

PREDICCION DE OXIGENO DISUELTO EN CULTIVOS ACUICOLAS POR MEDIO DE REDES NEURONALES ARTIFICIALES

J. R. Navarro Ruiz*, N. García-Díaz, J. A. Verduzco, J. García Virgen, A. C. Ruiz Tadeo
División de Estudios de Posgrado e Investigación, Tecnológico Nacional de México campus
Colima, Av. Tecnológico No. 1, C.P. 28976, Villa de Álvarez, Colima.

*g1946022@colima.tecnm.mx, ngarcia@colima.tecnm.mx, averduzco@colima.tecnm.mx,
jgarcia@colima.tecnm.mx, aruiz@colima.tecnm.mx

Área de participación: *Sistemas Computacionales*

Resumen

El cultivo de tilapia ha ido creciendo hasta convertirse en una de las actividades acuícolas más importantes en el mundo, en México se cultiva en 31 estados, Colima tiene granjas de crianza en sus 10 municipios. El agua que se utiliza en los cultivos debe contener ciertos parámetros de calidad para el crecimiento y desarrollo de las tilapias, los principales parámetros son: oxígeno disuelto (OD), conductividad y pH, siendo el OD quien afecta directamente el desarrollo y salud del cultivo, por ello es importante mantener ese parámetro controlado. El objetivo de este trabajo es desarrollar un modelo de red neuronal feed forward para predecir el OD, utilizando las variables de entrada conductividad y pH. Este modelo permite contribuir con la disminución de la mortalidad y mayor crecimiento de las tilapias mejorando su producción.

Palabras Clave: *Oxígeno Disuelto, Redes Neuronales, Predicción, Tilapia*

Abstract

The cultivation of tilapia has grown to become one of the most important aquaculture activities in the world, in Mexico it is grown in 31 states, Colima has breeding farms in its 10 municipalities. The water used in the crops must contain certain quality parameters for the growth and development of the tilapia, the main parameters are: dissolved oxygen (DO), conductivity and pH, being the DO that directly affects the development and health of the crop. Therefore, it is important to keep this parameter under control. The objective of this work is to develop a feed forward neural network model to predict DO, using the input variables conductivity and pH. This model makes it possible to contribute to the decrease in mortality and greater growth of tilapia, improving their production.

Key Words: *Dissolved Oxygen, Neural Networks, Prediction, Tilapia*

Introducción

La acuicultura es una actividad que se ha desempeñado con mucho éxito en los últimos tiempos, está se encarga de la producción controlada de organismos acuáticos por medio de diferentes técnicas de crianza. La Organización de las Naciones Unidas para la alimentación y agricultura (FAO) dice que se cultivan diversas especies de tilapia a nivel comercial, pero la tilapia del Nilo es la predominante mundialmente (FAO, 2009). En México es una de las principales especies producidas en granjas acuícolas de acuerdo con Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA) existen 9230 granjas acuícolas en el país. Debido a lo anterior México se posiciona en noveno lugar a nivel mundial en producción de tilapia y entre los principales consumidores de este producto (Comisión Nacional de Acuacultura y Pesca, 2017). Por otra parte, Colima registró en el año 2015 una producción de 788 toneladas en sistemas controlados y pesquerías acuaculturales, esto generó una derrama económica de 14.5 millones de pesos en la región (Comisión Nacional de Acuacultura y Pesca, 2016).

Oxígeno Disuelto (OD) para una buena calidad de cultivo

Los parámetros fisicoquímicos son quienes determinan la calidad del agua en los estanques de cultivo de tilapia, entre los más importantes se encuentran: oxígeno, pH, conductividad, temperatura, entre otros, destacando de entre ellos el OD, este se encuentra disuelto en el agua en forma de gas y es fundamental para la vida de los peces (Gutierrez S., Figueredo, & Spotorno, 2020). De acuerdo con el manual de cultivo de tilapia se menciona que a menor concentración de oxígeno el consumo de alimento se reduce, por lo que afecta el crecimiento de los peces (Saavedra Martínez, 2006). Los niveles aceptables de OD deben ser mayores a 5 mg/l, ya que niveles inferiores pueden presentar lento crecimiento, incluso un alto índice de mortalidad en el cultivo. En la tabla 1 se describen los niveles de los parámetros principales.

Tabla 1. Parámetros fisicoquímicos ideales para cultivo de tilapia (Instituto Nacional de Pesca, 2018).

PARAMETRO	OPTIMO	LIMITES
Temperatura	24 °C-29 °C	< 22 < 32°C
OD	< 5 mg/l	> 3 mg/l
pH	7.5	< 6.5 -≤ 8.5
CO2	< 30	< 50
Amonio	0.1	< 0.1 mg/l
Nitritos	4.6	< 5 mg/l
Salinidad	< 20*	< 20
Turbidez	25	< 30

Recolección de datos

Con el fin de probar el modelo, se recolectaron los parámetros mediante un grupo de sensores puestos directamente en el nacimiento de agua, la cual corre hasta los estanques para que estos sean llenados dentro del parque ecoturístico el Cahuite, ubicado en la comunidad de Acatitán, Colima, México. Para este primer estudio solo se incluyeron los 59 registros obtenidos mediante un censo de los datos por cada 10 minutos en el intervalo de las 8:15 a las 17:55 horas; Estos 59 registros fueron utilizados como el conjunto de datos, el cual incluye las variables OXÍGENO DISUELTO (mg/L), pH, CONDUCTIVIDAD (μS/cm), SÓLIDOS DISUELTOS (ppm) y TEMPERATURA (°C). En la figura 1 se ilustra la ubicación de los cultivos de tilapia.



Figura1. ubicación parque ecoturístico el Cahuite

Trabajos relacionados

En la actualidad la Inteligencia Artificial (IA) está dando solución a muchos problemas en el sector acuícola, la IA crece con un ritmo significativo abriendo camino a nuevas investigaciones con la finalidad de desarrollar nuevas tecnologías que brinden beneficios a la industria, en el trabajo realizado por (Quiñones Huatangar, y otros, 2020) desarrollaron un modelo de Redes Neuronales Artificiales (RNA) para estimar un índice de calidad de agua empleando el software Matlab, por otra parte, en el trabajo de (COPPO, JAKOMIN, & DELRIEUX, 2017) emplearon una RNA para la predicción de concentración de Boro. A su vez, (Carbajal Hernández & Sánchez Fernández, 2013) desarrollaron un sistema de Evaluación y predicción de la calidad del agua en el cultivo del camarón mediante un sistema de inferencia difusa. En el trabajo de (Meza Castellar & Gozales Salcedo, 2020) se elaboró un modelo neuronal artificial para la estimación de la demanda bioquímica de oxígeno en aguas marinas, por su parte (Rico, Paredes, & Fernández, 2009) utilizaron métodos referidos a RNA para la modelación y predicción del estado ecológico y calidad del agua.

Estos trabajos presentados anteriormente son resultados satisfactorios, pero cabe mencionar que ninguno de ellos presenta una solución al problema en cuestión del modelo de RNA propuesto en este trabajo que permite realizar una predicción del OD y anticipar con mayor precisión el estado futuro de la calidad del agua en los cultivos de tilapia.

Metodología

Materiales y métodos

De acuerdo con (Haykin, 2009) una RNA, es un procesador distribuido masivamente paralelo compuesto de unidades de procesamiento simples que tiene una propensión natural para almacenar el conocimiento experiencial y ponerlo a disposición para su uso, por lo cual, dados los requerimientos del proyecto, se decidió utilizar esta técnica para el desarrollo de este trabajo de

investigación. El presente trabajo fue desarrollado haciendo uso de la herramienta MATLAB, debido al amplio tipo de redes neuronales que incluye, así como la facilidad de empleo y desarrollo que ofrece por medio de su caja de herramientas. En la implementación de la RNA se optó por el uso de una red tipo Feed-Forward (newff).

Utilizando el software estadístico StatGraphics se llevó a cabo un análisis de regresión lineal múltiple y un ANOVA con el objetivo de determinar si el modelo puede simplificarse, tomando como variable dependiente OD (mg/L) y como variables independientes: SÓLIDOS DISUELTOS (ppm), TEMPERATURA (°C), pH y CONDUCTIVIDAD ($\mu\text{S/cm}$). Una vez realizado el análisis de regresión lineal múltiple y el ANOVA se observó que el modelo fue simplificado y solo se consideraron las variables pH y conductividad, puesto que el valor- p es menor que 0.05, hay una indicación de una posible correlación serial con un nivel de confianza del 95%.

La ecuación del modelo ajustado es:

$$\text{OD (mg/L)} = -78.2891 + 16.3936 \cdot \text{pH} - 0.0519755 \cdot \text{CONDUCTIVIDAD } (\mu\text{S/cm}) \quad (1)$$

La ecuación 1 describe la relación entre la variable dependiente (OD) y las variables independientes (pH y CONDUCTIVIDAD).

Un valor aceptable de coeficiente de determinación es $r^2 \geq 0.5$ (Humphrey, 1995). En este caso, el r^2 de la ecuación fue 0.7685. El ANOVA para esta ecuación tuvo una relación estadísticamente significativa entre las variables a un 95% del nivel de confianza

Una vez realizado el análisis de regresión se observó que el modelo fue simplificado y solo se consideraron las variables pH y conductividad puesto que el valor- p es menor que 0.05, ese término es estadísticamente significativo con un valor de confianza del 95%.

Feed-Forward Neural Network (newff)

Una red neuronal feed forward consiste en un conjunto de neuronas organizadas en capas por donde la información fluye en un único sentido desde las neuronas de entrada a la capa o capas de procesamiento, para los casos de redes mono capas y multi capa, respectivamente hasta llegar a la capa de salida de la red neuronal (Caicedo Bravo, 2009). En la figura 2 se muestra la arquitectura de una red Feed-Forward.

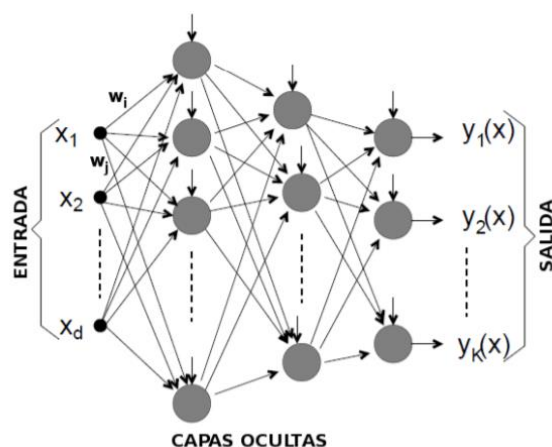


Figura2. Diagrama de la estructura típica de una red neuronal artificial tipo Feed-Forward (Vásquez López, 2014).

En la implementación de este proyecto se usó un método de validación conocido como "Hold-out". Los datos han sido divididos en dos particiones estratificadas, es decir, conservando la distribución

original de clases en cada una de las particiones. Para el conjunto de entrenamiento (train_*.dat = 41) se han utilizado un 70% de los patrones y para el conjunto de prueba (test_*.dat = 18) un 30%. Esto se utilizó con el propósito de establecer predicciones (Postiguillo-García, Ribal, & Blasco Ruíz, 2018).

Los datos se guardaron en 2 vectores: el primero corresponde a los datos de entrenamiento y el segundo a los datos de prueba, posteriormente ya con los datos almacenados se procedió al entrenamiento y clasificación de la RNA.

En la figura 3 podemos ver un fragmento del código para leer los datos de entrenamiento, se crearon los vectores donde se guarda el error y el MRE, utilizando un ciclo for el cual se encarga de crear, entrenar y calcular el tiempo DT con la red neuronal Feed-Forward para cada numero de neuronas especificado, y dentro de este ciclo cada interacción se calcula el MRE para cada uno de los 59 datos después se guarda el promedio de estos vectores de error y finalmente se calcula el error mínimo obteniendo en toda la corrida dando como resultado el error mínimo con el numero de neurona.

```

for i=1:neuronas
    net = feedforwardnet(i);
    net = train(net,input_real,dt_real);
    dt_calculado = sim(net,input_test);

    for j=1:size(data_Test,1)
        mre(j) = abs((dt_test(j) - dt_calculado(j))/dt_test(j));
    end

    error(i) = mean(mre);

    if(error(i)<error_min)
        error_min = error(i);
        neuronas_min = i;
        calculos = dt_calculado;
    end
end
fprintf('MMRE_MIN = %f\n',error_min);
fprintf('NEURONAS = %d\n',neuronas_min);

```

Figura 3. Código utilizado para almacenar y crear la RNA

Criterio de evaluación

Para evaluar los diferentes modelos de estimación de aprendizaje maquina es utilizada la MRE (por sus siglas en inglés, Magnitude of Relative Error) cual es la métrica de precisión más popular cuando es comparada en modelos de estimación (J. Wen, 2012).

La MRE es definida como sigue:

$$MRE_i = \frac{|Actual_i - Estimado_i|}{Actual_i} \quad (2)$$

El valor de MRE es calculado por cada observación i . La suma de múltiples observaciones de MRE se puede lograr a través de la Media MRE (MMRE) como se muestra en la ecuación 3.

$$MMRE = \frac{1}{n} \sum_{i=1}^n MRE \quad (3)$$

Como referencia, para modelos de predicción, un valor de $MMRE \leq 0.25$ es considerado aceptable (Conte S.D., 1986).

Resultados y discusión

El presente trabajo tuvo como resultado un modelo de red neuronal desarrollado en Matlab el cual permite predecir los niveles de oxígeno disuelto en el agua. En la tabla 2 podemos observar los resultados del valor del error mínimo obtenido y el número de neurona en el que se obtuvo ese valor.

Tabla 2. Resultados

MMRE_MIN	0.022
NEURONAS	4

Al mismo tiempo haciendo una comparativa de los datos actuales contra los datos estimados por las 4 neuronas de la red, podemos observar que los rangos son muy similares y esto indica que el nivel de precisión es aceptable. En la tabla 3 se muestran los resultados de dicha comparativa.

Tabla 3. Comparativa de datos

Conductividad	pH	Oxigeno Actual	Oxigeno Estimado	MRE
871.70	7.81	4.40	4.216	0.042
862.30	7.79	4.35	4.400	0.011
856.10	7.79	3.98	4.671	0.174
849.10	7.78	5.97	5.513	0.076
839.50	7.78	5.92	5.860	0.010
836.20	7.77	5.97	5.942	0.005
834.70	7.77	5.92	5.949	0.005
831.60	7.76	6.06	5.966	0.016
828.50	7.76	5.96	5.972	0.002
823.10	7.74	5.98	5.984	0.001
821.80	7.74	5.92	5.988	0.011
820.80	7.74	5.87	5.991	0.021
817.00	7.74	6.01	6.008	0.000
813.80	7.73	6.06	6.025	0.006
809.40	7.72	6.07	6.051	0.003
801.60	7.69	6.11	6.084	0.004
799.60	7.68	6.09	6.085	0.001
798.30	7.66	6.17	6.077	0.015
MMRE				0.022

Debido a que el valor obtenido (0.022) de **MMRE es menor o igual que 0.25** se considera el modelo de predicción propuesto en este trabajo de investigación como aceptable.

Trabajo a futuro

Se recomienda como trabajo futuro desarrollar una investigación con un tipo diferente de red neuronal, con objetivo de hacer un comparativo entre el trabajo futuro y el presente, para ver cual tiene mayor precisión.

Realizar un comparativo entre el resultado obtenido de MMRE con redes neuronales versus la ecuación de regresión lineal múltiple mostrada en la ecuación 1 para determinar la mejor precisión.

Además, en el futuro inmediato se sugiere usar una muestra de datos mucho más grande a la utilizada en el presente estudio.

Conclusiones

En este trabajo de investigación fue creada una red neuronal tipo Feed-Forward, la cual permite predecir el nivel de oxígeno disuelto en los estanques de tilapia.

El uso de las redes neuronales contribuye con excelente precisión en los estanques de tilapia, logrando con ello una mejor producción y un mayor crecimiento de la especie en cultivo.

Agradecimientos

Los autores del presente trabajo agradecen al Consejo Nacional de Ciencia y Tecnología (CONACYT), al Programa para el Desarrollo Profesional Docente (PRODEP), así como al Tecnológico Nacional de México Campus Colima por el apoyo brindado durante el desarrollo de esta investigación.

Referencias

1. Caicedo Bravo, E. F. (2009). *Una aproximación práctica a las redes neuronales artificiales*. ISBN: 9789586707671 9586707679 Cali, Colombia
2. Carbajal Hernández, J. J.; Sánchez Fernández, L. P., (2013). Centro de Investigación en Computación, Instituto Politécnico Nacional Computación y Sistemas Vol. 17 No.3, pp. 435-455 ISSN 1405-5546 <http://www.scielo.org.mx/pdf/cys/v17n3/v17n3a14.pdf>
3. Comisión Nacional de Acuacultura y Pesca. (2016). *Se reorganizan acuacultores de tilapia en Colima; buscan mejorar productividad y comercialización*. <https://www.gob.mx/conapesca/prensa/se-reorganizan-acuacultores-de-tilapia-en-colima-buscan-mejorar-productividad-y-comercializacion#:~:text=Se%20reorganizan%20acuacultores%20de%20tilapia%20en%20Colima%3B%20buscan%20mejorar%20productividad%20y%20comercial>
4. Comisión Nacional de Acuacultura y Pesca. (2017). *México entre los diez primeros lugares a nivel mundial en producción de tilapia: SAGARPA*. <https://www.gob.mx/conapesca/prensa/mexico-entre-los-diez-primeros-lugares-a-nivel-mundial-en-produccion-de-tilapia-sagarpa-103606?idiom=es>
5. Coppo, Ricardo J., Jakomin Mariana L., Y Delrieux, Claudio A. (2017). *Proimca – Prodeca Predicción de Parámetros de calidad de agua mediante redes neuronales en la cuenca del río Pilcomayo Argentina* http://www.edutecne.utn.edu.ar/prodeca-proimca/actas-prodeca-2017/DCA04_Prediccion-de-Parametros.pdf
6. Conte S.D., Dunsmore H.E., Shen V.Y., (1986) *Software Engineering Metrics and Models*. Benjamin-Cummings Publishing Co., Inc. Redwood City, CA, USA
7. FAO. (2009). *Oreochromis niloticus*. In *Cultured aquatic species fact sheets*. Text by Rakocy, J. E. Edited and compiled by Valerio Crespi and Michael New. CD-ROM (multilingual).
8. Haykin, Simon. (2009). *Neural networks and learning machines / Simon Haykin.—3rd ed. p. cm* <http://dai.fmph.uniba.sk/courses/NN/haykin.neural-networks.3ed.2009.pdf>
9. Humphrey W., (1995). *A Discipline for Software Engineering* Wesley.
10. Instituto Nacional de Pesca (2018) *Parámetros físico-químicos ideales para cultivo de tilapia* <https://www.gob.mx/inapesca/acciones-y-programas/acuacultura-tilapia>
11. Quiñones Huatangar, L., Ochoa Toledo, L., Kemper Valverde, N., Gamarra Torres, O., Bazán Correa, J., Delgado Soto, J. (2020) *Artificial neural network to estimate an index of water quality*. <http://ingenieria.ute.edu.ec/enfoqueute/> e-ISSN: 1390-6542 DOI: 10.29019/enfoque.v11n2.633

12. Meza Castellar, R., & Gonzalez Salcedo, L. (2020). *Elaboración de un modelo neuronal artificial para la estimación de la demanda bioquímica de oxígeno en aguas marinas*. *Revista De Investigación Agraria Y Ambiental*, 11(2), 147 - 156. <https://doi.org/10.22490/21456453.3441>
13. Postiguillo-García, D.; Ribal, J.; Blasco Ruíz, A. (2018). *Caso de estudio: Modelización de la vida útil de maquinaria destinada a proyectos de obra pública a través de la aplicación de regresiones mínimo cuadráticas y validación cruzada*. *Finance, Markets and Valuation*. 4(1):57-79. <http://hdl.handle.net/10251/122884>
14. Rico, C., Paredes, M., Fernández, N., (2009) *Modelacion de la estructura jerárquica de macroinvertebrados bentónicos a traves de redes neuronales artificiales* *Acta Biológica Colombiana*, 14(3),71-95.[fecha de Consulta 16 de Septiembre de 2020]. ISSN: 0120-548X. Disponible en: <https://www.redalyc.org/articulo.oa?id=3190/319028005006>
15. Gutierrez S., Figueredo, & Spotorno, (2020) *Revista Tecnología y Ciencia - Universidad Tecnológica Nacional DOI:https://doi.org/10.33414/rtyc.38.67-82.2020 - ISSN 1666-6933 Reconocimiento-NoComercial 4.0 Internacional*.
16. Saavedra Martínez, M.A., (2006). *Manejo del cultivo de Tilapia*. [https://www.crc.uri.edu/download/ Manejo del cultivo de Tilapia.pdf](https://www.crc.uri.edu/download/Manejo%20del%20cultivo%20de%20Tilapia.pdf)
17. Vásquez López, J. P., (2014). *Instituto Mexicano del Transporte*, <https://www.imt.mx/archivos/Publicaciones/PublicacionTecnica/pt406.pdf>
18. Wen, J., Li, S., Lin, Z., Hu, Y., Huang,C., (2012). "Systematic literature review of machine learning based software development effort estimation models," *Information and Software Technology*, 54, 41–59.