacterias, fitoplancton y otros organismos, o por los químicos vertidos por el suelo o por la materia vegetal en descomposición.

1.1.1.4 Conductividad

La conductividad eléctrica de una solución es una medida de la capacidad de la misma para transportar la corriente eléctrica y permite conocer la concentración de especies iónicas presentes en el agua. Como la contribución de cada especie iónica a la conductividad es diferente, su medida da un valor que no está relacionado de manera sencilla con el número total de iones en solución. Depende también de la temperatura. Está relacionada con el residuo fijo por la expresión:

$$\text{Conductividad } (\mu\text{S/cm}) \times f = \text{residuo fijo (mg/L)}$$

El valor de f varía entre 0.55 y 0.9.

1.1.1.5 Sólidos totales disueltos

Corresponden al residuo remanente después de secar una muestra de agua. Equivalen a la suma del residuo disuelto y suspendido. El residuo total del agua es secado a una temperatura de 103 – 105 °C.

$$\text{Sólidos totales} = \text{Sólidos suspendidos} + \text{sólidos disueltos}$$

$$\text{Sólidos totales} = \text{Sólidos fijos} + \text{sólidos volátiles}$$

1.1.2 Parámetros Químicos

Los parámetros químicos afectan las características fisicoquímicas de un cuerpo de agua, entre los parámetros químicos necesarios para determinar la calidad de un cuerpo de agua y de interés para el estudio, se tienen los siguientes:

1.1.2.1 Alcalinidad

Es la capacidad del agua de neutralizar ácidos. Sin embargo, aniones de ácidos débiles (bicarbonatos, carbonatos, hidróxido, sulfuro, bisulfuro, silicato y fosfato) pueden contribuir a la alcalinidad.

La alcalinidad está influenciada por el pH, la composición general del agua, la temperatura y la fuerza iónica. Por lo general, está presente en las aguas naturales como un equilibrio de carbonatos y bicarbonatos con el ácido carbónico, con tendencia a que prevalezcan los iones de bicarbonato. De ahí que un agua pueda tener baja alcalinidad y un pH relativamente alto o viceversa.

El Departamento de Protección Ambiental de los Estados Unidos (USEPA, por sus siglas en inglés United States Environmental Protection Agency) no hace recomendaciones respecto a la alcalinidad en fuentes de agua, ya que ésta se liga a factores como el pH y la dureza, pero concluye que una fuente no debe mostrar cambios bruscos o repentinos en el contenido de la alcalinidad, pues esto podría indicar un cambio en la calidad del agua. (USEPA, 2000).

1.1.2.2 Cloruros

Las aguas superficiales normalmente no contienen cloruros en concentraciones tan altas como para afectar el sabor, excepto en aquellas fuentes provenientes de terrenos salinos o de acuíferos con influencia de corrientes marinas.

A partir de ciertas concentraciones, los cloruros pueden ejercer una acción disolvente sobre ciertas sales presentes en el agua y también sobre algunos componentes del cemento, al impartirles una acción corrosiva y erosionante, especialmente a pH bajo.

Cuando las concentraciones de cloruros presentes en el agua pasan el nivel de 250 ppm, concentración permisible en aguas para consumo humano, los iones Cl-

pueden reaccionar con el ion H^+ presente en el cuerpo de agua y formar HCl, el cual alteraría los procesos biológicos del cuerpo de agua.

1.1.2.3 Oxígeno Disuelto (OD)

Su presencia es esencial en el agua; proviene principalmente del aire. Niveles bajos o ausencia de oxígeno en el agua puede indicar contaminación elevada, condiciones sépticas de materia orgánica o una actividad bacteriana intensa; por ello se le puede considerar como un indicador de contaminación.

La presencia de oxígeno disuelto en el agua cruda depende de la temperatura, la presión y la mineralización del agua. Las aguas superficiales no contaminadas, si son corrientes, suelen estar saturadas de oxígeno y a veces incluso sobresaturadas; su contenido depende de la aireación, de las plantas verdes presentes en el agua, de la temperatura y de la hora del día (mañana o tarde). Se ha demostrado la existencia de una estrecha relación entre la distribución de oxígeno y la productividad de materia orgánica, viva o muerta. Por otro lado, la cantidad de OD en un cuerpo de agua está relacionada con su capacidad de autodepuración. (Wetzel, 2001).

1.1.2.4 Nitratos

En los nitratos está presente en anión NO_3^- , el nitrógeno es estado de oxidación +5 se encuentra en el centro de un triángulo formado por los tres oxígenos. Los nitratos constituyen parte de los nutrientes esenciales para muchos organismos autótrofos o fotosintéticos y en este sentido, se presencia en el agua puede ocasionar fenómenos de eutrofización en ríos y lagos. La eutrofización es el crecimiento desmedido de algunas especies vegetales que cubren con un manto vegetal la superficie del agua impidiendo de esta forma su oxigenación natural. (Sierra, 2011).

1.1.2.5 Fosfatos

Son nutrientes para las plantas. Tienen aplicaciones industriales diversas y como fertilizantes. Los vestidos de PO_4^{-3} , a las aguas naturales pueden causar eutrofización. De la utilización de los fosfatos y nitratos, presentes en la materia orgánica de la basura de los detergentes hechos a base fosfatos, ocurre el proceso eutrofización, ya que estos arrojados a los ríos y lagos. Durante las épocas cálidas, la

sobrecarga a estos productos químicos, que sirven de nutrientes, generan el crecimiento acelerado de vegetales como algas, cianobacterias, lentejas de agua, los cuales al morir y ser descompuestos por las bacterias aeróbicas provocan el agotamiento del oxígeno disuelto en la capa superficial de agua y causan la muerte de los diferentes tipos de organismos acuáticos que consumen oxígeno, en las aguas de los lagos y ríos. (Romero, 2004).

1.1.2.6 Clorofila-a

La reproducción de las algas guarda estrecha relación con la naturaleza de los distintos hábitats, caracterizados a su vez por diferentes factores ecológicos como la luz, la temperatura, los nutrientes como los nitratos y los fosfatos, el oxígeno, el anhídrido carbónico y las sales minerales. El incremento anormal de las algas se produce por el exceso de nutrientes y cambios en la temperatura. Cuando las algas traspasan ciertos valores por unidad de volumen valores que dependen de la especie de alga predominante, la temperatura del agua, el tipo de tratamiento, etcétera, causan problemas en los cuerpos de agua y las plantas de tratamiento. Estos problemas son los siguientes (Sierra, 2011).

Sabor y olor. Se ha detectado que algunas algas producen olor a pescado, tierra y pasto, entre otros.

Color. La abundancia de las algas clorófitas produce un color verde en el agua; otras, como la *Oscillatoria Rubens*, originan un color rojo.

Toxicidad. Algunos tipos de algas azul-verdes, actualmente denominadas *cyanobacterias*, causan disturbios gastrointestinales en los seres humanos.

Alteración del pH. Esta alteración se produce debido al consumo de CO₂ con precipitación de CaCO₃, lo que aumenta el pH.

La concentración de clorofila-a se utiliza para estimar en forma indirecta la biomasa de las comunidades fitoplanctónicas, debido a que es el principal pigmento fotosintético presente en las algas (Gregor y Marsálek, 2004). La clorofila-a también es un indicador del grado de contaminación de los ecosistemas acuáticos y un importante índice del estado fisiológico del fitoplancton (Pinto *et al.*, 2001). Para estimar la clorofila existen diferentes métodos, procedimientos y modelos

matemáticos, desarrollados con el objeto de minimizar los errores ocasionados por la presencia de otros pigmentos fotosintéticos y de algunos compuestos químicos en el agua (APHA, 1998).

1.2 Redes Neuronales Biológicas

Una neurona (nombre que se les da a las células del cerebro), es una célula biológica especial que procesa información. Este tipo de células puede presentarse en múltiples formas, aunque muchas de ellas muestran un aspecto similar (Figura 1).

Cada neurona tiene tres componentes principales: el cuerpo celular o soma, las dendritas y el axón. El cuerpo de la célula (con un tamaño que varía entre los 10 y 80 μm), tiene un núcleo donde se almacena información acerca de la herencia y un plasma que permite suplir las necesidades materiales de la célula. Las dendritas son los receptores de la información proveniente de otras neuronas, mientras que el axón es una fibra larga (su longitud varía desde 100 μm hasta 1 m en el caso de las neuronas motoras), que también se ramifica en el extremo y transmite la información generada en la neurona hacia otras (en el caso de las neuronas motoras, la información se transmite directamente al músculo). Existe otro tipo de neuronas (las sensoras), que en lugar de recibir la información de otras neuronas, la reciben desde el exterior (por ejemplo, en la retina). Se estima que una neurona de la corteza cerebral recibe en promedio información de otras 10000 neuronas y envía impulsos a varios cientos de ellas (Chapra, 1977).

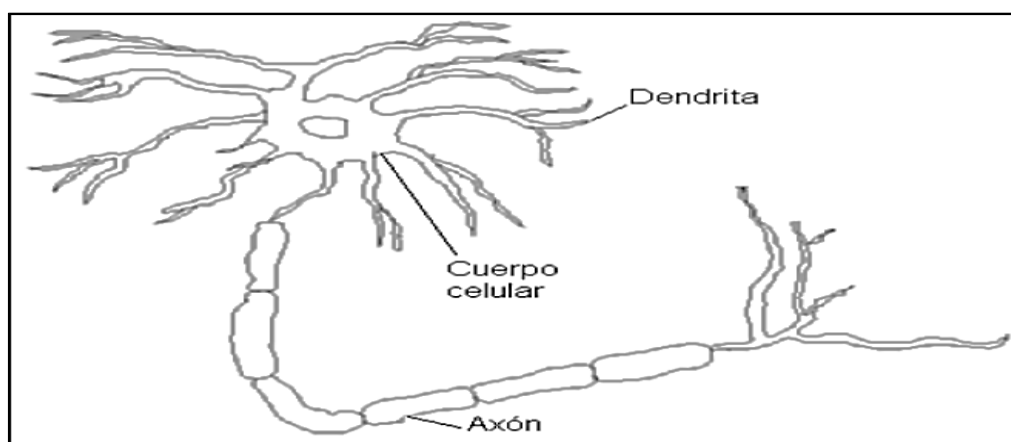


Figura 1. Estructura básica de una Neurona

Fuente: Chapra, (2001).

1.3 Redes Neuronales Artificiales (RNA)

La alta capacidad que presenta el cerebro humano en la realización de tareas cognitivas complejas como el reconocimiento de formas (visión) y sonidos (audición), fue una gran motivación para intentar modelar su comportamiento. De esta motivación histórica es de donde deriva el nombre de este tipo de modelos: redes neuronales artificiales. Sin embargo, la mayoría de las arquitecturas de redes neuronales que se emplean en la actualidad no tratan de imitar detalladamente el modelo biológico, sino que simplemente son algoritmos que en lugar de trabajar secuencialmente, lo hacen en paralelo. (Almeida, 2002).

En los últimos años las redes neuronales artificiales se han convertido en una herramienta muy utilizada en diversas áreas científicas, como la econometría y las finanzas, la quimiometría, la automática, los recursos hídricos y las ciencias Medio ambientales. Deben principalmente su creciente popularidad a la versatilidad de sus estructuras y algoritmos, que además de ser sencillos y fáciles de implementar, permiten abordar problemas complejos, y proporcionan en muchos casos soluciones adecuadas a la vez que robustas a dichos problemas. (Boznar, 2004).

1.3.1 La Neurona Artificial

La neurona artificial es un procesador simple de cálculo que, a partir de un vector de entrada procedente del exterior o de otras neuronas, genera una respuesta única o salida. Como puede observarse en las Figuras 2 y 3, consta de varias partes:

- El conjunto de entradas X_i , que representan las señales procedentes de otras neuronas o del exterior y que son capturadas por las dendritas;
- Los pesos sinápticos, que representan la intensidad de la interacción entre la neurona presináptica (o el exterior) y la neurona postsináptica;
- El umbral representa el nivel mínimo que debe alcanzar el potencial postsináptico para que la neurona se active;
- La regla de propagación ó función de activación, que permite obtener el valor del potencial postsináptico es decir, proporciona la manera de obtener el valor de activación a partir de la regla de las entradas, los pesos sinápticos y del umbral;

- La función de transferencia o de salida, que se aplica al valor de activación y proporciona la salida de la neurona;

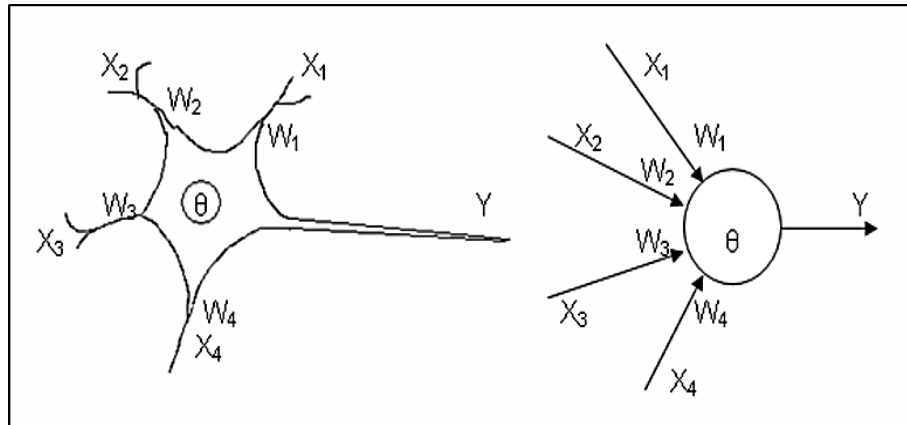


Figura 2. Neurona biológica y neurona artificial.

Fuente: Choi, (2001)

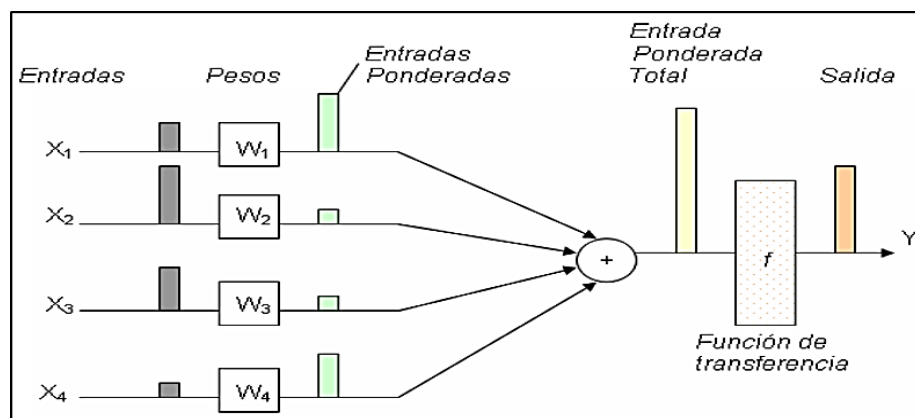


Figura 3. Proceso en una neurona artificial.

Fuente: Dedecker, (2004)

La salida de una neurona se conecta a la entrada de la siguiente neurona. Mediante este tipo de conexiones se puede realizar una red neuronal tan compleja como se quiera. Las entradas y salidas pueden ser binarias o continuas. Dependiendo del tipo de salida, las neuronas reciben nombres específicos. Así por ejemplo, aquellas neuronas que sólo pueden entregar como resultado un 0 o un 1, reciben el nombre de neuronas de tipo Mc Culloch – Pitts. Si tiene como salidas los valores +1 y -1, se las llama neuronas tipo Ising.

Las neuronas tipo Potts son aquellas que pueden adoptar diversos valores discretos como salidas. Los valores que entregan las neuronas de salida continua pueden ir de $-\alpha$ a $+\alpha$, pero generalmente se las limita a un intervalo definido (por ejemplo, $[-1,+1]$). (Gupta & Rao, 1994)

Respecto a la notación, aún no ha sido establecida una notación matemática estándar para las redes neuronales. En este Proyecto se adopta la siguiente convención:

- Valores escalares: se representan mediante una letra minúscula itálica;
- Vectores: se representan mediante una letra minúscula en negrita;
- Matrices: se representan mediante una letra mayúscula en negrita;
- Parámetros: se representan mediante letras minúsculas con un supra índice que indica la capa a la que corresponde el parámetro y uno o dos subíndices que indican la/s neurona/s que participa/n del proceso. En caso de ser un peso sináptico, la notación sería (Gupta & Rao, 1994)

$$W_{n,m}^C \quad (1)$$

Esto indicaría que se trata del peso correspondiente a la capa C y que la conexión es entre la neurona n y la neurona m . Los números n y m corresponden a la neurona postsináptica y presináptica, respectivamente. Si bien existen neuronas con reglas de propagación de tipo no lineal (por ejemplo, aquellas que implican una interacción de tipo multiplicativo entre las entradas y los pesos; o la distancia euclídea, que tiene en cuenta el cuadrado de la distancia que existe entre el vector de entrada y los pesos), su uso es limitado y las reglas más usadas son de tipo lineal (sumas ponderadas de pesos y entradas equivalente al producto escalar de los vectores pesos y entradas). Para cualquier capa, la notación sería (omitiendo la indicación de capa) (Hong, S. 2007)

$$h_i = \sum_{k=1}^n w_{ik} x_k + \theta = \mathbf{w}_k \cdot \mathbf{x} + \theta \quad (2)$$

h_i : salida de la neurona i

θ_i : umbral de la neurona i

n: número de señales de entradas a la neurona i Se suele incorporar el valor del umbral dentro del producto escalar. Para ello, se supone una entrada adicional, de valor 1, y con un peso sináptico θ . Con ello:

$$h_i = \sum_{k=1}^{n+1} w_{ik} x_k \tag{3}$$

1.3.2 Análisis Del Perceptrón Simple

Para entender el comportamiento del Perceptrón, se principia analizando una red constituida por una neurona con dos entradas, como la mostrada en la figura 4

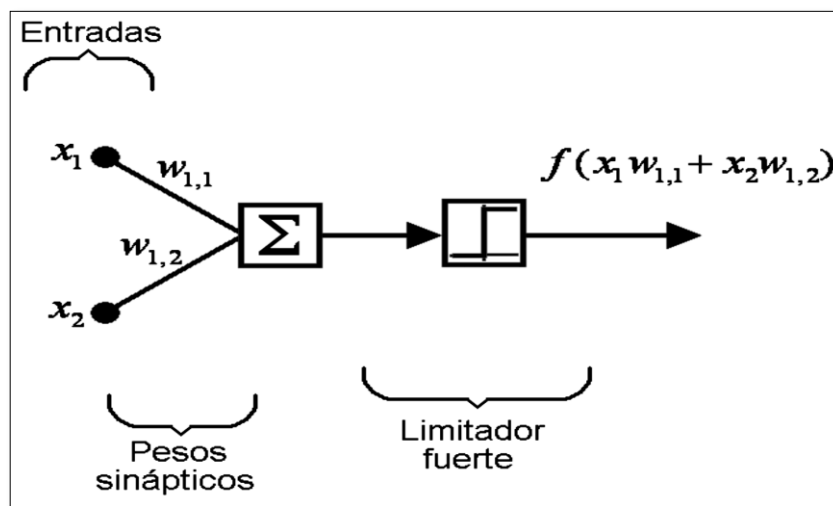


Figura 4. Perceptrón de una neurona con dos entradas.

El bloque con el símbolo de sumatoria indica la suma de las entradas multiplicadas por sus pesos sinápticos y el último bloque representa la función de activación limitador fuerte. El valor de salida de la red está determinado por la siguiente función:

$$Salida = \begin{cases} 1 & \text{si } (w_{1,1}x_1 + w_{1,2}x_2) \geq 0 \\ 0 & \text{si } (w_{1,1}x_1 + w_{1,2}x_2) < 0 \end{cases} \tag{4}$$

Donde x_1, x_2 son las entradas y $w_{1,1}, w_{1,2}$ son los pesos sinápticos correspondientes. La ecuación del umbral de decisión es la siguiente:

$$w_{1,1}x_1 + w_{1,2}x_2 = 0 \quad (5)$$

Al graficar esta ecuación en un plano coordenado donde los valores de x_1 se representan en el eje horizontal y los valores de x_2 en el eje vertical, se obtiene una línea recta con pendiente $m = -w_{1,1}/w_{1,2}$ que pasa por el origen como la mostrada en la figura 5

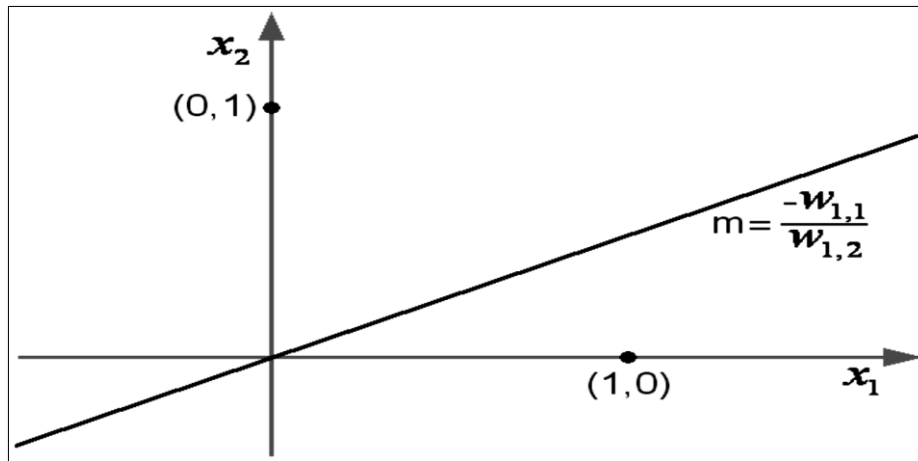


Figura 5. Gráfica x_2 vr. x_1 .

En esta gráfica se observa que la línea divide al plano x_2 - x_1 en dos regiones, los puntos (x_1, x_2) en la región sobre la línea darán una salida, mientras que los puntos bajo la línea darán otra salida; por ejemplo, los puntos $(1,0)$ y $(0,1)$, mostrados en la figura 5, darán una salida diferente. Los valores de los pesos sinápticos $w_{1,1}$ y $w_{1,2}$ se ajustan durante el aprendizaje para obtener la pendiente adecuada según las salidas deseadas.

1.3.3 El Algoritmo Backpropagation

El algoritmo *Backpropagation* es el método de entrenamiento más empleado en redes con conexión hacia delante. Es un método de aprendizaje supervisado de gradiente descendente, en el que se distinguen claramente dos fases: primero se aplica un patrón de entrada, el cual se propaga por las distintas capas que componen la red hasta producir la salida de la misma. Esta salida se compara con la salida deseada y se calcula el error cometido por cada neurona de salida. Estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de las capas intermedias (Fritsch, 1996).

Cada neurona recibe un error que es proporcional a su contribución sobre el error total de la red. Basándose en el error recibido, se ajustan los errores de los pesos sinápticos de cada neurona.

1.4 Perceptrón Multicapa

El Perceptrón multicapa (MLP) se basa en la presente comprensión del sistema nervioso biológica. Es un sistema paralelo masivo compuesto de muchos elementos de proceso conectados por enlaces de peso variable. Entre los muchos paradigmas MLP, la red de propagación inversa es de lejos el más popular (Haykin 1998).

La red consiste en capas de unidades de procesamiento en paralelo o neurona. Cada capa está conectada a la capa de procedimiento por puntos fuertes de interconexiones o pesos, W . Fig.1 ilustra una red de tres capas MLP que consiste en capas de i , j , y k , con los pesos con interconexiones w_{ij} y w_{jk} entre las capas de las neuronas. Los pesos asignados iniciales se corrigen progresivamente durante el proceso de formación. En este proceso, las salidas predichas por MLP se comparan con las salidas conocidas, y los errores se propagan hacia atrás (de izquierda a derecha en la figura 9) Para determinar los ajustes de peso apropiadas necesarias para minimizar los errores. En este estudio, la Levenberg Marquardt (Marquardt 1963) Algoritmo se utiliza para ajustar los pesos MLP (Marquardt 1963; kisi 2004; Kisi et al, 2011). La información detallada sobre MLP se puede encontrar en Haykin (1998).

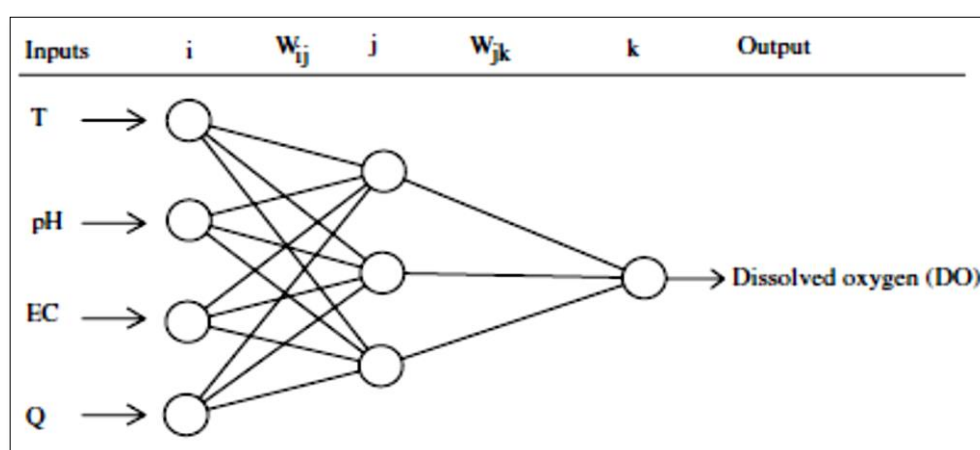


Figura 6. Diagrama esquemático de la topología MLP.

1.5 Aplicaciones de las redes neuronales artificiales a la calidad del agua

Las redes neuronales artificiales están siendo empleadas en una variedad de disciplinas con el fin de resolver problemas complejos o buscar una nueva forma de resolverlos; incluyendo optimización, predicción, control o reconocimiento de patrones; algunas de las aplicaciones de dichas redes en el campo de las ciencias ambientales principalmente para control de la contaminación de los cuerpos de agua.

Hoy en día, el oxígeno disuelto (OD) de control de concentración es el método más ampliamente utilizado. El oxígeno es un sustrato clave; Por lo tanto, el seguimiento de su consumo es uno de los métodos más eficaces de seguimiento del proceso de tratamiento (Shen, Chen, y Corriou, 2008).

El nivel de OD en un reactor aerobio voluntad influir en el comportamiento y la actividad de los microorganismos heterótrofos y autótrofos que viven en el lodo activado. La concentración de OD durante el proceso de lodo activado debería ser suficiente para suministrar oxígeno a los microorganismos en el lodo y por lo tanto degradar las sustancias orgánicas y convertir de amonio en nitrato alta. Sin embargo, un nivel excesivamente alto de la OD, lo que requiere un alto en el aire velocidad de flujo, dará lugar a un alto nivel de consumo de energía y puede dañar la calidad de lodos. La investigación muestra que un alto nivel de OD en agua recirculada internamente también disminuye la eficiencia del proceso de Nitrificación (Shen et al., 2008).

El desarrollo y el progreso actual en la integración de las diversas técnicas de inteligencia artificial (sistema basado en el conocimiento, algoritmos genéticos, redes neuronales artificiales, y del sistema de inferencia difusa) en, por ejemplo, modelización de la calidad del agua, los sedimentos de transporte, la evaporación, la concentración de OD, estimación profunda e integral de la OD ha sido estudiado por muchos investigadores. Por ejemplo, las redes neuronales artificiales (RNA) se han utilizado con éxito en varios campos de la estimación y los métodos de previsión (Chatterjee y Hadi 1986; Broomhead y Lowe 1988; Hagan y Menhaj 1994; Haykin 1994; Lerner et al. 1994; Haykin 1998; kisi 2004; kisi 2005; Cigizoglu y Kisi 2006; kisi 2007; Hanbay et al. 2008; Dogan et al. 2009; kisi 2009; Ranković et al. 2010; Akkoyunlu et al. 2011).

Calidad del agua superficial es una de las cuestiones más importantes hoy en día debido a sus efectos sobre la salud humana y los ecosistemas acuáticos. Se determina en gran medida por ambos procesos naturales, incluyendo la litología de la cuenca, entradas atmosféricas y condiciones climáticas, y por las entradas antropogénicas (Wen et al. 2012). Mediciones de la calidad del agua incluyen una variedad de parámetros biológicos físicas, químicas y. problema básico en el caso del monitoreo de la calidad del agua es la complejidad asociada con el análisis de la gran cantidad de variables (Rankovi do et al. 2010). El oxígeno disuelto (OD) representa la cantidad de Oxígeno existentes en el cuerpo de agua. Es un parámetro crítico utilizado con frecuencia y de forma continua para determinar la calidad del agua de los ríos y uno de los parámetros clave cuando se analiza la calidad del agua (Pengorur et al. 2006). El nivel de OD es también un mea-seguro de la salud del sistema acuático. Se requiere un cierto nivel mínimo de OD en agua para la vida acuática para sobrevivir (Basant et al. 2010).

Modelización de la calidad del agua se reconoce cada vez más como una herramienta útil para adquirir información valiosa para la gestión de la calidad del agua óptima (Vieira et al. 2013). Dado que un gran número de factores, tales como la temperatura, la descarga de agua, sedimentación, aireación, la descomposición, la nitrificación y la fotosíntesis, afecta a la OD en agua, una relación no lineal entre las variables de entrada y la de OD se puede esperar.

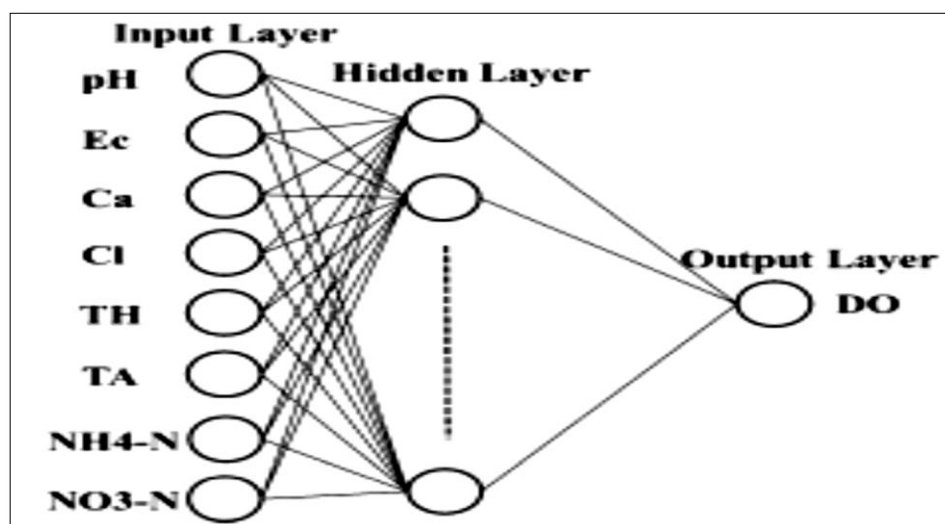


Figura 7. Red neuronal conceptual general para el DO en el río Heihe.

1.5.1 Topología de red neuronal artificial para la clorofila alfa

La eficacia de la RNA como herramienta predictiva no se ha abordado plenamente en algunas de las obras. La dinámica de algas (por ejemplo, como se representa por la biomasa de algas) en el momento actual se une mediante el modelo con otras variables ambientales, al mismo tiempo (Recknagel et al., 1997; Whitehead et al., 1997; French et al., 1998; Karul et al., 2000). Si bien este enfoque podría ofrecer una idea de las variables de control, no puede ser utilizado como una herramienta de pronóstico. Por Ejemplo, en un caso, se utilizó el recuento de células de una especie de algas como la salida de red deseada, con la concentración de clorofila a como un nodo de entrada. En una situación realista, ambos son exactamente las incógnitas críticas que estamos buscando. De manera similar, la nutriente concentración en el momento actual a menudo se ha utilizado como un nodo de entrada; esto sólo se puede obtener a través de un sistema de monitorización continua. Sin embargo, incluso en aguas costeras eutróficas a menudo existen las concentraciones de nutrientes (por ejemplo, amoniacó nitro-GEN) (Yabunaka et al., 1997; Maier et al., 1998).

La capa de salida ajusta de forma incremental los pesos de conexión entre la salida y las capas ocultas y de entrada, en ese orden. La actualización de peso se realiza de acuerdo con el algoritmo de retro propagación como sigue:

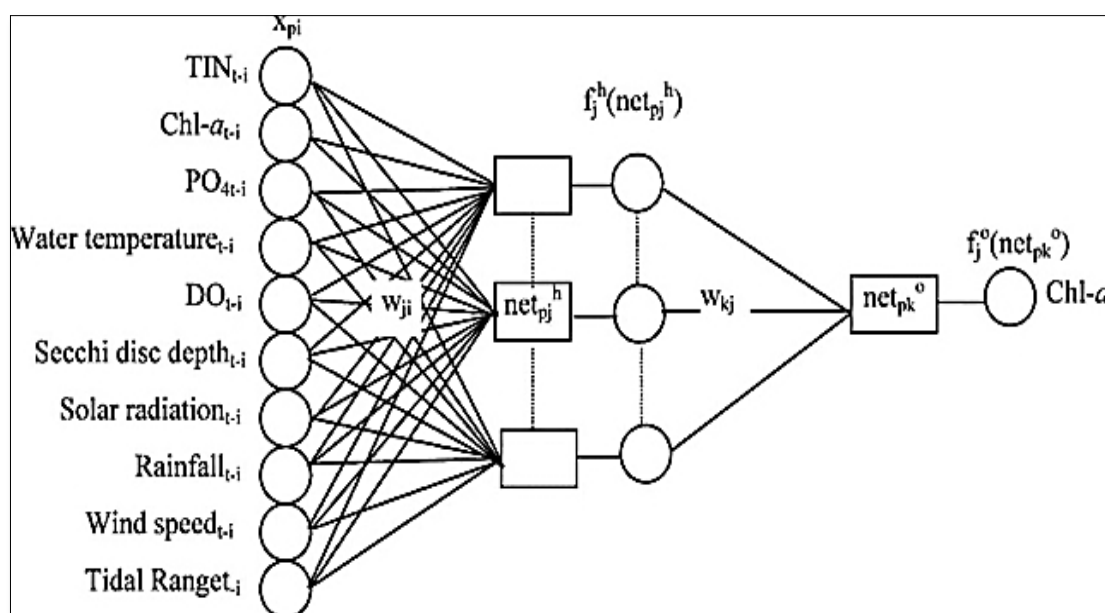


Figura 8. Red general de predicción de 7 días de floraciones de algas en Todo Harbor

Donde h y a son coeficientes conocidos, respectivamente, como el aprendizaje de tasa y el impulso plazo, k es contraproducente época y E_p es una función objetivo (error). La tasa de aprendizaje es un indicador de la tasa de convergencia. Si es demasiado pequeña, la tasa de convergencia será lenta, debido al gran número de pasos necesarios para alcanzar el mínimo error. Si es demasiado grande, la convergencia inicialmente será rápida, pero va a producir oscilaciones excesivas, y puede no alcanzar el mínimo error. El impulso plazo es un parámetro que se suma a la inercia del cambio de peso en función de la dirección de su cambio anterior. Tiene la tendencia a aumentar la velocidad de convergencia. También mejora la capacidad de aprendizaje de tops cuando los datos de entrenamiento contienen ruido y su valor se encuentra entre cero y la unidad. Un valor próximo a cero da más peso al error actual de los errores anteriores, y viceversa.

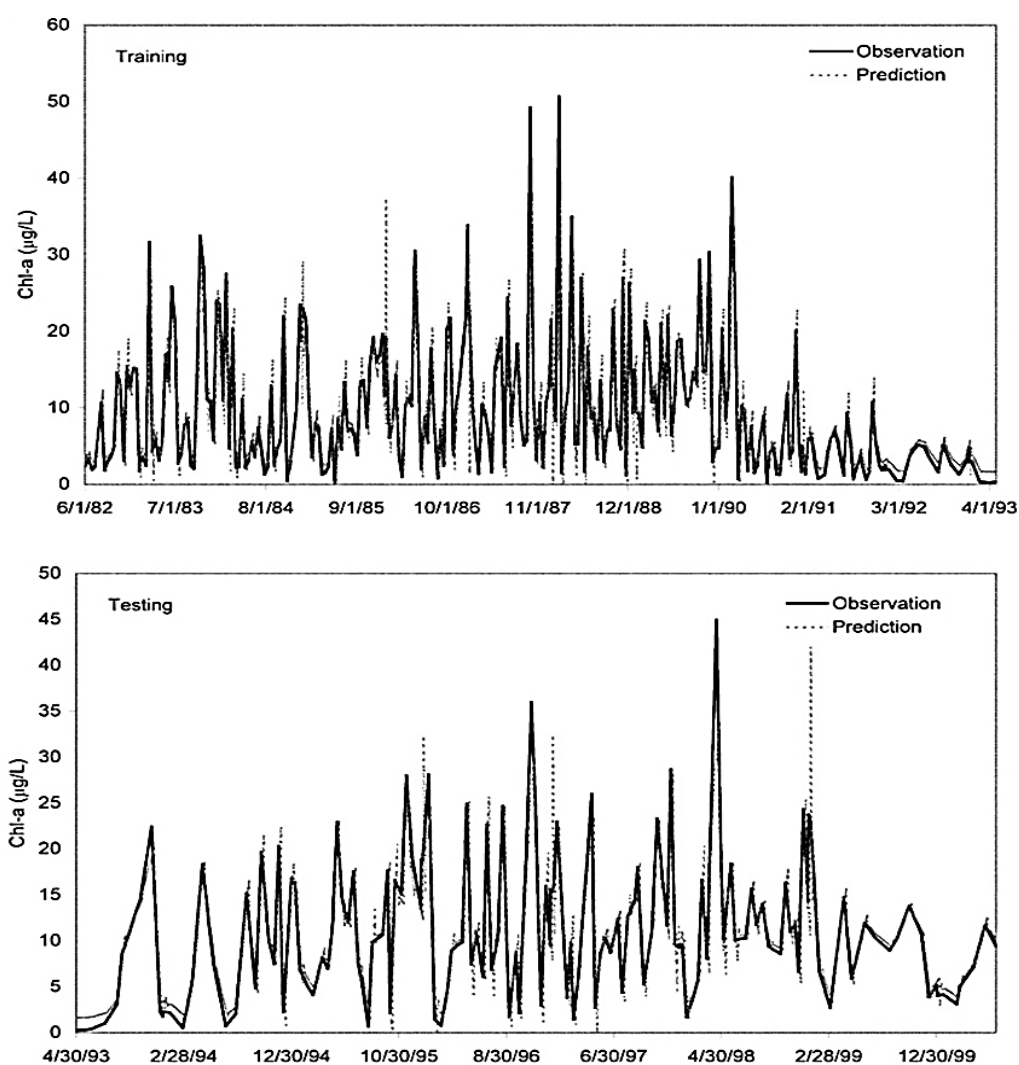


Figura 9. Comparación del valor predicho (plomo tiempo de 7 días) y clorofila-a

1.5.1 Función De Activación

La función de activación determina el estado de activación actual de la neurona en base al potencial resultante h_i y al estado de activación anterior de la neurona $a_i(t-1)$. El estado de activación de la neurona para un determinado instante de tiempo t puede ser expresado de la siguiente manera:

$$a_i(t) = f_i(a_i(t-1), h_i(t)) \quad (6)$$

Sin embargo, en la mayoría de los modelos se suele ignorar el estado anterior de la neurona, definiéndose el estado de activación en función del potencial resultante h_i :

$$a_i(t) = f_i(h_i(t)) \quad (7)$$

A continuación, se muestra las funciones de activación más utilizadas en los distintos modelos de red neuronales artificiales.

1.5.1.1 Función tangente hiperbólica

La función de activación tangente hiperbólica se emplea en los casos que presentan variaciones suaves de valores positivos y negativos de la señal a clasificar. Como se puede ver en su descripción es una de las funciones más empleadas en entrenamientos supervisados, como en el caso del entrenamiento de retropropagación del error.

Debe de tenerse cuidado de emplear esta figura entre los umbrales positivos y negativos antes de la saturación, de otra forma la salida siempre generará valores saturados iguales a 1 y -1.

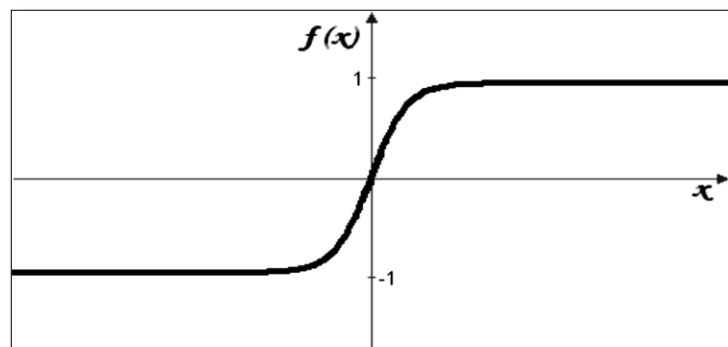


Figura 10. Función hiperbólica.

Esta función puede representarse mediante la siguiente expresión:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (8)$$

1.5.1.2 Función sigmoïdal logarítmica

Esta función se asemeja a la función limitador fuerte, ya que su salida está comprendida entre cero y uno, pero a diferencia de ésta, es continua en todo su dominio y por tanto, es diferenciable. Debido a que muchos algoritmos de aprendizaje, necesitan hacer uso de la derivada de la función de activación, esta función tiene mayor utilidad que la función limitador fuerte, la cual no es diferenciable en cero.

En la función sigmoïdal logarítmica, los valores de entrada total un poco mayores que cero, poseen una salida muy cercana a la unidad y entradas un poco menores que cero, tienen una salida cercana a cero. En la figura 7 se muestra la respuesta de esta función.

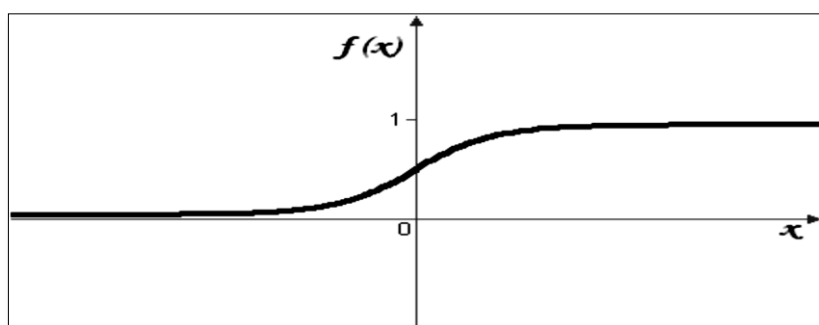


Figura 11. Función sigmoïdal logarítmica.

Esta función puede representarse mediante la siguiente expresión:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

1.5.1.3 Función lineal o purelin

Para obtener los valores correspondientes de concentraciones relativas, tomando como entradas a la función los valores emitidos por las neuronas de la capa oculta. Esta función lo que hace es calcular los valores de salida de la(s) neurona(s) de la siguiente manera:

$$\text{Purelin}(n) = \text{purelin}[(\mathbf{Wp}) + \mathbf{b}] = (\mathbf{Wp}) + \mathbf{b}$$

Donde W es la matriz de pesos, p son los valores de entrada y b es el vector de sesgos.

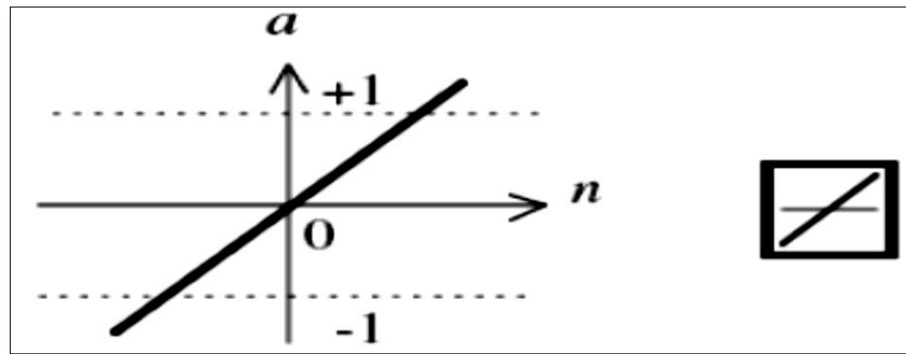


Figura 12. Función de transferencia tipo lineal.

Fuente: Demuth *et al.*, (2005)

1.5.1.4 Algoritmo de Levenberg Marquardt

El algoritmo de Levenberg Marquardt puede aplicarse en cualquier problema donde se necesite encontrar los valores de las raíces de una función; en el caso de las redes neuronales artificiales, la función es el error cuadrático medio de las salidas de la red y las raíces de esta función son los valores correctos de los pesos sinápticos desarrollados. (Rankovic et al, 2010)

Para encontrar los valores de los pesos sinápticos utilizando este algoritmo deben realizarse operaciones matemáticas con matrices; la matriz de error en los pesos sinápticos se calcula de la siguiente forma:

$$W_{\text{nuevo}} = W_{\text{actual}} - [J^T \cdot J + \mu \cdot I]^{-1} \cdot J^T \cdot E \quad (10)$$

Donde W es una matriz constituida por valores de los pesos sinápticos, I es la matriz de identidad, μ es un valor escalar que varía de acuerdo al cambio en el error; E es la matriz de errores en la red y J es la matriz Jacobiana, ésta es calculada de la siguiente forma:

$$J = [\nabla E \cdot E^{-1}]^T \quad (11)$$

Donde ∇E es la matriz de gradientes de error en valores numéricos.

El algoritmo de Levenberg Marquardt converge a la solución en menos iteraciones que la regla delta generaliza, pero cada iteración requiere más tiempo, ya que se realizan más operaciones. Cuando se entrena a la red utilizando una gran cantidad de ejemplos, es mejor utilizar el algoritmo de Levenberg Marquardt, ya que permite obtener la solución en un menor tiempo. (Rankovic *et al.*, 2010)

1.6 Antecedentes

Según, Obregón N., (2001), en su publicación titulada *Principios de Hidrodinámica Urbana* considera que actualmente se tiene conocimiento de métodos numéricos que han sido utilizados dentro de las ciencias ambientales e hidráulicas como son la regresión multivariable, diferencias finitas, elementos finitos, volúmenes finitos, que han posibilitado el estudio y la investigación en áreas de la ingeniería en Perú; sin embargo se tiene un limitado conocimiento del uso de las RNA en distintas áreas de investigación en Perú.

La necesidad de la explotación de los recursos naturales, en conjunto con la complejidad de los sistemas y de la aceptación de nuestras limitaciones de medición biofísica ha presionado por soluciones aceptables, pero no óptimas en el desarrollo de la infraestructura, aprovechamiento y conservación de los recursos naturales. Por esta razón las aproximaciones para los estudios de modelación han demandado, y demandan en la actualidad, nuevas técnicas que permitan mejorar las capacidades de predicción acerca de los procesos que tienen lugar en la naturaleza y por consiguiente buscar un desarrollo sustentable

Las ventajas que presentan las RNA hacen que este tipo de tecnología se esté aplicando en múltiples áreas como son: Bioquímica (Fagundes-Klen *et al.*, 2007), Ecología Acuática (Goethals *et al.*, 2007), Química (Kalach, 2006; Abbas y Al-Bastaki, 2005), Medio Ambiente (Shu *et al.*, 2006; Dimopoulos *et al.*, 2004; Tóth-Nagy *et al.*, 2006; Doğan *et al.*, 2007; Nagendra y Khare, 2005; Kuo *et al.*, 2007; Cherkassky *et al.*, 2007; Vladimir y Frederic, 2003; Çinar y Merdun, 2008), principalmente en la predicción de contaminantes como el Ozono (O₃) (García, 2003; Brunelli *et al.*, 2007; Wirtz *et al.*, 2005; Melas *et al.*, 2000) y partículas menores a 10 y 2.5 micrómetros, PM₁₀ y PM_{2.5}, respectivamente (Chaloulakou *et al.*, 2003; McKendry, 2002; Chelani *et al.*, 2002; Jiang *et al.*, 2004; Thomas y Jacko, 2007), Medicina (Gamito y Crawford, 2004; Tägil *et al.*,

2008), Militar (Hilera y Martínez, 2000) e Hidráulica (Coppola Jr. *et al.*, 2007; Altunkaynak, 2007; Ahmed y Sarma, 2007), por mencionar algunas.

En lo que respecta a la predicción de las concentraciones de clorofila a nivel internacional se tiene conocimiento de estudios realizados con el propósito de conocer la situación ecológica de algunos cuerpos de agua utilizando las RNA; como ejemplo, el desarrollado por Jan-Tai Kuo y colaboradores (2006), para predecir el grado de eutroficación de un cuerpo de agua. Dicho estudio ofrece un enfoque alternativo de la relación causa-efecto del proceso de eutroficación utilizando las RNA; el cuerpo de agua estudiado fue el Te-Chi, catalogado como el cuarto cuerpo de agua más importante en Taiwán.

Otro estudio de importancia similar es el realizado por Zeng Yong y colaboradores (2006), en el cual se utilizó una base de datos que tenía una frecuencia semanal para realizar dos modelos; uno utilizando las RNA y el segundo utilizando la regresión multivariable; con el objetivo de pronosticar las concentraciones de clorofila.

Un estudio realizado por R.O. Strobl y colaboradores (2007), en los Estados Unidos de América; con apoyo de investigadores pertenecientes a la Agencia de Protección Ambiental de los Estados Unidos (USEPA). El objetivo que persiguió dicho trabajo fue el de investigar qué parámetros fisicoquímicos tienen mayor importancia en determinados estados de eutroficación en los lagos que se encuentran en la zona Este del país.

De acuerdo con la bibliografía consultada el algoritmo de aprendizaje utilizado en estudios referidos a las ciencias ambientales es el método que en 1986, Rumelhart, Hinton y Williams (Hilera y Martínez, 2000), basándose en los trabajos de otros investigadores formalizaron una regla para que una RNA aprendiera la asociación que existe entre los patrones de entrada y los valores reales, utilizando más niveles de neuronas que los que utilizó Rosenblat para desarrollar el perceptrón; dicho método es conocido en general como retropropagación o propagación del error hacia atrás (backpropagation) y cuya importancia radica en su capacidad de autoadaptar los pesos de las neuronas de las capas intermedias para aprender la relación que existe entre un conjunto de patrones dados como ejemplo y sus salidas correspondientes para poder aplicar esa misma relación, después del entrenamiento, a nuevos datos de entrada.

Hong *et al.* (2011). En este trabajo, se propone un controlador adaptativo basado en una red neural estructura dinámica para controlar la concentración de oxígeno disuelto (OD) en un proceso de tratamiento de aguas residuales (PTAR). La propuesta incorpora una red de estructura feedforward variable de neural (FNN), donde el FNN puede determinar su estructura en línea de forma automática. La estructura de la FNN está adaptada para hacer frente a los cambios en las características de funcionamiento, mientras que los parámetros de peso se actualizan para mejorar la precisión del controlador. Una característica particularmente fuerte de este método es que la precisión control se puede mantener durante la adaptación, y por lo tanto el rendimiento de control no será degradado cuando el carácter de los cambios en el modelo varíe.

Saint-Gerons (2004). Indica que el oxígeno disuelto constituye una variable fundamental en la evaluación de la calidad del agua de los ríos, ya que determina la diversidad de organismos presentes en su seno. Para monitorear su valor, se ha desarrollado un sensor software haciendo uso de la técnica de redes neuronales. La red neuronal elegida es de tipo *backpropagation*, con una arquitectura sigmoideal/lineal de dos capas, que utiliza como entradas tres variables medibles: temperatura, turbidez y potencial redox. La solución propuesta se ha mostrado idónea en términos de error de estimación, complejidad y entradas necesarias. La red permite además, reducir costes en instrumentación, al ser los datos requeridos como entradas deducibles en una estación depuradora de agua residual (E.D.A.R.). Las pruebas de validación, realizadas sobre datos tomados de las estaciones medioambientales de Navarra, demuestran la efectividad de la solución propuesta.

Xiaohu *et al.* (2012). Identifican y cuantifican los perfiles de oxígeno disuelto (OD) de río es una de las principales preocupaciones para los gestores de recursos hídricos. En esta investigación, una red neuronal artificial (RNA) fue desarrollado para simular las concentraciones de OD en el río Heihe, noroeste de China. Una de tres capas de retropropagación RNA se utilizó con el algoritmo de entrenamiento regularización bayesiana. Las variables de entrada de la red neuronal pH eran, conductividad eléctrica, cloruro (Cl⁻), calcio (Ca²⁺), la alcalinidad total, la dureza total, nitrógeno de nitrato (NO₃-N), y el nitrógeno amoniacal (NH₄-N). La estructura RNA con 14 neuronas ocultas obtiene la mejor selección.

Schmid *et al.* (2006). Ellos tienden a ser particularmente útil cuando se aplica a procesos complejos, los detalles de los cuales no se conocen bien. El régimen de oxígeno disuelto en los estanques de humedales construidos es, a su vez, un proceso tan complejo, que se rige por un número considerable de los controles hidrológicos, hidrodinámicos y ecológicos que operan a una amplia gama de escalas espacio-temporales. Este documento informa sobre los resultados de un estudio llevado a cabo para probar la adecuación de las redes neuronales artificiales en el modelado de las concentraciones de cerca de fondo de oxígeno disuelto en el finlandesa humedal superficie de agua libre en Hovi. Se han desarrollado diversas redes diferentes del tipo perceptrón multicapa (MLP) de RNA.

La aplicación tuvo éxito, y en particular, se observó que MLP fueron capaces de “aprender” el mecanismo de transporte de oxígeno por convección bastante bien. La RNA también se utilizó para determinar la influencia relativa de la tasa de flujo y la cizalladura del viento sobre la saturación de oxígeno inferior cerca.

Murat *et al.* (2012). examinan la exactitud de dos técnicas diferentes de la red neural artificial (RNA), el perceptrón multicapa (MLP) y la red neural de base radial (RBNN), para estimar la concentración de oxígeno disuelto (OD). Los resultados RNA se comparan con el modelo de regresión multilínea (MLR). El modelo de red neural se desarrolla utilizando datos experimentales recogidos de la corriente arriba (USGS estación No: 07105530) y aguas abajo (USGS estación No: 07106000). Estaciones en Fundación Creek, las variables de entrada utilizados para los modelos de RNA son pH del agua, temperatura, conductividad eléctrica, y la descarga. El coeficiente de determinación (R^2), Error absoluto medio (MAE), y la raíz del error cuadrático medio (RMSE) estadísticas se utilizan para la evaluación de los modelos aplicados. Los modelos MLP y RBNN también se comparan con el modelo de MLR en la estimación de la OD de la estación aguas abajo mediante el uso de los parámetros de entrada de la estación aguas arriba. Resultados de la comparación indican que el modelo RBNN se comporta mejor que los modelos MLP y MLR.

CAPÍTULO II

PLANTEAMIENTO DEL PROBLEMA

2.1 Definición del Problema

En los últimos años la aplicación de las RNA ha ganado popularidad en la predicción y pronóstico en un gran número de áreas, incluyendo finanzas generación de energía, medicina recursos acuáticos y ciencias ambientales. Aunque el concepto de neurona artificial fue introducido desde el año 1943, y se ha hecho diferentes investigaciones al respecto de sus aplicaciones en distintos campos, una RNA puede ser considerada como una herramienta nueva en campo de los pronósticos y las predicciones. Debido a la constitución y fundamentos de las RNA, presentan un gran número de características semejantes a las del cerebro (Abbas y Al-Bastaki, 2005). Por ejemplo, son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos, de separar características esenciales a partir de entradas que representan información irrelevante, etc.; esto hace que ofrezcan numerosas ventajas como son: (Maren *et. al.*, 1990)

Auto organización. Las RNA usan su capacidad de aprendizaje adaptativo para auto organizar la información que reciben durante el aprendizaje y/o la operación. Mientras que el aprendizaje es la modificación de cada elemento procesal, la auto organización consiste en la modificación de la RNA completa para llevar a cabo un objetivo específico. Estas peculiaridades de la metodología a emplear, son elementales para poder modelizar y predecir el grado de contaminación con un grado de confiabilidad muy alto en la BIP.

2.1.1 Problema General

¿En qué medida la formulación de un modelo pronóstico, utilizando RNA permitirá analizar la calidad de un cuerpo de agua enfocado principalmente en oxígeno disuelto y clorofila?

2.1.1.1 Problemas específicos

¿En qué medida el modelo de RNA, a partir de una base de datos anteriores de variables físicas, químicas y biológicas predice satisfactoriamente el nivel de contaminación de aguas de la BIP?

¿En qué grado los modelos obtenidos podrán ser valorados mediante la significancia o prueba estadística P?

¿En qué medida se podrá evaluar el desempeño de los modelos obtenidos, mediante la variación de las condiciones de una base de datos?

2.2 Justificación

Uno de los problemas que más interés y preocupación han suscitado en los últimos años, es la contaminación del medio ambiente por productos generados por actividades antropogénicas. En gran parte del mundo, principalmente en las ciudades donde se concentra la mayor cantidad de población, la cantidad y calidad de agua disponible para satisfacer las necesidades humanas es limitada. La brecha entre el suministro de agua y la demanda aumentará en el transcurso del presente siglo como resultado del cambio climático y el incremento en el consumo de agua por la población, se incrementará considerablemente. En los próximos 25 años, por ejemplo, es improbable que la escorrentía accesible de agua dulce aumente más del 10%, sin embargo, se espera que la población mundial crezca una tercera parte. A menos que los humanos utilicen el agua de modo más eficiente, los impactos de este desequilibrio entre el suministro y la demanda disminuirán los servicios provistos por los ecosistemas de agua dulce, aumentarán el número de especies acuáticas que se enfrenten con la extinción, y se fragmentarán aún más los humedales, los ríos, los deltas y los estuarios (Conagua, 2008).

La calidad se ve afectada por problema de contaminación que son tanto característicos de países en vías de desarrollo (ocasionada por organismos patógenos) como por la presencia de tóxicos sintéticos frecuentes en países industrializados. Dichos problemas se generan tanto a partir de las descargas municipales como industriales, ya que el cumplimiento de la normatividad de las mismas es muy bajo; aun cuando los límites permisibles de la normatividad en el Perú. (ANA, 2014).

El Perú enfrenta problemas de disponibilidad de agua sino también por la calidad deficiente de varias de sus fuentes, como es el caso de la bahía interior de Puno, de

acuerdo con los informes del ANA, ALT Y PELT; es uno de las bahías más contaminadas del país, dicha laguna se encuentra eutrofizada lo cual limita sus usos. (ALT- Suma Marka, 2012).

En los últimos años a nivel internacional se ha desarrollado estudios que pretenden diagnosticar y determinar el grado de contaminación de algunos cuerpos de agua empleando las bondades de las RNA, sin embargo, en el Perú no se ha aplicado hasta el momento dicha tecnología. (Rodríguez, 2006).

Las RNA son una herramienta computacional que presentan un gran número de características semejantes a las del cerebro, es decir, son capaces de aprender de la experiencia; de generalizar de casos anteriores a nuevos casos; de abstraer características esenciales a partir de entradas que representan información irrelevante; cuentan con un aprendizaje adaptativo, ésta es una de las características más atractivas de las RNA, ya que es la capacidad de aprender a realizar tareas basadas en un entrenamiento o una experiencia inicial; las RNA no necesita un algoritmo para resolver un problema, ya que ella puede generar su propia distribución de los pesos de los enlaces mediante el aprendizaje y son de fácil inserción dentro de la tecnología existente. (Rodríguez, 2006).

Las características y ventajas que representa el uso de las RNA, aunado a la problemática ambiental en cuanto a los recursos hídricos en México abren una ventana de posibilidades a las aplicaciones de esta tecnología no sólo en cuestiones hidráulicas como especialmente se ha observado en la bibliografía al ser aplicadas en la determinación de flujos, caudales, escorrentía y Las características y ventajas que representa el uso de las RNA, aunado a la problemática ambiental en cuanto a los recursos hídricos en México abren una ventana de posibilidades a las aplicaciones de esta tecnología no sólo en cuestiones hidráulicas como especialmente se ha observado en la bibliografía al ser aplicadas en la determinación de flujos, caudales, escorrentía y precipitaciones; si no también a la aplicación en determinación de concentraciones de contaminantes presentes y que sólo es posible determinarlas mediante análisis fisicoquímicos, eliminando con esto, tiempos de determinación, errores de medición y ya que es un modelo computacional de fácil aplicación y de fácil inserción dentro de la tecnología existente las RNA ofrecen un alto grado de confianza en cuanto a la obtención de resultados en un plazo corto de estimación. (Comrie, 1997).

2.3 Objetivos

2.3.1 Objetivo General

Formular un modelo pronóstico, utilizando RNA que permita analizar la calidad de un cuerpo de agua enfocado principalmente en oxígeno disuelto y clorofila.

2.3.2 Objetivos Específicos

- Desarrollar un modelo de RNA, a partir de una base de datos física y química.
- Demostrar estadísticamente mediante la prueba P de significancia los modelos obtenidos de RNA.
- Evaluar el desempeño de los modelos obtenidos, mediante la variación de las condiciones de una base de datos.

2.4 Hipótesis

2.4.1 Hipótesis General

De acuerdo con la bibliografía consultada, las RNA han sido utilizadas satisfactoriamente en múltiples áreas, entre ellas las Bioquímica, Ecología acuática, Geología, Química, Militar, Hidráulica y Medio Ambiente; especialmente en el Perú se encuentra aplicada en aeronáutica, entonces es factible construir un modelo de RNA para pronóstico de oxígeno disuelto y clorofila presentes en cuerpos de agua superficiales.

2.4.2 Hipótesis Específico

- Los modelos de RNA, a partir de una base de datos anteriores de variables físicas y química predecirán satisfactoriamente el nivel de contaminación de aguas de la BIP.
- La demostración estadística de los modelos obtenidos podrá ser valorada mediante la prueba P.
- La evaluación del desempeño de los modelos obtenidos se realizará mediante las variaciones de las condiciones de una base de datos.

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1 Lugar de Estudio

El presente trabajo de modelización con red neuronal artificial RNA se desarrolló con los datos monitoreados por el Instituto del Mar del Perú “IMARPE” datos de la Bahía Interior Puno, perteneciente a la capital de la Región Puno. Cuyas coordenadas del punto de monitoreo son: Latitud S: 15° 50` 0,1” Longitud W: 70° 00` 9,0” de importancia poblacional y turística, cuyos estudios se centran en el pronóstico mediante redes neuronales artificiales la variación de oxígeno disuelto y clorofila. Parámetros que evidencian el grado de eutrofización de la Bahía interior Puno.

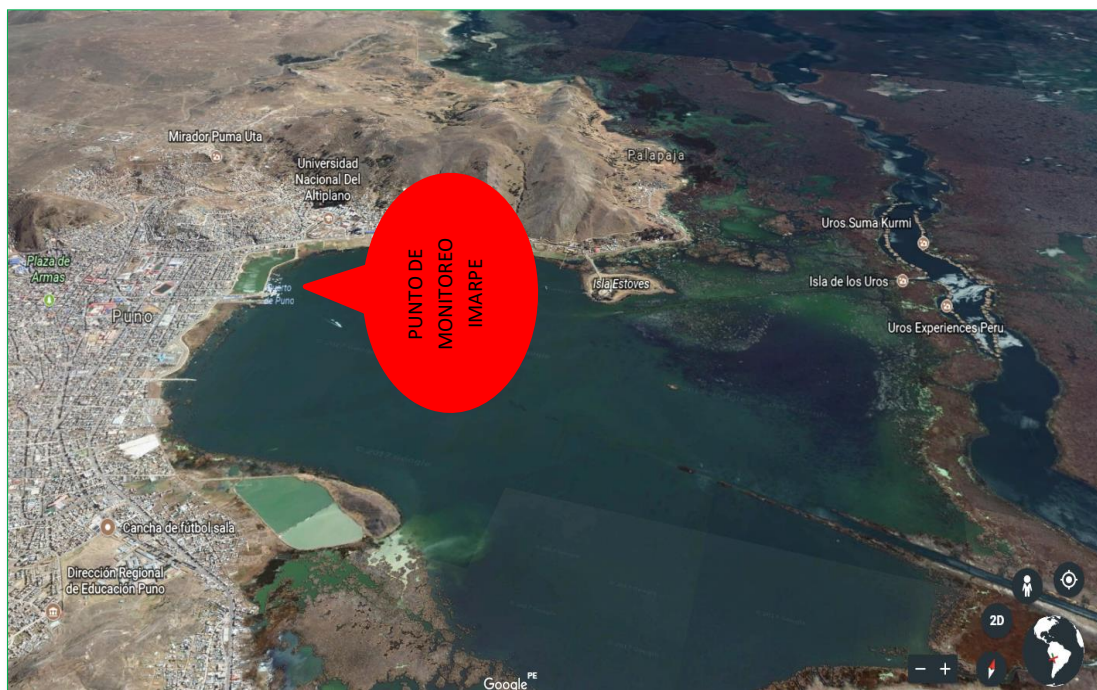


Figura 13. Mapa de localización geográfica del punto de monitoreo IMARPE.

3.2 Población

En el presente trabajo la población está representada por el cuerpo de 2 agua esencialmente por las variables fisicoquímicas, incluyendo las variables respuesta en este caso la concentración de oxígeno disuelto y la concentración de clorofila alfa presentes en este estudio.

3.3 Muestra

La muestra representativa está determinada por las siguientes variables; Sales, Temperatura, Clorofila alfa, Oxígeno disuelto, Conductividad, Nitratos y fosfatos.

3.4 Método

Se trata de una investigación de tipo cuantitativa, descriptiva y correlacional porque se busca predecir cómo se puede comportar la variable dependiente (OD y Chl - alfa) conociendo el comportamiento de las variables independientes relacionadas (algoritmos de entrenamiento, números de neuronas, números de iteraciones, número de observaciones y porcentaje de muestra).

Tabla 2

Variables suministradas para el modelo de la Bahía interior Puno.

Variable	Unidad de medición
Oxígeno Disuelto	mg/L
Potencial de Hidrogeno	pH
Conductividad Eléctrica	µs/Cm
Temperatura	°C
Nitratos	mg/L
Fosfatos	mg/L
Clorofila alfa	mg/L

3.5 Modelamiento con redes Neuronales Artificiales

Para el modelamiento con redes neuronales artificiales, se clasificaron variables como Oxígeno disuelto, potencial de hidrogeniones, conductividad eléctrica, temperatura, clorofila alfa, nitratos y fosfatos. Datos monitoreados por el IMARPE, con una data de 5 años contados desde el año 2013 hasta el 2016 realizando dos bloques de entrenamiento

una para OXÍGENO DISUELTO, cuyas variables de entrenamiento: fueron oxígeno disuelto, potencial de hidrogeniones, conductividad eléctrica y temperatura. CLOROFILA ALFA, cuyas variables fueron oxígeno disuelto, potencial de hidrogeniones, conductividad eléctrica y temperatura, nitratos y fosfatos para la predicción de ambos bloques. Para dicho fin se utilizó el programa MATLAB, versión 2014 a, dicho programa permite la construcción, entrenamiento, prueba y validación de modelos de redes neuronales artificiales. A continuación, se describe el procedimiento seguido en el modelamiento con redes neuronales.

3.5.1 Formulación De Modelos De Redes Neuronales Artificiales

De acuerdo a los objetivos planteados en esta investigación, se han formulado 6 modelos de redes neuronales artificiales, evaluando su performans para cada uno con un conjunto propio de variables de entrada y de salida denominada entrenamiento supervisado.

3.5.1.1 Determinación de la topología y parámetros de los modelos de redes neuronales artificiales

La determinación de la topología para cada uno de los modelos de redes neuronales artificiales se efectuó en forma anual.

En la práctica no existe una regla definida para establecer el número inicial de neuronas y capas intermedias u ocultas (Baganha, 2000). Por lo que la metodología seguida en esta investigación es la de Prueba – Error.

Para la selección adecuada del número de capas y neuronas intermedias de la red, se utilizó un conjunto de arquitecturas candidatas y a través de la técnica de validación cruzada se estimaron los parámetros, verificando el desempeño de los modelos sobre el conjunto de validación, siendo elegido el modelo que presentó el mejor desempeño en relación al error cuadrático medio.

3.5.1.2 Fase de entrenamiento de los modelos de redes neuronales artificiales

Este proceso consistió en evaluar distintas opciones de entrenamiento empleando tres algoritmos de entrenamiento diferente, Levenberg Marquardt, Regularización Bayesiana, Scalet Conjugate con el objetivo de minimizar el error en el pronóstico.

El algoritmo que se utilizó para hacer dichos modelos, fue el de Retropropagación o propagación del error hacia atrás (Backpropagation). El algoritmo de propagación del

error hacia atrás, es una regla de aprendizaje que se puede aplicar en modelos de redes con más de dos capas. Una característica importante de este algoritmo es la representación interna del conocimiento que es capaz de organizar en la capa intermedia de las neuronas para conseguir cualquier relación entre la entrada y la salida de la red.

El algoritmo de aprendizaje utilizado en el entrenamiento de todos los modelos de redes neuronales artificiales es el algoritmo de retro-propagación (back – propagation), que según Hilera y Martínez (2000) mencionado por Ovando et al. (2005) tiene el siguiente procedimiento.

Paso 1:

Inicializar los pesos de la red con valores aleatorios.

Paso 2:

Presentar un patrón de entrada $X_p : (x_{p1}, x_{p2}, \dots, x_{pN})$ y especificar la salida deseada que debe generar la red: (d_1, d_2, \dots, d_M) , si la red se utiliza como clasificador, todas las salidas serán cero, salvo una, la que sea de la clase a la que pertenece el patrón de entrada.

Paso 3:

Calcular la salida actual de la red, para la entrada presentada, se van obteniendo los valores de las respuestas que presenta cada capa hasta llegar a la capa de salida. Los subpasos que lleva este procedimiento son los siguientes:

- a) Se calculan las entradas netas (*net*) para las neuronas ocultas procedentes de las neuronas de entrada. Para una neurona oculta j (O_j):

$$net_{pj}^h = \sum_{i=0}^N w_{ji}^h x_{pi} + \theta_j^h$$

Donde el subíndice p corresponde al p -ésimo vector de entrenamiento, j a la j -ésima neurona oculta, w_{ji} es el peso de la conexión entre E_i y O_j y el término θ_j^h corresponde a un término de umbral mínimo a alcanzar por la neurona para su activación. A partir de estas entradas se calculan las salidas (y) de las neuronas ocultas, utilizando una función de activación f .

$$y_{pj}^h = f_j^h(net_{pj}^h)$$

- b) Se realizan los mismos cálculos para obtener los valores de resultado de cada neurona k de la capa de salida.

$$net_{pk}^o = \sum_{j=1}^L w_{kj}^o y_{pj}^h + \theta_k^o$$

$$y_{pk}^o = f_k^o(net_{pk}^o)$$

Para minimizar el error (fórmula en el Paso 6) la función f debe ser derivable, lo que implica la imposibilidad de utilizar la función escalón. En general se utilizan dos tipos de función de activación: la función lineal ($f(net_{jk}) = (net_{jk})$) y la función sigmoidea definida por la ecuación:

$$f(net_{jk}) = \frac{1}{1 + e^{-net_{jk}}}$$

La elección de esta función depende de la forma en que se decida representar los datos: si se desea que las neuronas de salida sean binarias, se utiliza la función sigmoidea, puesto que la función es casi biestable y además, derivable.

Paso 4:

Después que todas las neuronas de la red tienen un valor de activación asociado para un patrón de entrada dado, el algoritmo continúa encontrando el error que se presenta para cada neurona, excepto las de la capa de entrada. Para la neurona k de la capa de salida, si la respuesta es (y_1, y_2, \dots, y_M) , dicho error (δ) se puede escribir como:

$$\delta_{pk} = (d_{pk} - y_{pk}) f'_k(net_{pk})$$

Y para la función sigmoidea en particular:

$$\delta_{pk} = (d_{pk} - y_{pk}) y_{pk} (1 - y_{pk})$$

Si la neurona j no es de salida, entonces la derivada parcial del error no puede ser evaluada directamente. Por tanto, se obtiene el desarrollo a partir de valores que son conocidos y otros que pueden ser evaluados. La expresión obtenida en este caso es:

$$\delta_{pj}^h = x_{pi} (1 - x_{pj}) \sum_k \delta_{pk} w_{kj}$$

Donde observamos que el error en las capas ocultas depende de todos los términos del error de la capa de salida. De aquí el nombre de propagación hacia atrás.

El error que se produce en una neurona oculta es proporcional a la suma de los errores conocidos que se producen en las neuronas a las que está conectada la salida de la misma, multiplicando cada uno por el peso de la conexión. Los umbrales internos de las neuronas se adaptan de forma similar, considerando que están conectados con pesos desde entradas auxiliares de valor constante.

Paso 5:

Para la actualización de los pesos se utiliza el algoritmo recursivo, comenzando por las neuronas de salida y trabajando hacia atrás hasta llegar a la capa de entrada, ajustando los pesos de la forma siguiente:

- a) Para los pesos de las neuronas de la capa de salida.

$$w_{kj}(t+1) = w_{kj}(t) + \Delta w_{kj}^o(t+1)$$

$$\Delta w_{kj}^o(t+1) = \alpha \delta_{pk} y_{pj}^h$$

- b) Para los pesos de las neuronas de la capa oculta:

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji}^h(t+1)$$

$$\Delta w_{ji}^h(t+1) = \alpha \delta_{pj}^h y_{pi}$$

En ambos casos, para acelerar el proceso de aprendizaje, se incluye una tasa de aprendizaje que varía entre 0 y 1, dependiendo del problema a solucionar. También se puede añadir un término de momento, para corregir la dirección del error, de valor $\gamma(w_{kj}(t) - w_{kj}(t-1))$ en el caso de una neurona de salida, y $\gamma(w_{ji}(t) - w_{ji}(t-1))$, cuando se trata de una neurona oculta; la constante γ es la tasa de momento; y un tercer parámetro β para acelerar la convergencia del proceso.

Paso 6:

El proceso se repite hasta que el error medio cuadrático (E_p).

$$E_p = \frac{1}{2} \sum_{k=1}^M (\delta_{pk})^2$$

Resulta aceptablemente pequeño para cada uno de los patrones aprendidos. Este estadístico se utiliza para la evaluación objetiva del desempeño de la RN.

3.5.1.3 Modelo de Levenberg–Marquardt

Una vez fijo el conjunto de patrones con el que vamos a entrenar la red la fórmula del error de la red E tan sólo depende de los pesos y los umbrales. A partir de ahora llamaremos w al vector que contiene a los pesos y umbrales de la red.

El algoritmo de Retropropagación calcula el gradiente del error con respecto a los pesos y umbrales, es decir, $\nabla E(w)$ y actualiza los pesos e incrementos en función de ese vector. Con la nueva notación las Ecuaciones pueden resumirse en las dos siguientes:

$$\Delta w(t+1) = \alpha \Delta w(t) - \eta \nabla E(w(t)) \quad (12)$$

$$w(t+1) \leftarrow w(t) + \Delta w(t+1) \quad (13)$$

Hagamos por un momento $\alpha = 0$ en la Ecuación 12. Entonces tenemos la regla delta y el incremento en el vector de pesos se calcula mediante la expresión $\Delta w = -\eta \nabla E(w)$.

Esta expresión procede de una aproximación de la función de error mediante el primer término del desarrollo en series de Taylor esto es:

$$E(w + \Delta w) \approx E(w) + \nabla E(w)^T \Delta w \quad (14)$$

Ecuación se elige un valor para Δw que tenga la misma dirección y sentido opuesto a $\nabla E(w)$ para que el error disminuya.

Si en lugar de realizar una aproximación lineal de la función de error realizamos una cuadrática usando los dos primeros términos del desarrollo en series de Taylor obtenemos la expresión:

$$\Delta E(\mathbf{w}) \approx \nabla E(\mathbf{w})^T \Delta \mathbf{w} + 1/2 \Delta \mathbf{w}^T H(\mathbf{w}) \Delta \mathbf{w} \quad (15)$$

Donde:

$H(\mathbf{w})$ es la matriz Hessiana de E con respecto a \mathbf{w} y $\Delta E(\mathbf{w}) = E(\mathbf{w} + \Delta \mathbf{w}) - E(\mathbf{w})$.

Diferenciando con respecto a $\Delta \mathbf{w}$, igualando a cero y despejando tenemos:

$$\Delta \mathbf{w} = -H^{-1} \nabla E(\mathbf{w}) \quad (16)$$

La Ecuación anterior nos permite calcular el valor de $\Delta \mathbf{w}$ que minimiza el cambio $\Delta E(\mathbf{w})$. Este es el método de Newton, que consigue en muchos casos mejores resultados que la Retropropagación pero requiere calcular derivadas de segundo orden.

Existe una aproximación al método de Newton que no requiere dicho cálculo y que es conocido como Levenberg–Marquardt. Es más potente que el método del gradiente descendiente pero requiere más memoria. Llamaremos $e^p(\mathbf{w})$ al vector de error de la red para el patrón p cuando los pesos tienen valor \mathbf{w} , esto es:

$$e_i^p(\mathbf{w}) = t_i^p - o_i^p(\mathbf{w}) \quad (17)$$

Nótese que ahora indicamos explícitamente que el valor de la salida de la red depende del valor de los pesos sinápticos (el hecho de que antes no se indicara esta dependencia no significa que no existiera). Podemos escribir lo siguiente:

$$E(\mathbf{w}) = \frac{1}{2} \sum_{p=1}^P \mathbf{e}^p(\mathbf{w})^T \mathbf{e}^p(\mathbf{w}) \quad (18)$$

$$\nabla E(\mathbf{w}) = \sum_{p=1}^P J^p(\mathbf{w})^T \mathbf{e}^p(\mathbf{w}) \quad (19)$$

$$H(\mathbf{w}) = \sum_{p=1}^P (J^p(\mathbf{w})^T J^p(\mathbf{w}) + S^p(\mathbf{w})) \quad (20)$$

Donde:

$H(\mathbf{w})$ es la matriz Hessiana de E evaluada en \mathbf{w} , $S^p(\mathbf{w})$ es una matriz dependiente de las segundas derivadas de $e(\mathbf{w})$ y $J^p(\mathbf{w})$ es la matriz Jacobiana de e^p evaluada en \mathbf{w} , la cual puede expresarse como:

$$J^p(\mathbf{w}) = \begin{bmatrix} \nabla e_1^p(\mathbf{w}) \\ \vdots \\ \nabla e_S^p(\mathbf{w}) \end{bmatrix} = \begin{bmatrix} \frac{\partial e_1^p}{\partial w_1}(\mathbf{w}) & \cdots & \frac{\partial e_1^p}{\partial w_Q}(\mathbf{w}) \\ \vdots & \ddots & \vdots \\ \frac{\partial e_S^p}{\partial w_1}(\mathbf{w}) & \cdots & \frac{\partial e_S^p}{\partial w_Q}(\mathbf{w}) \end{bmatrix} \quad (21)$$

Donde:

Q es el número de pesos más el de umbrales. Ahora tomamos

$$\Delta \mathbf{w} = -\alpha \mathbf{M}(\mathbf{w}) \nabla E(\mathbf{w}) \quad (22)$$

Para modificar los pesos. El valor de la matriz M usado por el método es $\mathbf{M}(\mathbf{w}) = [\mu \mathbf{I} + \mathbf{H}(\mathbf{w})]^{-1}$ donde μ es algún valor no negativo. Además el valor de la Hessiana es sustituido por:

$$\sum_{p=1}^P J^p(\mathbf{w})^T J^p(\mathbf{w}) \quad (23)$$

Asumiendo que $S^p(\mathbf{w}) \cong 0$. Con todo esto llegamos al método de Levenberg Marquardt (LM) que queda resumido en la siguiente expresión:

$$\Delta \mathbf{w} = -\alpha \left[\mu \mathbf{I} + \sum_{p=1}^P J^p(\mathbf{w})^T J^p(\mathbf{w}) \right]^{-1} \nabla E(\mathbf{w}) \quad (24)$$

El parámetro μ se incrementa o disminuye en cada paso. Si $E(\mathbf{w}(t+1)) \leq E(\mathbf{w}(t))$ entonces en la siguiente iteración μ se divide por un factor β . En caso contrario en la siguiente iteración el parámetro μ se multiplica por β . En el que se sugiere tomar $\beta = 10$ y $\mu = 0.01$ al comienzo, sin embargo, nosotros hemos usado $\mu = 0.001$ tal y como hace MATLAB.

El algoritmo Levenberg–Marquardt actúa sobre los P patrones a la vez como sigue:

1. Calcula la salida de la red y los vectores de error $e^p(\mathbf{w})$ para cada uno de los patrones.
2. Calcula la matriz Jacobiana $J^p(\mathbf{w})$ para cada patrón usando entre otras cosas el vector de error $e^p(\mathbf{w})$.
3. Calcula $\Delta \mathbf{w}$ usando la Ecuación 24 y los resultados anteriores.

4. Vuelve a calcular el error usando $w + \Delta w$ como pesos de la red. Si el error ha disminuido divide μ por β , hace la asignación $w \leftarrow w + \Delta w$ y vuelve al paso 1. Si el error no disminuye multiplica μ por β y vuelve al paso 3.
5. $\left[\mu I + \sum_{p=1}^P J^p(\mathbf{w})^T J^p(\mathbf{w}) \right]$. El algoritmo acaba cuando la norma del gradiente es menor que un valor predeterminado o el error se ha reducido por debajo de un objetivo.

3.5.1.4 Aprendizaje en redes neuronales usando Regularización Bayesiana

Este método de aprendizaje se basa en la optimización de la función de densidad de probabilidad de los hiperparámetros α y β , los cuales explican la varianza del error, que se expresa como la diferencia entre la salida esperada y la salida estimada de la red neuronal.

Mediante la aplicación del teorema de Bayes, Se observa que se extrae información de los datos, para determinar a θ . Esta expresión lleva a la siguiente forma de error:

$$S(\theta) = s(\theta^*) + \frac{1}{2!}(\theta - \theta^*)^T H(\theta - \theta^*) \quad (25)$$

Nótese que esta es una expansión aproximada del verdadero término de error, la cual se halla mediante una aproximación gaussiana. Para encontrar el θ_{MP} (más probable) se procede a la minimización de la función de error $S(\theta)$ por medio de un proceso de optimización no lineal, en este caso Levenbert-Marquardt. Para el cálculo de los hiperparámetros α y β , se fijan como constantes conocidas en el primer paso del proceso iterativo y después mediante la aproximación Bayesiana se infieren los valores más apropiados para estos hiperparámetros a partir de los datos de entrenamiento. Tales valores deben maximizar la función de distribución a posterior $p(\theta / D)$.

Llegamos a obtener la varianza de la red que se expresa de la siguiente manera:

$$\sigma_t^2 = \frac{1}{\beta} + (\nabla_{\theta y} / \theta_{MP})^T H^{-1}(\nabla_{\theta y} / \theta_{MP}) \quad (26)$$

Básicamente, los hiperparámetros explican la varianza de la red neuronal, la cual se desea minimizar.

Estas cantidades nos permiten estimar los valores correspondientes a los parámetros α y β más probables, así:

$$\alpha_{MP} = \frac{\sum_{i=1}^{NM} \left(\frac{\lambda_i}{\lambda_i + \alpha} \right)}{2E_{\theta}^{MP}} \alpha_{MP} \quad (27)$$

$$\beta_{MP} = \frac{N_y - \sum_{i=1}^{NM} \left(\frac{\lambda_i}{\lambda_i + \alpha} \right)}{2E_D^{MP}} \quad (28)$$

Siendo λ_i cada uno de los valores propios de la matriz G, la cual se obtiene como $\nabla \nabla E_D$

3.5.1.5 Algoritmos basados en método de Gradiente Conjugado Escalado (SCG)

1. Seleccionar el vector de pesos para la primera interacción w_1 si los

$$\text{Escalares } 0 < \sigma \leq 10^{-4}, \quad 0 < \lambda_1 \leq 10^{-6}, \quad \bar{\lambda}_1 = 0 \quad (29)$$

Establecer $P_1 = g_1 = E'(w_1)$, $k = 1$ y éxito = verdadero.

2. Si éxito = verdadero entonces calcular la información de segundo orden:

$$\sigma_k = \frac{\sigma}{|P_k|}, \quad (30)$$

$$S_k = \frac{E'(w_k + \sigma_k P_k) - E'(w_k)}{\sigma_k}, \quad (31)$$

$$\delta_k = P_k^T S_k, \quad (32)$$

3. Escalar $\delta_k = \delta_k + (\lambda_k - \bar{\lambda}_k) |P_k|^2$, (33)

4. Si $\delta_k \leq 0$ entonces transformar la matriz hessiana en una matriz definida positiva:

$$\lambda_k = 2 \left(\lambda_k - \frac{\delta_k}{|P_k|^2} \right), \quad (34)$$

$$\delta_k = -\delta_k + \lambda_k |P_k|^2, \quad \lambda_k = \bar{\lambda}_k. \quad (35)$$

5. Calcular el tamaño de la tasa de aprendizaje:

$$\alpha_k = \frac{P_k^T g_k}{|\delta_k|}, \quad (36)$$

6. Determinar el parámetro de comparación:

$$\Delta_k = \frac{2\delta_k (E(w_k) - E(w_k + \alpha_k P_k))}{(P_k^T g_k)^2} \quad (37)$$

7. Si $\Delta_k \geq 0$ entonces:

$$w_{k+1} = w_k + \alpha_k P_k, \quad (38)$$

$$g_{k+1} = -E'(w_{k+1}), \quad (39)$$

$\lambda_k = 0$, Éxito = verdadero.

Si $k \bmod N = 0$ entonces reiniciar el algoritmo:

$$P_{k+1} = g_{k+1} \quad (40)$$

Si no crear una nueva dirección conjugada:

$$\beta_k = \frac{|g_{k+1}|^2 - g_{k+1}^T g_k}{|g_k|^2}, \quad (41)$$

$$P_{k+1} = g_{k+1} + \beta_k P_k \quad (42)$$

Si $\Delta_k \geq 0,75$ entonces reducir el escalar: $\lambda_k = \frac{1}{4} \lambda_k$

Si no, $\bar{\lambda}_k = \lambda_k$ éxito = falso.

8. Si $\Delta_k < 0,25$ entonces incrementar el escalar:

$$\lambda_k = \lambda_k \frac{\delta_k (1 - \Delta_k)}{|P_k|^2}, \quad (43)$$

9. Si la dirección del gradiente descendente $g_k \neq 0$, establecer $k = k + 1$ y volver al paso 2. Si no, terminar y devolver w_{k+1} como el mínimo deseado.

El algoritmo SCG trata de acelerar la convergencia de la red MLP y evitar su dependencia de parámetros externos mediante la definición de la tasa de aprendizaje:

$$\alpha_j = \frac{-E'(w_j)^T P_j}{P_j^T E''(w_j) P_j}, \quad (44)$$

Que minimiza los fallos de la red cuando se considera una función de error cuadrática. Si la función de error no es cuadrática, se introduce un escalar λ_k con el fin de controlar la definición positiva de la matriz hessiana, a través de la expresión:

$$\alpha_j = \frac{-E'(w_k) P_k}{P_k^T E''(w_k) P_k + \lambda_k |P_k|^2}, \quad (45)$$

Posteriormente se procede al ajuste de los pesos según la expresión:

$$w_{k+1} = w_1 + \alpha_k P_k \quad (46)$$

La matriz inicial H_0^{-1} suele ser la matriz identidad I, de forma que la primera iteración resulta equivalente a la actualización a través del gradiente descendente H_k^{-1} se actualiza en cada iteración de forma recursiva, a través de un método de actualización, entre los que destacan (Battiti, 1992).

3.5.1.6 Selección de los mejores modelos de redes neuronales artificiales

El criterio empleado para la elección del mejor modelo de redes neuronales artificiales, fue el de evaluar la capacidad de desempeño y generalización a través de las estadísticas de validación de los modelos tanto para la fase de entrenamiento como para la de validación, entre los datos calculados y observados para el periodo de validación.

3.5.1.7 Construcción de los modelos de red neuronal artificial

Las bases de datos a introducir para el entrenamiento de los modelos de RNA fueron las bases de datos globales, es decir, para el punto de monitoreo puerto muelle Puno ejecutado por el IMARPE, se manejaron 360 datos pertenecientes a 4 variables, con una data desde el 2011 - 2016 y con la finalidad de predecir Oxígeno disueltos, de igual modo se tiene 360 datos con 6 variables, con una data desde el 2011 - 2016.

Tiene un arreglo matricial para el primer caso 1460×4 en el segundo caso 1460×6 Con la que se determinaron el entrenamiento de los mejores modelos De acuerdo a las características de la RNA de retropropagación, se requiere el uso de un valor de salida (Target o valor objetivo), por lo que se introdujo una segunda matriz con los valores de la concentración Oxígeno disuelto y de clorofila para cada uno de los modelos con un arreglo de 1 (variable) x 1460 (mediciones) y 1 x 1460, respectivamente. El diseño se expone lo obtenido por Matlab Ver anexo 1.

3.6 Resultados del rendimiento de parámetros de determinación

En la realización del entrenamiento de los modelos de ANN construidos para la base de datos correspondientes a la bahía interior se realizó, como se explicó en los ítems anteriores, considerando el valor obtenido de los coeficientes de correlación para cada uno de ellos y el desempeño global.

Tabla 3

Parámetros de determinación del rendimiento para Oxígeno disuelto (OD)

MODELO	MODALIDAD	Nº NEURONAS	MSE	R
SCG I	Training	4	0.09709	0.9926
	Validation	4	0.10228	0.9917
	Testing	4	0.13645	0.9906
SCGII	Training	7	0.02382	0.9983
	Validation	7	0.02273	0.9982
	Testing	7	0.02317	0.9983
SCGIII	Training	10	0.01067	0.9992
	Validation	10	0.01071	0.9992
	Testing	10	0.01316	0.9991

Describiendo la tabla 3 en lo que respecta a las funciones de desempeño, la Media del Cuadrado del Error (*MSE*), presenta valores alejados con respecto a la meta establecida, que es el error global fijado por el usuario de 1×10^{-5} (García, 2003; Rodríguez y Tenorio, 2006). Utilizando la función *SSE*, se observó que el desempeño de dos de los tres modelos está por debajo de la meta (1×10^{-5}), el único modelo que tuvo un excelente desempeño fue el de 10 neuronas en la capa oculta, obteniéndose con dicho arreglo un desempeño por debajo de 1×10^{-5} , como se presenta en la parte derecha del cuadro 3, que corresponde al desempeño de los modelos de RNA.

Tabla 4

Parámetros de determinación del rendimiento para Clorofila – a

MODELO	MODALIDAD	Nº NEURONAS	MSE	R
SCG I	Training	4	1.57254	0.9926
	Validation	4	1.4861	0.9927
	Testing	4	1.5449	0.9935
SCGII	Training	7	0.8346	0.9959
	Validation	7	0.7149	0.9969
	Testing	7	2.2415	0.9900
SCGIII	Training	10	0.1700	0.9992
	Validation	10	0.0107	0.9994
	Testing	10	0.1402	0.9992

De la tabla 4, en lo que respecta al coeficiente de correlación, se observa que 3 de los modelos de RNA cuentan con un valor aceptable de coeficiente de correlación, es decir son muy cercanos a 1; sin embargo, en lo que respecta al modelo de 10 neuronas en la capa oculta, que utiliza la función de desempeño *SSE*, presenta un *r* próximo a 1; lo que indica que cada uno de los puntos observados sobre la recta de regresión. Siendo en este caso el modelo de RNA con 10 neuronas, el mejor para la Bahía interior Puno.

Describiendo la tabla 4, en lo que respecta a las funciones de desempeño, la Media del Cuadrado del Error (*MSE*), presenta valores alejados con respecto a la meta establecida, que es el error global fijado por el usuario de 1×10^{-5} (García, 2003; Rodríguez y Tenorio, 2006). Utilizando la función *SSE*, se observó que el desempeño de dos de los tres modelos está muy por arriba de la meta (1×10^{-5}), el único modelo que tuvo un excelente desempeño fue el de 10 neuronas en la capa oculta, obteniéndose con dicho arreglo un desempeño por debajo de 1×10^{-5} , como se presenta en la parte derecha de la tabla 4, que corresponde al desempeño de los modelos de RNA.

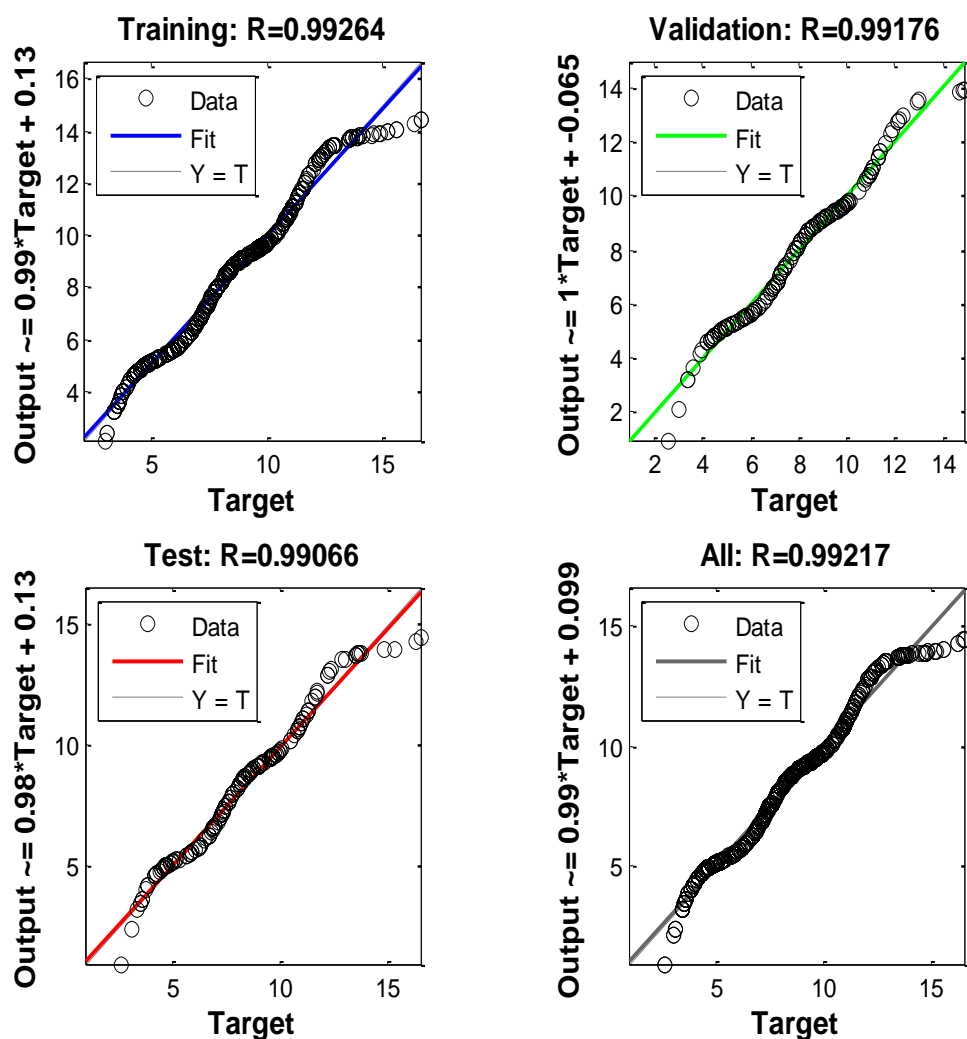


Figura 14. Gráficos de regresión durante el desempeño del entrenamiento para OD con 4 neuronas y una capa oculta.

Fuente: Matlab toolbox

En la figura 14, se observa los datos pronosticados generados por el optimizado CGS el conjunto del tren de ajustes, donde en el eje “X” se representa el objetivo (Target) y en el eje “Y” el objetivo de salida (Output Target). Cuya respuesta del coeficiente de correlación R^2 para cada caso es: en el entrenamiento 0.9926, validación 0.9917, prueba 0.9906 y regresión del conjunto total 0,9921 de izquierda a derecha, de manera que se puede ver que la convergencia del entrenamiento es positiva con un alto índice de ajuste.

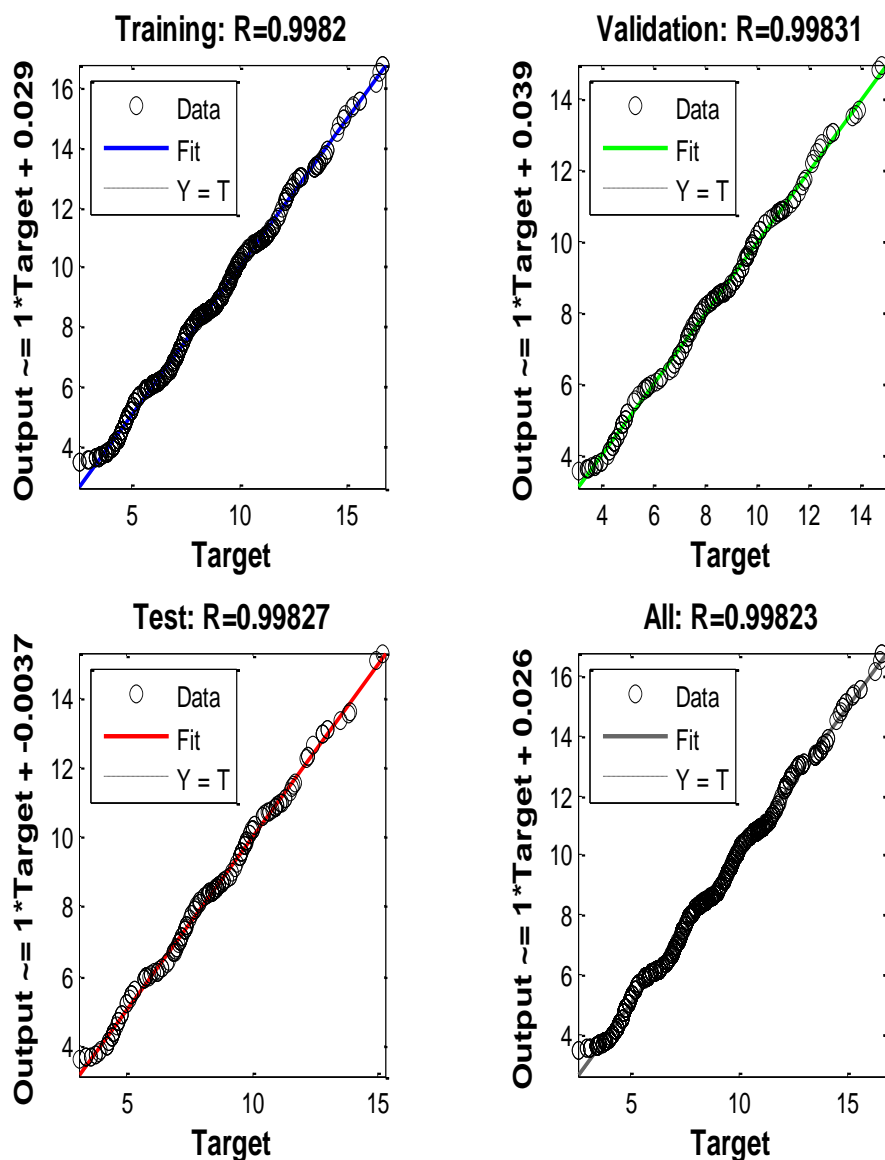


Figura 15. Gráficos de regresión durante el desempeño del entrenamiento para OD con 7 neuronas y una capa oculta.

Fuente: Matlab toolbox

Figura 15, observamos los datos pronosticados generados por el optimizado CGS con el conjunto del tren de ajustes, donde en el eje “X” se representa el objetivo (Target) y en el eje “Y” el objetivo de salida (Output Target). Cuya respuesta del coeficiente de correlación R^2 para cada caso es: en el entrenamiento 0.9982, validación 0.9983, prueba 0.9983 y regresión del conjunto total 0,9982 de izquierda a derecha, siendo la convergencia del entrenamiento positiva con un alto índice de ajuste.

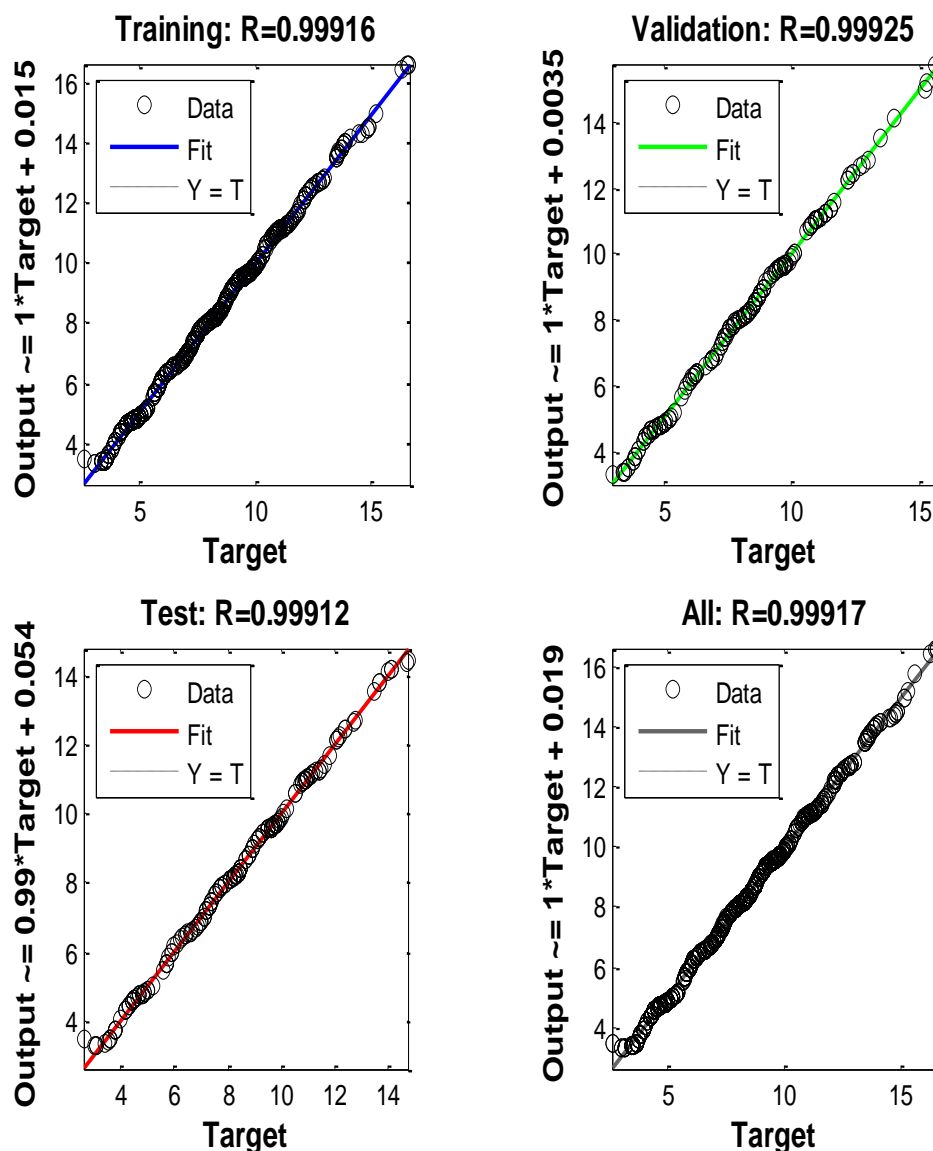


Figura 16. Gráficos de regresión durante el desempeño del entrenamiento para OD con 10 neuronas y una capa oculta.

Fuente: Matlab toolbox

En la figura 16, el optimizado CGS el conjunto del tren de ajustes, donde en el eje “X” se representa el objetivo (Target) y en el eje “Y” el objetivo de salida (Output Target). Cuya respuesta del coeficiente de correlación R^2 para cada caso es: en el entrenamiento 0.9991, validación 0.9993, prueba 0.9991 y regresión del conjunto total 0,9991de izquierda a derecha, de manera que se puede ver que la convergencia del entrenamiento es positiva con un alto índice de ajuste.

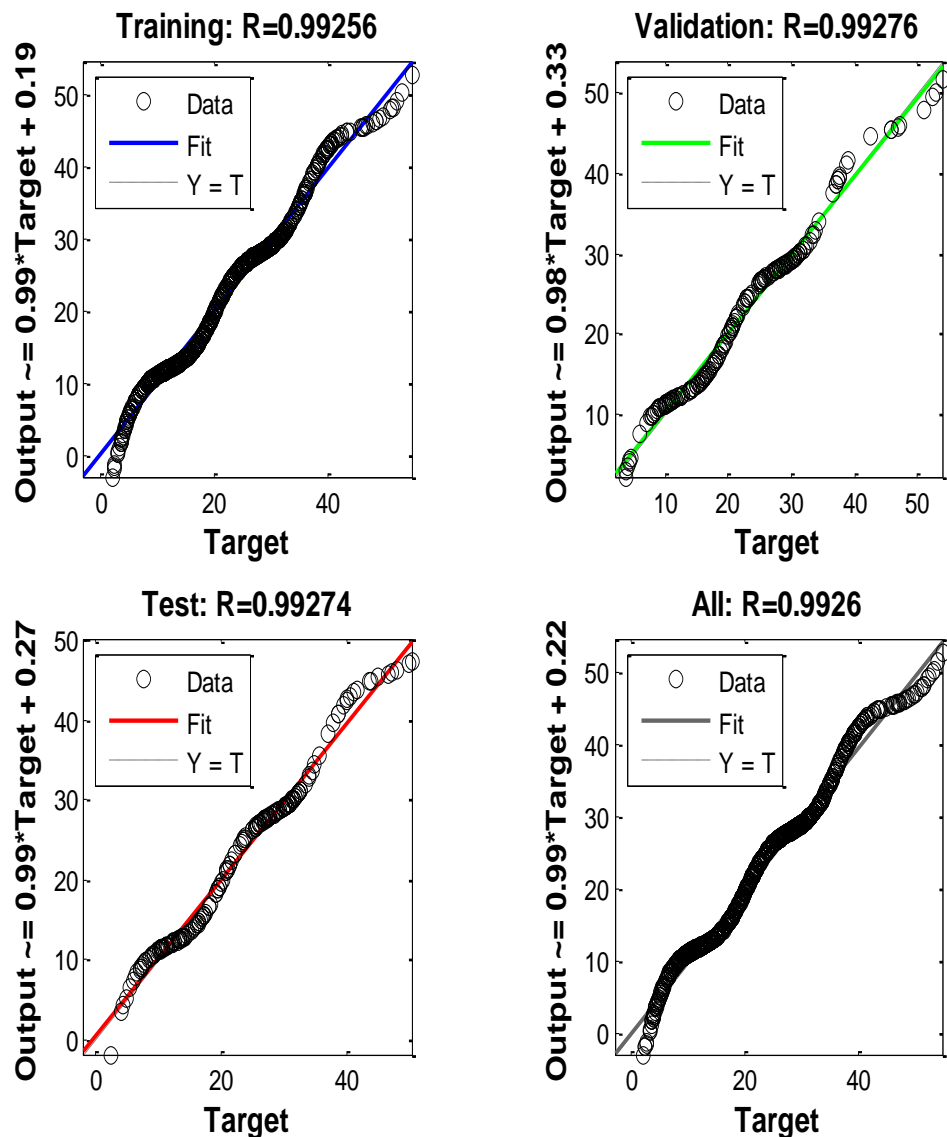


Figura 17. Gráficos de regresión durante el desempeño del entrenamiento para clorofila alfa con 4 neuronas y una capa oculta.

Fuente: Matlab toolbox

Figura 17, se observa los datos pronosticados por el optimizado CGS el conjunto del tren de ajustes, donde en el eje “X” se representa el objetivo (Target) y en el eje “Y” el objetivo de salida (Output Target). Cuya respuesta del coeficiente de correlación R^2 para cada caso es: en el entrenamiento 0.9925, validación 0.9927, prueba 0.9927 y regresión del conjunto total 0,9926 de izquierda a derecha, de manera que se puede ver que la convergencia del entrenamiento es positiva con un alto índice de ajuste.

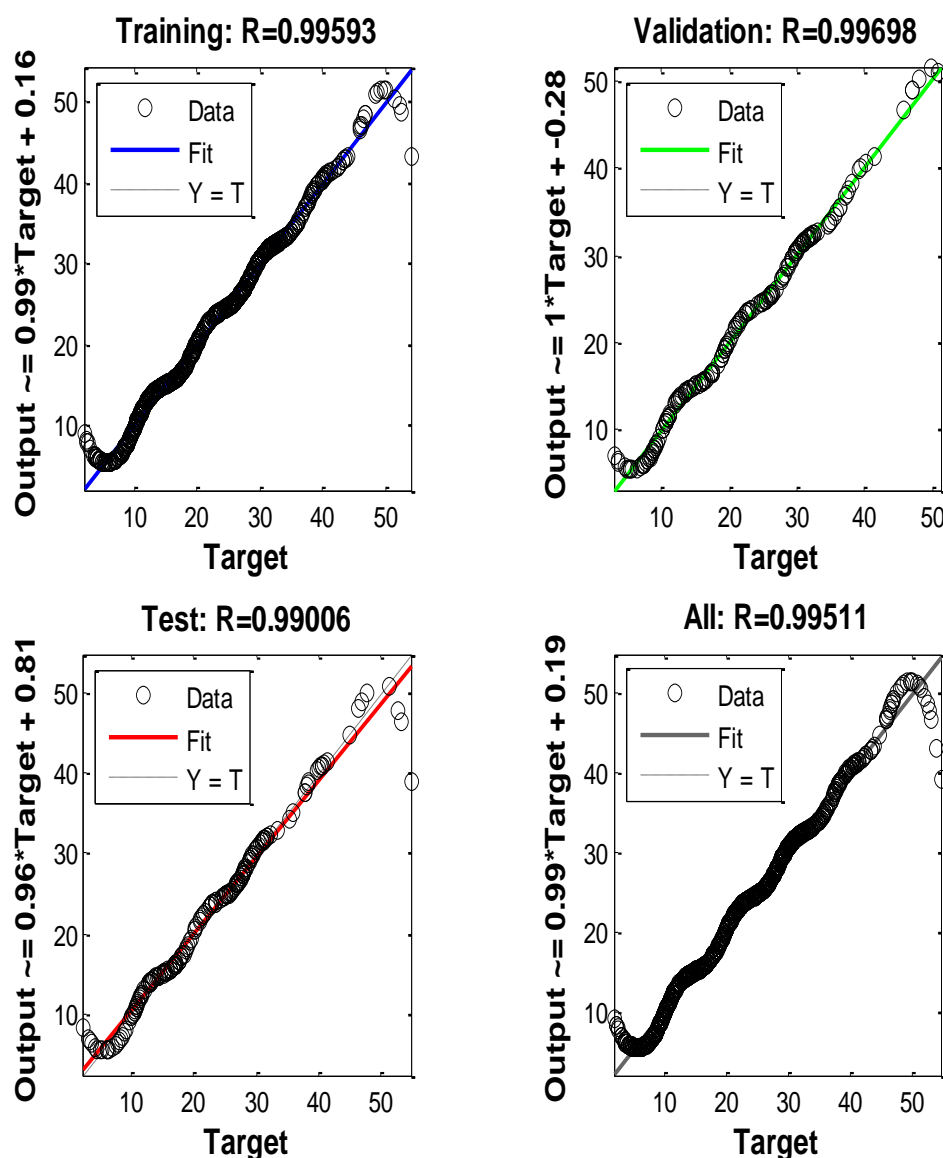


Figura 18. Gráficos de regresión durante el desempeño del entrenamiento para clorofila alfa con 7 neuronas y una capa oculta.

Fuente: Matlab toolbox

En la figura 18, se observa los datos pronosticados generados por el optimizado CGS el conjunto del tren de ajustes, donde en el eje “X” se representa el objetivo (Target) y en el eje “Y” el objetivo de salida (Output Target). Cuya respuesta del coeficiente de correlación R^2 para cada caso es: en el entrenamiento 0.9959, validación 0.9969, prueba 0.9901 y regresión del conjunto total 0,9951 de izquierda a derecha, de manera que se puede ver que la convergencia del entrenamiento es positiva.

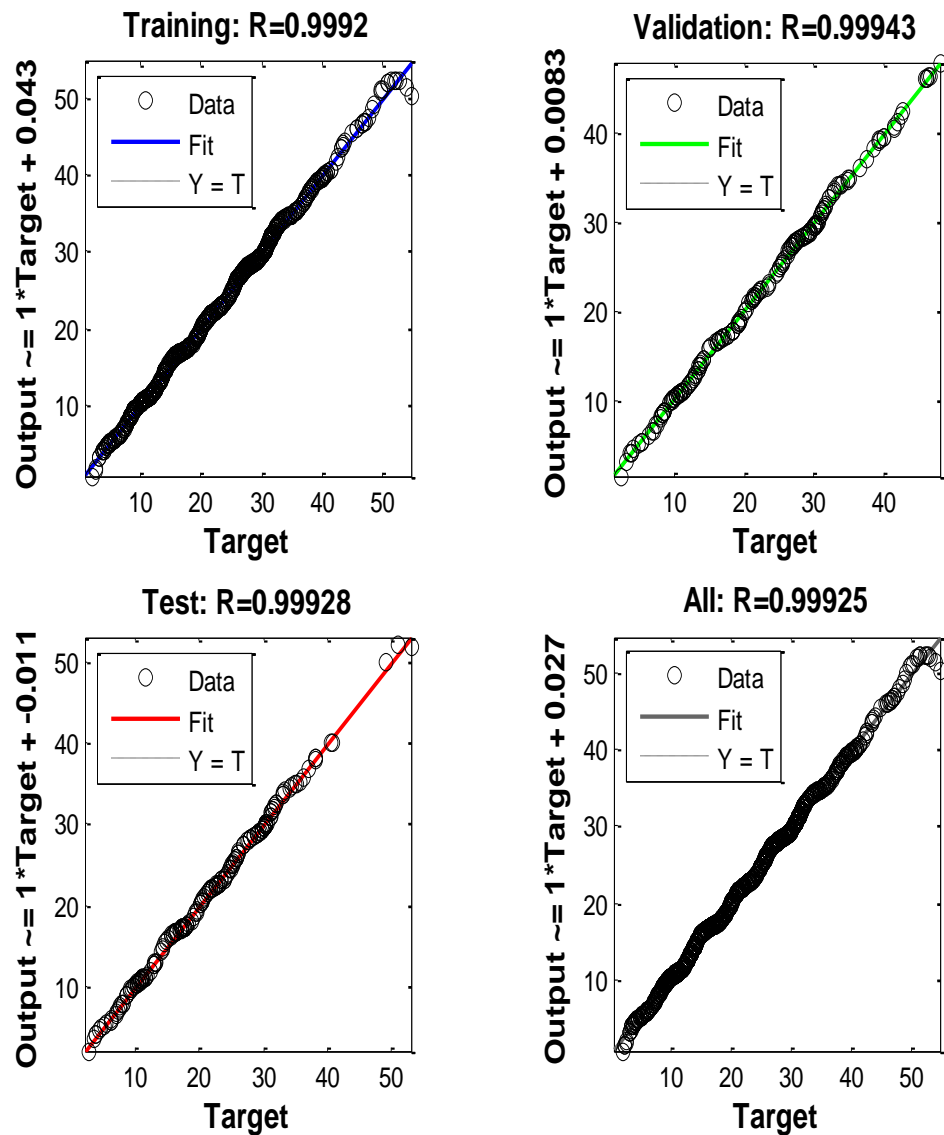


Figura 19. Gráficos de regresión durante el desempeño del entrenamiento para clorofila alfa con 10 neuronas y una capa oculta.

Fuente: Matlab toolbox

En la figura 19, se observa los datos generados por el optimizado CGS el conjunto del tren de ajustes, donde en el eje “X” se representa el objetivo (Target) y en el eje “Y” el objetivo de salida (Output Target). Cuya respuesta del coeficiente de correlación R^2 para cada caso es: en el entrenamiento 0.9992, validación 0.9994, prueba 0.9993 y regresión del conjunto total 0,9992 de izquierda a derecha, de manera que se puede ver que la convergencia del entrenamiento es positiva con un alto índice de ajuste.

3.7 Evaluación del Modelo

Realizado el entrenamiento de los modelos de RNA, es preciso realizar la evaluación o prueba de los modelos construidos, con el objetivo de observar qué tan favorable es el desempeño de dichos modelos al ser alimentados con nuevos datos. Antes de mostrar las RNA que han sido seleccionadas para los dos sitios de estudios, es de suma importancia tomar en cuenta el fenómeno de sobre entrenamiento, que se puede presentar en el desarrollo de un modelo de RNA.

En un proceso de entrenamiento se debe considerar, por una parte, un error de aprendizaje (MSE y/o SSE). Existe un error de generalización, que se puede medir empleando un conjunto representativo de patrones diferentes a los utilizados en el entrenamiento. De esta manera, podemos entrenar una red neuronal haciendo uso de un conjunto de entrenamientos, y comprobar su eficiencia real, o error de generalización, mediante un conjunto de prueba (test).

Un hecho experimental, fácilmente observable, es que si se entrena una red hasta alcanzar un error pequeño de aprendizaje, la eficacia real del sistema o generalización (medido como error en prueba) se degrada. Si representamos a la vez el error de aprendizaje y el error en test durante el transcurso del aprendizaje, se obtiene una gráfica como en la figura 20(a): tras una fase inicial, en la que pueden aparecer oscilaciones en el valor del error, el de aprendizaje tiende a disminuir monótonamente, mientras que el error de generalización a partir de cierto punto comienza a incrementarse, lo que indica una degradación progresiva del aprendizaje. (Martín y Sanz, 2007).

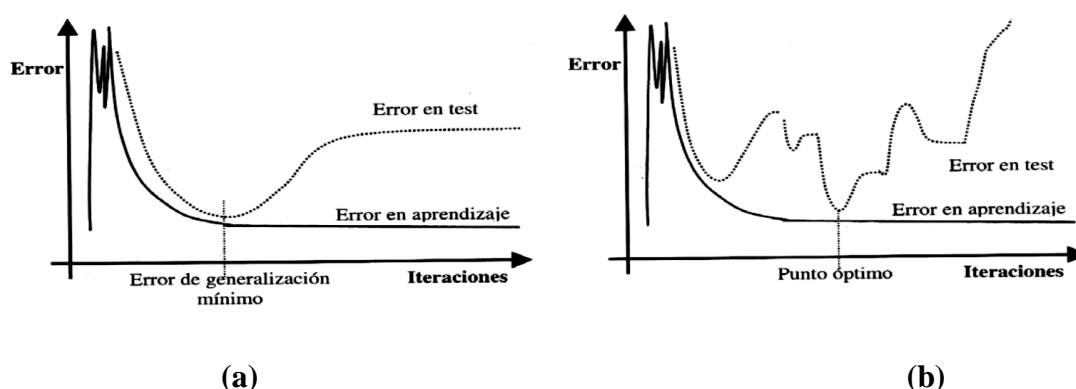


Figura 6. Evolución del error de aprendizaje y del error de generalización. En la figura 20 (a) se presenta una Situación idealizada, y en la figura 20 (b) una situación real.

Fuente: (Martín y Sanz, 2007).

La explicación de este fenómeno es el siguiente: al principio la red se adapta progresivamente al conjunto de aprendizaje, adaptándose al problema y mejorando la generalización. Sin embargo, en un momento dado el sistema se ajusta demasiado a las particularidades de los patrones empleados en el entrenamiento, por lo que crece el error que comete ante patrones diferentes a los empleados en el entrenamiento (error de generalización).

En este momento la red no ajusta correctamente el *mapping*, sino que simplemente está memorizando los patrones del conjunto de aprendizaje, lo que técnicamente se denomina sobre entrenamiento o sobreajuste (*overtraining* o *overfitting*, respectivamente), pues la red está aprendiendo demasiado. Idealmente, dada una arquitectura de red neuronal, esta debería entrenarse hasta un punto óptimo en el que el error de generalización sea mínimo, lo que indica una degradación progresiva del aprendizaje. (Martín y Sanz, 2007)

Una situación más realista sería la figura 3.8 (b), en realidad pueden presentarse varios mínimos para el conjunto de test, debiendo detener el aprendizaje en el punto óptimo de mínimo error de generalización. (Martín y Sanz, 2007) por lo tanto, para evitar el fenómeno de sobre entrenamiento, se crearon dos bases de datos (la primera para llevar a cabo el entrenamiento y la segunda para realizar la prueba o evaluación).

Por otra parte, el mínimo error de generalización, es decir el error que existe entre los patrones de entrenamiento y los patrones de prueba, el cual se explicó en párrafos anteriores fue la prueba decisiva para la elección de cada modelo de redes neuronales de cada sitio de estudio, ya que como se mencionó anteriormente, lo que se busca es que la RNA lleve a cabo un proceso de generalización, no una memorización de las variables suministradas.

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1 Presentación de resultados

Se logró implementar la RNA gracias a la aplicación de MATLAB (Toolbox), para evaluar el comportamiento del sistema. Se utilizó la aplicación de redes neuronales dentro del toolbox “Neural Networks” de Matlab R2014a.

4.1.1 Evaluación de resultados

De acuerdo a la fase de entrenamiento y prueba, se tiene que los modelos seleccionados de RNA diseñados para este estudio cuentan con las siguientes características:

- Red del tipo de retro propagación en multicapas (Backpropagation), con una capa de entrada (variables), una capa oculta (en la cual se varió el número de neuronas) y una capa de salida (con 1 nodo).
- Las variables para el modelo de la Bahía interior Puno fueron: oxígeno disuelto (mg/L), Potencial de hidrogeniones (pH), conductividad (mS/cm), temperatura (°C), nitratos (mg/L), fósforo total (mg/L), clorofila a (mg/L) (ver cuadro 3.1).
- El número de neuronas en la capa oculta fue de 10 para la Bahía interior Puno.
- Las funciones de desempeño fueron: *MSE* para la Bahía interior Puno.
- Función de entrenamiento en ambos modelos para oxígeno disuelto y clorofila a, fue la de Gradiente a escala conjugada (Lourakis, 2005).
- Y para ambos modelos, la función de transferencia para la capa oculta es *TANSIG* y *PURELIN* para la capa de salida lineal.

En la tabla 5, se presentan los modelos de RNA seleccionados para el lugar de estudio considerando el número de neuronas, las funciones de desempeño y los coeficientes de

correlación obtenidos durante cada fase de entrenamiento y prueba, para la predicción de las concentraciones de oxígeno disuelto y clorofila a.

Tabla 5

Modelo de RNA seleccionados para OD y Clorofila a para la Bahía interior Puno

Lugar de estudio	Neuronas en la Capa oculta	Función de desempeño	R Fase de entrenamiento	R Fase de Prueba
B. interior Puno	4	MSE	0.9997	0.9930
B. interior Puno	10	MSE	0.9998	0.9999

De acuerdo a lo presentado en la tabla 5, la función de desempeño que se obtiene en cada modelo es diferente. Lo anterior es debido a que dicha función evalúa el error global de diferente forma, es decir, la media del cuadrado del error (MSE), calcula el error promedio de la RNA. Aunque la forma de evaluación es diferente, el coeficiente de correlación obtenido en ambos modelos es muy próximo a 1, lo que indica que los modelos realizan una predicción igual a los datos obtenidos de los estudios proporcionados por el IMARPE, como se puede observar en las figuras 21 y 22 para la Bahía interior Puno, respectivamente.

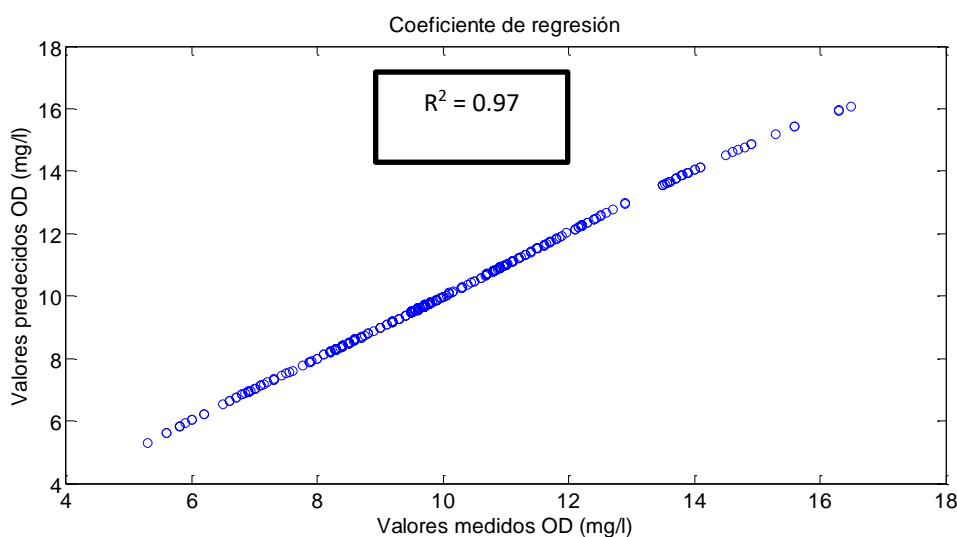


Figura 21. Coeficiente de regresión para el Oxígeno disuelto.

En la figura 21, observamos la regresión lineal de los valores predcidos de oxígeno disuelto en la ordenada y los valores medidos de Oxígeno Disuelto en la abscisa, cuyo coeficiente de correlación tiene una tendencia óptima que converge la predicción del oxígeno disuelto tal como se observa en la figura 21.

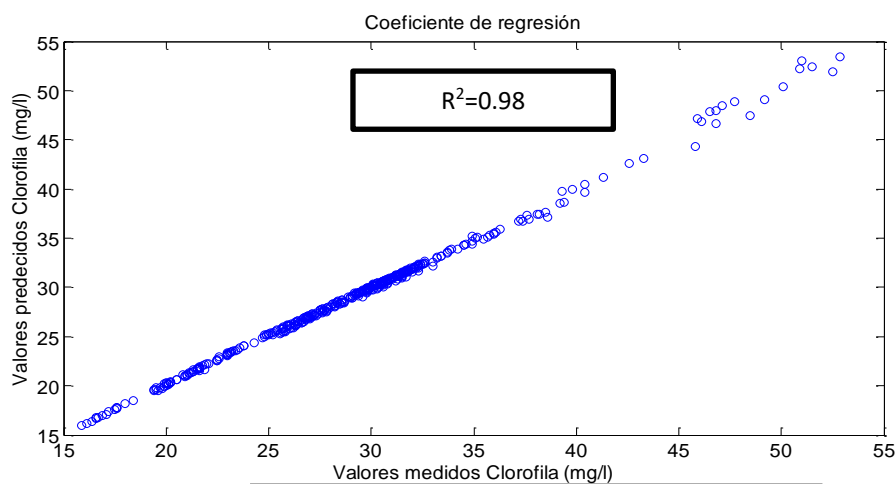


Figura 22. Coeficiente de regresión para la clorofila alfa

En la figura 22, también observamos una regresión de valores prededidos de clorofila alfa en mg/L versus valores medidos de clorofila en el campo, cuyo coeficiente de correlación converge con una tendencia óptima para el caso de clorofila alfa tal como se observa en la figura 22.

Por otro lado, aunque los dos modelos persiguen el mismo objetivo, esto es, realizar el pronóstico de concentraciones de oxígeno disuelto y clorofila a, las condiciones fisicoquímicas, tamaño de la base de datos y número de variables es distinto en cada sitio de estudio. Lo anterior se ve reflejado en que el número de neuronas en la capa oculta de cada cuerpo de agua sea diferente; esto concuerda con lo planteado por algunos autores (Anderson y McNeil, 1992; García, 2003; Rodríguez y Tenorio, 2006) que indican que en cada problema -entiéndase en este estudio, como problema a cada cuerpo de agua- se deben realizar diferentes arreglos, variando el número de neuronas en la capa oculta y seleccionar la que proporcione los resultados que mejor se adecuen a los objetivos planteados.

Debido a esta situación, en el capítulo 3, se plantean las reglas sugeridas por el autor (Goethals, 2007) para determinar el número de neuronas en la capa oculta, ya que, hasta este momento, no existe una regla que indique el número óptimo de neuronas para cada problema.

Las fases de entrenamiento y prueba, fueron realizadas de manera simultánea, con el objetivo de eliminar el fenómeno de sobre entrenamiento que se puede presentar en un modelo de RNA. Lo anterior se llevó a cabo de forma exitosa; para el caso de la Bahía anterior Puno, se cumplió con dos condiciones primordiales; la primera que fue el

obtener un error global menor al propuesto de 1×10^{-5} y la segunda de obtener un error de generalización mínimo, ya que las curvas de entrenamiento y prueba se encuentran muy cercanas, como se puede observar en la figura 23 y 24.

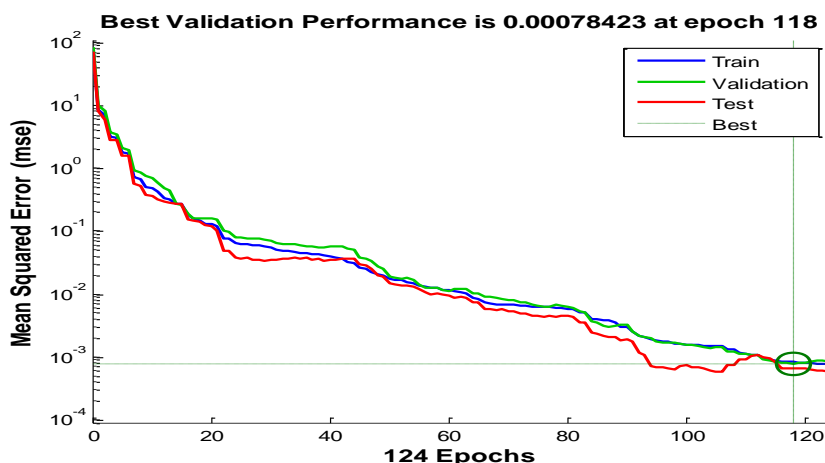


Figura 23. Entrenamiento y prueba de la RNA para OD con 4 neuronas en capa oculta y MSE como función de desempeño para la Bahía interior Puno.

Fuente: Matlab toolbox

Después de haber finalizado el entrenamiento, se observa tres gráficas verde, azul y rojo disponibles en la figura 23, se muestra el error cuadrático medio en función del número de iteraciones llevadas a cabo durante el entrenamiento. La red finalizó el entrenamiento después de 118 iteraciones, con un error MSE = 0,00078423. Luego de 95 iteraciones el error no disminuye y se detiene el entrenamiento.

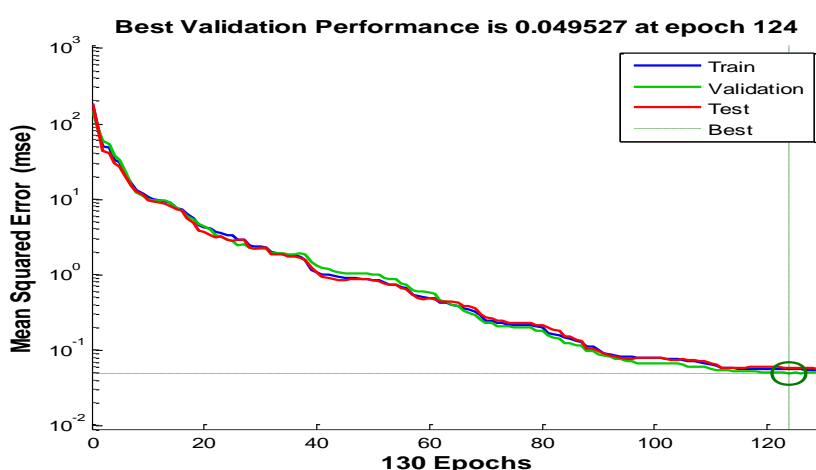


Figura 24. Entrenamiento y prueba de la RNA para Clorofila a con 10 neuronas en capa oculta y MSE como función de desempeño para la Bahía interior Puno.

Fuente: Matlab toolbox

Después de haber finalizado el entrenamiento, se observa tres gráficas verde, azul y rojo disponibles en la figura 24, se muestra el error cuadrático medio en función del número de iteraciones llevadas a cabo durante el entrenamiento. La red finalizó el entrenamiento después de 124 iteraciones, con un error $MSE = 0,049527$. Luego de 112 iteraciones el error no disminuye y se detiene el entrenamiento.

Para el caso de la Bahía interior Puno, se muestra en la figura 24 un nulo error de generalización de la iteración 10 a la 130, aproximadamente; sin embargo, el proceso de entrenamiento no es detenido en ese punto porque no se había obtenido el error global establecido; el proceso se finalizó cuando se cumplieron las dos condiciones, es decir, se cumple con el error global y se obtiene un mínimo error de generalización.

4.1.2 Resultados de la prueba estadística

Tabla 6

Análisis de Varianza para Oxígeno Disuelto

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
A:OD	2,1025	1	2,1025	8,43	0,0440
B:pH	0,49	1	0,49	1,96	0,2336
C:CE	0,81	1	0,81	3,25	0,1459
D:T	0,0	1	0,0	0,00	1,0000
AB	0,01	1	0,01	0,04	0,8511
AC	0,36	1	0,36	1,44	0,2958
AD	1,96	1	1,96	7,86	0,0486
BC	0,2025	1	0,2025	0,81	0,4185
BD	0,5625	1	0,5625	2,26	0,2075
CD	0,5625	1	0,5625	2,26	0,2075
bloques	1,96	1	1,96	7,86	0,0486
Error total	0,9975	4	0,249375		
Total (corr.)	10,0175	15			

$R^2 = 90,0424 \%$, Error estándar del est. = 0,499375, R^2 (ajustada por g.l.) = 70,1273 %.

Error absoluto medio = 0,209375, Estadístico Durbin - Watson = 1,82832 (P=0,3876)

Modelo Matemático: **OD = 5,03476- 0,104405pH+0,000801786CE- 0,0488095T+0,0000642857pHCE+0,00535714pHT- 0,0000375CET**

La tabla ANOVA particiona la variabilidad de OXÍGENO DISUELTO en piezas separadas para cada uno de los efectos. Entonces prueba la significancia estadística de cada efecto comparando su cuadrado medio contra un estimado del error experimental. En este caso, 3 efectos tienen un valor-P menor que 0,05, indicando que son significativamente diferentes de cero con un nivel de confianza del 95,0%. El estadístico R^2 indica que el modelo, así ajustado, explica 90,0424% de la variabilidad en OXÍGENO DISUELTO. El estadístico R-cuadrada ajustada, que es más adecuado

para comparar modelos con diferente número de variables independientes, es 70,1273%. El error estándar del estimado muestra que la desviación estándar de los residuos es 0,499375. El error medio absoluto (MAE) de 0,209375 es el valor promedio de los residuos. El estadístico de Durbin-Watson (DW) prueba los residuos para determinar si haya alguna correlación significativa basada en el orden en que se presentan los datos en el archivo. Puesto que el valor-P es mayor que 5,0%, no hay indicación de auto correlación serial en los residuos con un nivel de significancia del 5,0%.

Tabla 7

Análisis de Varianza para Clorofila Alfa

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
A:OD	0,545752	1	0,545752	0,02	0,8822
B:pH	10,2961	1	10,2961	0,42	0,5208
C:CE	50,5343	1	50,5343	2,06	0,1589
D:T	83,9285	1	83,9285	3,42	0,0716
E:NO	209,055	1	209,055	8,52	0,0057
F:PO	316,084	1	316,084	12,88	0,0009
AB	6,84476	1	6,84476	0,28	0,6003
AC	0,00113906	1	0,00113906	0,00	0,9946
AD	22,4321	1	22,4321	0,91	0,3447
AE	0,443889	1	0,443889	0,02	0,8937
AF	6,23126	1	6,23126	0,25	0,6171
BC	18,6084	1	18,6084	0,76	0,3890
BD	75,1039	1	75,1039	3,06	0,0877
BE	0,0984391	1	0,0984391	0,00	0,9498
BF	12,8433	1	12,8433	0,52	0,4736
CD	26,6901	1	26,6901	1,09	0,3031
CE	58,3123	1	58,3123	2,38	0,1309
CF	0,234014	1	0,234014	0,01	0,9227
DE	69,16	1	69,16	2,82	0,1008
DF	1,55314	1	1,55314	0,06	0,8026
EF	0,871889	1	0,871889	0,04	0,8514
bloques	4305,49	1	4305,49	175,42	0,0000
Error total	1006,29	41	24,5437		
Total (corr.)	6281,66	63			

R-cuadrada = 83,9805 %, R-cuadrada (ajustada por g.l.) = 75,9707 %, Error estándar del est. = 4,95416, Error absoluto medio = 3,15339, Estadístico Durbin-Watson = 1,76472 (P=0,1281)

Modelo Matemático: **Chl-alfa = 24,7768 + 3,62841*NO + 1,19758*PO**

La tabla ANOVA particiona la variabilidad de CLOROFILA ALFA en piezas separadas para cada uno de los efectos. Entonces prueba la significancia estadística de cada efecto comparando su cuadrado medio contra un estimado del error experimental. En este caso, 3 efectos tienen un valor-P menor que 0,05, indicando que son significativamente diferentes de cero con un nivel de confianza del 95,0%. El estadístico R² indica que el modelo, así ajustado, explica 83,9805% de la variabilidad en CLOROFILA ALFA. El

estadístico R-cuadrada ajustada, que es más adecuado para comparar modelos con diferente número de variables independientes, es 75,9707%. El error estándar del estimado muestra que la desviación estándar de los residuos es 4,95416. El error medio absoluto (MAE) de 3,15339 es el valor promedio de los residuos. El estadístico de Durbin-Watson prueba los residuos para determinar si haya alguna correlación significativa basada en el orden en que se presentan los datos en el archivo. Puesto que el valor-P es mayor que 5,0%, no hay indicación de auto correlación serial en los residuos con un nivel de significancia del 5,0%.

4.1.3 Resultado pronóstico y validación de los modelos entrenados con data IMARPE

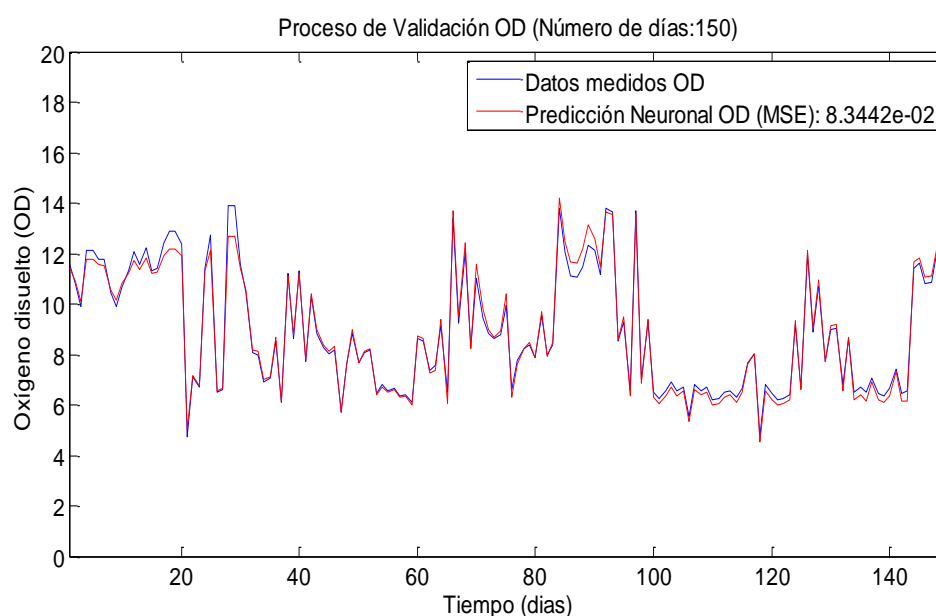


Figura 25. Gráfico de validación con data de 1 año del IMARPE.

De acuerdo con la figura 25 se observa la predicción de los datos de OD del IMARPE, donde su coeficiente de regresión es del 0.97 como se observa en la figura 26, con una predicción para 150 días y concentraciones de picos de 14 mg/L como máximo y mínimo 2.5 aproximadamente. Esta tendencia podría mejorarse empleando una mejor selección de variables de predicción o mayor número de variables (Olyai *et al* 2017).

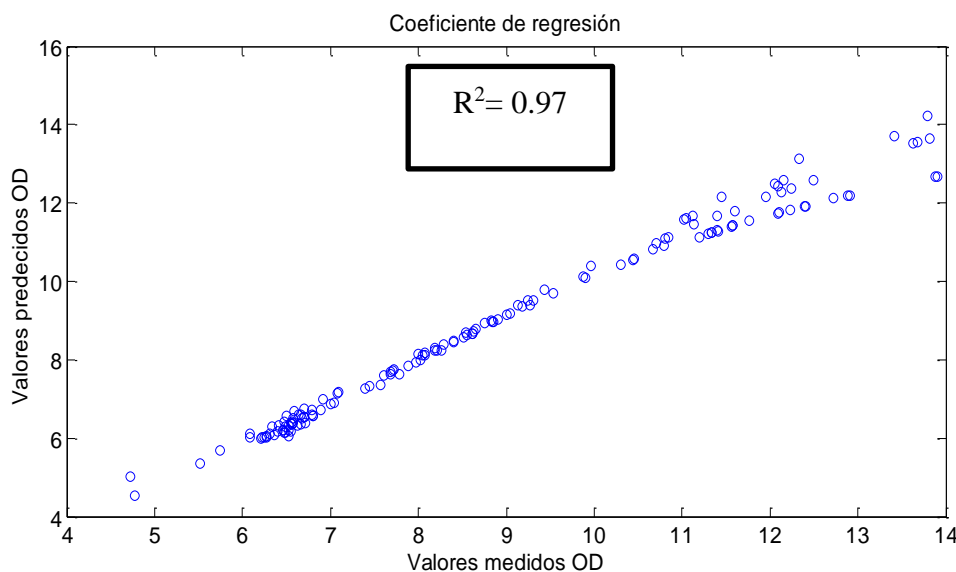


Figura 26. Coeficiente de regresión para los datos de OD.

La figura 26, muestra el comportamiento de los valores predcidos frente a los valores monitoreados de oxígeno disuelto, evidencia que el coeficiente de determinación es bueno el porcentaje de error es mínimo.

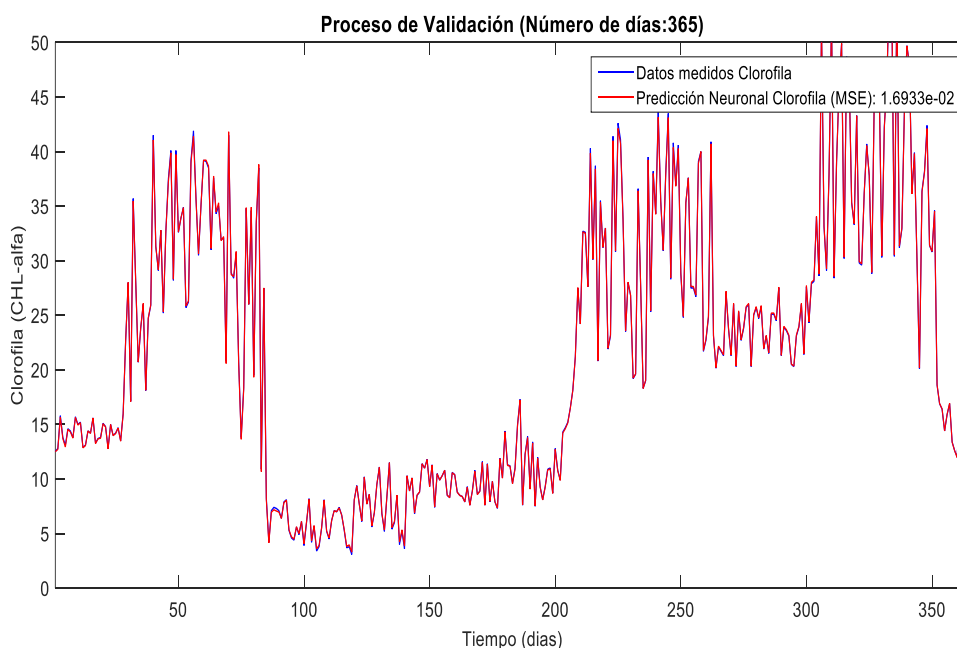


Figura 27. Gráfico pronóstico de Clorofila alfa con data de 1 año del IMARPE.

En la figura se observa el entrenamiento con datos reales de clorofila alfa y su pronóstico con una data de 1 año del IMARPE, cuyo performance fue óptimo tal como se observa en la figura 28, con un $R^2 = 0.98$, en este caso el modelo matemático fue el

de gradiente a escala conjugada. El modelo entrenado con los datos de predicción es casi superpuesto de forma dinámica en función del tiempo (Mahewaran y Khosa 2015).

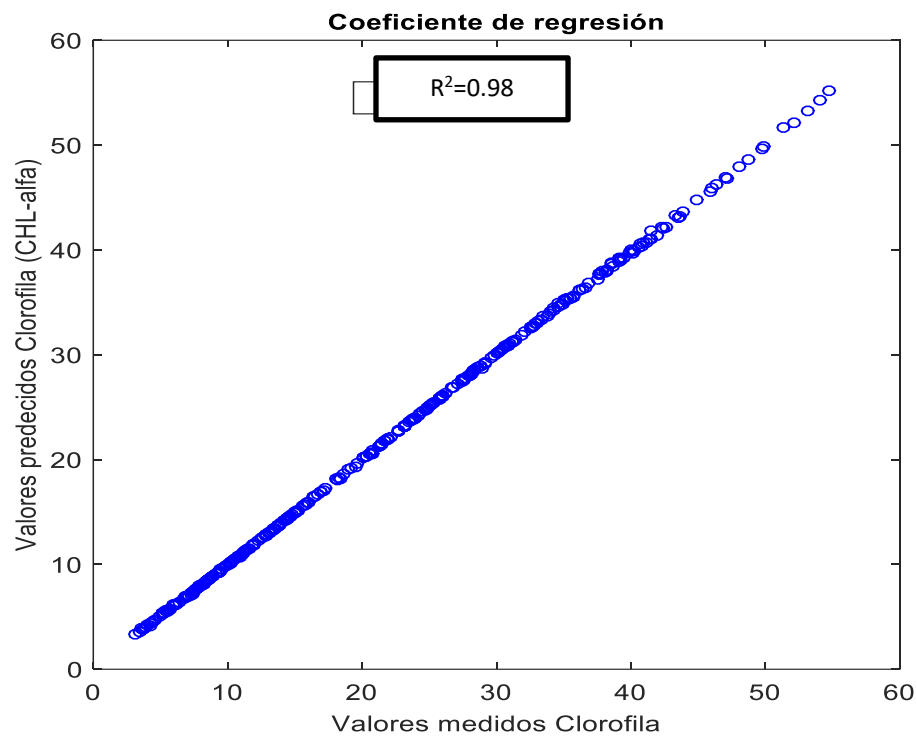


Figura 28. Coeficiente de regresión para los datos de Clorofila alfa

En la figura 28, observamos el comportamiento de la clorofila alfa luego de una regresión valores predichos frente a valores medidos de clorofila alfa, cuyo coeficiente de determinación es muy buena dando una confiabilidad del 98%.

CONCLUSIONES

De acuerdo a los resultados obtenidos, se cumplió el objetivo general planteado para la realización de este trabajo, ya que se obtuvo un modelo de RNA que permitirá el pronóstico de la concentración de oxígeno disuelto y clorofila alfa presentes en un cuerpo de agua.

Con respecto al desempeño del modelo, se logró la eliminación del fenómeno de Sobreentrenamiento en los modelos estudiados de RNA desarrollados para el cuerpo de agua de bahía interior Puno; esto a partir de la creación de dos bases de datos (base de datos de entrenamiento y base de datos de prueba), obteniendo así el mínimo error de generalización entre los patrones de entrenamiento y los patrones de prueba siendo el algoritmo de entrenamiento más adecuado para este fin la de **Gradiente a Escala Conjugada** con un performance del 98% para la clorofila alfa y 97% para el oxígeno disuelto.

De acuerdo con los resultados satisfactorios obtenidos se concluye que para la Bahía interior Puno, las variables utilizadas para el caso de **clorofila alfa** fueron 6 se evidencian que son suficientes para entrenarlo y evaluarlo; dichas variables fueron: oxígeno disuelto (mg/L), potencial de hidrogeniones (pH), conductividad (mS/cm), temperatura (°C), nitratos (mg/L), fosfatos (mg/L).

Para efectos del **oxígeno disuelto** se utilizaron 4 variables cuyos elementos fueron: oxígeno disuelto (mg/L), potencial de hidrogeniones (pH), conductividad (mS/cm), temperatura (°C), para la predicción de este elemento aún podría mejorarse el modelo debido a que no fue muy bueno el performance, esto podría deberse a la forma de selección de las variables o seleccionar desde un punto de vista formal las variables que influyen directamente en la predicción de este elemento.

Se realizaron pruebas estadísticas a los modelos seleccionados como los más adecuados para el punto de monitoreo del IMARPE en el puerto muelle de Puno de acuerdo a la evaluación del desempeño mostrado por dichos modelos; para el caso de la bahía interior, el modelo seleccionado fue el de $4 \times 4 \times 1$; es decir 4 variables de entrada, 4 neuronas en la capa oculta y una de salida para el caso de OD. En el caso de Clorofila alfa el modelo seleccionado fue el que tuvo el arreglo de $6 \times 10 \times 1$; es decir 6 variables de entrada, 10 neuronas en la capa oculta y una de salida. Con el desempeño obtenido por los modelos especialmente de clorofila alfa, el resultado de pruebas estadísticas, se comprueba que las RNA son una herramienta sumamente útil, de bajos requerimientos computacionales y con la suficiente eficiencia para pronosticar concentraciones de clorofila y oxígeno disuelto a partir de otros parámetros presentes en un cuerpo de agua.

La prueba estadística empleada fue significativa, tanto para el oxígeno disuelto como para la clorofila alfa, siendo los más relevantes para el caso de oxígeno disuelto y temperatura, en el caso de la clorofila alfa las variables significativas fueron los nitratos y los fosfatos cuyos elementos influyen en el modelado de red neuronal para la clorofila alfa. Con un nivel de significancia del 95%.

Estos modelos pueden ayudar a alertar a la población y autoridades en escenarios no deseados de calidad del agua, tomando medidas precautorias para salvaguardar la salud de la comunidad, por ello es importante que los resultados obtenidos sean presentados de una forma más simplificada que sea entendible por la sociedad en general, como es el caso de un Índice de Eutroficación.

Los modelos desarrollados de RNA pueden ser empleados para cuerpos de agua que presenten condiciones similares a los que presentan los de la bahía interior Puno; respecto a las características fisicoquímicas, esto es, que presenten un índice de eutroficación dentro del rango mesotrófico y eutrófico (Tenorio 2009).

Con los resultados obtenidos con anterioridad, se comprobó que las redes neuronales pueden ser empleadas satisfactoriamente en la predicción de concentraciones de clorofila y oxígeno disuelto en aguas superficiales, por lo que son una herramienta sumamente útil y alterna a los métodos ya existentes.

RECOMENDACIONES

Una recomendación muy importante es que a partir de los modelos presentados se realicen nuevos modelos con una base de datos más amplia, es decir; con un mayor número de datos; esto con el fin de que el modelo sea más robusto.

A partir del presente trabajo, se pretende que se desarrollen nuevas líneas de investigación en el área de pronóstico de contaminantes en cuerpos de agua superficiales y considerando otros fenómenos que puedan afectar la calidad del agua; como las reacciones químicas presentes en un cuerpo de agua.

Se sugiere utilizar estos modelos para un pronóstico rápido aproximado de la concentración de clorofila y oxígeno disuelto en aguas superficiales, que se entrenen y prueben con datos específicos del área de estudio, con el objetivo de obtener mejores desempeños.

Los buenos resultados obtenidos en esta investigación recomiendan que en trabajos futuros se haga la investigación de la aplicación de los modelos de redes neuronales artificiales a eventos horarios, diarios, semanales y anuales, la investigación de otros algoritmos de optimización, predicciones con horizonte corto, medio y largo plazo. (Redes neuronales recurrentes) y modelos que consideran las características hidrológicas de las series.

BIBLIOGRAFÍA

- Abbas, A., & Al-Bastaki, N. (2005). Modeling of an RO water desalination unit using neural networks. *Chemical Engineering Journal*, 114(1–3), 139–143. <https://doi.org/10.1016/j.cej.2005.07.016>
- Almeida, J. S. (2002). Predictive non-linear modeling of complex data by artificial neural networks. *Current Opinion in Biotechnology*, 13(1), 72–76. [https://doi.org/10.1016/S0958-1669\(02\)00288-4](https://doi.org/10.1016/S0958-1669(02)00288-4)
- Altunkaynak, A. (2007). Forecasting surface water level fluctuations of lake van by artificial neural networks. *Water Resour. Manage.*, 21, 399-408.
- Ahmed J. A; Sarma A. K.; (2007). Artificial neural network model for synthetic streamflow generation. *Water Resour. Management*, vol. 21, 1015-1029.
- Akkoyunlu, A., Altun, H., & Cigizoglu, H. K. (2011). Depth-Integrated Estimation of Dissolved Oxygen in a Lake. *Journal of Environmental Engineering*, 137(10), 961–967. [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000376](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000376)
- Ay, M., & Kisi, O. (n.d.). Modeling of Dissolved Oxygen Concentration Using Different Neural Network Techniques in Foundation Creek, El Paso County, Colorado. [https://doi.org/10.1061/\(ASCE\)EE](https://doi.org/10.1061/(ASCE)EE)
- Basant, N., Gupta, S., Malik, A., & Singh, K. P. (2010). Linear and nonlinear modeling for simultaneous prediction of dissolved oxygen and biochemical oxygen demand of the surface water - A case study. *Chemometrics and Intelligent Laboratory Systems*. <https://doi.org/10.1016/j.chemolab.2010.08.005>
- Bamias, A., Aravantinos, G., Deliveliotis, C., Bafaloukos, D., Kalofonos, C., Xiros, N., Dimopoulos, M. A. (2004). Docetaxel and Cisplatin With Granulocyte Colony-Stimulating Factor (G-CSF) Versus MVAC With G-CSF in Advanced Urothelial Carcinoma: A Multicenter, Randomized, Phase III Study From the Hellenic Cooperative Oncology Group. *Journal of Clinical Oncology*, 22(2), 220–228. <https://doi.org/10.1200/JCO.2004.02.152>
- Broomhead, D. S. and Lowe, D. (1988). Multivariable Functional Interpolation and Adaptive Networks. *Complex Systems*, 2, 321–355.

<https://doi.org/10.1126/science.1179047>

- Brown, R.; McClelland, N.; (1973). "*Water Quality Index. Application in the Kansas River Basin*", 46th. Conf., Water Poll. Fed., Cleveland, Ohio.
- Brunelli, A.; Piazza, V.; Pignato, L.; Sorbello, F.; Vitabile, S., (2007). Two-days ahead prediction of daily maximum concentrations of SO₂, O₃, PM₁₀, NO₂, CO in the urban area of Palermo, Italy, Atmospheric Environment, 41, 2967-2995.
- Carlsson, B., Lindberg, C.F., Hasselblad, S., & Xu, S. (1994). On-line estimation of the respiration rate and the oxygen transfer rate at Kungsängen wastewater plant in Uppsala. Water Science and Technology, 30(4), 255–263.
- Carlsson, B., & Rehnström, A. (2002). Control of an activated sludge process with nitrogen removal—a benchmark study. Water Science and Technology, 45(4-5), 135–142.
- Castro de Esparza, M. L., (1987). *Parámetros físico-químicos que influyen en la calidad y en el tratamiento del agua*. Lima, CEPIS.
- Cigizoglu, H. K., and Kisi, O. (2006). “Methods to improve the neural network performance in suspended sediment estimation.” J. Hydrol. (Amsterdam), 317(34), 221-238.
- Çinar, O.; Merdun, H.; (2008). Application of an unsupervised artificial neural network technique to multivariate surface water quality data. Ecol. Res., 10.1007/s11284-008-0495.
- Cox, B. A. (2003). “A review of dissolved oxygen modeling techniques for lowland rivers.” Sci. Total Environ., 314-316, 303-334.
- Coppola Jr., E. A.; Szidarovszky, F.; Davis, D.; Spayd, S.; Poulton, M. M.; Roman, E.; (2007). Multiobjective Analysis of a Public Wellfield Using Artificial Neural Networks. Ground Water, 45, 53-61.
- Chaloulakou, A.; Grivas, G.; Spyrellis, N.; (2003). Neural Networks and Multiple Regression Models for PM₁₀ Prediction in Athens: A Comparative Assessment. J. Air & Waste Manage. Assoc. 53, 1183-1190

- Chapra, S.C., (1997). Surface Water-Quality Modeling. McGraw-Hill, NY, p. 844.
<https://trove.nla.gov.au/work/13884244?q&versionId=46590707>
- Chatterjee, S., & Hadi, A. S. (1986). Influential Observations, High Leverage Points, and Outliers in Linear Regression. *Statistical Science*, 1(3), 379–393.
<https://doi.org/10.1214/ss/1177013622>
- Chenali, A. B.; Gajghate, D. G.; Hasan, M. Z.; (2002). Prediction of Ambient PM10 and Toxic Metals Using Artificial Neural Networks. *J. Air & Waste Manage. Assoc.*, 52, 805- 810.
- Cherkassky, V.; Hsieh, W.; Krasnopolsky, V.; Solomatine, D.; (2007). “Computational intelligence in earth and environmental sciences”, Elsevier, Neural Networks 20, UNESCO-IHE Institute for Water Education, Julio Valdes, pág. 433.
- Dedecker, A. P., Goethals, P. L. M., Gabriels, W., & De Pauw, N. (2004). Optimization of Artificial Neural Network (ANN) model design for prediction of macroinvertebrates in the Zwalm river basin (Flanders, Belgium). *Ecological Modelling*, 174(1–2), 161–173. <https://doi.org/10.1016/j.ecolmodel.2004.01.003>
- Dimopoulos I. F; Tsiros I. X; Konstantinos S; Chronopoulou A; (2004). Combining Neural Networks Models to Predict Spatial Patterns of Airborne Pollutant Accumulation in Soils around an Industrial Point Emission Source. *J. Air & Waste Manage. Assoc.*, 54, 1506-1515
- Dogan, E., Sengorur, B., & Koklu, R. (2009). Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. *Journal of Environmental Management*, 90(2), 1229–1235.
<https://doi.org/10.1016/j.jenvman.2008.06.004>
- Dolcine, L., Andrieu, H., & French, M. N. (1998). Evaluation of a conceptual rainfall forecasting model from observed and simulated rain events. *Hydrology and Earth System Sciences*, 2(2–3), 173–182.
- Fagundes-Klen, M. R., Ferri, P., Martins, T. D., Tavares, C. R. G., & Silva, E. A. (2007). Equilibrium study of the binary mixture of cadmium-zinc ions biosorption by the *Sargassum filipendula* species using adsorption isotherms models and neural

- network. *Biochemical Engineering Journal*, 34(2), 136–146.
<https://doi.org/10.1016/j.bej.2006.11.023>
- J. Fritsch. Modular Neural Networks for Speech Recognition. PhD thesis, Carnegie Mellon University, 1996.
- García, I., (2003). Aplicación de Modelos Semiempíricos para el Análisis y pronóstico de la Calidad del Aire en el área metropolitana de Monterrey, N. L., tesis de maestría, ITESM, Monterrey, México
- Gamito, E.J. y Crawford, E.D., (2004) Artificial Neural Networks for predictive modeling in prostate cancer, *Current Oncology Reports*, Springer, vol. 6, número 3, págs. 216-221
- Goethals, P. L. M.; Dedecker, A. P.; Gabriels, W.; Lek, S. y De Pauw, N; (2007). Applications of artificial neural networks predicting macroinvertebrates in freshwaters. *Aquat. Ecol.*41, 491 – 508.
- Gupta, M. M., & Rao, D. H. (1994). On the principles of fuzzy neural networks. *Fuzzy Sets and Systems*, 61(1), 1–18. [https://doi.org/10.1016/0165-0114\(94\)90279-8](https://doi.org/10.1016/0165-0114(94)90279-8)
- EPA United State Environmental Protection Agency (1994). Water Quality Standards Handbook, 2nd ed. <http://www.epa.gov/waterscience/standardshandbook>.
- Hagan, M. T., & Menhaj, M. B. (1994). Training Feedforward Networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, 5(6), 989–993. <https://doi.org/10.1109/72.329697>
- Hanbay, D., Turkoglu, I., and Demir, Y. (2008). Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks. *Expert Syst. Appl.*, 34(2), 1038-1043
- Haykin, S. (2008). *Neural Networks and Learning Machines*. Pearson Prentice Hall New Jersey USA 936 pLinks. <https://doi.org/978-0131471399>
- Haykin, S. (1994). *Neural network: A comprehensive foundation*, MacMillan, New York.
- Haykin, S. (1998). *Neural networks: A comprehensive foundation*, 2nd Ed., Prentice-

- Hall, Upper Saddle River, NJ.
- Haykin, Simon; (1999). “*Neural Networks: a comprehensive foundation*”, Prentice-Hall.
- Hilera, J. R., Martínez, V. J., (2000). *Redes Neuronales Artificiales; fundamentos, modelos y aplicaciones*, Alfaomega, pp. 390
- Ji, Zhen-Gang; (2008). “*Hydrodynamics and water quality: modeling rivers, lakes and estuaries*”, Wiley-Interscience, USA. págs. 247 – 406.
- Jiang, D; Zhang, Y; Hu, X; Zeng, Y; Tan, J; Shao, D; (2004). Progress in developing an ANN model for air pollution index forecast. *Atmospheric Environment*, 38, 7055-7064.
- Kalach, A. V. (2006). Using artificial neural networks for prediction of organic acid partition coefficients. *Russian Chemical Bulletin, International Edition*, 55(2), 212–217.
- Karul, C., Soyupak, S., Çilesiz, A. F., Akbay, N., & Germen, E. (2000). Case studies on the use of neural networks in eutrophication modeling. *Ecological Modelling*, 134, 145–152. [https://doi.org/10.1016/S0304-3800\(00\)00360-4](https://doi.org/10.1016/S0304-3800(00)00360-4)
- Kalff, J. (2002). *Limnology: Inland water ecosystems*, Prentice-Hall, Upper Saddle River, NJ.
- Kisi, Ö. (2004). Multi-layer perceptrons with Levenberg- Marquardt training algorithm for suspended sediment concentration prediction and estimation. *Hydrological Sciences–Journal*, 49(6), 1025–1040. <https://doi.org/10.1623/hysj.49.6.1025.55720>
- Kisi, O. (2005). “Suspended sediment estimation using neuro-fuzzy and neural network approaches.” *Hydrol. Sci. J.*, 50(4), 683-696.
- Kisi, O. (2007). “Streamflow forecasting using different artificial neural network algorithms.” *J. Hydrol. Eng.*, 12(5), 532-539.
- Kisi, O. (2009). “Modeling monthly evaporation using two different neural computing techniques.” *Irrig. Sci.*, 27(5), 417-430.

- Kuo, J. T., Lung, W. S., Yang, C. P., Liu, W. C., Yang, M. Der, & Tang, T. S. (2006). Eutrophication modelling of reservoirs in Taiwan. *Environmental Modelling and Software*, 21(6), 829–844. <https://doi.org/10.1016/j.envsoft.2005.03.006>
- Kuo, J. T., Hsieh, M. H., Lung, W. S., & She, N. (2007). Using artificial neural network for reservoir eutrophication prediction. *Ecological Modelling*. <https://doi.org/10.1016/j.ecolmodel.2006.06.018>
- Lerner, B., Levinstein, M., Rosenberg, B., Guterman, H., Dinstein, I., and Romen, Y. (1994). “Feature selection and chromosomes classification using a multilayer perceptron neural network.” IEEE Int. Conf. on Neural Networks, Orlando, FL, 3540-3545.
- Martín del Brío B. y Sanz A; (2007). Redes Neuronales y Sistemas Borrosos, 3ª edición, Ra-ma. págs. 71 – 75.
- McKendry, I. G; 2002. Evaluation of Artificial Neural Networks for Fine Particulate Pollution (PM10 and PM2.5) Forecasting. *J. Air & Waste Manage. Assoc.*, 52, 1096-1101.
- Melas, D; Kioutsioukis, L; Ziomas, L. C; (2000). Neural Networks Model for Predicting Peak Photochemical Pollutant Levels. *Air & Waste Management Association*, 50, 495- 501.
- Mulholland, P J., Houser, J. N., & Maloney, K. O. (2005). Stream diurnal dissolved oxygen profiles as indicators of in-stream metabolism and disturbance effects: Fort Benning as a case study. *Ecological Indicators*, 5(3), 243–252. <https://doi.org/10.1016/j.ecolind.2005.03.004>
- Nagendra, S. M. S; Khare, M; (2005). Modelling urban air quality using artificial neural network. *Clean Techn. Environ. Policy*, 7, 116-126.
- Quinn, N. W. T., Jacobs, K., Chen, C. W., & Stringfellow, W. T. (2005). Elements of a decision support system for real-time management of dissolved oxygen in the San Joaquin River Deep Water Ship Channel. *Environmental Modelling and Software*, 20(12 SPEC. ISS.), 1495–1504. <https://doi.org/10.1016/j.envsoft.2004.08.014>

- Obregón N., (2001). Principios de hidroinformática urbana”, Seminario Internacional de Hidrología Urbana, Bogotá, Universidad javeriana
- Radwan, M., Willems, P., El-Sadek, A., & Berlamont, J. (2003). Modelling of dissolved oxygen and biochemical oxygen demand in river water using a detailed and a simplified model. *International Journal of River Basin Management*, 1(2), 97–103. <https://doi.org/10.1080/15715124.2003.9635196>
- Ranković, V., Radulović, J., Radojević, I., Ostojić, A., & Čomić, L. (2010). Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia. *Ecological Modelling*. <https://doi.org/10.1016/j.ecolmodel.2009.12.023>
- Recknagel, F., French, M., Harkonen, P., & Yabunaka, K.-I. (1997). Artificial neural network approach for modelling and prediction of algal blooms. *Ecological Modelling*, 96(1–3), 11–28. [https://doi.org/10.1016/S0304-3800\(96\)00049-X](https://doi.org/10.1016/S0304-3800(96)00049-X)
- Saint-gerons, A. I., & Navarra, N. (2004). Desarrollo de una red neuronal para estimar el oxígeno disuelto en el agua a partir de instrumentación DE E.D.A.R.
- SAFE Strategic assesment of Florida’s environmental (1995). Florida stream quality index, statewide summary. <http://www.pepps.fsu.edu./safe/environ/sqw1.html>.
- Şengörür, B., Doğan, E., Koklu, R., and Ve Samandar, A. (2006). “Dissolved oxygen estimation using artificial neural network for the water quality control.” *Fresen. Environ. Bull.*, 15(9a),1064-1067.
- Soares, J., Oliveira, A. P., Božnar, M. Z., Mlakar, P., Escobedo, J. F., & Machado, A. J. (2004). Modeling hourly diffuse solar-radiation in the city of São Paulo using a neural-network technique. *Applied Energy*, 79(2), 201–214. <https://doi.org/10.1016/j.apenergy.2003.11.004>
- Schmid, B. H., and Koskiaho, J. (2006). Artificial neural network modeling of dissolved oxygen in a wetland pond: The case of Hovi, Finland.*J. Hydrol. Eng.*, 11(2), 188-192.
- Shen, W., Chen, X., & Corriou, J. P. (2008). Application of model predictive control to the BSM1 benchmark of wastewater treatment process. *Computers and Chemical Engineering*, 32(12), 2849–2856.

- <https://doi.org/10.1016/j.compchemeng.2008.01.009>
- Shu, H; Lu, H; Fan, H; Chang, M; Chen, J; (2006). Prediction for Energy Content of Taiwan Municipal Solid Waste Using Multilayer Perceptron Neural Networks, *Air & Waste Manage. Assoc.* 56, 852–858
- Singh, K. P., Malik, A., Mohan, D., & Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) - A case study. *Water Research*, 38(18), 3980–3992. <https://doi.org/10.1016/j.watres.2004.06.011>
- Spanou, M & Chen, D. (1993). *Transactions on Ecology and the Environment* vol 2, © 1993 WIT Press, www.witpress.com, ISSN 1743-3541. *Environment*, 2(1), 14328–14336. <https://doi.org/10.2495/AIR990911>
- Suen, J. P., & Eheart, J. W. (2003). Evaluation of neural networks for modeling nitrate concentrations in rivers. *Journal of Water Resources Planning and Management-Asce*, 129(6), 505–510. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2003\)129:6\(505\)](https://doi.org/10.1061/(ASCE)0733-9496(2003)129:6(505))
- Tägil, K.; Marving, J.; Lomsy, M.; Hesse, B. y Edenbrant, L., (2008). Use the neural networks to improve quality control of interpretation in myocardial perfusion imaging, *The International Journal of Cardiovascular Imaging*, Springer, vol. 24, número 8, págs. 841-848.
- Tóth-Nagy, C., Conley, J. J., Jarrett, R. P., & Clark, N. N. (2006). Further validation of artificial neural network-based emissions simulation models for conventional and hybrid electric vehicles. *Journal of the Air and Waste Management Association*, 56(7), 898–910. <https://doi.org/10.1080/10473289.2006.10464513>
- Thomas, S; Jacko, R. B; (2007). Model for Forecasting Expressway Fine Particulate Matter and Carbon Monoxide Concentration; Application of Regression and Neural Network Models. *Air & Waste Management Association*, 57, 480-488.
- United States Environmental Protection Agency –USEPA, (2000). Estándares del Reglamento Nacional Primario de Agua Potable. EPA 815-F-00-007.

- Vieira, J., Fonseca, A., Vilar, V. J. P., Boaventura, R. A. R., & Botelho, C. M. S. (2013). Water quality modelling of Lis River, Portugal. *Environmental Science and Pollution Research*, 20(1), 508–524. <https://doi.org/10.1007/s11356-012-1124-5>
- Vladimir, M. K; Frederic, C; (2003). Some neural network applications in environmental sciences. Part II: advancing computational efficiency of environmental numerical models, *Neural Networks*, 16, 3-4
- Wen, X. H., Fang, J., Diao, M. N., & Zhang, C. Q. (2013). Artificial neural network modeling of dissolved oxygen in the Heihe River, Northwestern China. *Environmental Monitoring and Assessment*, 185(5), 4361–4371. <https://doi.org/10.1007/s10661-012-2874-8>
- Williams, R. J., White, C., Harrow, M. L., & Neal, C. (2000). Temporal and small-scale spatial variations of dissolved oxygen in the Rivers Thames, Pang and Kennet, UK. *Science of the Total Environment*, 251–252, 497–510. [https://doi.org/10.1016/S0048-9697\(00\)00401-0](https://doi.org/10.1016/S0048-9697(00)00401-0)
- Wirtz, D. S; El-Din, M. G; El-Din, A.G; Idriss, A; (2005). Systematic Development of an Artificial Neural Networks Model for Real-Time Prediction of Ground-Level Ozone in Edmonton, Alberta, Canada. *J. Air & Waste Manage. Assoc.*, 55, 1847-1857.
- Whitehead, P.G. and Hornberger, G.M., (1984). Modelling algal behavior in the River Thames. *Water Res.*, 18(8): 945-953,



ANEXOS

Anexo 1. Pasos secuenciales de entrenamiento de neuronas

Welcome to the Neural Fitting app.
Solve an input-output fitting problem with a two-layer feed-forward neural network.

Introduction
In fitting problems, you want a neural network to map between a data set of numeric inputs and a set of numeric targets.
Examples of this type of problem include estimating house prices from such input variables as tax rate, pupil/teacher ratio in local schools and crime rate (*house_dataset*); estimating engine emission levels based on measurements of fuel consumption and speed (*engine_dataset*); or predicting a patient's bodyfat level based on body measurements (*bodyfat_dataset*).
The Neural Fitting app will help you select data, create and train a network, and evaluate its performance using mean square error and regression analysis.

Neural Network
A two-layer feed-forward network with sigmoid hidden neurons and linear output neurons (*fitnet*), can fit multi-dimensional mapping problems arbitrarily well, given consistent data and enough neurons in its hidden layer.
The network will be trained with Levenberg-Marquardt backpropagation algorithm (*trainlm*), unless there is not enough memory, in which case scaled conjugate gradient backpropagation (*trainscg*) will be used.

To continue, click [Next].
Neural Network Start Welcome Back Next Cancel

Select Data
What inputs and targets define your fitting problem?

Get Data from Workspace
Input data to present to the network.
Inputs: data
Target data defining desired network output.
Targets: data
Samples are: Matrix columns Matrix rows

Summary
Inputs 'data' is a 1460x1 matrix, representing static data: 1460 samples of 1 element.
Targets 'data' is a 1460x1 matrix, representing static data: 1460 samples of 1 element.

To continue, click [Next].
Neural Network Start Welcome Back Next Cancel

Neural Fitting (nftool)

Validation and Test Data

Set aside some samples for validation and testing.

Select Percentages

Randomly divide up the 1460 samples:

Training:	70%	1022 samples
Validation:	15%	219 samples
Testing:	15%	219 samples

Explanation

Three Kinds of Samples:

- Training:** These are presented to the network during training, and the network is adjusted according to its error.
- Validation:** These are used to measure network generalization, and to halt training when generalization stops improving.
- Testing:** These have no effect on training and so provide an independent measure of network performance during and after training.

Change percentages if desired, then click [Next] to continue.

Neural Network Start Welcome Back Next Cancel

Neural Fitting (nftool)

Network Architecture

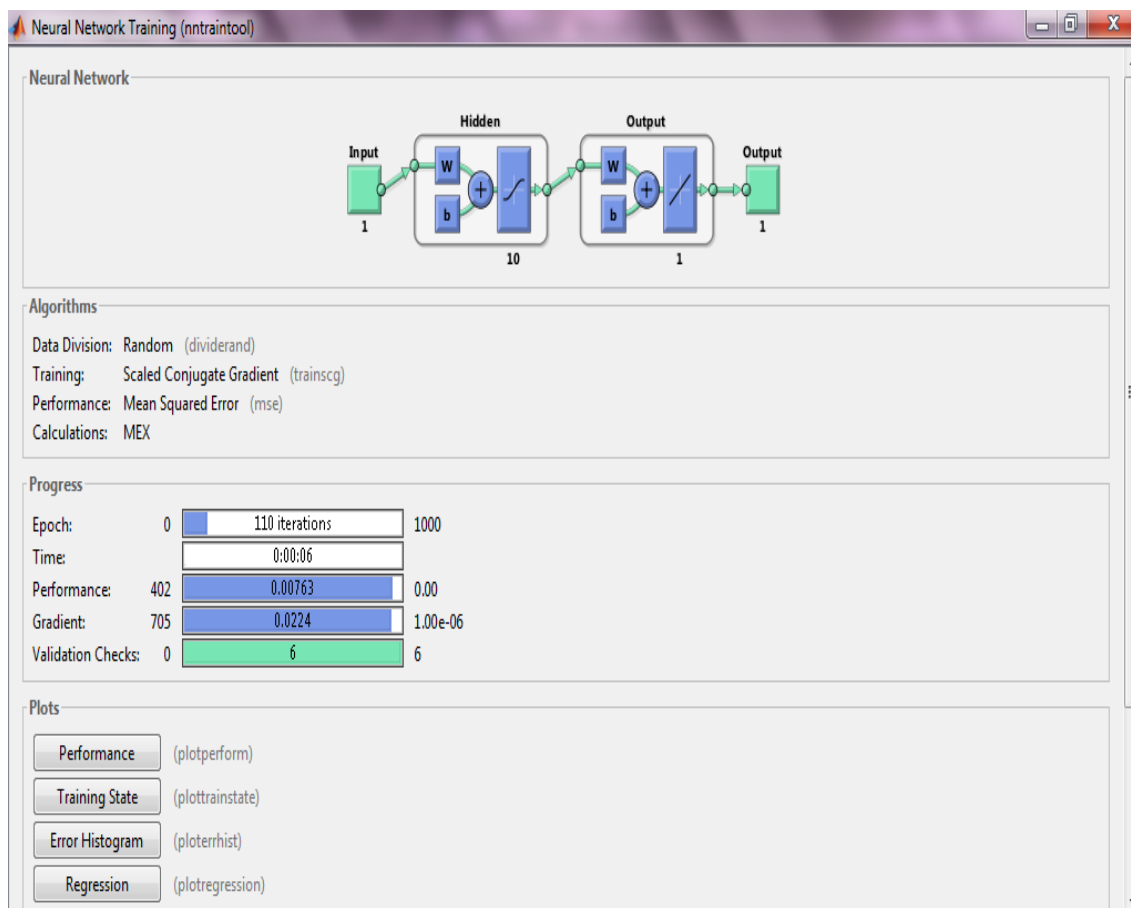
Set the number of neurons in the fitting network's hidden layer.

Restore Defaults

Neural Network

Change settings if desired, then click [Next] to continue.

Neural Network Start Welcome Back Next Cancel



Anexo 2. Pesos óptimos: con 4 variables de entrada 10 neuronas una capa oculta

```
function [y1] = myNeuralNetworkFunction(x1)

%MYNEURALNETWORKFUNCTION neural network simulation function.

% Generated by Neural Network Toolbox function genFunction, 01-Jan-2018 07:36:43.

% [y1] = myNeuralNetworkFunction(x1) takes these arguments:

% x = 1xQ matrix, input #1

% and returns:

% y = 1xQ matrix, output #1

% where Q is the number of samples.

%#ok<*>RPMTO>

% ===== NEURAL NETWORK CONSTANTS =====

% Input 1

x1_step1_xoffset = 2.6;

x1_step1_gain = 0.142857142857143;

x1_step1_ymin = -1;

% Layer 1

b1 = [-13.999662051054273;-11.170974436249514;8.0282120681148967;-
4.8071777252096419;-1.5677853994250068;-1.561398843366177;-
4.5957686774267179;7.686955974137291;10.763497961074513;14.0898169529623];

IW1_1 = [13.994975865307735;13.751829815679351;-
13.845173002367414;13.937053274346114;13.97686263383739;13.96602168642726-
13.994835642145173;14.028387665143086;14.08741215179872;13.90820726707949;

% Layer 2

b2 = -0.091201285441147278;

LW2_1 = [0.049030268801278122 0.1201457974948214 -0.1178472910237197
0.12137026043235767 0.10746000500023965 -0.10982032336332244
0.10916688785904051 0.11536901466220689 0.1014127222117697
0.17399165042500742];

% Output 1
```

```

y1_step1_ymin = -1;

y1_step1_gain = 0.142857142857143;

y1_step1_xoffset = 2.6;

% ===== SIMULATION =====

% Dimensions

Q = size(x1,2); % samples

% Input 1

xp1 = mapminmax_apply(x1,x1_step1_gain,x1_step1_xoffset,x1_step1_ymin);

% Layer 1

a1 = tansig_apply(repmat(b1,1,Q) + IW1_1*xp1);

% Layer 2

a2 = repmat(b2,1,Q) + LW2_1*a1;

% Output 1

y1 = mapminmax_reverse(a2,y1_step1_gain,y1_step1_xoffset,y1_step1_ymin);

end

% ===== MODULE FUNCTIONS =====

% Map Minimum and Maximum Input Processing Function

function y = mapminmax_apply(x,settings_gain,settings_xoffset,settings_ymin)

    y = bsxfun(@minus,x,settings_xoffset);

    y = bsxfun(@times,y,settings_gain);

    y = bsxfun(@plus,y,settings_ymin);

end

% Sigmoid Symmetric Transfer Function

function a = tansig_apply(n)

    a = 2 ./ (1 + exp(-2*n)) - 1;

end

```

```
% Map Minimum and Maximum Output Reverse-Processing Function
```

```
function x = mapminmax_reverse (y, settings_gain,settings_xoffset,settings_ymin)
```

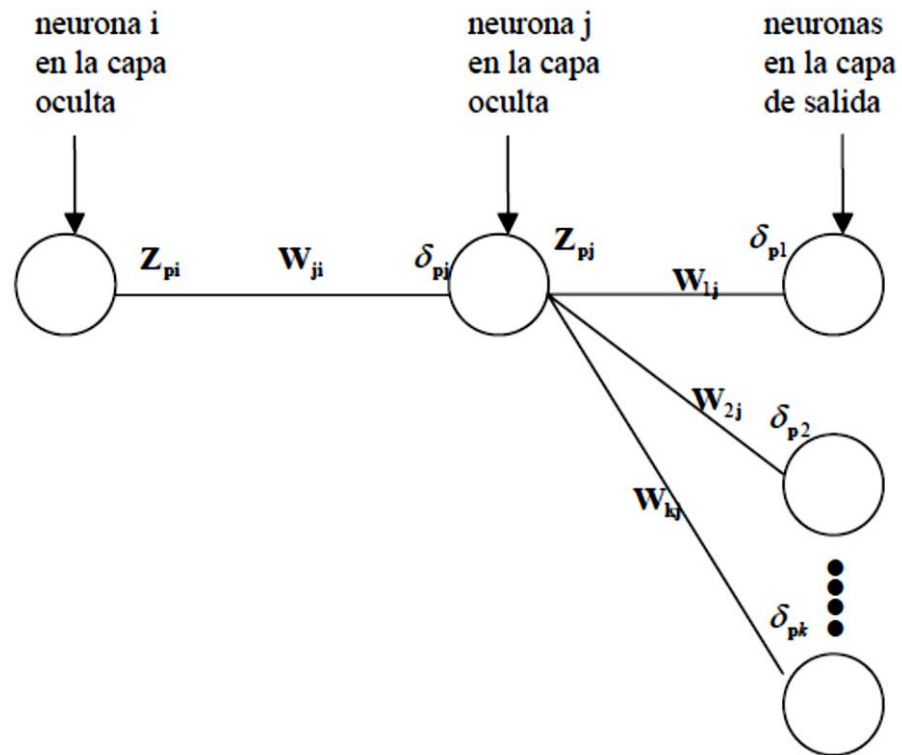
```
    x = bsxfun(@minus,y,settings_ymin);
```

```
    x = bsxfun(@rdivide,x,settings_gain);
```

```
    x = bsxfun(@plus,x,settings_xoffset);
```

```
end
```


Anexo 3. Como ajusta los pesos entre capas ocultas



Por otro lado, la actualización de los pesos entre las capas ocultas sería calculando el δ propagando hacia atrás desde la capa de salida.

$$\delta_{pj} = Z_{pj}(1 - Z_{pj}) \left(\sum_k \delta_{pk} W_{kj} \right)$$

$$\Delta_p W_{ji} = \eta \delta_{pj} Z_{pi}$$

Este cálculo se repite en las capas ocultas de la red.

Anexo 4. Derivación de las formulas

1. Backpropagation es un método gradiente descendente.
2. Minimiza el cuadrado de las diferencias entre los valores calculados y deseados de las salidas sumadas sobre todas las unidades de salida y todos los pares, entradas y salidas, del conjunto de entrenamiento.
3. Para un par p de entrenamiento (entrada y salidas) y n unidades de salida,

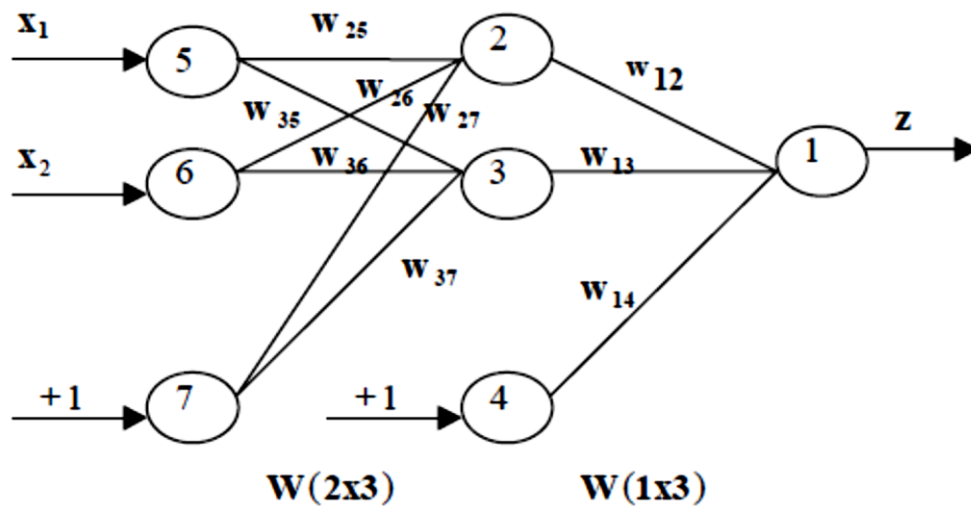
$$E_p = \frac{1}{2} \sum_{j=1}^n (T_{pj} - Z_{pj})^2$$

4. El error total sobre todo los patrones sería

$$E = \sum_p E_p$$

UN EJEMPLO USANDO BACKPROPAGATION

Sea la red:



Asuma una función sigmoide tal que $f'(n) = Z_{pi}(1-Z_{pi})$

Asuma $\eta = 1$ y todos los pesos inicialmente a 1.0.

El conjunto de entrenamiento es: $0\ 0 \rightarrow 1$
 $0\ 1 \rightarrow 0$

¿Cómo sería un ciclo de aprendizaje?

Para el primer patrón de entrada: $0\ 0 \rightarrow 1$

Pasos hacia delante:

$$n_2 = w_{25} x_1 + w_{26} x_2 + w_{27} (+1) = 1*0 + 1*0 + 1*1 = 1$$

$$n_3 = w_{35} x_1 + w_{36} x_2 + w_{37} (+1) = 1*0 + 1*0 + 1*1 = 1$$

$$z_2 = 1 / (1 + e^{-n_2}) = 1 / (1 + e^{-1}) = 1 / (1 + 0.368) = 0.731$$

$$z_3 = 0.731$$

$$n_1 = w_{12} z_2 + w_{13} z_3 + w_{14} (+1) = 1*.731 + 1*.731 + 1*1 = 2.462$$

$$z_1 = 1 / (1 + e^{-n_1}) = 1 / (1 + e^{-2.462}) = \mathbf{0.921}$$