Prediction of Dissolved Oxygen from pH and Water Temperature in Aquaculture Prawn Ponds

Joel Janek Dabrowski
Data61, CSIRO
Brisbane, Queensland, Australia
Joel.Dabrowski@data61.csiro.au

Ashfaqur Rahman
Data61, CSIRO
Hobart, Tasmania, Australia
Ashfaqur.Rahman@data61.csiro.au

Andrew George
Data61, CSIRO
Brisbane, Queensland, Australia
Andrew.George@data61.csiro.au

ABSTRACT

In this study, dissolved oxygen is predicted from pH and water temperature. A linear regression model, a feed forward artificial neural network (ANN), a linear dynamical system (LDS), and a long short term memory (LSTM) recurrent neural network are compared. The LSTM obtains the highest levels of accuracy, whereas the ANN performs poorly in comparison. A key observation is that time dependent models provide superior results. This is significant as the ANN, which does not explicitly model time dependency, is the primary model used in most related studies. Furthermore, a key contribution is a novel approach to applying the LDS for prediction of dissolved oxygen.

KEYWORDS

Water quality, DO prediction, Long short term memory (LSTM), Recurrent neural network (RNN), Artificial neural network (ANN), Linear dynamical system (LDS)

ACM Reference Format:
Joel Janek Dabrowski, Ashfaqur Rahman, and Andrew George. 2018. Prediction of Dissolved Oxygen from pH and Water Temperature in Aquaculture Prawn Ponds. In *Proceedings of AIW'18.* ACM, New York, NY, USA, 5 pages.

1 INTRODUCTION

A crucial activity in prawn farming is monitoring prawn pond water quality. Variables such as dissolved oxygen (DO), pH, temperature, and salinity are commonly monitored using sensors [6, 18]. Sensor monitoring comes with challenges such as: purchasing and maintaining sensors is costly, gathering readings with hand-held sensors over many ponds is time consuming, readings can be incorrectly logged, and sensors can fail. Such problems can be mitigated by reducing the number of sensors required. A sensor can be made redundant if its associated variable can be predicted from other sensor data. In this study, a long short term memory (LSTM) recurrent neural network (RNN), a feed forward artificial neural network (ANN), a linear regression model, and a linear dynamical system (LDS) are used to predict dissolved oxygen from pH and water temperature sensor readings.

An additional aim of this study is to demonstrate the effectiveness of models that incorporate time dependence. The LSTM and LDS model time dependence, whereas the linear regression and ANN models do not. These time dependent models are shown to provide significant improvement over the static models.

A key contribution of this study is a novel approach to modelling and predicting water quality parameters using the LDS. All the water quality variables and their first derivatives are modelled together in a latent space. This provides the means for the LDS to model linear dependence over time as well as over the variables. These dependencies are learned using the expectation maximisation (EM) algorithm. For prediction, the DO is treated as a missing variable. The generative properties of the LDS allow for it to infer the "missing" DO based on the observations of pH and temperature.

2 BACKGROUND AND RELATED WORK

This study fits within a broader context of the development of a aquaculture prawn farm decision support system. Several sensors have been deployed on prawn farms for monitoring water quality related parameters. The sensors include water quality, spectral reflectance, hydrophones, and weather sensors. The sensor data is uploaded to a central cloud based system. Various inference and data analytics tasks are performed on the data. The framework of the decision support system is illustrated in Figure 1. In this study, the modelling and estimation of water quality parameters in prawn ponds is considered.

Several studies have been conducted where DO is predicted from other water quality parameters. The majority of the studies are conducted on rivers such as the Tualatin River, USA [19], the Karoon River, Iran [9], the Danube River, Republic of Serbia [3], the Surma River, Bangladesh [2], the Algeria river, Brazil [8], and the

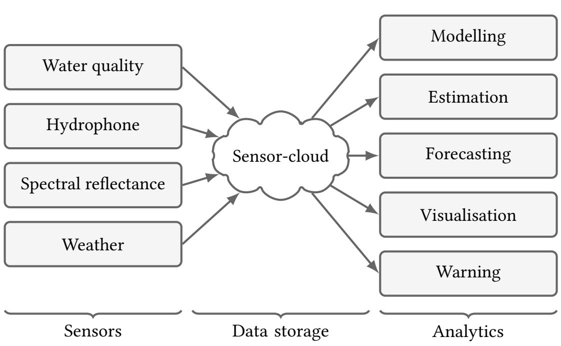


Figure 1: Aquaculture prawn farm decision support system.

AIW'18, December 11, 2018, Wellington, New Zealand

ACM ISBN 978-1-4503-6548-2…$15.00
DOI: 10.1145/3314487.3314488

Bow River, Canada [11]. All of these studies use some form of a neural network based model. None of the models incorporate any explicit form of time dependency.

Various water quality related variables have been used to predict DO. Rounds [19] predict DO from stream-flow, air temperature, solar radiation, and rainfall. In an additional test, water temperature and specific conductance were included in the inputs. Emamgholizadeh et al. [9] predict DO, biological oxygen demand, and chemical oxygen demand from electrical conductivity, pH, Ca, Mg, Na, Turbidity, PO_4 , NO_3 and NO_2. Antanasijević et al. [3] predict DO from temperature, pH, conductivity, CO_2, HCO_3^-, alkalinity, chemical oxygen demand, total suspended solids, NO_3-N, Cl^-, SO_4^{2-}, PO_4^{3-}, P, Ca, Mg, hardness, Na, K, and biological oxygen demand. Ahmed [2] predict DO from biochemical oxygen demand and chemical oxygen demand. de Araujo Schtz et al. [8] predict DO from biochemical oxygen demand, chemical oxygen demand, pH, electrical conductivity, temperature, nitrite, nitrate, ammonia nitrogen, flow, and distance from the launch. Finally, Khan and Valeo [11] predict DO from water temperature and flow rate.

Rivers differ from prawn ponds in various ways. Prawn ponds are smaller bodies of water and do not maintain flow. Highly intensive prawn ponds are stocked with prawn at high densities. Density can reach in excess of 75 prawn/m^2 [16]. Respiration and excretion of the crustaceans can significantly affect water quality. Through respiration, oxygen is absorbed and carbon dioxide is released. Through excretion, various nutrients and nitrogen based compounds are released into the water.

With higher flow rates in rivers, the algae blooms may not be as dense as prawn ponds. A dense algae bloom provides food for postlarvae and is a key driver of various water quality parameters in prawn ponds [6]. During the day, oxygen is produced by the algae through photosynthesis. During the night, photosynthesis ceases. The biotic component of the pond produce carbon dioxide through respiration. Carbon dioxide reacts with the water forming carbonic acid, which affects the pH levels in the water. The result is that dissolved oxygen and pH follow diurnal fluctuations. As the algae bloom is dense in a prawn pond and there is little flow in water, these fluctuations can be significant.

Applications of modelling DO in aquaculture farming exist but are limited in the literature. Several studies have been conducted in one-step-ahead forecasting of dissolved oxygen in crab culture ponds [13, 14, 22]. The studies propose hybrid machine learning algorithms that take DO and other weather and water quality variables as inputs. The models forecast the next DO sensor reading. These crab culture studies are related to the current study through the modelling of DO. The studies however use DO as an input variable and perform forecasting. In the current study, prediction of the current DO level is performed using only pH and temperature. Forecasting is not considered.

3 MODELS

The LSTM, LDS, ANN, and linear regression models are trained on a given dataset. The LSTM and LDS models are dynamical models as they take into account the time dependency of the variables. No explicit time dependency in the data is contained in the ANN and linear regression models. The LDS is a stochastic model, whereas the other models are deterministic.

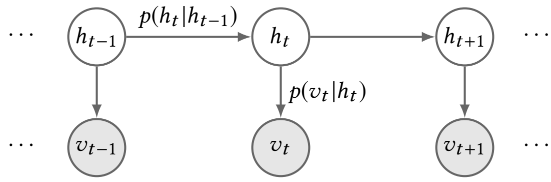


Figure 2: Graphical model representation of the LDS.

3.1 Linear Dynamical System

Dabrowski et al. [7] proposed a linear and nonlinear state space model for modelling and forecasting dissolved oxygen and pH in prawn ponds. The linear model is extended in this study to model multiple variables. The linear model is in the form of a linear dynamical system (LDS).

The LDS includes a latent or hidden variable h_t that evolves over time, $t = 0, \ldots, T$. The latent variable emits an observable variable v_t from which measurements can be made. The latent variable is assumed to follow a first order Markov process. The graphical model describing this system is illustrated in Figure 2. The edges between the latent variables describe the transition distribution $p(h_t|h_{t-1})$. The edges between the latent and observable variables describe the emission distribution $p(v_t|h_t)$. Linear-Gaussian assumptions in the LDS result in the following state space equations [4, 5, 15]

$$h_t = Ah_{t-1} + \eta_t^h \quad (1)$$

$$v_t = Bh_t + \eta_t^v \quad (2)$$

The variable h_t is the state vector, A is the state transition matrix, and $\eta_t^h \sim \mathcal{N}(0, \Sigma^h)$ is the state noise vector. The variable v_t is the observation vector, B is the emission or measurement matrix, and $\eta_t^v \sim \mathcal{N}(0, \Sigma^v)$ is the measurement noise vector.

The observations of a water quality variable are modelled with a seasonal, trend, and noise component as follows [7]

$$v_t = \alpha_t \sin(\omega t) + \gamma_t + \eta_t^v \quad (3)$$

The seasonal component, $\alpha_t \sin(\omega t)$ is modelled with a sinusoid. The trend component γ_t is modelled as a continuous local linear trend model (constant velocity process). The parameters of this function are modelled in the LDS latent space. Let $\psi_t = \alpha_t \sin(\omega t)$ such that the state vector $h(t)$ is defined as [7]

$$h(t) = \begin{bmatrix} \gamma_t & \dot{\gamma}_t & \psi_t & \dot{\psi}_t \end{bmatrix}^T \quad (4)$$

To extend this model for multiple variables, a state vector given in (4) is defined for each water quality variable. These vectors are then stacked into a single state vector. With the three variables of DO, pH, and temperature, each containing 4 vector-elements, the complete state vector has 12 elements. Similarly, the observation vector comprises the models given by (3) for each variable. The observation vector thus has 3 elements.

The state transition matrix A defines the interaction of the state vector over time as well as over the state space. With the linear-Gaussian assumptions, the interaction that is modelled is assumed to be linear. The matrix A is initialised as

$$A = \begin{bmatrix} A' & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A' \end{bmatrix} \quad (5)$$

where $\mathbf{0}$ denotes a 4×4 matrix of zeros and A', in continuous time form, is given by [7]

$$A' = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \omega^2 & 0 \end{bmatrix}$$

The form of (5) assumes that the three water quality variables are independent. To learn the linear dependence between the variables, the state transition matrix is learned using the expectation maximisation (EM) algorithm [4, 5]. The form given in (5) serves as an initial starting point for the EM algorithm.

The emission matrix maps the elements from the latent variable space to the observed variable space according to (3). The emission matrix is thus initialised as

$$B = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

As for the state transition matrix, the observation matrix is refined using the EM algorithm.

The model is trained using a complete dataset of DO, pH and temperature. The EM algorithm learns the relationship between the variables through the A and B matrices. For prediction, only the pH and temperature observations are presented to the model. The DO elements in the model states are treated as missing variables and are inferred. As the relationships between the variables have been learned, the model is able to infer the missing values. That is, the DO is inferred based on the learned linear relationship between pH, temperature and DO.

3.2 Features

A set of features are defined for the linear regression and neural network models. The features include the mean trend and the envelope of the water quality variable signal. The trend is computed in the γ_t component in the LDS model. The envelope of some signal $x(t)$ can be computed as the absolute value of the analytic signal. The analytic signal is given by [20]

$$x_a(t) = x(t) + i\mathcal{H}(x(t))$$

where $\mathcal{H}(x(t))$ is the Hilbert transform of $x(t)$. The envelope and mean trend together provide an indication of the amplitude of the water quality variable signal. A more dense algae bloom is likely to produce higher signal amplitudes in the DO and pH variables. The difference between the envelope and the mean is thus additionally included as a feature.

3.3 Linear Regression

A linear regression model is applied for DO prediction. The input comprises pH, temperature, and the corresponding features of these variables described in Section 3.2. The model is implemented with the Scikit-learn [17] python package with default settings.

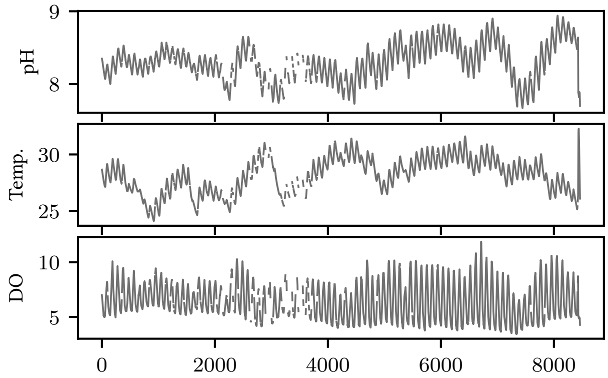


Figure 3: Plots of the pH, temperature (°C), and DO (mg/L) datasets.

3.4 Artificial Neural Network

A standard feed forward ANN with a hidden layer is applied for DO prediction. The input layer comprises pH, temperature, and the corresponding features of these variables described in Section 3.2. The hidden layer comprises 32 neurons. The hidden neurons use the hyperbolic tangent as an activation function. The output layer comprises a single neuron with a linear activation function. The Adam [12] algorithm is used for training the ANN with the recommended hyperparameters of $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$. Tuning and refinement is performed over the number of hidden units and the learning rate. The optimal learning rate is $\eta = 0.01$. The model is implemented with the Scikit-learn [17] python package.

3.5 LSTM Recurrent Neural Network

A standard LSTM [10] with 128 hidden units and a fully connected output layer is employed for predicting DO. The inputs comprises pH, temperature, and the corresponding features of these variables discussed in Section 3.2. The output layer comprises a single linear output neuron. The Adam [12] optimisation algorithm is used to train the LSTM by reducing the mean squared error. The default hyperparameters of $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$ are used. Tuning and refinement is performed over the number of hidden units in each hidden layer and the learning rate. A learning rate of $\eta = 0.001$ is used for training the final model. The model is implemented in the TensorFlow [1] python package.

4 RESULTS

4.1 Dataset

The dataset used in this study comprises of DO, pH, and water temperature readings taken from a 0.18ha prawn pond. The YSI EXO2 Multiparameter Sonde sensor [21] was used to acquire the readings. Samples were taken at 15 minute intervals over a period of 88 days. The dataset comprises a total of 8463 samples. Plots of the dataset are presented in Figure 3.

There is missing data in the dataset; especially in the sample range between 2000 and 4000. For the linear regression, ANN, and LSTM models, the missing data is interpolated using a linear model.

Interpolation is not required for the LDS as it is a generative model. A simple linear interpolation method is chosen to contrast the LDS's missing data handling ability with the other models.

4.2 Prediction Results

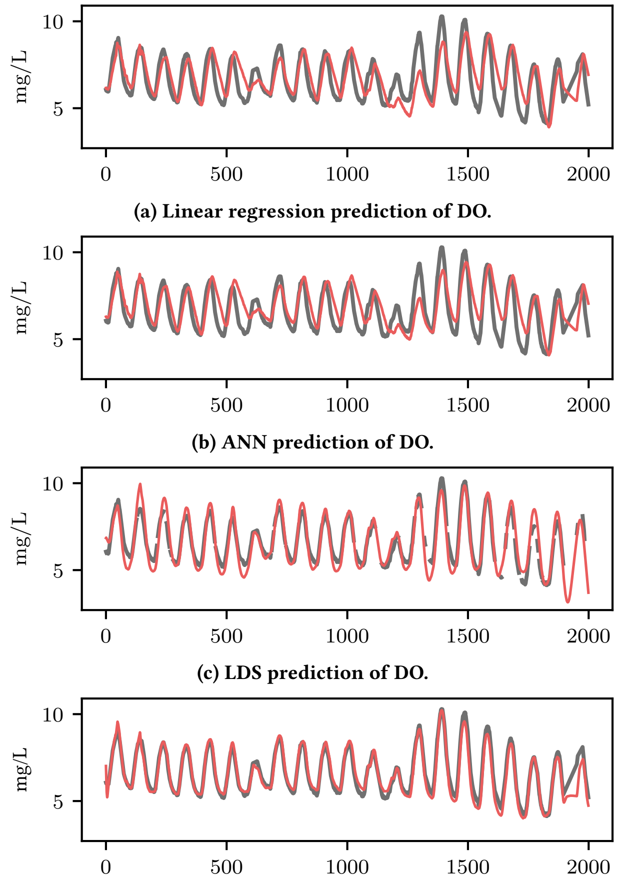
The predicted DO time series for the three models is presented in Figure 4. These predictions are made for samples 1000 to 3000 of the dataset. Over this region, there are various artefacts in the data that make predictions challenging. For example, there are several unexpected changes in the trend and envelope. This results in the time series being highly non-stationary. Furthermore, there are several examples of missing data over this region.

The normalised root mean squared error is used to provide an evaluation of the error between the predicted values and the measured data. Let y_t denote a sample from the dataset corresponding to time t. Let $\hat{y}_t$ denote the prediction of y_t at time t. For a set of N predictions, the normalised root mean squared error (NRMSE) is given by

$$\epsilon_{\text{nrmse}} = \frac{\sqrt{\frac{1}{N}\sum_{i=0}^{N}(y_i - \hat{y}_i)^2}}{y_{\max} - y_{\min}} \times 100\% \tag{6}$$

The predictions of the linear regression, ANN, LDS, and LSTM produce NRMSEs of 14.67%, 15.79%, 7.46%, and 5.79% respectively. The LSTM produces the optimal results with the LDS not far behind. A limitation of the LDS model is that it assumes a sinusoidal form of the data. There is an inherent error in the results as the measured DO has a slightly flattened trough. This error could be reduced by including additional harmonics in the form of a Fourier series into the LDS model. Despite the inherent error, the LDS still significantly outperforms the ANN and linear regression models.

Data for samples 1895 to 1951 are missing. The linear interpolation of the data over these points is a poor approximation. The linear regression, ANN, and LSTM models do not handle the error well. This is a key limitation of these models. No interpolation is required for the LDS model. The LDS provides a suitable approximation for these samples. This is a key feature of state-space based models.



(a) Linear regression prediction of DO.

(b) ANN prediction of DO.

(c) LDS prediction of DO.

(d) LSTM prediction of DO.

Figure 4: Prediction of DO from pH, temperature, and features for the linear regression, ANN, LDS, and LSTM models. The red curve is a plot of the predictions. The grey curve is a plot of the sensor data. Note that the missing sensor data has been interpolated in Figures 4a, 4b, and 4d.

4.3 Cross-Validation

To demonstrate statistical significance of the results, a repeated random sampling validation approach is used to test the models. A test set of 2000 sequential samples are drawn from the dataset. Each model is trained on the remaining samples and tested on the test set. This process is repeated 20 times. The start index of each of the 20 test sets is randomly selected.

The mean and median values of the NRMSE results over the 20 folds are presented in Table 1. Box-whisker plots of the results are presented in Figure 5. The upper and lower box edges extend from the upper to the lower quartiles of the data. The red line in the box indicates the median. The whiskers extend from the box indicating the range of the data. The LSTM provides the optimal results with the lowest median and mean values. The box height in the box-whisker plot is narrow indicating consistently low errors. The LDS is surprisingly effective given that it is a linear model. It provides a significant improvement to the ANN and linear regression models. The worst performing result of the LDS (14.04%), is better than the ANN's best result (14.10%). These are indicated by the LDS's upper whisker and the ANN's lower whisker in the plots.

The ANN and the linear regression models have very similar average performance. This indicates a linear relationship between the input and output variables in the static case. The ANN has less variation in its results compared to the linear regression model. This is indicated by it's smaller box in the box-whisker plot. Less variation is preferred.

5 SUMMARY AND CONCLUSION

A LSTM, a LDS, an ANN, and a linear regression model are trained to predict DO from pH and water temperature. Many related studies in the literature use a feed forward ANN for prediction purposes and do not consider time dependency in the data. It is shown that dynamic models (LSTM and LDS) are significantly more accurate than static models (ANN and linear regression) in the application presented in this study.

Model	Mean (%)	Median (%)
REG	17.77	17.61
ANN	17.84	17.28
LDS	10.22	10.12
LSTM	6.83	6.21

Table 1: Normalised root mean squared error (NRMSE) statistics for the cross validation results.

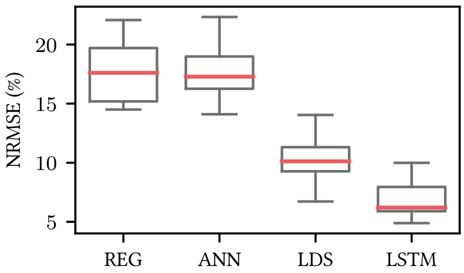


Figure 5: Box plots of the NRMSE for the cross validation results. The 'REG' label denotes the linear regression model results.

The LDS model performs extremely well. This is despite the fact that it is a linear model applied to model a complex system. The stochastic nature of the LDS provides a means to handle missing data. This highlights a key benefit of the LDS as the deterministic models do not have any capability to handle the missing data. Furthermore, the deterministic models show poor performance handling an interpolation of the missing data.

In future work, the ability of the models to generalise over different ponds may be investigated. Additionally, a study into the predictions of various other water quality related variables may be conducted. Such a study may additionally incorporate data from other sensors such as weather sensors.

REFERENCES

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. https://www.tensorflow.org/ Software available from tensorflow.org.

[2] A.A. Masrur Ahmed. 2017. Prediction of dissolved oxygen in Surma River by biochemical oxygen demand and chemical oxygen demand using the artificial neural networks (ANNs). *Journal of King Saud University - Engineering Sciences* 29, 2 (2017), 151 – 158. https://doi.org/10.1016/j.jksues.2014.05.001

[3] Davor Antanasijević, Viktor Pocajt, Aleksandra Perić-Grujić, and Mirjana Ristić. 2014. Modelling of dissolved oxygen in the Danube River using artificial neural networks and Monte Carlo Simulation uncertainty analysis. *Journal of Hydrology* 519 (2014), 1895 – 1907. https://doi.org/10.1016/j.jhydrol.2014.10.009

[4] D. Barber. 2012. *Bayesian Reasoning and Machine Learning*. Cambridge University Press. https://books.google.com.au/books?id=yxZtddB_Ob0C

[5] Christopher Bishop. 2006. *Pattern Recognition and Machine Learning*. Springer-Verlag New York. https://www.springer.com/gp/book/9780387310732

[6] Claude E Boyd and Craig S Tucker. 1998. *Pond aquaculture water quality management*. Springer US. https://doi.org/10.1007/978-1-4615-5407-3

[7] Joel Janek Dabrowski, Ashfaqur Rahman, Andrew George, Stuart Arnold, and John McCulloch. 2018. State Space Models for Forecasting Water Quality Variables: An Application in Aquaculture Prawn Farming. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18)*. ACM, New York, NY, USA, 177–185. https://doi.org/10.1145/3219819.3219841

[8] Fabiana Costa de Araujo Schtz, Vera Lucia Antunes de Lima, Eduardo Eyng, Adriano de Andrade Bresolin, and Fernando Schtz. 2015. Simulation of the concentration of dissolved oxygen in river waters using Artificial Neural Networks. In *2015 11th International Conference on Natural Computation (ICNC)*. 1252–1257. https://doi.org/10.1109/ICNC.2015.7378171

[9] S. Emamgholizadeh, H. Kashi, I. Marofpoor, and E. Zalaghi. 2014. Prediction of water quality parameters of Karoon River (Iran) by artificial intelligence-based models. *International Journal of Environmental Science and Technology* 11, 3 (01 Apr 2014), 645–656. https://doi.org/10.1007/s13762-013-0378-x

[10] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.

[11] Usman T. Khan and Caterina Valeo. 2017. Optimising Fuzzy Neural Network Architecture for Dissolved Oxygen Prediction and Risk Analysis. *Water* 9, 6 (2017), 381. https://doi.org/10.3390/w9060381

[12] D Kinga and J Ba Adam. 2015. A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, Vol. 5.

[13] Shuangyin Liu, Haijiang Tai, Qisheng Ding, Daoliang Li, Longqin Xu, and Yaoguang Wei. 2013. A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction. *Mathematical and Computer Modelling* 58, 3 (2013), 458 – 465. https://doi.org/10.1016/j.mcm.2011.11.021 Computer and Computing Technologies in Agriculture 2011 and Computer and Computing Technologies in Agriculture 2012.

[14] Shuangyin Liu, Longqin Xu, Yu Jiang, Daoliang Li, Yingyi Chen, and Zhenbo Li. 2014. A hybrid WA–CPSO-LSSVR model for dissolved oxygen content prediction in crab culture. *Engineering Applications of Artificial Intelligence* 29 (2014), 114 – 124. https://doi.org/10.1016/j.engappai.2013.09.019

[15] Kevin P Murphy. 2012. *Machine learning: a probabilistic perspective*. MIT press.

[16] Department of Primary Industries. accessed (September, 2018). Prawns - aquaculture prospects. https://www.dpi.nsw.gov.au/fishing/aquaculture/publications/species-saltwater/prawns.

[17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.

[18] Chris Robertson (Ed.). 2006. *Australian prawn farming manual: health management for profit*. Queensland Department of Primary Industries and Fisheries (QDPI&F).

[19] Stewart A Rounds. 2002. Development of a neural network model for dissolved oxygen in the Tualatin River, Oregon. In *Proceedings of the second federal interagency hydrologic modeling conference, Las Vegas, Nevada*. 1 – 13.

[20] Julius O. Smith. accessed (September 2018). *Mathematics of the Discrete Fourier Transform (DFT)*. `http://ccrma.stanford.edu/~jos/mdft/`. online book, 2007 edition.

[21] YSI. accessed (February, 2018). EXO2 Multiparameter Sonde. https://www.ysi.com/EXO2

[22] Huihui Yu, Yingyi Chen, ShahbazGul Hassan, and Daoliang Li. 2016. Dissolved oxygen content prediction in crab culture using a hybrid intelligent method. *Scientific reports* 6 (2016), 27292.