

Article

# Application of Regression Analysis to Achieve a Smart Monitoring System for Aquaculture

Wei-Chih Hsu <sup>1</sup>, Pao-Yuan Chao <sup>1</sup> , Chia-Sui Wang <sup>2,\*</sup>, Jen-Chieh Hsieh <sup>1</sup> and Wesley Huang <sup>2</sup>

<sup>1</sup> Department of Computer and Communication Engineering, National Kaohsiung University of Science and Technology, Kaohsiung City 80778, Taiwan; weichih@nkust.edu.tw (W.-C.H.); I107109103@nkust.edu.tw (P.-Y.C.); poan@nkust.edu.tw (J.-C.H.)

<sup>2</sup> Department of Information Management, Chia Nan University of Pharmacy and Science, Tainan 71710, Taiwan; wasener@mail.cnu.edu.tw

\* Correspondence: box1025@mail.cnu.edu.tw

Received: 24 June 2020; Accepted: 4 August 2020; Published: 7 August 2020



**Abstract:** The consumption awareness of people in recent years has increased, with food safety becoming more and more important. While non-toxic products can be achieved by avoiding using too much antibiotics to control growth factors in a water environment, the measurement tools for dissolved oxygen on the market are very expensive and a great economic burden to fishermen. Thus, the purpose of this study is to design more economical measurement modules and algorithms for monitoring ponds. The research collected pond data through Oxidation-Reduction Potential (ORP), pH and temperature sensors, used regression analysis to infer Dissolved Oxygen (DO) by ORP and pH, and employed a real-time pond monitoring data map to figure out pond conditions. Compared with traditional equipment, findings show our approach reduces costs by about 20%, and increases production capacity and output value.

**Keywords:** dissolved oxygen; regression analysis; smart monitoring; aquaculture

## 1. Introduction

A report by the Food and Agriculture Organization (FAO) of the United Nations estimated that total global seafood production is expected to increase 18% by 2030, reaching 2.01 million tons, with aquaculture as the main output [1]. Taiwan's geographical environment is surrounded by sea, with a coastline spanning over 1100 km. Innately, it has favorable conditions for the development of offshore and coastal fisheries. With such an excellent environment, its aquaculture fishery industry has great potential for development.

People have paid great attention to food safety issues alongside the improvement of quality of life. An investigation showed that 82% of respondents mistrusted current food safety, which highlights the important of this issue [2]. Nevertheless, the quality of water is influenced by economic activities, such as waste water disposal from factories and aquaculture industries; the discharged residue has a great influence on the aqua ecosystem. For the reason mentioned above, water quality monitoring is very much in demand for giving an excellent environment for aquaculture [3–7].

In the era of artificial intelligence, big data monitoring and the analysis of smart farming are now trends. The popularity of networked applications has made various online upload service systems more and more convenient. The combination of simple transmission methods and various sensing systems allows users to monitor the status of culture ponds and adjust the water quality environment in a timely manner. Factors affecting the external environment of the fish population can be judged and determined through the long-term monitoring of trend charts, so as to minimize any adverse

effects, improve the health of the fish, enhance the quality and output of the culture ponds, and reduce human resources and time costs.

Internet of things (IOT) technology is universally applied in several industries. Prabha et al. designed an aquaponics system and monitored the growth of fish as well as vegetables [8]. Purwandoko traced each parameter in the production process and established related knowledge and information with a rice traceability system [9]. Rajalashmi et al. established a system for monitoring and controlling water quality in a sewage treatment plant [10]. Abinaya et al. carried out a monitoring and control system for aquaculture [11]. In this study, an IOT-based smart monitoring system is carried out by analyzing water quality from the collected values of pH and ORP of water for aquaculture.

The ORP measures the potential difference between the gold or platinum electrode and the reference electrode. The survey instrument is in a pond fish, and discharged via the metabolites of protein in the form of organic nitrogen and ammonia. The water formed during nitrification comprises the effect and denitrification for measuring the ORP through redox chemical reactions [12].

Because the amount of DO influences the growth of soil microorganisms, DO represent an environmental growth indicator. When DO rises, rainbow trout fingerlings' weight gain rates, specific growth rates and feed conversion rates all increase, thus exhibiting a significant change in the growth index [13–18]. Based on previous research, this paper derives several issues, as follows.

- (1) DO measuring system: In general, the DO measuring system on the market costs NT\$ 600,000 on average. Their internal circuit diagram structure is too complicated to make products.
- (2) Restriction of transmission space: Farm fields cannot be covered with a full Wi-Fi environment. To set up a smart monitoring system, it becomes necessary to use a gateway and Bluetooth technology to receive each device's measured water quality data, which can then be uploaded to a cloud database via 4G or Wi-Fi.

The main purposes of this study are to use a low-cost hardware device to collect various environmental pond data, to transfer back the analysis of the estimated true values of DO, and to develop aquaculture ponds with intelligent environmental monitoring and control systems.

## 2. Methods

This paper used polynomial linear and non-linear regression analyses, where  $i$  is a sample of collected data, the pH value is set to  $X_{1i}$ , ORP is set to  $X_{2i}$ , and the amount of DO is set to  $Y_i$ . Analysis operations are conducted on the data.

### 2.1. Polynomial Regression Analysis

$$Y_i = \alpha + \beta_1 X_{1i} - \beta_2 X_{2i} \quad (1)$$

If there is no linear relationship between  $X_{1i}$  and  $X_{2i}$ , then the above equation can solve for  $\alpha$ ,  $\beta_1$ , and  $\beta_2$  as:

$$\alpha = \bar{Y} - \hat{\beta}_1 \bar{X}_1 - \hat{\beta}_2 \bar{X}_2 \quad (2)$$

$$\beta_1 = \frac{(\sum_{i=1}^n y_i x_{1i})(\sum_{i=1}^n x_{2i}^2) - (\sum_{i=1}^n y_i x_{2i})(\sum_{i=1}^n x_{1i} x_{2i})}{(\sum_{i=1}^n x_{1i}^2)(\sum_{i=1}^n x_{2i}^2) - (\sum_{i=1}^n x_{1i} x_{2i})^2} \quad (3)$$

$$\beta_2 = \frac{(\sum_{i=1}^n y_i x_{2i})(\sum_{i=1}^n x_{1i}^2) - (\sum_{i=1}^n y_i x_{1i})(\sum_{i=1}^n x_{1i} x_{2i})}{(\sum_{i=1}^n x_{1i}^2)(\sum_{i=1}^n x_{2i}^2) - (\sum_{i=1}^n x_{1i} x_{2i})^2} \quad (4)$$

The three variables are replaced by Equation (1) to get the regression model.

## 2.2. Non-Linear Polynomial Regression Analysis

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \beta_4 X_1^2 + \beta_5 X_2^2 \quad (5)$$

where  $Y$  is a return value (that is, DO quantity,  $X_1$  and  $X_2$ ) compared with two independent variables (namely, pH and ORP value) to make up Equation (5). The equation converts into a five-variable linear model set up according to the principle of the least square method, which is:

$$Y = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} \quad (6)$$

where,  $\alpha$ ,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$  and  $\beta_5$  are obtained from the above equation, and the regression model can be realized by substituting back into the original Equation (6).

## 3. System Design

### 3.1. System Architecture

In this study, an IOT-based system was designed and carried out, the architecture of which is shown in Figure 1, and which uses Raspberry PI as the gateway to receive two or more sensor module means, including temperature, pH meter and ORP. From the repeated judgement filtration of the received data, the filtered data are transferred into regression data analysis and uploaded to the database to be stored and checked for whether they exceed the recommended indicators. The information obtained goes back into regression analysis. Figure 2 shows the system push message activity diagram, including estimating the DO amounts by the pH values and the ORP values, sending data to GCM, and enabling the push broadcast service. When sending the push service, each mobile App device's REGID in the database can be captured and accurately sent to the device through this set of codes, and can successfully trigger a warning reminder, as shown in Figure 3.

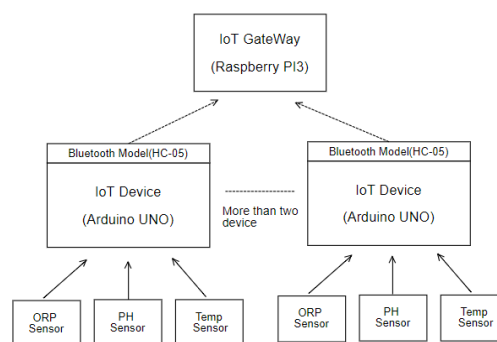


Figure 1. System architecture.

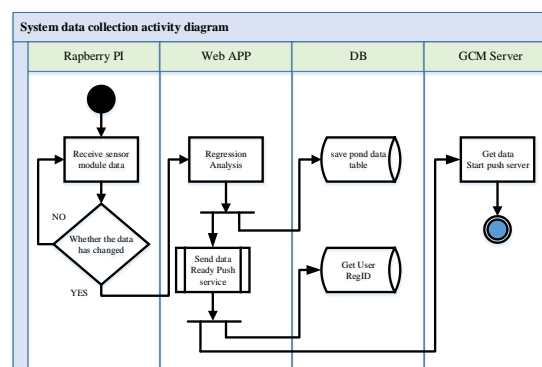
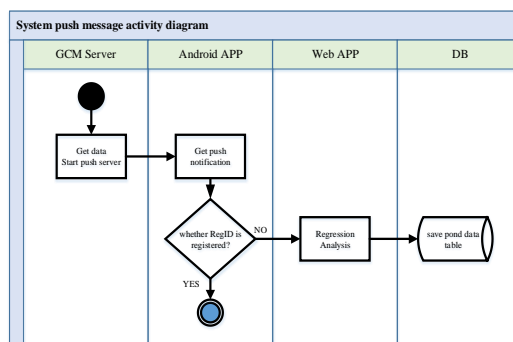


Figure 2. System data collection activity diagram.



**Figure 3.** System push message activity diagram.

### 3.2. Intelligent Monitoring Hardware for Aquaculture Ponds

Figure 4 shows the sensing module mounted on a styrofoam plate, which is located at the edge of the fish pond. The Arduino UNO microchip set and Bluetooth module are set nearby the pond. By Bluetooth communication, measured signals are transferred to Raspberry Pi, which is as an information gateway that receives the detected data from the pond. By this approach, multiple sensor devices set up over multiple fish ponds can be utilized under one integrated system. In this study, the total cost of equipment, including PH and ORP as well as supplies, is about USD 360. For individual sensors, the ORP sensor is USD 130, the PH sensor is USD 35 and Raspberry Pi is USD 70, separately. The offline version of the DO meter on the market costs an average of USD 1500.

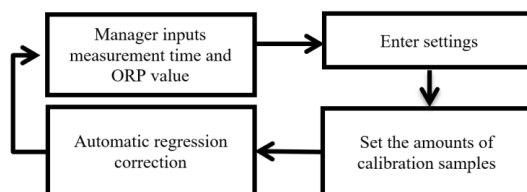


**Figure 4.** Intelligent monitoring hardware for aquaculture pond.

### 3.3. Sensor Correction Module

Sensors are put in the culture pond. Nevertheless, the accuracy of the sensors is affected greatly by the position on the sensor's surface of fish feed, as well as impurities in the water. After a period of time, the ORP and pH sensors will change the captured values, affecting the accuracy of DO. Therefore, a sensor calibration module is established.

Based on the problems mentioned above, this study developed a DO field as a management system for the active inputs of the manager. As shown in Figure 5, when management needs to use the automatic calibration system, the automatic calibration button is pressed to enter the set page. The amounts of samples are processed and calibrated, and the confirmation button is pressed to start automatic regression correction.



**Figure 5.** Auto-correction.

## 4. Results and Discussion

### 4.1. Dissolved Oxygen Prediction Analysis

Over-fitting in regression analysis produces inaccurate prediction, and comes down to insufficient training data and too many input parameters. In this study, the training data set is divided into two parts: the training and the validation data set. The new training data set is still used for the training model, and the new validation data set is used to observe the over-fitting trend.

Two ponds were targeted for the experiment data information, and 60 records were taken as a set. In the set, 50 records were used for training regression and 10 records were for validation. The study took a control approach for polynomial regression analysis and non-linear polynomial regression analysis.

Under the polynomial regression analysis method, 50 sets of training data, the predicted DO values and the real DO values were plotted as a line chart, as in Figure 6. The real values and the non-linear polynomial regression analysis prediction results are shown as Figure 7.

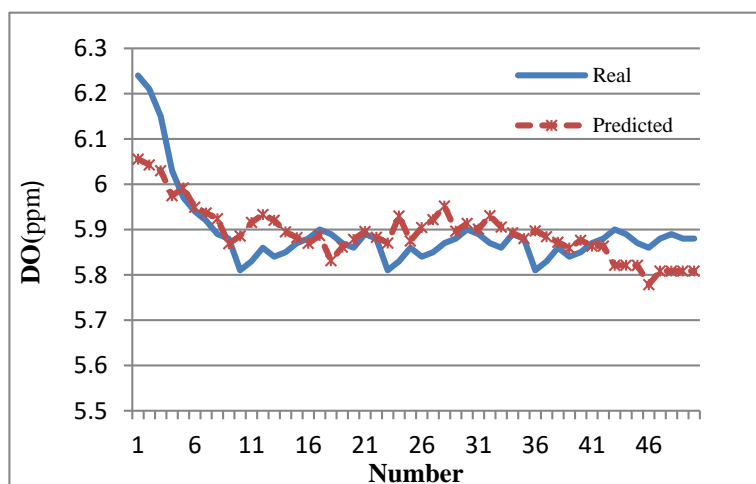


Figure 6. Polynomial regression analysis.

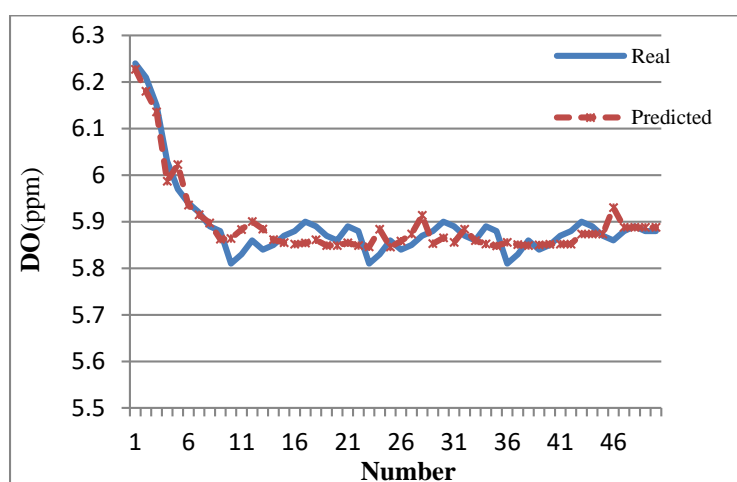


Figure 7. Non-linear polynomial regression analysis.

The prediction results shown in Figure 6; Figure 7 did not show large differences between the training data set and the validation data set. The predictions of both polynomial regression analysis and non-linear polynomial regression analysis did not indicate any over-fitting condition.

The study first calculates the sum of squared errors (SSE) for the two regression approaches, the takes the difference between real dissolved oxygen ( $DO_{real}$ ) and predicted dissolved oxygen ( $DO_{predict}$ ), and divides this by  $DO_{real}$ , so as to get the calculated variation rate manner, as follows:

$$SSE = (DO_{real} - DO_{predict})/DO_{real} * 100\% \quad (7)$$

Equation (7) allows one to estimate the SSE of the polynomial regression analysis of DO within a 0.82% error in Table 1, which complies with a reasonable range. Additionally, the SSE of the non-linear polynomial regression analysis predicts DO values within a 0.42% error. The result shows a lower error than the linear regression model.

**Table 1.** SSE of two modes.

Modes	SSE (%)
PRA	0.82
NLPRA	0.42

#### 4.2. Web Page Management Mode

In Figure 8, the manager enters the management page of the website to observe current aquaculture ponds and also to monitor the measured data of each pond. Adding, modifying or deleting other farmed fish tanks can be conducted. In fact, the real-time management of the detected information of fish ponds can be achieved more conveniently.

Current pond number

Pond	Log	Kind
Pool1	Record1	Striped bass
Pool2	Record1	Shrimp

• ADD Pond :

Select Kind:

**Figure 8.** Showing the current number of aquaculture ponds.

Given the different growth environments for each aquaculture pool, each aqua environment parameter should be set individually, such as the fish category, pool name, ORP limitation, pH limitation, temperature limitation and DO limitation. The user interface is shown in Figure 9.

Suggested indicators

ID	Kind	ORP (mV)	pH(mol/L)	temperature(°C)	DO(ppm)
1	Striped bass	100–200	6.0–7.0	18–25	6–8
2	Shrimp	100–200	6.0–7.0	18–25	6–8
3	Milkfish	100–200	6.0–7.0	18–25	6–8

• ADD Kind Water quality:

Kind name:

ORP Range (1–2000):  ~  (mV)

pH Range (1–14):  ~  (mol/L)

temperature Range (1–50):  ~  (°C)

DO Range (1–100):  ~  (ppm)

• Revise Kind Water quality:

Kind name:

• Delete Kind Water quality:

Kind name:

**Figure 9.** Management of aquaculture ponds.

The power durability of the remote control module should be studied closely. The saving of power can be achieved by a partial shutdown module. For instance, the large part of the system can

be shut down if there is no transformation data. Additionally, the entire system can be booted up if the transformation data are ready [19,20]. Therefore, we schematized an operation mode for the administrator which can set the remote control manually, including the maximum items stored in the repository, the refreshed APP cycle data and each parameter's flexible range, and filter repetitive information to avoid data duplication, as seen in Figure 10. This approach can improve the waste of power and upgrade the durability of the remote module.

Raspberry Pi Settings (Gateway):

- Config

Maximum number (number/page)	Cycle(s/n)	ORP Range (mV)	pH Range (mol/L)	temperature Range (°C)	DO Range (ppm)
200	5	1	1	1	1

Pond data table maximum number:  (number)

dataset update cycle:  (s)

Data is save in the database:

ORP  (mV) · pH  (mol/L)

temperature  (°C) · DO  (ppm)

OK

Figure 10. Database form settings.

In this study, the Raspberry Pi was connected to a 100 mAh rechargeable battery to measure the energy consumption of both the transformation period and the stop working period. When a detector transferred 100 data per second, the energy consumption was about 720 mA, and the module could keep working for 1.3 h. Meanwhile, when the transferring of data stopped, the energy consumption was about 350 mA, and the module could keep working for 4 h. For this result, the amount of transferring data can be set on web pages to carry out the purpose of power saving.

#### 4.3. APP Monitoring Mode

In this study, a mobile APP was designed and developed to monitor the water quality in an aquaculture pool. If the detected water quality is over the set limitation of the index parameters, the warning words in red are shown. When the mobile APP is opened up, the current condition information of the ponds will be displayed on the main page. If the values of the aquaculture pond exceed the recommended indicators, a notification will be shown on the main page to remind the manager, as shown in Figure 11a. The manager presses the menu of the aquaculture ponds interface. According to the currently paired Bluetooth device, the mobile APP will then dynamically display on the screen the amounts of aquaculture ponds, as well as the corresponding current parameters, as shown in Figure 11b.

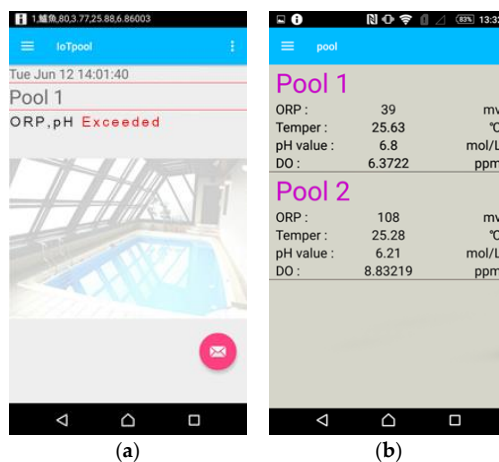


Figure 11. APP monitoring mode for (a) warning attention and (b) aquaculture pond status.

## 5. Conclusions

This study issues a low-cost monitoring system for aquaculture. Farmers can monitor the information of culture ponds, and historical data. The cost of the used equipment in this study is about USD 350, but the entire commercial system is about USD 2000 on the market place.

The amount of DO influences the growth of soil microorganisms, i.e., DO results indicate the quality of the aquatic environment. Because the variation trend of the DO value is related to similar pH and ORP values, DO increases with the increase in pH and ORP values. As mentioned above, we validated the feasibility of using pH and ORP values to evaluate the DO value via the regression approach. In this study, a linear polynomial regression approach and a non-linear regression approach were carried out to predict the value of DO. By comparing the SSE of approaches and the actual results, we determined that the non-linear regression approach performed more accuracy than the linear one.

This study was carried out through HTML in combination with PHP to establish the managers' user interface. The advantages of this foundation are that it makes it convenient for management to set up the parameters of the aquaculture pond and the Raspberry PI in an orderly manner. Additionally, the functions of the mobile APP are programmed through the Android operation system. The main purpose is to receive information about the aquaculture pond and to send notifications to users. The APP is responsible for monitoring the pond's environmental status, in a range of standard values, anytime and anywhere for the express purpose of utilizing human resources more efficiently.

**Author Contributions:** Conceptualization, W.-C.H., P.-Y.C. and W.H.; experiment, W.-C.H., P.-Y.C. and J.-C.H.; software, W.-C.H., P.-Y.C. and J.-C.H.; validation, C.-S.W.; writing—review and editing, C.-S.W. and W.H.; investigation, C.-S.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Lafont, M.; Dupont, S.; Cousin, P.; Vallauri, A.; Dupont, C. Back to the Future: IoT to Improve Aquaculture: Real-Time Monitoring and Algorithmic Prediction of Water Parameters for Aquaculture Needs. In Proceedings of the 2019 Global IoT Summit (GIoTS), Aarhus, Denmark, 17–21 June 2019; pp. 1–6.
2. Ma, Y.; Ding, W. Design of Intelligent Monitoring System for Aquaculture Water Dissolved Oxygen. In Proceedings of the 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 12–14 October 2018; pp. 414–418.
3. El-Mezayen, M.M.; Rueda-Roa, D.T.; Essa, M.A.; Muller-Karger, F.E.; Elghobashy, A.E. Water quality observations in the marine aquaculture complex of the Deeba Triangle, Lake Manzala, Egyptian Mediterranean coast. *Environ. Monit. Assess.* **2018**, *190*, 436. [[CrossRef](#)] [[PubMed](#)]
4. Kang, P.; Xu, S. The impact of mariculture on nutrient dynamics and identification of the nitrate sources in coastal waters. *Environ. Sci. Pollut. Res.* **2016**, *23*, 1300–1311. [[CrossRef](#)] [[PubMed](#)]
5. Ma, Z.; Li, H.; Ye, Z.; Wen, J.; Hu, Y.; Liu, Y. Application of modified water quality index (WQI) in the assessment of coastal water quality in main aquaculture areas of Dalian, China. *Mar. Pollut. Bull.* **2020**, *157*, 111285. [[CrossRef](#)] [[PubMed](#)]
6. Parmar, K.S.; Bhardwaj, R. Water quality index and fractal dimension analysis of water parameters. *Int. J. Environ. Sci. Technol.* **2013**, *10*, 151–164. [[CrossRef](#)]
7. Shareef, Z.; Reddy, S.R.N. Design and wireless sensor Network Analysis of Water Quality Monitoring System for Aquaculture. In Proceedings of the 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 27–29 March 2019; pp. 405–408.
8. Prabha, R.; Saranish, R.S.; Sowndharya, S.; Santhosh, A.C.; Varsha, R.; Sumathi, K. IoT Controlled Aquaponic System. In Proceedings of the 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Erode, India, 6–7 March 2020; pp. 376–379.
9. Purwandoko, P.B.; Seminar, K.B.; Mardjan, S.S. Development of a Smart Traceability System for the Rice Agroindustry Supply Chain in Indonesia. *Information* **2019**, *10*, 288. [[CrossRef](#)]

10. Rajalashmi, K.; Yugathian, N.; Monisha, S.; Jeevitha, N. IoT based water quality management system. *Mater. Today Proc.* **2020**. [\[CrossRef\]](#)
11. Abinaya, M.; Survenya, S.; Shalini, R.; Sharmila, P. Design and Implementation of Aquaculture Monitoring and Controlling System. In Proceedings of the 2019 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC), Tamil Nadu, India, 27–28 March 2019; pp. 232–235.
12. Lu, C.-H. Study on Nitrification and Denitrification of Wastewater Containing Organic Nitrogen. Master's Thesis, Tatung University, Taipei, Taiwan, 2001.
13. Dabrowski, J.J.; Rahman, A.; Pagendam, D.E.; George, A. 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, London, UK, 19–23 August 2018; pp. 177–185.
14. Ma, Z.; Song, X.; Wan, R.; Gao, L. A modified water quality index for intensive shrimp ponds of *Litopenaeus vannamei*. *Ecol. Indic.* **2013**, *24*, 287–293. [\[CrossRef\]](#)
15. Zhang, Y.; Wei, W. Finite-Time Extended State Observer-based PI Control for Dissolved Oxygen. In Proceedings of the 2019 19th International Conference on Control, Automation and Systems (ICCAS), Jeju, Korea, 15–18 October 2019; pp. 40–43.
16. Zhu, C.; Wang, R.; Tong, Q. Prediction Model of Dissolved Oxygen Based on SADE-RVM. In Proceedings of the 3rd International Conference on Robotics, Control and Automation, Chengdu, China, 11–13 August 2018; pp. 131–134.
17. Ho, B.-C.C. Development of a Real-Time Aeration Control Strategy for Low DO Level Biological Nutrient Removal (BNR) Process Control of a Continuous-Flow Sequencing Batch Reactor. Ph.D's Thesis, National Central University (NCU), Taoyuan City, Taiwan, 2003.
18. Dabrowski, J.J.; Rahman, A.; George, A. Prediction of Dissolved Oxygen from pH and Water Temperature in Aquaculture Prawn Ponds. In Proceedings of the Australasian joint conference on artificial intelligence-workshops, Wellington, New Zealand, 11–14 December 2018; pp. 2–6.
19. Ciuffoletti, A. Stateless IoT. *Information* **2020**, *11*, 85. [\[CrossRef\]](#)
20. Haruo, Y.; Yamamoto, H.; Arakawa, M.; Naka, I. Development and Evaluation of Environmental/Growth Observation Sensor Network System for Aquaponics. In Proceedings of the 2020 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 4–6 January 2020; pp. 1–6.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).