Big Data Analysis on Water Quality Condition in a White Shrimp Farming Environment

Wu-Chih Hu[1], Hsin-Te Wu[2], Jun-We Zhan[1], Jing-Mi Zhang[1]

[1] Department of Computer Science Information Engineering, National Penghu University of Science and Technology, Taiwan

[2] Department of Computer Science and Information Engineering, National I-lan University, Taiwan

wuchih.hu@gmail.com, hsinte@niu.edu.tw, asd8867521@gmail.com, o0925331806@gmail.com

Abstract

In order to prevent food shortage in the future, human kind must rely on aquafarming to compensate for shortage in marine resources. Our proposed scheme monitors water quality via the Internet of Things (IoT). Given that the survival rate of white shrimp is highly dependent on water quality, this study collects water quality-related data through the IoT sensors, including data on temperature, oxygen content, and more. The main goals of the proposed big data analysis include the following: (1) to analyze a variety of environmental factors of a culture pond and determine whether it is a suitable environment for culturing white shrimp, and (2) to analyze the correlation between a single environmental factor against other environmental factors. The above analysis should help aquafarmers examine whether a culture pond is suitable for culturing white shrimp; moreover, aquafarmers will also learn how to, when water quality deteriorates, adjust a single factor to improve the overall water quality. Experimental results indicate that the analysis performed can indeed effectively help us better understand the living environment of the white shrimp as well as how to adjust one single environmental factor in order to elevate the overall water quality. The results of this study will aid aquafarmers in obtaining a better grasp of the overall culture environment. With the water quality analysis and monitoring system to initiate relevant equipment, the livability of white shrimp has reached 37%, which is higher than that of general breeding approaches. The approach developed in this article can effectively reduce the waste of water resources and enhance the livability of white shrimp.

Keywords: Big data, White shrimp, Aquaculture, IoT

1 Introduction

Changes in fishing industry resources has long been determined by climate and its fickleness due to the fact that changes in climate directly or indirectly result in changes in the oceanic environment. Currently, the aquafarming industry is facing several problems: (1) ageing population in aquafarming labor force; (2) the fact that most aquafarmers rely on experience and rules of thumb, leaving no record or teaching method behind; (3) drastic changes in water quality and temperature caused by climate change that in turn lead to low survival rate of white shrimp; and (4) the difficulty brought by demand for heavy manpower to care for water quality and farming of a high-priced aquatic product like the white shrimp. Islam et al. [1] mentioned that, despite the efforts made by many developing countries towards actively promoting fish farming, many areas are faced with declining productivity of aquatic products due to fish disease; hence, they are in need of better approaches to monitor water quality and analyze disease symptoms. Piplani et al. [2] discussed how, in India, more than 14.5 million fish farmers make their living through aquafarming; between 2014 and 2015, inland fisheries increased by 7.9% and saw 5.5 billion US dollars input in foreign exchange. However, aquafarming calls for extensive care on factors such as the monitoring of weather, water quality, and fish feed; therefore, in order to achieve and make use of the best water quality environment, it becomes even more important to conduct big data analysis on environmental factors. Han et al. [3] raised discussion on the topic of gradually declining natural resources. China has long remained highly dependent on aquatic products as an essential source of protein. The Han's research [3] investigates disease distribution of essential fish in China. Given the above, we can see how it is crucial in aquafarming to monitor water quality and select the suitable fish kind for culturing based on the water quality condition of a specific environment; it follows naturally that fish farmers must have a firm grasp of the environmental factors of their culture pond. Meanwhile, they must also be capable to utilize adjustments of a single factor to elevate water quality whenever there is change in environmental factors.

*Corresponding Author: Hsin-Te Wu; E-mail: hsinte@niu.edu.tw
DOI: 10.53106/160792642021122207010

Lv et al. [4] examined how to access data from the IoT and store it. The study also takes a close look into how IoT data can satisfy the "5V" requirements of bid data – volume, velocity, variety, value, and veracity. The study also discusses confidentiality and privacy protection ; meanwhile, their study can benefit from its discussion on how to integrate IoT data and big data analysis to achieve the 5V when establishing our own scheme. On a different note, Xu et al. [5] employed hierarchical k-means for fast-speed processing of big data clustering while Jacobs and Bean [6] opted for spectral ensemble clustering and matrix computations to reduce computational complexity, which is hugely advantageous towards big data analysis. Ke et al. [7] focused on the tools needed for real-time analysis in big data analysis systems; factors that must be taken into consideration include expandability, flexibility of adjustment, and programming.

In order to enhance the production rate of culturing white shrimp as well as help aquafarmers better deal with water quality, our study aims at investigating two target questions using big data analysis: (1) whether a culture environment is suitable for farming white shrimp, and (2) the impact of a single environmental factor on other environmental factors. Prior to running big data analysis, data cleaning must be completed; on this note, our study adopts Teager energy operator (TEO) [8] and adaptive threshold [9] for inaccurate data cleaning. Since environmental factors increase or decrease gradually, drastic change in signals indicate a problematic issue such as network transmission error; given the above, TEO can highlight an issue like drastic signal change while adaptive threshold removes signals with too drastic a change. Our study utilizes Gaussian distribution and fuzzy sets for environmental factor analysis of white shrimp. Our study divides environmental factors into five scales using a Likert scale [10]. If an environmental factor falls within the scale of 1, it is indication that the environment in question is not appropriate for culturing white shrimp, and aquafarmers will be advised to select a different fish kind for culturing. Moreover, our study adopts a multiple linear regression approach to analyze the impact a single environmental factor casts on other environmental factors; meanwhile, the analysis results help aquafarmers obtain better grasp of a single environmental factor to elevate the overall water quality. Our study creates an intelligent aquafarming platform that monitors environmental factors of a culture pond; the webpage platform also provides statement analysis. Experimental results show that our experiment environment suits white shrimp farming; additionally, the analysis results of the impact of a single environmental factor towards other factors can help aquafarmers learn how to use a single environmental factor to improve the overall environment. Our research facilitates production rate and water quality control, which increases farming profits and reduces labor costs. The suggested approach has proven the livability can be up to 37%.

2 Related Works

Fabregas et al. [11] proposed screening for shrimp diseases using an artificial intelligence-based approach. Shrimp are among the high-priced aquatic products, and yet white spot syndrome virus (WSSV) can affect shrimp production. Dabrowski et al. [12] established a water quality monitoring system that utilizes Bayesian wave filters to conduct water quality prediction, which informs of changes in water quality and helps prevent further deterioration in water quality. The error rate of this approach is roughly 11%. Meanwhile, Caparida et al. [13] offered real-time water quality monitoring that employs fuzzy logic to perform water quality evaluation and takes a step further to control surrounding equipment so as to optimize water quality condition. In Konovalov et al. [14], underwater recognition is used for measuring fish size; the study runs positioning of all areas in a culture pond and then uses that information to calculate distance and measure the size of fish. Wang et al. [15] suggested employing robotic fish to obtain visual data on water quality and fish energy, which facilitates in-depth exploration of water quality. On a different note, Jayanthi et al. [16] mentioned how India is seeing vast increase in aquafarming, especially with shrimp farming, but white spot syndrome virus (WSSV) is causing decline in shrimp production.

Bharill et al. [17] focused on an Apache Spark system that features a clustering algorithm based on fuzzy logic. This proposed approach runs faster than other algorithms. Liu et al. [18] proposed using parallel fuzzy c-means for image recognition, which runs faster than the traditional fuzzy c-means approach. The proposed system in Liu et al. [18] is implemented on an Apache Spark system, enabling the function of parallel fuzzy c-mean processing. Meanwhile, Wu et al. [19] recommend the application of fuzzy algorithm towards big data for the purpose of clustering. Fuzzy algorithm can not only be effectively applied to different data analysis but also works with an Apache Spark system. Segatori et al. [20] suggested applying fuzzy decision trees towards big data analysis, which should facilitate the clustering of same-type data while allowing for expansion of the tree based on clustering needs. Małysiak-Mrozek et al. [21] examined the benefits adding fuzzy algorithm to big data analysis. Fuzzy algorithm offers flexibility of adjusting data range, which facilitates big data computational applications.

Hu [22] designed an IoT-based monitoring aquafarming alarm system. Most aquafarming systems monitor only a single factor; in order to monitor dissolved oxygen content, Hu the proposed scheme must obtain sensor data on soil temperature, moisture, atmospheric pressure, and CO2 concentration; moreover, it must detect these data every 3 minutes. The IoT detects multiple environmental factors to run analysis. Chen et al. [23] established a fish farm with automatic monitoring system. The acquired data is transmitted back to the remote server via ZigBee wireless sensors. The system monitors environmental factors of the culture pond while remaining real-time grasp and control of said factors, including temperature, oxygen content, pH level, and water level, all through the sensor modules and monitoring system. Rajeswari et al. [24] discussed aquafarming environment monitoring, including the monitoring of dissolved oxygen content in the water, temperature, and power consumption circumstances. When the main power source is cut off, the system activates a solar power battery, which prevents fish kills caused by lack of oxygen. In the method proposed by Vatrapu et al. [25], it is argued that traditional databases are incapable of storing large amounts of IoT data, which brings about the need for setting up cloud services for data storage and big data analysis. Chiang et al. [26] has proposed the method of water salinity difference detector, primarily controlling the real-time changes in the fish pool to avoid excessive salinity in the water, which causes the death of the fish. The Li et al. [27] suggests a way to predict oxygen levels because this factor is critical to culture ponds; accurately predicting the oxygen levels can increase the livability of the breeding species.

3 The Proposed Scheme

3.1 Subsection

Figure 1 is an illustration of the intelligent aquafarming system proposed in this study. We have implemented intelligent aquafarming sensors by the culture ponds; the sensors are for salinity, temperature, oxygen content, pH level, and oxidation-reduction potential (ORP). To reduce the wasting of power, the sensors are only activated every 5 minutes. The proposed system transmits data via 4G networks, which ensures network quality and reduces packet loss. We have also established an IoT XMPP platform that stores data in the database via XML approach. The platform can present data in the form of charts and statements; it can even unify the format of IoT data and export it for normalization. This study proposes applying big data analysis to conduct further analysis of the IoT data and, from there, obtain understanding of the environmental conditions of the white shrimp culture ponds as well as the relationship between different environmental factors.

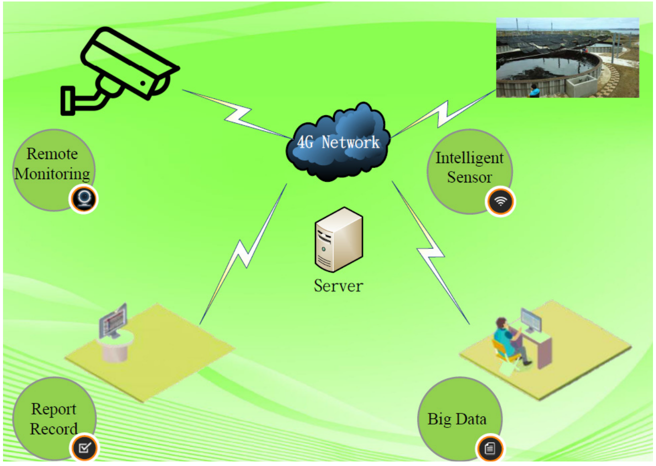

Figure 1. System illustration

3.2 Teager Energy Operator

First of all, prior to running data analysis, the data must be cleaned and normalized. Since environmental factors of culture ponds increase or decrease only gradually, any data that exhibits harsher fluctuation must be filtered. This study applies TEO for calculation, excluding any data with that stands out for its fluctuation. As shown in Figure 2, the TEO computation is as follows:

$$T\left[D_s(i)\right] = D_s(i)^2 - D_s(i+1) * D_s(i-1) \qquad \textbf{(1)}$$

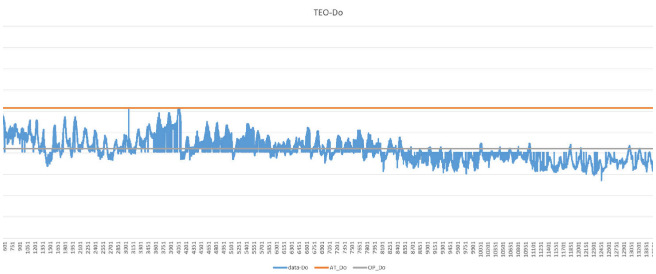

Figure 2. TEO curve diagram

Here, $D_s(i)$ stands for the original data, s represents different sensors, i stands for sensor data i=1~n, and $T\left[D_s(i)\right]$ is the result after TEO computation.

3.3 Adaptive Threshold Computation

When the system computes and obtains the TEO value, it will then compute the adaptive threshold (AT). If TEO is higher than AT, it indicates that the value is seeing great fluctuation; hence, the corresponding original data will be excluded. The computation is as follows:

Step 1: For the initial value, let k=1, $T_s(i)^k = T\left[D_s(i)\right]$.

Step 2: Define $T_s(n)^{k+1}$ as follows:

$$T_s(i)^{k+1} = \begin{Bmatrix} T_s(i)^k, if \ \ T_s(i)^k < E\left[T_s(n)^k\right] \\ E\left[T_s(i)^k\right], otherwise \end{Bmatrix} \quad \textbf{(2)}$$

In which $E\left[T_s(n)^k\right]$ stands for the average of $T_s(i)^k$.

Step 3: Repeat Step 2 until $E\left[T_s(n)^{k+1}\right] = E\left[T_s(n)^k\right]$, then stop the computation. $E\left[T_s(n)^k\right]$ is AT.

Step 4: Use AT to identify values with greater fluctuation, and then proceed to exclude them. The computation is as follows:

$$D_s(i) = \begin{Bmatrix} D_s(i), if \ \ T\left[D_s(i)\right] < AT \\ none, otherwise \end{Bmatrix} \quad \textbf{(3)}$$

As shown in Figure 2, the red line represents the AT value.

3.4 Evaluating the Environmental Factors for White Shrimp

This study aims at analyzing whether the current farming environment is suitable for culturing white shrimp. This study applies multiple factors to compute whether at a certain given moment the environment is suitable for culturing white shrimp. The key environmental factors to watch include salinity, temperature, oxygen content, pH level, and ORP level. When it comes to these five environmental factors, each organism has its own range of adaption. Suppose the upper limit threshold is $Thr_{u,i}$ and set the lower limit to $Thr_{l,i}$. This study employs normal distribution to determine the probability value for each environmental factor to fall between $Thr_{u,i}$ and $Thr_{l,i}$. The equation is as follows:

$$f\left(D(i);u,\sigma\right) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{\left(D_s(i)-u\right)^2}{2\sigma^2}\right) \quad \textbf{(4)}$$

σ stands for standard deviation, u stands for the mean value, and σ^2 stands for the variance. Following the above, environmental factors at all times are input to a multiple factor fuzzy logic. This study employs 5 scales for environment evaluation; the probability of 100% is divided equally into 5 sets. However, taking environmental factors and dividing them into clear-cut sets will lead to problematic results such as "bad" environment or indeterminacy of adaptiveness. For instance, if the probability falls at 20%, then it is scale 1; nevertheless, the system will deem the environment to be not adaptive. Therefore, this study employs fuzzy logic to help with determination while using Π-type membership function. Each scale is defined as shown in Figure 3, with $P_s(i) = f\left(D_s(i);u,\sigma\right)$. Next, the system computes data of each sensor and identifies which scale cluster the probability of that sensor will fall into. The equation is as follows:

$$EF_s(i) = \begin{Bmatrix} 5, P_s(i) >= PE_5 \\ j, P_s(i) < PE_j \text{ and } P_s(i) > PE_{j-1} \\ 1, P_s(i) <= PE_1 \end{Bmatrix} \quad \textbf{(5)}$$

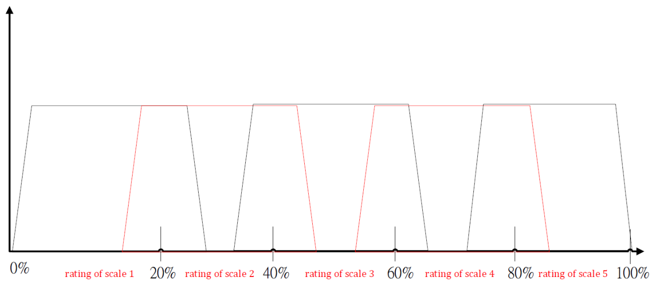

Figure 3. Determining the 5 scales of environmental factors

Following the above, the system computes the rating of all environmental factors; if any environmental factor rates at scale 1, it is indication that that particular environmental factor is not suitable for culturing white shrimp and that environmental factor is given a rating of scale 1. The algorithm is as follows:

Algorithm 1. Determine whether the body temperature and heart rate falls within the normal range.

Algorithm for overall environmental rating
if EF_i >= Bisection then
$EP = E\ [EF_i]$
else
EP =rating of scale 1
end if

EP represents the result of overall environmental factor rating while $E\left[EF_i\right]$ is the mean value of EF_i.

3.5 Multiple Linear Regression Analysis

This study first runs single linear regression to determine how the environmental factors affect each other; for instance, fish feed casts impact on the ORP value. The model for computing single linear regression is as follows:

$$D_s(n) = a + bD_{s-1}(n) \quad \textbf{(6)}$$

Next, the system finds the value of a and b with the following equation:

$$S_{x,x} = \sum\left(D_{s-1}(n) - \bar{D}_{s-1}(n)\right)^2 \quad (7)$$

$$S_{y,y} = \sum\left(D_s(n) - \bar{D}_s(n)\right)^2 \quad (8)$$

$$S_{x,y} = \sum\left(D_{s-1}(n) - \bar{D}_{s-1}(n)\right)\left(D_s(n) - \bar{D}_s(n)\right) \quad (9)$$

$\bar{D}_{s-1}(n)$ stands for the mean value of $D_{s-1}(n)$ while $\bar{D}_s(n)$ stands for the mean value of $D_s(n)$. Then, the system identifies the values of a and b using the following algorithm:

$$a = \bar{D}_{s,t}(n) - b\bar{D}_{s-1}(n) \quad (10)$$

$$b = \frac{S_{x,y}}{S_{x,x}} \quad (11)$$

When the above is completed, the system proceeds to identify the correlation coefficient *corr*. When *corr*>0, it is considered positive linear correlation; when *corr*<0, it is considered negative linear correlation. The equation is as follows:

$$corr = \frac{S_{x,y}}{\sqrt{S_{x,x} \cdot S_{y,y}}} \quad (12)$$

The next step is to run multiple linear regression analysis to compute the correlation between multiple environmental factors. The equation is as follows:

$$D_s(n) = b^T D_{s-1}(n) \quad (13)$$

Finally, the system computes the optimal solution for b^T and identifies the correlation coefficient. When *corr*>0, it is considered positive linear correlation; when *corr*<0, it is considered negative linear correlation. This linear regression algorithm can not only help with analysis of correlation between different environmental factors but also help us understand environmental changes caused by fish feed, which can serve useful for aquafarmers to control how much fish feed they should cast in order to maintain an adaptive water quality for the whiteleg shrimp.

4 Experiment Results

4.1 Intelligent Aquafarming Platform

The intelligent aquafarming platform created in this study can be divided into two parts: (1) the IoT monitoring system, and (2) the intelligent aquafarming webpage platform. Figure 4 is an illustration of the IoT monitoring system; the sensors are for salinity, temperature, oxygen content, pH level, and ORP level. Our proposed IoT system is power efficient; it only activates every 5 minutes to run sensor detection. The IoT monitoring system utilizes 4G network for packet transmission and the sensor-detected data is sent back to the intelligent aquafarming platform. As shown in Figure 5, the intelligent aquafarming platform adopts XML format for data transmission; meanwhile, the intelligent aquafarming platform features chart and statement presentation with the function of outputting data. The proposed system provides charts and statements that offer a curve diagram view of environmental factor trends.

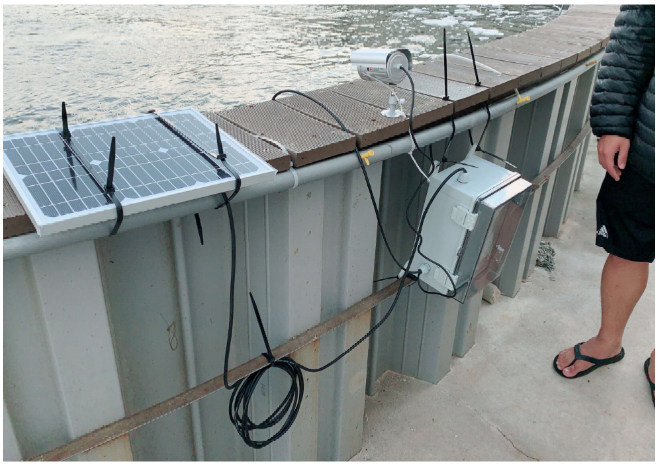

Figure 4. IoT monitoring system

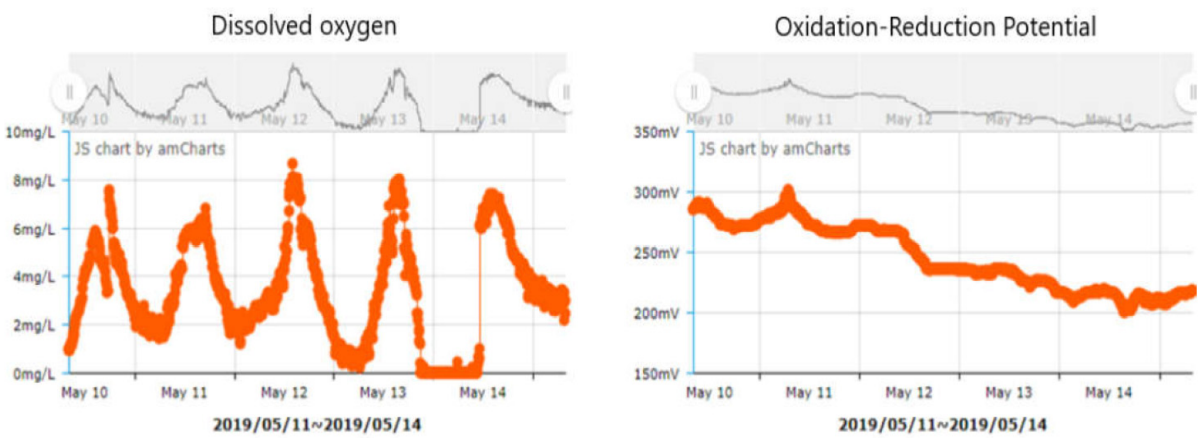

Figure 5. Intelligent aquafarming platform

4.2 Big Data Experimental Results

This study includes over 15,000 entries of data on white shrimp; for determination of white shrimp environmental factors, the study analyzed data on overall environmental conditions dating from December 2018 to March 2019, as shown in Figure 6 to Figure 9. Table 1 demonstrates the five scales to indicate the water quality. Experimental results show that the overall water quality dropped to below 0.2 in December 2018, which classified it as a scale 1 environment that warranted subsequent procedures such as water relief and change of water. Afterwards, the water quality returned to the standard above scale 2.

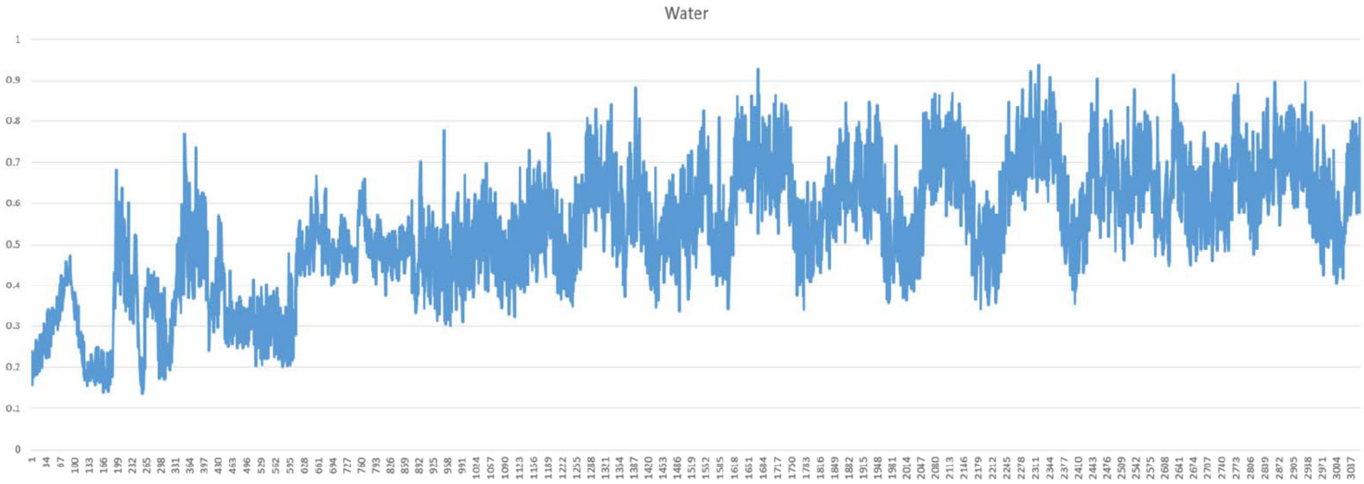

Figure 6. The mean value of overall environmental factors in December 2018

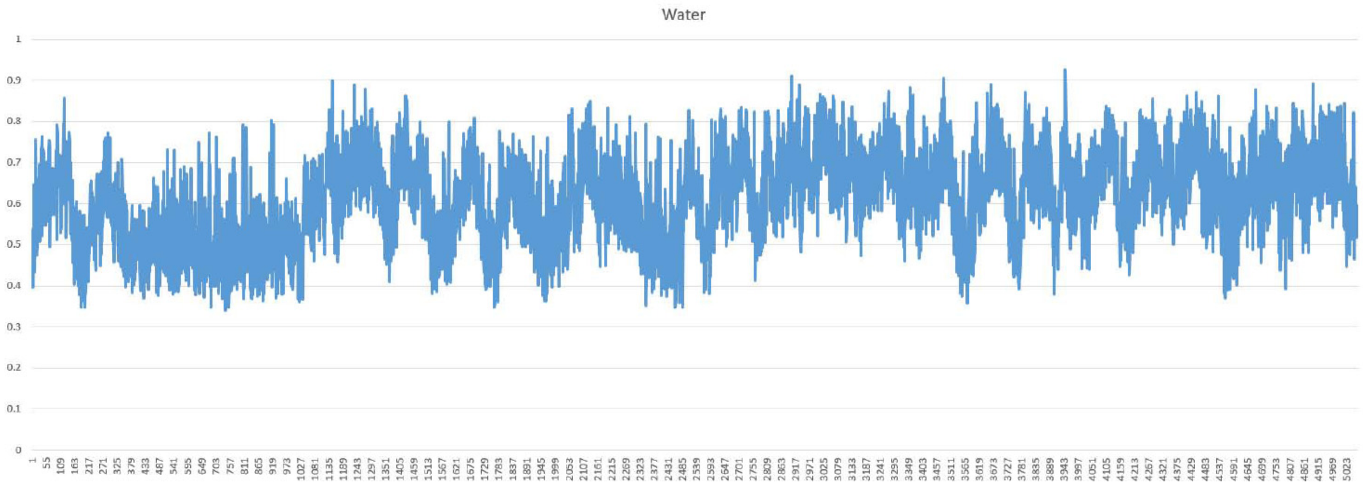

Figure 7. The mean value of overall environmental factors in January 2019

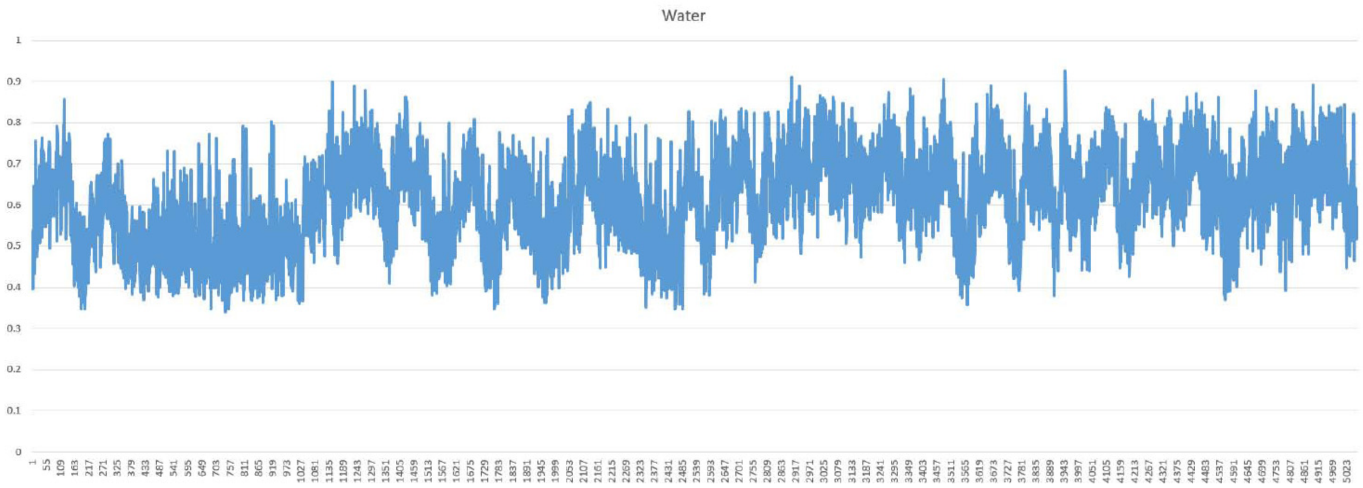

Figure 8. The mean value of overall environmental factors in February 2019

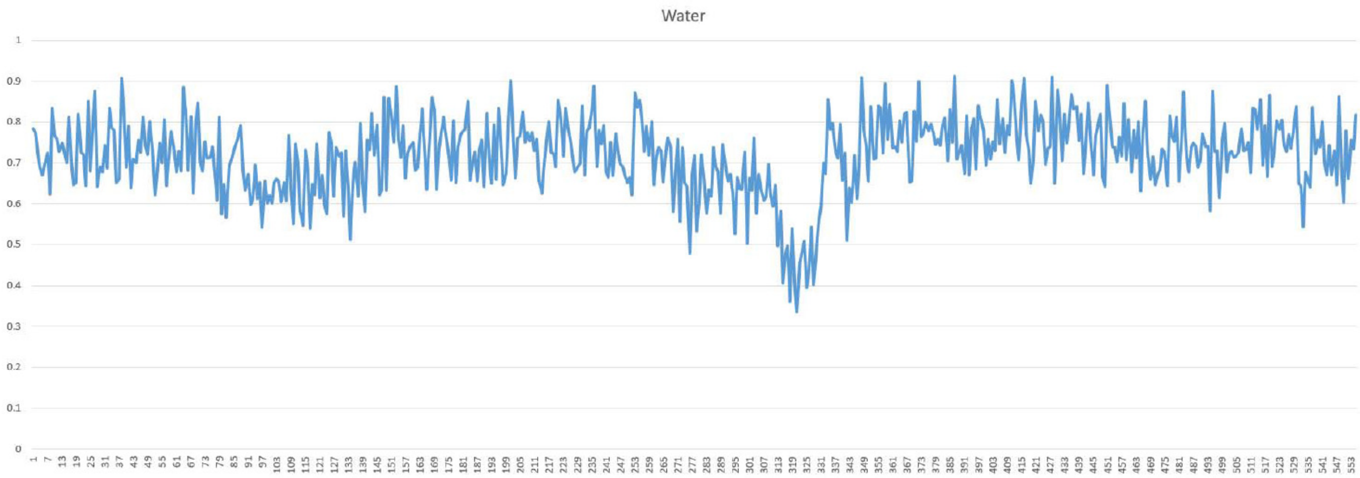

Figure 9. The mean value of overall environmental factors in March 2019

Table 1. The definitions of the five water-quality Scale

Five-scale Ranges	Definition
>0.9	Excellent
0.7~0.9	Good
0.5~0.7	Neutral
0.25~0.5	Bad
<0.25	Awful

This study utilized big data analysis to effectively obtain knowledge on the overall conditions of an aquafarming environment. Next, this study used multiple linear regression analysis to examine the impact of a single environmental factor on other environmental factors. Figure 10 to Figure 14 illustrate experimental results on the impact of a single environmental factor on other environmental factors. Figure 10 shows the influence of dissolved oxygen on other environmental factors. It can be seen from the figure how dissolved oxygen bears negative correlation to all other environmental factors. Meanwhile, Figure 11 exhibits how ORP affects other environmental factors. The figure illustrates ORP casting positive correlation to all other environmental factors with the exception of salinity. In Figure 12, we can see the impact of pH level on other environmental factors. As seen in the figure pH level exhibits positive correlation to ORP but little correlation to other factors. What Figure 13 demonstrates is the effect of salinity on other environmental factors; salinity only exhibits positive, low correlation towards ORP and dissolved oxygen while remaining low correlation to all other factors. Figure 14, on the other hand, demonstrates the effect of temperature on other environmental factors. Temperature shows positive correlation only to pH level and salinity; it bears low correlation to all other factors.

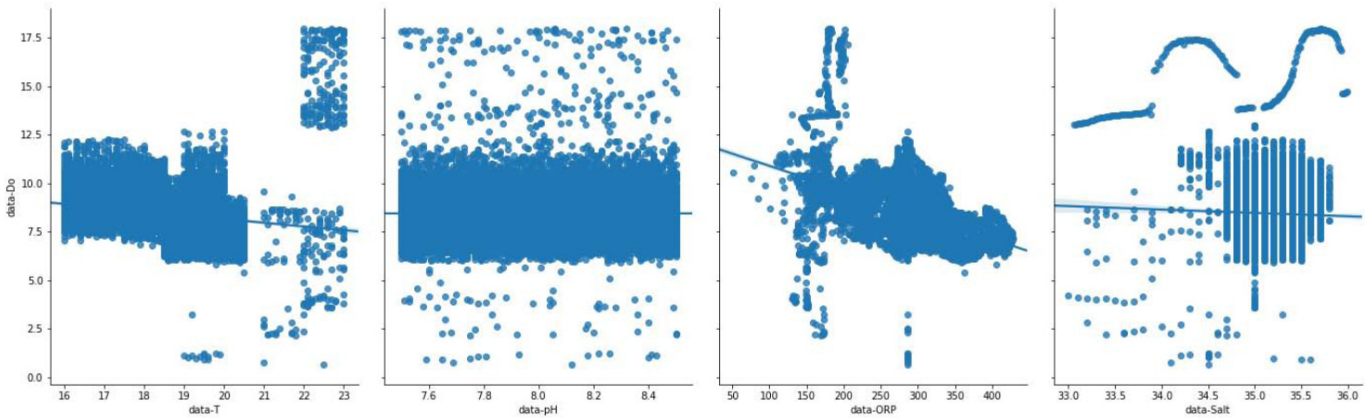

Figure 10. The effect of dissolved oxygen on other environmental factors

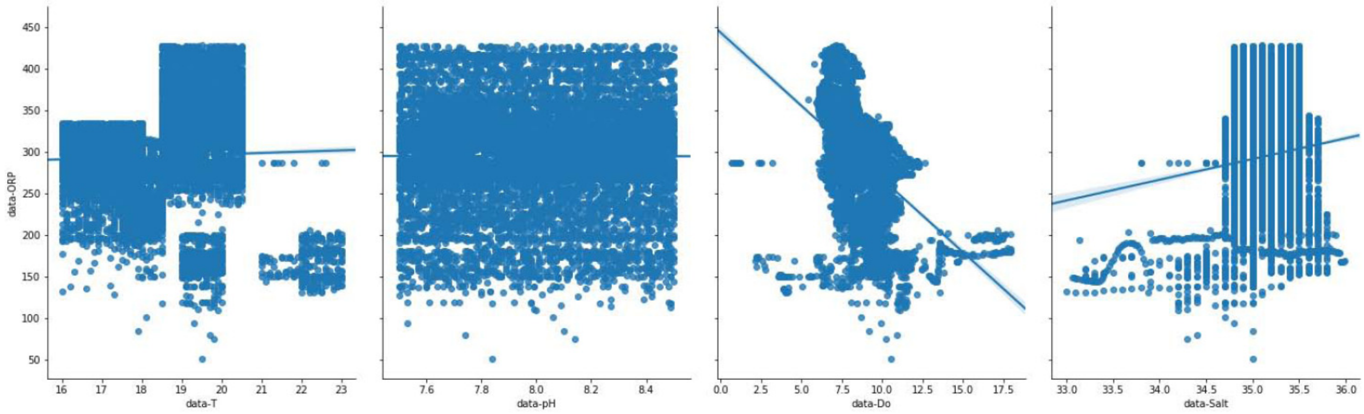

Figure 11. The effect of ORP on other environmental factors

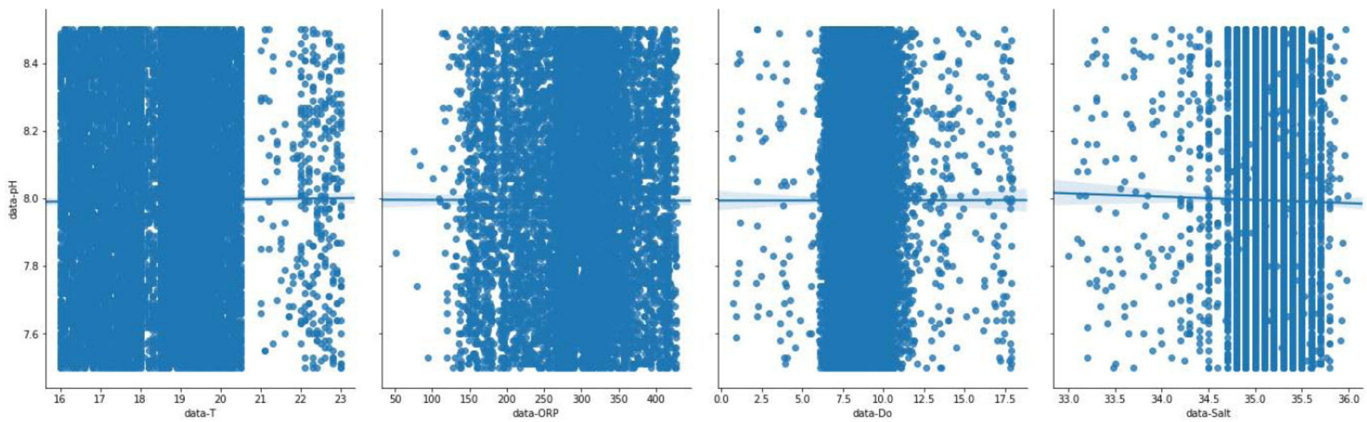

Figure 12. The effect of pH level on other environmental factors

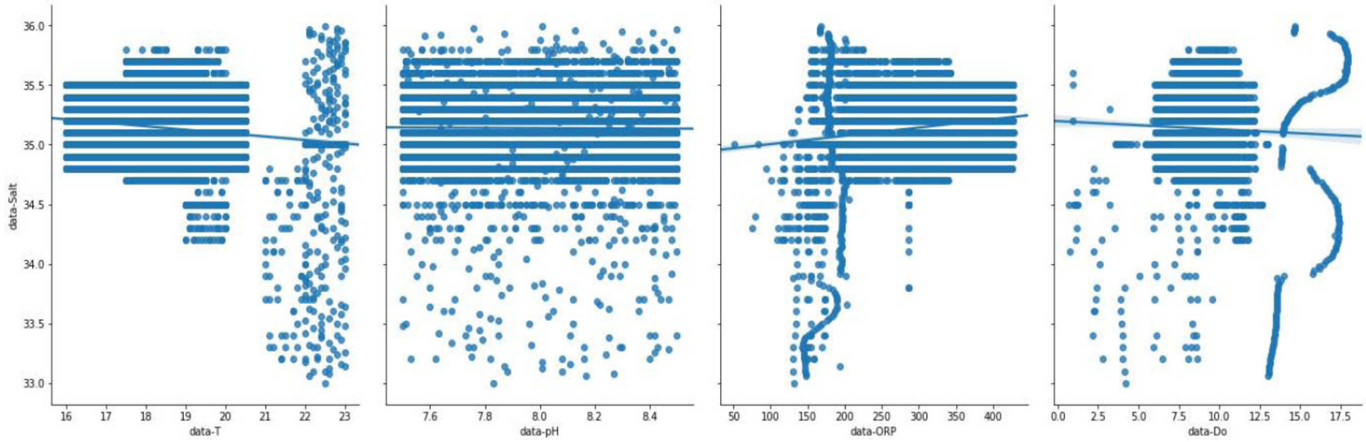

Figure 13. The effect of salinity on other environmental factors

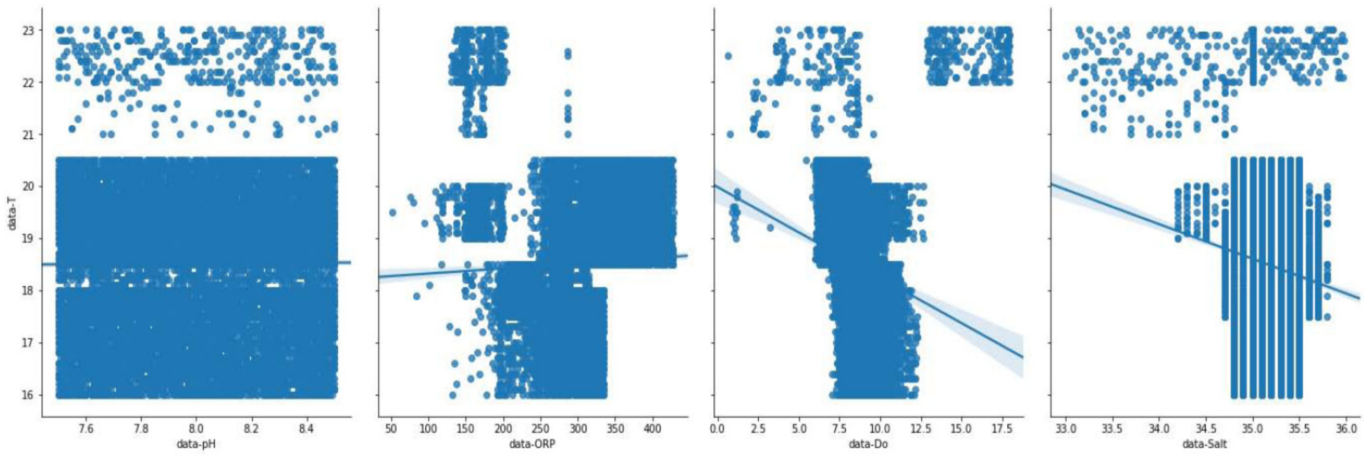

Figure 14. The effect of temperature on other environmental factors

Table 2 shows the impact of each and every environmental factor on other environmental factors. We can see from the chart that ORP bears positive influence on all factors except for salinity; hence, we can conclude and confirm the importance of ORP to aquafarming. However, for traditional aquafarmers, in order to monitor ORP values, they must purchase water quality sensors and invest manpower for regular measurement-taking, which lead to increase in labor costs. With this in mind, this study designed the intelligent aquafarming platform so that it can automatically detect and monitor ORP values, effectively reducing labor costs while safeguarding the overall environmental condition of a culture pond.

Table 2. The effect of each environmental factor on other environmental factors.

Correlation	T	Salt	PH	DO	ORP
T	None	Modestly correlated	Moderately correlated	Modestly correlated	Modestly correlated
Salt	Modestly correlated	None	Modestly correlated	Modestly correlated	Modestly correlated
PH	Modestly correlated	Modestly correlated	None	Modestly correlated	Highly correlated
DO	Modestly correlated	Modestly correlated	Modestly correlated	None	Modestly correlated
ORP	Highly correlated	Modestly correlated	Highly correlated	Highly correlated	None

This study utilized multiple linear regression analysis to investigate the effect of each environmental factor towards other factors. When environmental factors are reduced, this allows aquafarmers to adjust a single environmental factor to the effect of improving the overall water quality. For instance, since ORP affects other factors, the aquafarmer can improve water quality by changing the water. Our proposed big data analysis can not only help aquafarmers maintain a firm grasp of whether a certain culture environment is suitable for farming white shrimp but also enable them to maintain a desirable living environment for white shrimp by adjusting a single factor.

This research conducts the white shrimp experiment in the school's culture pond. The pond capacity is 200 tons. There are fifty thousand white shrimp larvae initially; four months later, the numbers of harvested white shrimp are 18,500, about 280 kilograms, and the average weight is 15 grams with an average length of 14 centimeters. The overall breeding rate is about 37%. Figure 15 demonstrates the white shrimp larvae, and Figure 16 illustrates the length of the shrimp. Consequently, the research presents that the white shrimp livability has reached 37%, which is higher than that of other general breeding approaches.

Figure 15. The length of the white shrimp

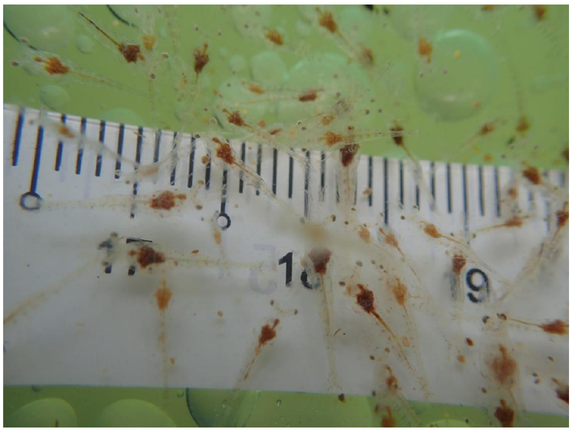

Figure 16. The white shrimp larvae

5 Conclusion

The proposed scheme makes use of TEO and adaptive threshold to filter out inaccurate information; meanwhile, it applies Gaussian distribution and fuzzy algorithm to compute and determine environmental factors in white shrimp farming. Experimental results indicate that our proposed method can successfully monitor water quality conditions of the living environment of white shrimp; the results also confirm that our marine culture pond is suitable for farming white shrimp. Moreover, our research on the effect of a single environmental factor on other environmental factors can help aquafarmers learn about environmental factors through water quality monitoring. The experimental results can serve to help aquafarmers improve the overall water quality condition by adjusting a single environmental factor, in turn elevating the production rate and quality of their aquatic products. In the future, our proposed scheme can incorporate deep learning so that when the water quality condition approaches scale 2, the system is capable of determining which environmental factor is exhibiting abnormality before adjusting that single environmental factor to improve the overall water quality, eventually realizing the goal of intelligent automatic farming.

Acknowledgments

This paper was supported by the National Science Council, Taiwan, under grant no. MOST107-2221-E-346-007-MY2 and MOST 110-2622-E-197-015. The authors also wish to express their appreciation to Prof. Jhy-Yun Shy for his help with the experiments of culturing white shrimp.

References

[1] M. M. Islam, A. Barman, G. K. Kundu, M. A. Kabir, B. Paul, Vulnerability of inland and coastal aquaculture to climate change: Evidence from a developing country, *Aquaculture and Fisheries*, Vol. 4, No. 5, pp. 183-189, September, 2019.

[2] D. Piplani, D. K. Singh, K. Srinivasan, N. Ramesh, A. Kumar, V. Kumar, Digital platform for data driven aquaculture farm management, *The 7th International Conference on HCI*, Guwahati, India, 2015, pp. 95-101.

[3] S. Han, Y. Wang, S. Nie, Research on fish diseases along coastal waters in china based on web gis database, *2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, Lyon, France, 2011, pp. 94-97.

[4] Z. Lv, H. Song, P. Basanta-Val, A. Steed, M. Jo, Next-generation big data analytics: State of the art, challenges, and future research topics, *IEEE Transactions on Industrial Informatics*, Vol. 13, No. 4, pp. 1891-1899, August, 2017.

[5] T.-S. Xu, H.-D. Chiang, G.-Y. Liu, C.-W. Tan, Hierarchical k-means method for clustering large-scale advanced metering infrastructure data, *IEEE Transactions on Power Delivery*, Vol. 32, No. 2, pp. 609-616, April, 2017.

[6] I. S. Jacobs, C. P. Bean, Fine particles, thin films and exchange anisotropy, *Research Information Section*, 1963.

[7] H. Ke, P. Li, S. Guo, M. Guo, On traffic-aware partition and aggregation in mapreduce for big data applications, *IEEE Transactions on Parallel and Distributed Systems*, Vol. 27, No. 3, pp. 818-828, March, 2016.

[8] J. Kaiser, On teager's energy algorithm and its generalization to continuous signals, *The 4th IEEE Digital Signal Processing Workshop*, Mohonk, 1990.

[9] S.-H. Chen, H.-T. Wu, C.-H. Chen, J.-C. Ruan, T. K. Truong, Robust voice activity detection algorithm based on the perceptual wavelet packet transform, *2005 International Symposium on Intelligent Signal Processing and Communication Systems*, Hong Kong, China, 2005, pp. 45-48.

[10] R. Likert, A technique for the measurement of attitudes, *Archives of Psychology*, 1932.

[11] A. C. Fabregas, D. Cruz, M. D. Marmeto, Sugpo: A white spot disease detection in shrimps using hybrid neural networks with fuzzy logic algorithm, *2018 International Conference on Information Technology*, Hong Kong, China, 2018, pp. 199-203.

[12] J. J. Dabrowski, A. Rahman, A. George, S. J. Arnold, J. Mcculloch, State space models for forecasting water quality variables: An application in aquaculture prawn farming, *The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London, United Kingdom, 2018, pp. 177-185.

[13] P. B. Bokingkito Jr., L. T. Caparida, Using fuzzy logic for real - time water quality assessment monitoring system, *The 2018 2nd International Conference on Automation, Control and Robots*, Bangkok, Thailand, 2018, pp. 21-25.

[14] D. A. Konovalov, J. A. Domingos, C. Bajema, R. D. White, D. R. Jerry, Ruler detection for automatic scaling of fish images, *International Conference on Advances in Image Processing*, Bangkok, Thailand, 2017, pp. 90-95.

[15] Y. Wang, R. Tan, G. Xing, X. Tan, J. Wang, R. Zhou, Spatiotemporal aquatic field reconstruction using robotic sensor swarm, *2012 IEEE 33rd Real-Time Systems Symposium*, San Juan, PR, United States, 2012, pp. 205-214.

[16] M. Jayanthi, S. Thirumurthy, M. Muralidhar, P. Ravichandran, Impact of shrimp aquaculture development on important ecosystems in India, *Global Environmental Change*, Vol. 52, pp. 10-21, September, 2018.

[17] N. Bharill, A. Tiwari, A. Malviya, Fuzzy based scalable clustering algorithms for handling big data using apache spark, *IEEE Transactions on Big Data*, Vol. 2, No. 4, pp. 339-352, December, 2016.

[18] B. Liu, S. He, D. He, Y. Zhang, M. Guizani, A spark-based parallel fuzzy c-means segmentation algorithm for agricultural image big data, *IEEE Access*, Vol. 7, pp. 42169-42180, March, 2019.

[19] J. Wu, Z. Wu, J. Cao, H. Liu, G. Chen, Y. Zhang, Fuzzy

consensus clustering with applications on big data, *IEEE Transactions on Fuzzy Systems*, Vol. 25, No. 6, pp. 1430-1445, December, 2017.

[20] A. Segatori, F. Marcelloni, W. Pedrycz, On distributed fuzzy decision trees for big data, *IEEE Transactions on Fuzzy Systems*, Vol. 26, No. 1, pp. 174-192, February, 2018.

[21] B. Małysiak-Mrozek, A. Lipin´ska, D. Mrozek, Fuzzy join for flexible combining big data lakes in cyber-physical systems, *IEEE Access*, Vol. 6, pp. 69545-69558, November, 2018.

[22] S. Hu, Dynamic monitoring based on wireless sensor networks of iot, *2015 International Conference on Logistics, Informatics and Service Sciences*, Barcelona, Spain, 2015, pp. 1-4.

[23] J.-H. Chen, W.-T. Sung, G.-Y. Lin, Automated monitoring system for the fish farm aquaculture environment, *2015 IEEE International Conference on Systems, Man, and Cybernetics*, Hong Kong, China, 2015, pp. 1161-1166.

[24] S. Rajeswari, K. Suthendran, K. Rajakumar, A smart agricultural model by integrating iot, mobile and cloud-based big data analytics, *2017 International Conference on Intelligent Computing and Control*, Coimbatore, India, 2017, pp. 1-5.

[25] R. Vatrapu, R. R. Mukkamala, A. Hussain, B. Flesch, Social set analysis: A set theoretical approach to big data analytics, *IEEE Access*, Vol. 4, pp. 2542-2571, April, 2016.

[26] C.-T. Chiang, T.-Y. Chen, Y.-T. Wu, Design of a Water Salinity Difference Detector for Monitoring Instantaneous Salinity Changes in Aquaculture, *IEEE Sensors Journal*, Vol. 20, No. 6, pp. 3242-3248, March, 2020.

[27] D. Li, J. Sun, H. Yang, X. Wang, An Enhanced Naive Bayes Model for Dissolved Oxygen Forecasting in Shellfish Aquaculture, *IEEE Access*, Vol. 8, pp. 217917- 217927, December, 2020.

Biographies

Wu-Chih Hu received his Ph.D. degree in Electrical Engineering from the National Taiwan University of Science and Technology, Taiwan, in 1998. From 1998, he worked at the National Penghu University of Science and Technology for 20 years. He is currently a Professor in Engineering and the Department of Computer Science and Information Engineering. His current research interests include image processing, pattern recognition, visual surveillance, IoT, and Deep learning.

Hsin-Te Wu received the Ph.D. Degree in Department of Computer Science and Engineering from National Sun Yat-Sen University, Taiwan, in 2013. Hsin-Te Wu is an Assistant Professor of Department of Computer Science and Information Engineering from National Ilan University, Taiwan.

Jun-We Zhan is a student of Department of Computer Science and Information Engineering from National Penghu University of Science and Technology, Taiwan.

Jing-Mi Zhang is a student of Department of Computer Science and Information Engineering from National Penghu University of Science and Technology, Taiwan.