Journal of Fundamental and Applied Sciences

ISSN 1112-9867

Available online at http://www.jfas.info

Research Article

Special Issue

ESCHERICHIA COLI GROWTH MODELING USING NEURAL NETWORK

S. N. Shamsudin[1], M. H. F. Rahiman[1,*], M. N. Taib[1], A. H. Ahmad[2] and W. R. W. A. Razak[2]

[1]Faculty of Electrical Engineering, UniversitiTeknologi MARA, 40450 Shah Alam, Selangor, Malaysia

[2]Faculty of Applied Sciences, UniversitiTeknologi MARA, 40450 Shah Alam, Selangor, Malaysia

Published online: 05 October 2017

ABSTRACT

The assessment of water microbial quality is normally performed by verification of Escherichia coli where the growth is in nonlinearity. NARX is computational tools that have extensive utilization in solving nonlinear time series problems. It is well known as one of the technique that has the ability to predict with efficient and good performance. Using NARX, a highly accurate model was developed to predict the growth of Escherichia coli (E. coli) based on pH water parameter. The multiparameter portable sensor and spectrophotometer data were used to build and train the neural network. The selection of neural network structure for pH and optical density modelling was optimized and also the training and validation were analyzed. The result exhibited that NARX modeling was able to predict the growth of E. coli based on pH water parameter with overall regression is 0.99956.

Keywords: neural network; NARX; prediction; Escherichia coli; pH; optical density.

Author Correspondence, e-mail: hezrif@ieee.org

doi: http://dx.doi.org/10.4314/jfas.v9i4s.42

1. INTRODUCTION

An extensive microbial waterborne disease outbreak among peoples who used water for drinking, domestic purpose and recreation [1-3]. Faecal coliforms that include Escherichia coli is the most common risk associated with microbial contamination in water [2]. Escherichia coli commonly referred to as E. coli is found in the digestive systems of human and warm-blooded animals. E. coli get in the water during rainfalls and snow melt, washed into creeks, rivers, lakes, stream, sea, beaches, recreational water from the land surfaces [4].

The presence of E. coli in water is associated with the microbial growth itself. Microbial growth phases under specific environment conditions represent a principal process in microbiology. The microbial growth phases consists of lag, exponential, stationary and death for typical growth curve [5]. In water quality application areas, the verification of microbial growth is more focused at the stationary phases [6].

Nowadays, modeling has become significant mechanism for expanding our understanding of microbial growth. In water quality, the models are developed to describe water parameter changes. It must be acknowledged that water is very complex and many interactions may occur. For instance, the growth of microbes may lead to pH changes which in turn may have consequences for chemical reactions if they are acid-catalyzed. Nevertheless, models can help in controlling and predicting water quality attributes and their changes.

A wide range of nonlinear system can be interpreted by neural network modeling technique for use in microbial growth [7-10]. Nonlinearity is essential to the living microbial culture and limits strongly the use of traditional deterministic modeling techniques to represent the growth of microbes as a function of time [11]. Neural network have been engaged in recent years as an alternative to traditional regression models due to strength of describing complex and nonlinear problems [12].

The realization of microbial growth modeling is based on an Autoregressive Network with an Exogenous Input (NARX) model. NARX neural network is a model that is based on the linear ARX model. It is commonly used to predict future especially in difficult time series prediction tasks [13].

In recent years, there has been an increasing amount of literature on pH data were used for determining the microbial water quality [14-18]. In 2001, data for aerobic growth of microbe in response to change in pH were used to develop a prediction model [19]. In a different study in 2005, a model was developed to predict bacterial spore inactivation based on combined effect of pH as a function of heat resistance [20].

With this in mind, in this study, the effect between E. coli and pH at various times was examined. To this end, NARX neural network was used as an alternative approach to conventional methods of microbial growth prediction.

2. MATERIAL AND METHOD

2.1. Multiparameter Portable Sensor Unit

The data of pH were collected from the experiment of water physical parameter using multiparameter portable sensor unit. The HORIBA multiparameter, U-50 Series was soaked in water sample containing E. coli culture. The pH sensor that builds in the multiparameter was read the changes of pH value every 1 hour interval within 8 hours. The pH readings were recorded manually and those data will be prepared for the analysis.

2.2. Spectrophotometer

The data of optical density, OD was collected from the experiment of E. coli growth using a spectrophotometer. An amount of 0.1 ml of water sample was taken using micropippete and then placed in a cuvette. A light beam passing through the cuvette will be scattered more or less by the cells, depending on the cell density (=turbidity). The wavelength of the optical density (OD) was set to 600 nm [21]. The OD readings were recorded manually and those data will be prepared for the analysis.

2.3. Significant Correlation between E. Coli Growth and pH

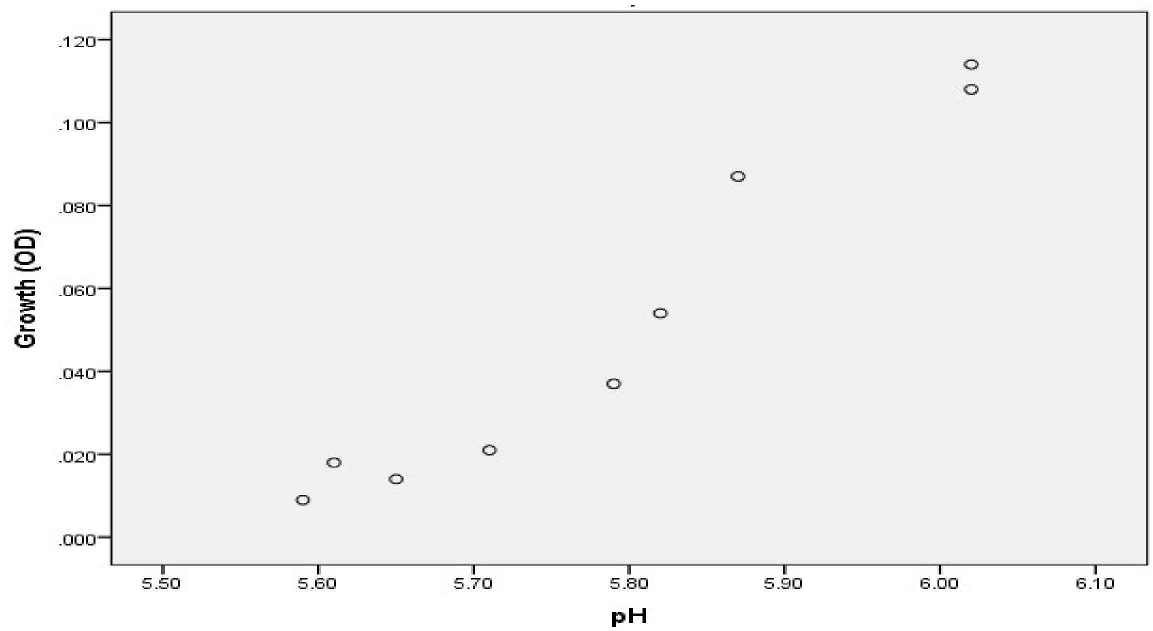


Fig.1. Statistical analysis using Pearson correlation between E. coli growth and pH

From the previous study [22], pH was identified as significant parameter that correlates with the growth of E. coli as seen in Fig. 1. The statistical analysis technique used was Pearson correlation. The relationship between E. coli growth (≈OD)and pH gave the correlation coefficient, r = 0.971.

2.4. NARX Experimental Design

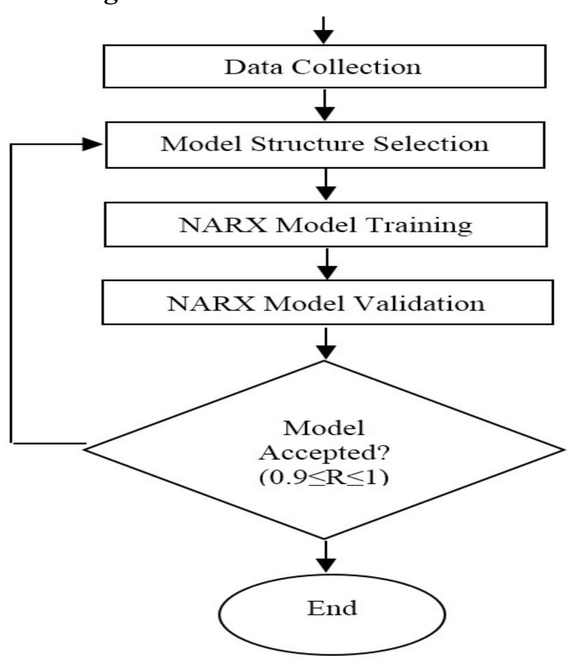


Fig.2. Flowchart of NARX experimental design

Fig. 2 shows the experimental design of the NARX modeling developed by using MATLAB 2015. The network consists of five steps i.e. the data collection, the model structure collection, NARX model training, NARX model validation and modelling acceptance. First step, the pH and OD data were collected and were normalized. Second step, the number of hidden neurons and input delays were selected based on optimization of NARX training and validation under model structure selection. When the number of hidden neurons and input delays were selected with the nearest regression value, R = 1, the input data were divided by block so that 70% of samples were assigned to the training set and the remaining 30% to the validation set in the third and fourth step. All the data set were used interpolation technique [23]. Lastly, if the regression value, R for the training, validation and testing data set is acceptable, the number of hidden neurons and input delays can be used to evaluate the NARX model.

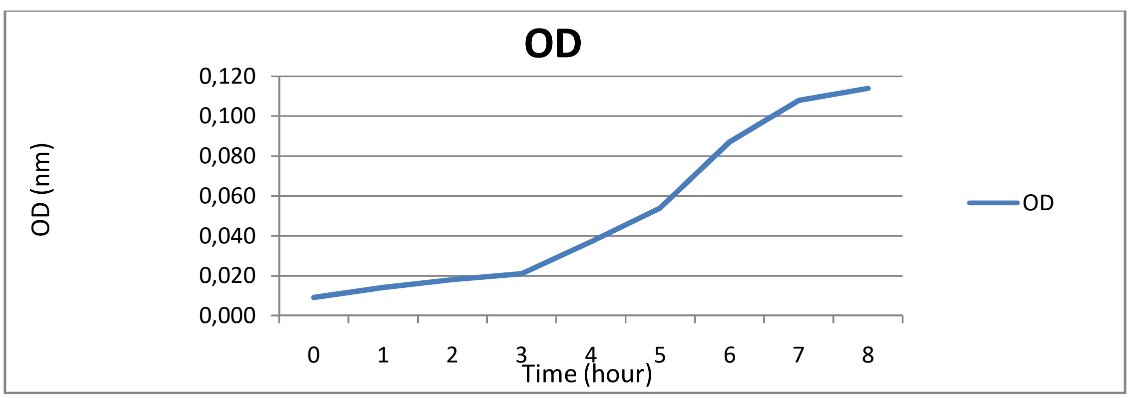
The definition equation for the NARX model is:

$$y(t) = f\big(y(t-1)\big), y(t-2), \ldots, y\big(t-n_y\big), u(t-1), u(t-2), \ldots, u(t-n_u) \quad (1)$$

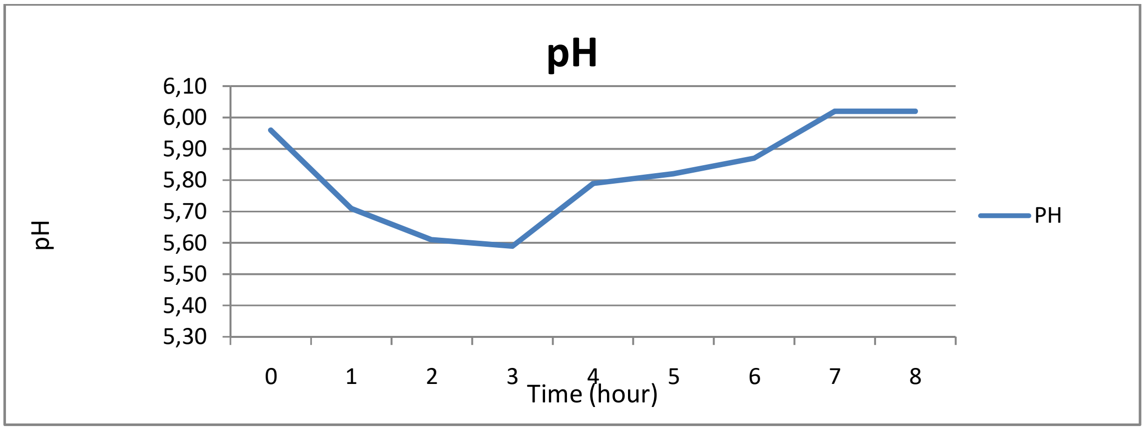
where the next value of the dependent output signal $y(t)$ is regressed on previous values of the output signal and previous values of an independent (exogenous) input signal.

3. RESULTS AND DISCUSSION

3.1. Data Collection for OD and pH



(a)



(b)

Fig.3. Observed data on the input signal (a) OD versus time and (b) pH versus time of E. coli growth recorded during an experiment

The experimental study yielded an input and output data as illustrated in Fig. 3. Data recorded by manual monitoring are pH as an input and OD of E. coli growth as an output. For each parameter, 2 curves were generated at 1 hour interval for 8 hours. Fig. 3 (a) shows the relationship between OD and time,while Fig. 3 (b) shows the relationship between pH and time. The pH shows some decrement pattern for the first 3 hours. This unexplained behaviour may be further investigated. However, the statistic shows significant correlation exists between OD and pH (Fig. 1).

3.2. NARX Structure Selection

3.2.1. Hidden Neuron Number Selection

A quantitative observation for hidden neuron number selection based on number of delay 1 using NARX modelling is presented in Table 1. The ratio of training and validation used are 70% for training (23 data) and 30% for validation (10 data). This table shows that regression calculated using both training and validation data sets were trained until 10th number of hidden neurons. The results indicate that the number of hidden neurons 2 gives the best regression value, R of 0.99914, which is closest to 1.

Table 1. Selection of hidden neuron number based on regression value

No. of Hidden Neuron	No. of Delay	Regression Training	Regression Validation	Regression All
1	1	-0.97953	0.47772	-0.30499
2*	1	0.99882	0.96605	0.99914
3	1	0.92089	0.79351	0.97182
4	1	0.99922	0.21497	0.97767
5	1	0.99884	0.56714	0.99245
6	1	0.99640	0.34913	0.98139
7	1	0.99875	0.90371	0.99399
8	1	0.99862	0.52664	0.98724
9	1	0.99938	-0.92001	0.94077
10	1	0.99937	-0.00402	0.95054

*Indicate the best overall regression value.

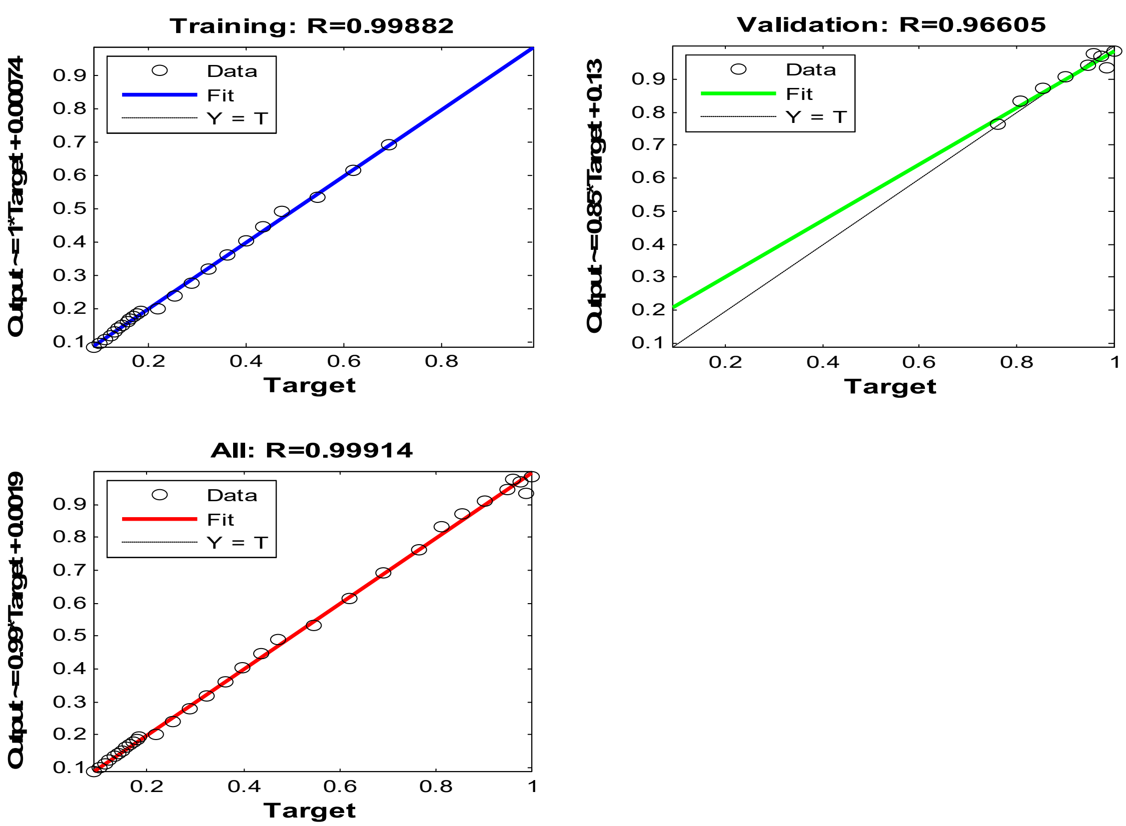


Fig.4. Result of training and validation data

Fig. 4 shows the regression value, R of training and validation during the selection of number of hidden neurons 2. The result indicates that the regression value, R = 0.99914 which is near to 1 when the input delay is 1.

3.2.2. Number of Delays Selection

The numerical results of the identified model parameters and the number of delays based on the number of hidden neurons 2 are shown in Table 2. The ratio of training and validation were 70% (23 data) and 30% (10 data) respectively. This table validates the accuracies of the NARX model work reasonably well during 7 numbers of delays. The result of least value among the overall regression, R was 0.99882 for performing the best model structure.

Table 2. Selection of delays number based on regression

No. of Hidden Neuron	No. of Delay	Regression Training	Regression Validation	Regression All
2	1	0.99888	0.94519	0.99844
2	2	0.96496	-0.46384	-0.19256
2	3	0.99922	0.39226	0.98858
2	4	0.99909	-0.38825	0.91396
2	5	0.99883	0.94145	0.99823
2	6	-0.78170	-0.45950	-0.34176
2	7*	0.99915	0.94531	0.99882
2	8	0.99934	0.93860	0.99872
2	9	-0.97418	0.61180	-0.56236
2	10	0.99376	-0.63632	0.96788

*Indicate the best overall regression value.

3.3. Result of NARX Model

3.3.1. Regression

The NARX model consists of two features which are pH and OD over 1 hour time interval within 8 hours. The regression values achieved by NARX model using interpolation data are shown in Fig. 5. This NARX model signifies a very good linear regression correlation between measured and predicted data. From the results, it can be seen that NARX model gives good performance over the training and validation models with all regression, R = 0.99882.

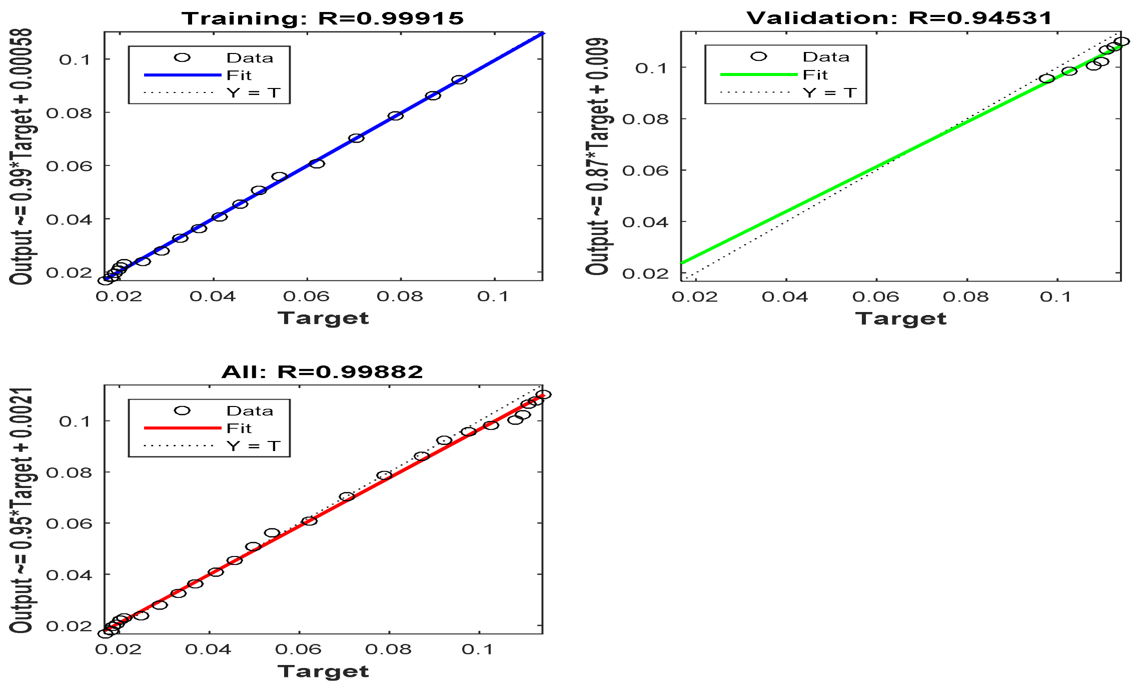


Fig.5. Best regression of NARX model

3.3.2. Time Series Response

In this section, the time series response using data testing is shown in Fig. 6. The training and validation output were fitted with the targets. This shows that time series response can identify novel relationships and patterns in microbial data.

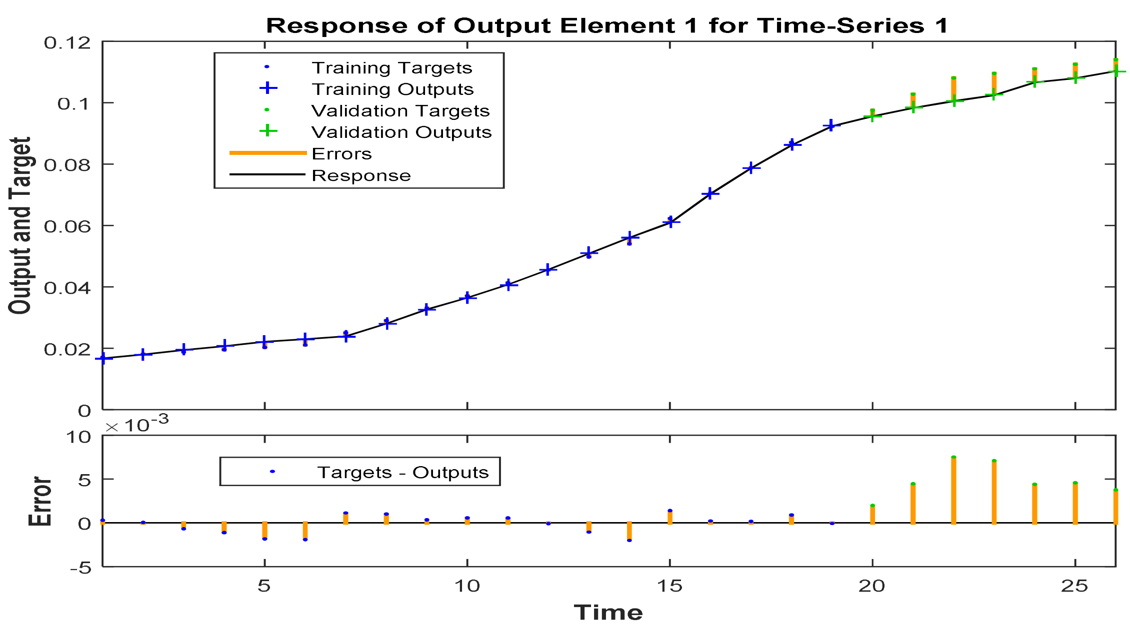


Fig.6. Time series response of output and error for structure selection of NARX model

4. CONCLUSION

Using simple yet reasonable input combination, the proposed NARX [24] model is able to perform real time E. coli growth with good accuracies. This paper also investigates the number of hidden neurons and delays selection to predict E. coli growth based on pH parameter using NARX modelling [25]. The selection criteria are based on regression values during testing. The number of hidden neurons and delays affected the performance of prediction technique. This model shows a good agreement between target and output values.

Overall, the prediction of E. coli based on pH parameter performs the best result with 2 hidden neurons and 7 numbers of delays with regression value, R = 0.99882 which is close to 1. NARX [26] is one part of the efficient model to predict E. coli growth based on pH parameter.

5. ACKNOWLEDGEMENTS

The financial support by Research Management Institute (RMI), UniversitiTeknologi MARA and MyPhD are gratefully acknowledged. Faculty of Applied Sciences (FSG), Faculty of Civil (FKA), Faculty of Electrical Engineering (FKE) UniversitiTeknologi MARA and PICO/PICON members are also acknowledged for their technical support in conducting the field experiment.

6. REFERENCES

[1] Soller JA, Schoen ME, Varghese A, Ichida AM, Boehm AB, Eftim S, Ashbolt NJ, Ravenscroft JE. Human health risk implications of multiple sources of faecal indicator bacteria in a recreational waterbody. Water Research, 2014, 66:254-264

[2] Khan MJ, Ali MF, Hassan M. Microbial evaluation of drinking water and frequency of bacterial isolates from Rawalpindi Pakistan. Pakistan Journal of Pathology, 2017, 28(1):28-32

[3] Uejio CK, Christenson M, Moran C, Gorelick M. Drinking-water treatment, climate change, and childhood gastrointestinal illness projections for northern Wisconsin (USA) communities drinking untreated groundwater. Hydrogeology Journal, 2017, 25(4):969-979

[4] Sampson RW, Swiatnicki SA, McDermott CM, Kleinheinz GT. The effects of rainfall on Escherichia coli and total coliform levels at 15 Lake Superior recreational beaches. Water

Resources Management, 2006, 20(1):151-159

[5] Peleg M, Corradini MG. Microbial growth curves: What the models tell us and what they cannot. Critical Reviews in Food Science and Nutrition, 2011, 51(10):917-945

[6] Esser DS, Leveau JH, Meyer KM. Modeling microbial growth and dynamics. Applied Microbiology and Biotechnology, 2015, 99(21):8831-8846

[7] Oscar T. Neural network models for growth of Salmonella serotypes in ground chicken subjected to temperature abuse during cold storage for application in HACCP and risk assessment. International Journal of Food Science and Technology, 2017, 52(1):214-221

[8] Oladunjoye AO, Oyewole SA, Singh S, Ijabadeniyi OA. Prediction of Listeria monocytogenes ATCC 7644 growth on fresh-cut produce treated with bacteriophage and sucrose monolaurate by using artificial neural network. LWT-Food Science and Technology, 2017, 76:9-17

[9] Abdollahzadeh E, Ojagh SM, Hosseini H, Irajian G, Ghaemi EA. Predictive modeling of survival/death of Listeria monocytogenes in liquid media: Bacterial responses to cinnamon essential oil, ZnO nanoparticles, and strain. Food Control, 2017, 73:954-965

[10] Choi SW, Bae HK. Daily prediction of total coliform concentrations using artificial neural networks. KSCE Journal of Civil Engineering, 2017, 1-8

[11] Geeraerd AH, Valdramidis VP, Devlieghere F, Bernaert H, Debevere J, Van Impe JF. Development of a novel approach for secondary modelling in predictive microbiology: Incorporation of microbiological knowledge in black box polynomial modelling. International Journal of Food Microbiology, 2004, 91(3):229-244

[12] Hajmeer MN, Basheer IA, Najjar YM. Computational neural networks for predictive microbiology II. Application to microbial growth. International Journal of Food Microbiology, 1997, 34(1):51-66

[13] MenezesJr JM, Barreto GD. A new look at nonlinear time series prediction with NARX recurrent neural network. In IX Brazilian Neural Networks Symposium, 2006, pp. 28-33

[14] Bakobie N, Awal FM, Ballu A. Water quality assessment of hand-dug wells in Janga Ghana.International Research Journal of Public and Environmental Health, 2015, 2(12):197-205

[15] Balamurugan S, Ahmed R, Gao A. Survival of Shiga toxin-producing Escherichia coli in

broth as influenced by pH, water activity and temperature. Letters in Applied Microbiology, 2015, 60(4):341-346

[16] HamzahAI, Kipli S H, Rahil I S. Microbiological study in coastal water of Port Dickson, Malaysia. SainsMalaysiana, 2011, 40(2):93-99

[17] Alonso V, Cavaglieri L, Ramos AJ, Torres A, Marin S. Modelling the effect of pH and water activity in the growth of Aspergillusfumigatus isolated from corn silage. Journal of Applied Microbiology, 2017, 122(4):1048-1056

[18] Sivasankar P, Kumar GS. Influence of pH on dynamics of microbial enhanced oil recovery processes using biosurfactant producing Pseudomonas putida: Mathematical modelling and numerical simulation. Bioresource Technology, 2017, 224:498-508

[19] Jeyamkondan S, Jayas DS, Holley RA. Microbial growth modelling with artificial neural networks. International Journal of Food Microbiology, 2001, 64(3):343-354

[20] Esnoz A, Periago PM, Conesa R, Palop A. Application of artificial neural networks to describe the combined effect of pH and NaCl on the heat resistance of Bacillus stearothermophilus. International Journal of Food Microbiology, 2006, 106(2):153-158

[21] Low SX, Aw ZQ, Loo BZ, Lee KC, Oon JS, Lee CH, Ling MH. Viability of Escherichia coli ATCC 8739 in nutrient broth, luria-bertani broth and brain heart infusion over 11 weeks. Electronic Physician, 2013, 5(1):576-581

[22] Shamsudin SN, Rahman MH, Taib MN, Razak WR, Ahmad AH, Zain MM. Analysis between Escherichia Coli growth and physical parameters in water using Pearson correlation. In7th IEEEControl and System Graduate Research Colloquium, 2016, pp. 131-136

[23] Whiting RC. Microbial modeling in foods. Critical Reviews in Food Science and Nutrition, 1995, 35(6):467-494

[24] Zabidi A, Tahir NM, Yassin IM, Rizman ZI. The performance of binary artificial bee colony (BABC) in structure selection of polynomial NARX and NARMAX models. International Journal on Advanced Science, Engineering and Information Technology, 2017, 7(2):373-379

[25] Yassin IM, Jailani R, Ali MS, Baharom R, Hassan AH, Rizman ZI. Comparison between cascade forward and multi-layer perceptron neural networks for NARX functional electrical stimulation (FES)-based muscle model. International Journal on Advanced Science,

Engineering and Information Technology, 2017, 7(1):215-221

[26] Yassin IM, Zabidi A, Ali MS, Tahir NM, Abidin HZ, Rizman ZI. Binary particle swarm optimization structure selection of nonlinear autoregressive moving average with exogenous inputs (NARMAX) model of a flexible robot arm. International Journal on Advanced Science, Engineering and Information Technology, 2016, 6(5):630-637

How to cite this article:
Shamsudin S N, Rahiman M H F, Taib M N, Ahmad A H, Razak W R W A. Escherichia coli growth modeling using neural network. J. Fundam. Appl. Sci., 2017, *9(4S), 759-771.*