https://doi.org/10.31436/iiumej.v23i2.2173

NEW APPROACH TO PREDICT FECAL COLIFORM REMOVAL FOR STORMWATER BIOFILTER APPLICATIONS

SAI HIN LAI[1*], CHUN HOOI BU[1], REN JIE CHIN[2*], XIANG TING GOH[3] AND FANG YENN TEO[4]

[1]*Department of Civil Engineering, Faculty of Engineering, Universiti Malaya, 50603 Kuala Lumpur, Malaysia.*
[2]*Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, 43000 Kajang, Malaysia.*
[3]*Department of Parasitology, Faculty of Medicine, Universiti Malaya, 50603 Kuala Lumpur, Malaysia.*
[4]*Department of Civil Engineering, Faculty of Engineering, University of Nottingham (Malaysia Campus), 43500 Semenyih, Malaysia.*

[*]Corresponding author: laish@um.edu.my; chinrj@utar.edu.my

(Received: 21st August 2021; Accepted: 26th March 2022; Published on-line: 4th July 2022)

ABSTRACT: Fecal coliform removal using stormwater biofilters is an important aspect of stormwater management. A model that can provide an accurate prediction of fecal coliform removal is essential. Therefore, feedforward backpropagation neural network (FBNN) and adaptive neuro-fuzzy inference system (ANFIS) models were developed using a range of input features, namely grass type, the thickness of biofilter, and initial concentration of E. coli, while the estimated final concentration of *E. coli* was the output variable. The ANFIS model shows a better overall performance than the FBNN model, as it has a higher R^2-value of 0.9874, lower MAE and RMSE values of 3.854 and 6.004 respectively, and a smaller average percentage error of 14.2%. Hence, the proposed ANFIS model can be served as an advanced alternative to replace the need for laboratory work.

ABSTRAK: Penyingkiran kolifom tinja menggunakan turas biologi (bioturas) air hujan merupakan aspek penting dalam pengurusan air hujan. Model yang dapat menunjukkan anggaran tepat tentang penyingkiran kolifom tinja adalah penting. Oleh itu, model rangkaian suapan neural perambatan belakang (FBNN) dan sistem adaptasi inferen neuro-fuzi (ANFIS) telah dibentukkan menggunakan pelbagai ciri input, iaitu jenis rumput, ketebalan bioturas dan kepekatan awal E. coli, manakala anggaran kepekatan akhir bagi *E. coli* merupakan hasil pembolehubah. Model ANFIS menunjukkan peningkatan keseluruhan yang lebih baik berbanding model FBNN, kerana ia mempunyai nilai R^2 yang lebih tinggi iaitu 0.9874, nilai MAE dan RMSE yang lebih rendah iaitu sebanyak 3.854 dan 6.004 masing-masing, dan ralat peratusan purata yang lebih kecil sebanyak 14.2%. Oleh itu, model ANFIS yang dicadangkan boleh dijadikan alternatif awal bagi menggantikan keperluan kerja makmal.

KEYWORDS: *artificial intelligence; biofilters; fecal coliform; neural network; stormwater*

1. INTRODUCTION

Biofiltration systems such as swale and bio-detention systems are increasingly popular low-energy treatment technologies for improved stormwater management, e.g. increase of infiltration, reduction of peak flow, improvement of water quality and increase of surrounding aesthetic value. Stormwater biofilters can be defined as vegetated vertical infiltration systems that can achieve runoff volumes and contaminant load reductions for urban environments [1]. They have shown promising yet variable removal of fecal microorganisms [2-4]. Fecal coliform bacteria are a group of bacteria that are passed through the fecal excrement of humans, livestock, and wildlife and they are the indicator bacteria. The most common member of fecal coliform bacteria is *Escherichia coli*.

Vegetation and filter media depth may affect the capability of stormwater biofilters in removing fecal coliform. In the past two decades, researchers from around the world have experimented with biofilters using different design elements to investigate the biofiltration system in removing fecal coliform [5-10]. These studies reported that vegetation type or filter media depth caused variable bacteria removal performance. A previous study in Australia reported that biofilters planted with native grasses (*Paspalum conjugatum* and *Buchloe dactyloides*) and shrubs (*Melaleuca incana*, *Leptospermum continentale*) showed improved *E. coli* removal, possibly due to reduced infiltration rates in vegetated biofilter systems. In addition, the leaf or seed extracts of *L. continentale* demonstrated potential antibacterial activity against *E. coli* [7]. With regards to filter media depth, it was reported that *E. coli* concentration decreased with increasing filter media depth [7]. Nevertheless, there is no data available for the influence of native vegetation and filter media depth on microbial removal by stormwater biofilters in Malaysia. Therefore, in this study, the effect of biofilter designs (i.e., vegetation type, media thicknesses), as well as the inflow concentration in fecal coliform removal, are investigated to fill the gap of knowledge.

Artificial intelligence (AI) appears as a popular tool in providing the solution to complex non-linear problems and its application on issues relevant to environmental and hydrological researches has been widely seen, i.e. application of artificial neural networks (ANN), fuzzy logic and adaptive neuro-fuzzy systems (ANFIS) for the solution of water/ wastewater and air pollution-related environmental problems [11], integration of ANN and genetic algorithms (GA) for water quality modelling [12], implementation of machine learning classification to detect simulated increases of de facto reuse and urban stormwater surges in surface water [13], performance prediction of stormwater biofilters in heavy metal removal and risk mitigation using multilinear regressions (MLR), neural network (NN), and random forest (RF) [14], etc. Therefore, this study aims to introduce the use of feedforward backpropagation neural network (FBNN) and adaptive neuro-fuzzy inference system (ANFIS) to predict the final concentration of fecal coliform for different conditions of stormwater biofilters. The proposed model can serve as an advanced method to replace the need for laboratory work. This study is innovative as it adds value to the current development of AI applications in improving stormwater management systems.

2. MATERIALS AND METHODS

2.1 Experimental Works

In this study, four native plants, namely Cow grass (*Axonopus compressus*), Pearl grass (*Axonopus compressus,* dwarf), Philippine grass (*Zoysia matrella*), and Japanese grass (*Microstegiumvimineum*) were selected. The biofilter columns were set up as shown in Fig. 1.

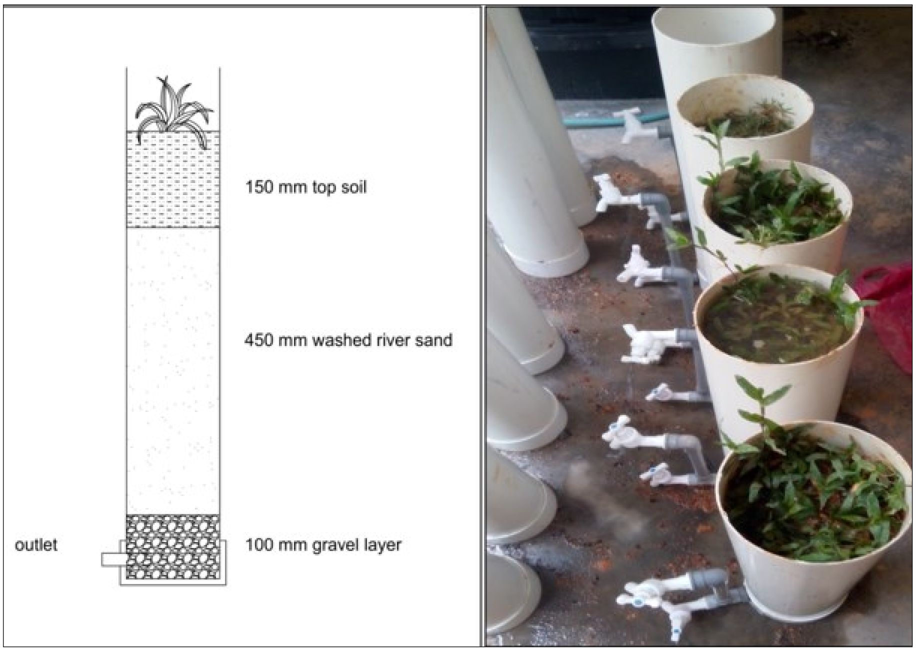

Fig. 1: The setup of biofilter columns used in this study.

The main constituent that forms the filter media was washed river sand. Four different depths were fixed in the sand columns, which are 150 mm, 250 mm, 350 mm and 450 mm. With respect to each depth, the recorded average hydraulic conductivity was 60.2, 47.1, 35.4 and 25.4 mm/hr, respectively. The values fell within the ranges recommended by Urban Stormwater Management Manual of Malaysia (MSMA) [15].

One month before conducting the experiments, the biofilter columns were planted with native vegetation for the plants to mature. An amount of 80 L of water was collected from a local pond to act as stormwater in the experiments.

Four liters of water were poured into every biofilter column and the filtered water was collected. The analysis was carried out for the number of remaining indicator bacteria. A vacuum pump was then used to further filter the collected water sample through sterile nitrocellulose membrane filters. The membrane filters had characteristics of 0.45 µm pore size and 47 mm diameter. The membrane was transferred to a sterile petri dish with an absorbent pad (Millipore, Bedford, MA, USA) containing lauryl sulfate membrane medium (Oxoid, Hampshire, UK) agar plates. The plate was sealed with parafilm and incubated at 30 °C for 4 hours to resuscitate the growth of bacteria before further incubation at 44.5 °C for 14 hours. Fecal coliform that formed yellow colonies were counted and expressed as colony-forming units (CFU) per 100 mL (CFU/100 mL).

The water samples collected before and after the filtration were termed inflow and outflow concentration respectively. The removal efficiency of the biofilter columns can be obtained using Eq. (1).

$$\log removal = \; log_{10} \frac{Influent\ pathogen\ concentration}{Effluent\ pathogen\ concentration} \qquad (1)$$

2.2 Architecture of the Feedforward Backpropagation Neural Network (FBNN) Model

Data used in this study can be retrieved from the authors' previous work [10]. Feedforward backpropagation neural network (FBNN), as shown in Fig. 2, has been commonly used in different fields of applications, particularly in developing non-linear mathematical/prediction models [16-17].

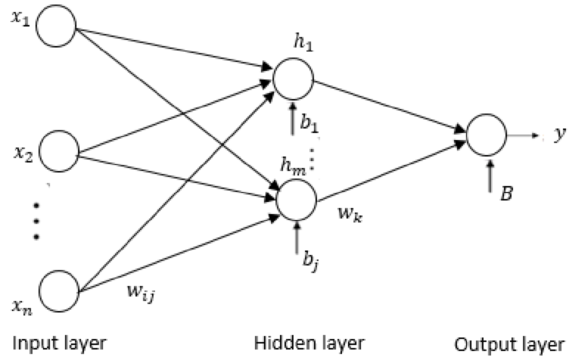

Fig. 2. The general architecture of the FBNN model.

The net values at each hidden neuron (with first pattern inputs and random weight and bias) are presented as [16]:

$$netvalue_j = \sum_{i=1}^{n} w_{ij}x_j + b_j \quad (2)$$

where $netvalue_j$ is net input to node i in hidden or output layer, x_j are the inputs to node i (or output of the previous layer), w_{ij} are the weights representing the power of the relationship between the *i*th node and *j*th node, n is the number of nodes and b_j is the bias related to node j.

The transfer function is required to activate the neurons. In this research study, the sigmoid function is chosen as the activation function.

$$h_j = \frac{1}{1+e^{-netvalue_j}} \quad (3)$$

where h_j is the output node of j and is an element of the inputs to the nodes in the next layer.

The net values at the output layer and output neuron values are calculated by Eqs. (4) and (5) respectively.

$$netvalue = \sum_{k=1}^{m} w_k h_k + B \quad (4)$$

$$output\ neuron\ values = \frac{1}{1+e^{-netvalue}} \quad (5)$$

where B is the bias.

In this study, the inputs for the FBNN model were grass type, the thickness of the biofilter, and the initial concentration of *E. coli*. The expected resulting output of the model is the final concentration of *E. coli*. The architecture of the feedforward backpropagation neural network (FBNN) for the final *E. coli* concentration prediction is shown in Fig. 3.

In addition, data sorting is one of the crucial procedures in developing any FBNN model. This is to ensure the smoothness of the overall process and to obtain a model with a respectively high level of accuracy. A proper size of training-testing data is required so that the model can learn enough possible input-output patterns [18-19]. There is no fixed guideline while setting the training to testing ratio. However, it was normally suggested to set the training dataset within the range of 60% to 80% while the remaining 20% to 40% becomes the testing dataset [20-21]. Since this study is considered as very first attempt to introduce the application of artificial intelligence in predicting the final concentration of fecal coliform with respect to different conditions of stormwater biofilters, the upper limit of 80% is selected so that the developed model will be provided with the most possible input-output patterns.

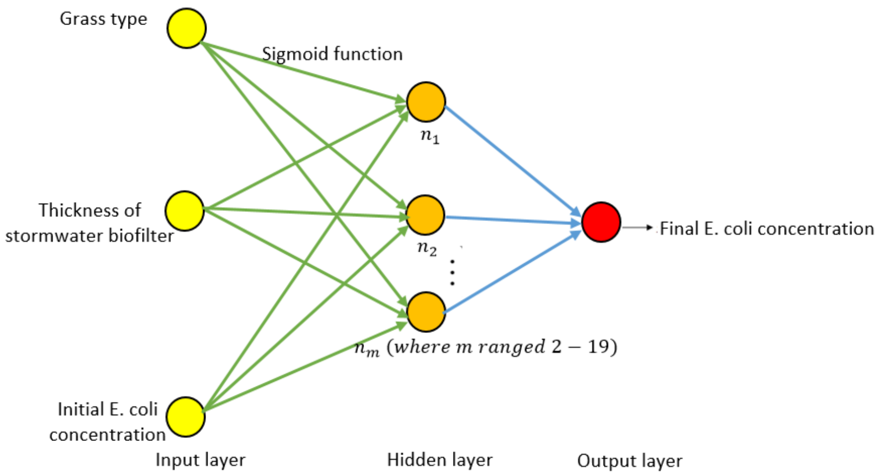

Fig. 3: The architecture of the FBNN model.

The number of hidden layers and the transfer function were set as one and a sigmoid function, respectively [21]. This is mainly due to its performance achievement in the prediction and forecasting model. Meanwhile, the training algorithm is selected as Levenberg-Marquardt (trainlm), since it is suited for function fitting (nonlinear regression) problems [22-24].

While designing the architecture of the FBNN, the determination of the number of hidden neurons is one of the main challenging tasks. This is due to the sensitivity of the networks to the number of hidden neurons. Underfitting problems may appear if there are too few neurons while overfitting issues may arise if there are too many neurons. Therefore, it is important to choose a proper number of neurons [25]. For this study, the hidden neurons were set within the ranges of 2 to 19.

2.3 Architecture of Adaptive Neuro-Fuzzy Inference System (ANFIS) Model

The integration of different techniques to form a hybrid AI model becomes the main trend of the development of AI applications. Adaptive neuro-fuzzy inference system (ANFIS) is a technique that integrates both neural networks and fuzzy logic principles within a single framework. This may strengthen the ability of the model to reach a higher level of accuracy [26-28]. A basic ANFIS architecture is presented in Fig. 4 [29-31].

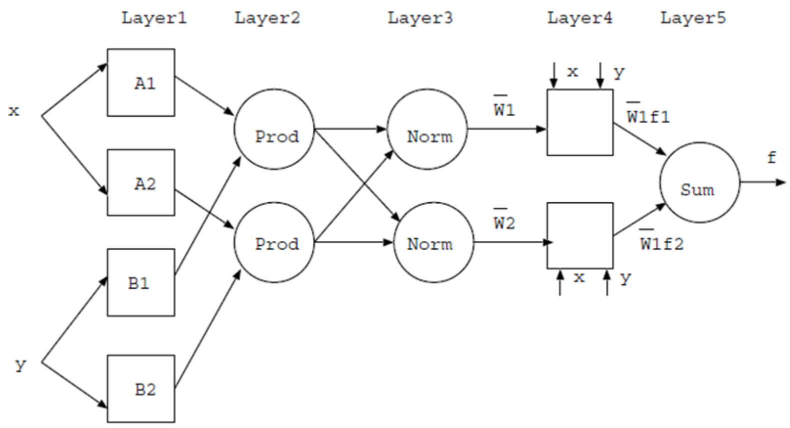

Fig. 4: The general architecture of ANFIS.

Two rules were used in the method of "If-Then" for Takagi-Sugeno fuzzy model, as shown in the following:

Rule 1: If x is A_1 and y is B_1, then $f_1 = p_1x + q_1y + r_1$ (6)

Rule 2: If x is A_2 and y is B_2, then $f_2 = p_2x + q_2y + r_2$ (7)

where A_1, A_2, B_1 and B_2 are the membership functions for each input x and y (part of the premises), f_1 and f_2 are the outputs within the fuzzy region specified by the fuzzy rule, while p_1, q_1, r_1, p_2, q_2 and r_2 are linear parameters in part- Then (consequent part) of Takagi-Sugeno fuzzy inference model [26].

ANFIS architecture consists of five layers excluding the input layer (layer 0). The description of each layer is shown as follows [26]:

1. Layer 0: It is an input layer that has n nodes, where n is the number of inputs to the system.
2. Layer 1: It is the fuzzification layer. Every node in this layer adapts to a function parameter. The output from each node is a degree of membership value that is given by the input of the membership functions. The typical membership function is shown below:

$$\mu_A(x) = \frac{1}{1+\left|\frac{x-c_i}{a_i}\right|^{2b_i}} \tag{8}$$

   where a_i, b_i and c_i are parameters for the function. The parameters in this layer are defined as premise parameters.
3. Layer 2: Every node in this layer is a fixed or nonadaptive node. The output is the product of all the incoming signals. Each node in this layer represents the fire strength for each rule. T-norm operator with general performance, such as the AND, is used to obtain the output:

$$O_{2i} = w_i = \mu_{Ai}(x) * \mu_{Bi}(y), i = 1, 2 \tag{9}$$

4. Layer 3: It is the normalization layer. Each node in this layer is fixed. Each node is a calculation of the ratio between the *i*-th rules firing strength and the sum of all rules' firing strengths. The result is known as the normalized firing strength. The strength of all rules is normalized by:

$$O_{3i} = \bar{w}_i = \frac{w_i}{\sum_i w_i} \tag{10}$$

5. Layer 4: It is a layer of adaptive nodes. Every node in this layer is an adaptive node to output, with a node function defined as:

$$O_{4i} = \bar{w}_i f_i = \bar{w}_i(p_ix + q_iy + r_i) \tag{11}$$

   where $\bar{w}_i$ is the normalized firing strength from the third layer and $(p_ix + q_iy + r_i)$ is a parameter in the node. The parameters in this layer are referred to as consequent parameters. It is assumed in Eq. (11) that all the universe of discourse for all input variables can be defined using the selected type of the membership functions, and the final output is computed using the regression parameters for each rule R. The regression parameters are the premise parameters in Eq. (11) which define the shape of the selected type of the membership function for each input variable. Consequently, the training process aims at tuning the premise and consequence parameters to achieve the desired output.

6. Layer 5: It is an output layer whose function is the summation of net outputs of the nodes in the fourth layer using the formula as shown:

$$\sum_i w_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{12}$$

While determining the input and output features as well as sorting data for ANFIS model development, a procedure that is similar to the FBNN model development was followed. The input features were made up of grass type, the thickness of stormwater, and the initial concentration of *E. coli*, while the output feature was the final concentration of *E. coli*, as depicted in Fig. 5. The number of membership function (mf) was set as three. In order to tune the patterns of the ANFIS network, the hybrid optimization method, which is the combination of backpropagation and least square-type approaches, was selected. The models were trained using different input membership functions, i.e. *trimf, trapmf, gbellmf, gaussmf, gauss2mf, pimf,* dsigmf and *psigmf,* and output membership function, i.e. constant and linear membership function.

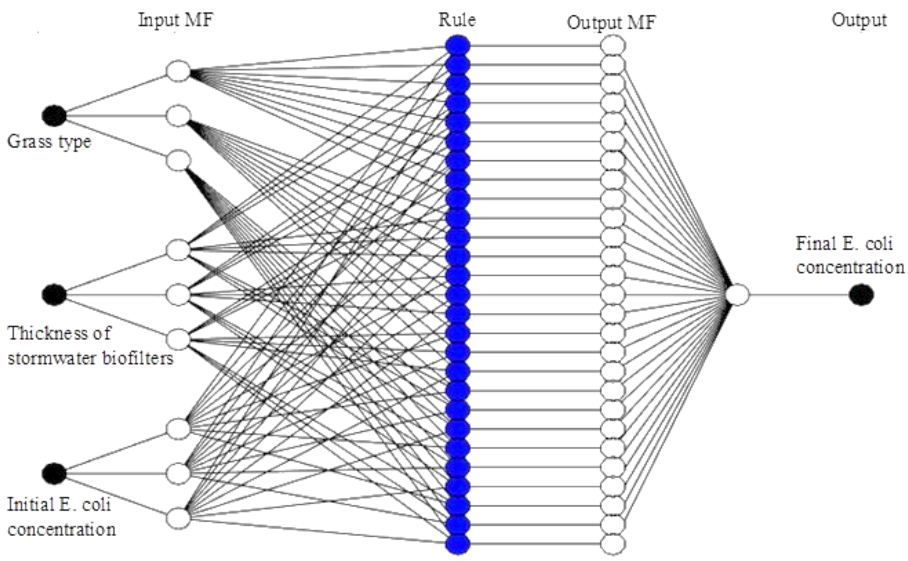

Fig. 5: The architecture of the ANFIS model.

2.4 Model Performance Evaluation

The commonly used analyses for model performance evaluation are coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and percentage error (% error). These are the relevant important indicators to show the suitability of the developed model in predicting the final concentration of *E. coli*.

$$\mathrm{R}^2 = \left(\frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}} \right)^2 \tag{13}$$

$$\mathrm{MAE} = \frac{\sum |y_i - x_i|}{n} \tag{14}$$

https://doi.org/10.31436/iiumej.v23i2.2173

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum(y_i - x_i)^2} \tag{15}$$

$$\text{Percentage error} = \frac{|True\ value - Predicted\ value|}{True\ value} \times 100\% \tag{16}$$

where n is the number of data pairs, x is the observed variable, y is the predicted variable.

3. RESULTS AND DISCUSSION

3.1 Experimental Works

As shown in Fig. 6, the removal capability of Cow grass on a 450 mm depth river sand column together with 150 mm of topsoil and 100 mm of gravel drainage layer revealed the highest fecal coliform mean log removal (2.4 *log*) compared to other biofilter columns. It agrees with the findings in Barrett et al. [32] and Chandrasena et al. [5] which indicated that vegetated biofilters improved FC removal.

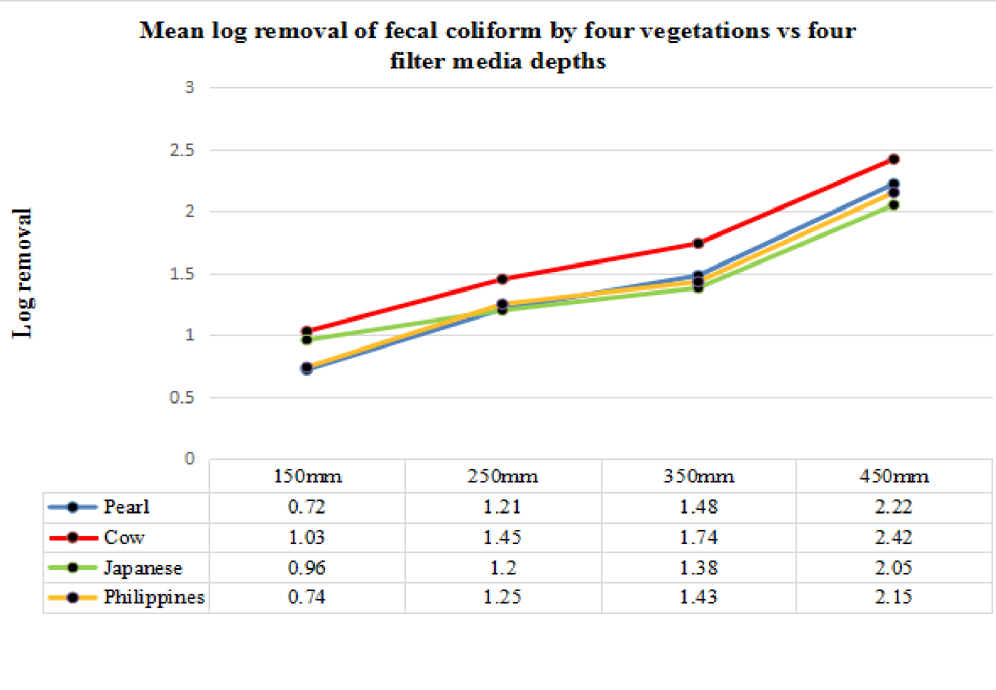

Fig. 6: Mean log removal of fecal coliform of four vegetations vs four filter media.

Planting of grasses, sedges, and shrubs in bioretention systems not only fulfils an esthetic purpose but also improves pathogen removal [33-35]. In the present study, Cow grass was found to be more suitable to use in stormwater biofilters because the survival rate of Cow grass was the highest compared to Pearl, Philippine, and Japanese grass. The bigger root mass of Cow grass improved the removal rates of FC due to its effect on biofilter retention time. It is hardy and able to grow with minimal to no fertilizer. The physical appearance of Pearl grass is similar to Cow grass but it has shorter, rounder, and thicker leaves. Pearl grass needs more water compared to Cow grass to grow and is less hardy. As for Philippine grass, it needs regular trimming about once every 2 weeks. Meanwhile, Japanese grass blades are softer, shorter, and compact. They grow rather slowly, require frequent watering, and in dry soil, they tend to die off.

https://doi.org/10.31436/iiumej.v23i2.2173

Apart from vegetation type, the physiochemical nature of filter media in biofilters plays a significant role in microbe removal [36-37]. In this study, the major component that constituted the filter media was washed river sand, offering an effective and low-cost means of treating stormwater. The finding is in-line with the previous studies which reported that sand filters showed satisfying outputs on fecal coliform removal in stormwater treatment [5]. Barrett et al. [32] reported that Austin sand filters achieved high FC (85%) and *E. coli* removal (97.1%) in an experiment to test for biofiltration performance.

In terms of depth, the biofilter with 150 mm media depth exhibited an inconsistent performance in removing fecal coliform. This may be mainly because the preferential flow was more prone to occur in lower media depth due to the intermittent wet-drying cycle. It was found that, in order to achieve > 1 *log* fecal coliform removal, 250 mm should be the minimum media depth required. On the other hand, the mean FC removal at 350 mm depth filter was slightly higher than the 250 mm depth filter but it exceeded 2 *log* for all types of biofilter columns at 450 mm depth filter.

3.2 Models Development

3.2.1 FBNN Models

The proposed FBNN models, which were developed with a range of hidden neurons from 2 to 19, were evaluated using the selected statistical analyses. However, since there were so many developed models, only the selected models were presented in this paper.

Table 1 depicts the performance of all the developed FBNN models from the aspect of R^2, MAE, RMSE and average percentage error (% error). Based on the common theory, a higher R^2-value indicates that the model has a higher ability to explain all the variance within the model. In this case, the highest R^2-value is 0.9285, as shown in model IV.

Table 1: Statistical performance of the selected FBNN models with respect to the different number of neurons

Model	Number of neurons	R^2	MAE	RMSE
I	14	0.8646	6.014	9.926
II	15	0.0650	27.850	65.771
III	16	0.8910	5.107	9.194
IV	17	0.9285	4.204	7.322
V	18	0.9249	4.512	7.423
VI	19	0.0630	57.487	122.191

While evaluating the performance of a FBNN model, a smaller error value is always favorable. This is because the smaller the calculated value, the better the accuracy of the estimated output. From the aspect of MAE and RMSE, model IV achieves the lowest values, displaying a value of 4.204 and 7.322, respectively.

All the above-discussed aspects indicate that model IV is the best-performed model. The appropriateness of model IV to predict the final *E. coli* concentration is further verified using the average percentage error. Percentage error is another common indicator. Fig. 7 contains the average percentage error of the selected FBNN models. Model IV shows the lowest average percentage error (27.5%), indicating that it has the highest level of accuracy among the examined models.

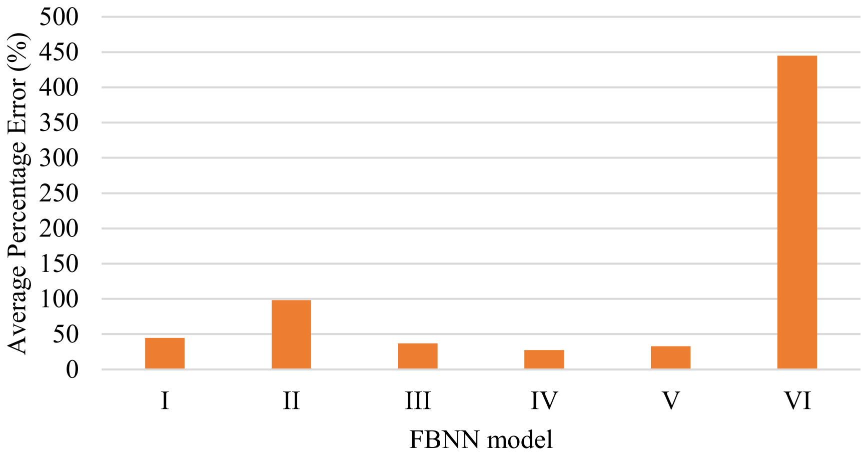

Fig. 7: The average error of the developed FBNN models.

Overall, model IV appears as the most suitable FBNN model to predict the final *E. coli* concentration for the stormwater biofilters application. This is because it has the highest R^2-value (0.9285), the lowest values of MAE (4.204) and RMSE (7.322), and the smallest average percentage error (27.5%).

3.2.2 ANFIS Models

A total number of 16 ANFIS models were developed, and their respective performances are contained in Table 2. In terms of R^2-value, if the value is near 1, indicating that the observed values and the predicted values have a strong linear relationship. In other words, the observed values and predicted values are almost similar if the R^2-value approximates 1. Referring to Table 2, the highest R^2-value is 0.9874.

On the other hand, the models with the output linear membership function (model II, IV, VI, VIII, X, XII, XIV, and XVI), in general, show a better performance than the models with constant output membership function. Therefore, it can be deduced that the output linear membership function is more suitable for the development of the ANFIS model to deal with the stormwater biofilters application. In this case, model IV displays the lowest MAE and RMSE values.

Table 2: Performance in terms of R^2, MAE and RMSE for the developed ANFIS models.

Models	NMFs	MFTI	MFTO	R^2	MAE	RMSE
I	3	trimf	constant	0.7994	11.397	23.885
II	3	trimf	linear	0.9112	7.820	15.866
III	3	trapmf	constant	0.7994	11.397	23.885
IV	3	trapmf	linear	0.9874	3.854	6.004
V	3	gbellmf	constant	0.8231	11.260	22.428
VI	3	gbellmf	linear	0.9217	7.665	14.892
VII	3	gaussmf	constant	0.7909	11.451	24.394
VIII	3	gaussmf	linear	0.9215	7.706	14.908
IX	3	gauss2mf	constant	0.7784	11.471	25.126
X	3	gauss2mf	linear	0.9216	7.626	14.910
XI	3	pimf	constant	0.7714	11.597	25.534
XII	3	pimf	linear	0.9214	7.688	14.927
XIII	3	dsigmf	constant	0.7709	12.328	25.502
XIV	3	dsigmf	linear	0.9215	7.630	14.915
XV	3	psigmf	constant	0.7726	12.253	25.406
XVI	3	psigmf	linear	0.9218	7.629	14.892

From the perspective of average percentage error, model IV has achieved the lowest value if compared with other investigated models, recording at 14.2%, as shown in Fig. 8. In other words, it can achieve an average accuracy of around 86%. In general, no guideline was set for the range of acceptable error in certain engineering applications. However, a smaller error is always preferable. As this is the first attempt to introduce the use of ANFIS in predicting the final *E. coli* concentration for stormwater biofilters application, such an average accuracy is encouraging.

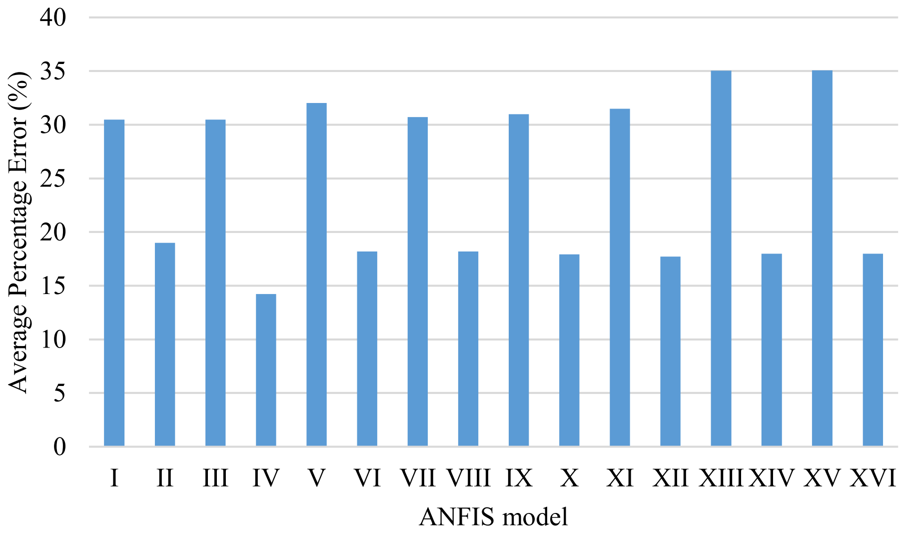

Fig. 8. Average percentage error of the ANFIS model.

In short, model IV is the model with the best performance while evaluating through the series of analytical analyses. It exhibits the highest R^2-value of 0.9874, the lowest MAE and RMSE of 3.854 and 6.004 respectively, and the smallest average percentage error of 14.2%.

3.2.3 AI Models Comparison

This study investigates both FBNN and ANFIS as the advanced methods to predict the final E. coli concentrations. After conducting the performance evaluation through a series of statistical analyses, the best-performed model for each approach was identified. Table 3 shows the comparison of the selected model in terms of R^2, MAE, RMSE and average percentage error.

Table 3: Comparison between the best-performed FBNN and ANFIS model

Model	R^2	MAE	RMSE	Average percentage error
FBNN model IV	0.9285	4.204	7.322	27.5
ANFIS model IV	0.9874	3.854	6.004	14.2

Among all the examined statistical indicators, ANFIS model IV achieved a better performance than that of FBNN model IV. Overall, it shows an improvement from FBNN to ANFIS. The most significant enhancement can be seen from the aspect of the average percentage error. The value has been reduced from 27.5% to 14.2%, showing an improvement of around 50%.

The architecture of the best-performed ANFIS model is therefore described as follows:

- Network inputs: Grass-type, the thickness of biofilters, initial *E. coli* concentration

- Network output: Final *E. coli* concentration
- Number of membership functions: 3
- Input membership function: trapmf
- Output membership function: Linear
- Optimization method: Hybrid

4. CONCLUSIONS

The main purpose of this study was to develop an artificial intelligence (AI) model to serve as an alternative to predict the final E. coli concentration in the application of stormwater biofilters in stormwater management practices. Both feedforward backpropagation neural network (FBNN) and adaptive neuro-fuzzy inference system (ANFIS) models have seen their application in different fields of study, especially while dealing with non-linear regression problems. Both techniques are appropriate for this task because it has the capability to learn the relationships between input-output variables for a complex physical relationship and hence provide an output with a considerably high level of accuracy.

In this study, it is found that a single-layer feedforward backpropagation neural network (FBNN) with 17 hidden neurons to be the most suitable model for the final E. coli concentration prediction. Meanwhile, the ANFIS model with the number of membership function of 3, input trapmf membership function and output linear membership function has shown the best performance among the examined models. The selection of the models was supported by the results of a range of statistical analyses. The selected FBNN model and ANFIS model were then further compared using the same series of statistical analyses to investigate their appropriateness to achieve the main goal of this study.

In conclusion, the ANFIS model appears as the more suitable model for the final E. coli concentration prediction after comparing it with the selected FBNN model. In short, ANFIS is an effective tool to provide a more accurate simulation of the non-linear behavior between the final E. coli concentration and the factors affecting it. With such a model, it allows the user to determine the final E. coli concentration and thereby the removal percentage of the biofilters application by inserting the relevant input parameters into the model. Since this study is seen as the first attempt to implement the artificial intelligence techniques in predicting the final coliform concentration under different stormwater biofilter conditions, two basic techniques (BPNN and ANFIS) were chosen for the model development. To further enhance the model performance in terms of accuracy and effectiveness, the integration of optimization algorithms such as genetic algorithm (GA), ant colony algorithm (ACO), etc. to the proposed model can be performed.

ACKNOWLEDGEMENTS

This work was supported by University of Malaya under Postgraduate Research Grant (PPP) (Grant No. 4576). The authors would like to thank the Faculty of Engineering, University of Malaya for providing space to conduct the biofilter column experiments and Department of Parasitology, Faculty of Medicine, University of Malaya for the laboratory facilities and technical assistance provided for their experiments.

https://doi.org/10.31436/iiumej.v23i2.2173

REFERENCES

[1] Payne EG, Fletcher TD, Cook PL, Deletic A, Hatt BE. (2014) Processes and drivers of nitrogen removal in stormwater biofiltration. Crit. Rev. Environ. Sci. Technol., 44: 796-846.

[2] Kanda R, Kishimoto N, Hinobayashi J, Hashimoto T, Tanaka S, Murakami Y. (2017) Influence of temperature and COD loading on biological nitrification-denitrification process using a trickling filter: An empirical modeling approach. Int. J. Environ. Res., 11: 71-82.

[3] Samhan SA, Al-Sa'ed RM, Mahmoud NJ. (2007) Removal of pathogenic microorganisms in pilot-scale UASB-septic tanks and Albireh urban wastewater treatment plant in Palestine. Water Int., 32: 798-809.

[4] Wang R, Li X, Feng Y, Tariq F, Li K, Wei Y, . . . Chen L. (2019) Removal of formaldehyde from the air with a suspended growth bioreactor. Int. J. Environ. Res., doi:https://doi.org/10.1007/s41742-019-00228-2

[5] Chandrasena GI, Pham T, Payne EG, Deletic A, McCarthy DT. (2014) *E. coli* removal in laboratory scale stormwater biofilters: Influence of vegetation and submerged zone. J. Hydrol., 519: 814-822

[6] Li Y, McCarthy DT, Deletic A. (2016) *Escherichia coli* removal in copper-zeolite-integrated stormwater biofilters: Effect of vegetation, operational time, intermittent drying weather. Ecol. Eng., 90: 234-243.

[7] Chandrasena GI, Shirdashtzadeh M, Li Y, Deletic A, Hathaway, JM, Mccarthy DT. (2017) Retention and survival of E. coli in stormwater biofilters: Role of vegetation, rhizosphere microorganisms and antimicrobial filter media. Ecol. Eng., 102: 166-177.

[8] Galbraith P, Henry R, McCarthy DT. (2019) Rise of the killer plants: investigating the antimicrobial activity of Australian plants to enhance biofilter-mediated pathogen removal. J. Biol. Eng., 13: 52.

[9] Hunt WF, Smith JT, Jadlocki SJ, Hathaway JM, Eubanks PR. (2008) Pollutant removal and peak flow mitigation by a bioretention cell in urban Charlotte, N.C. J. Environ. Eng., 134: 403-408.

[10] Bu CH, Lai SH, Goh XT, Chong WT, Chin RJ. (2021) Influence of filter media depth and vegetation on Faecal Coliform removal by stormwater biofilters. Water Environ. J., 35: 181-189.

[11] Yetilmezsoy K, Ozkaya B, Cakmakci M. (2011) Artificial intelligence-based prediction models for environmental engineering. Neural Netw. World, 3: 193-218.

[12] Chau, K-w. (2006) A review on integration of artificial intelligence into water quality modelling. Mar. Pollut. Bull., 52: 726-733.

[13] Thompson KA, Dickenson ER. (2021) Using machine learning classification to detect simulated increases of de facto reuse and urban stormwater surges in surface water. Water Res., 204: 117556.

[14] Fang H, Jamali B, Deletic A, Zhang K. (2021) Machine learning approaches for predicting the performance of stormwater biofilters in heavy metal removal and risk mitigation. Water Res., 200: 117273.

[15] DID Malaysia (2012) Urban stormwater management manual for Malaysia 2nd Edition. Kuala Lumpur, Department of Irrigation and Drainage Malaysia.

[16] Chin RJ, Lai SH, Shaliza I, Wan Zurina WJ, Ahmed Elshafie AH. (2019) New approach to mimic rheological actual shear rate under wall slip condition. Eng. Comput., 35: 1409-1418.

[17] Deng B, Chin RJ, Tang Y, Jiang C, Lai SH. (2019) New approach to predict the motion characteristics of single bubbles in still water. Appl. Sci., 9: 3981.

[18] Alimissis A, Philippopoulos K, Tzanis CG, Deligiorgi D. (2018) Spatial estimation of urban air pollution with the use of artificial neural network models. Atmos. Environ., 191: 205-213.

[19] Zhang Y, Chen H, Yang B, Fu S, Yu J, Wang Z. (2018) Prediction of phosphate concentrate grade based on artificial neural network modeling. Results Phys., 11: 625-628.

[20] Akter T, Desai S. (2018) Developing a predictive model for nanoimprint lithography usingartificial neural networks. Mater. Des., 160: 836-848.

[21] Chin RJ, Lai SH, Shaliza I, Wan Zurina WJ, Elshafie A. (2019) Rheological wall slip velocity prediction model based on artificial neural network. J. Exp. Theor. Artif. Intell., 31: 659-676.

[22] Pham DT, Sagiroglu S. (2001) Training multilayered perceptrons for pattern recognition: a comparative study of four training algorithms. Int. J. Mach. Tools Manuf., 41: 419-430.

[23] Demuth H, Beale M. (2014) Neural Network Toolbox for Use with MATLAB User's Guide Version 4. Natick, MA, The MathWorks Inc.

[24] Sharma B, Venugopalan K. (2014) Comparison of Neural Network Training Functions for Hematoma Classification in Brain CT Images. IOSR J. Comput. Eng., 16: 31-35.

[25] Alsmadi MK, Omar KB, Noah SA (2009) Back propagation algorithm: The best algorithm among the multi-layer perceptron. Int. J. Comput. Sci. Netw., 9: 378-383.

[26] Chin RJ, Lai SH, Shaliza I, Wan Zurina WJ, Elshafie A. (2020) ANFIS-based model for predicting actual shear rate associated with wall slip phenomenon. Soft Comput., 24: 9639-9649.

[27] Ghasemi E, Kalhori H, Bagherpour R. (2016) A new hybrid ANFIS–PSO model for prediction of peak particle velocity due to bench blasting. Eng. Comput., 32: 607-614.

[28] Korotkikh V, Korotkikh G (2008) On principles in engineering of distributed computing systems. Soft Comput., 12: 201-206.

[29] Jang JS (1993) Adaptive network-based Fuzzy Inference System. IEEE Trans. Syst. Man Cybern. Syst., 23: 665-685.

[30] Jang JS, Sun CT. (1997) Neuro-Fuzzy and soft computing: A computing approach to learning and machine intelligence, Englewood Cliffs, New Jersey, Prentice Hall.

[31] Jang JS, Sun CT, Mizutani E. (1997) A computational approach to learning and machine intelligence. Neuro-Fuzzy and Soft Computing. Upper Saddle River, New Jersey, Prentice-Hall.

[32] Barrett ME, Limouzin M, Lawler DF. (2013) Effects of media and plant selection on biofiltration performance. J. Environ. Eng., 139: 462-470.

[33] Read J, Fletcher TD, Wevill T, Deletic A. (2009) Plant traits that enhance pollutant removal from stormwater in biofiltration systems. Int. J. Phytoremediation, 12: 34-53.

[34] Kim MH, Sung CY, Li MC (2012). Bioretention for stormwater quality improvement in Texas: removal effectiveness of *Escherichia coli*. Sep. Purif. Technol., 84: 120-124.

[35] Li YL, Deletice A, Alcazar L, Bratieres K, Fletcher TD, McCarthy DT. (2012) Removal of *Clostridium perfringens*, *Escherichia coli* and F-RNA coliphages by stormwater biofilters. Ecol. Eng., 49: 137-145.

[36] Hermawan AA, Chang JW, Pasbakhsh P, Hart F, Talei A. (2018) Halloysite nanotubes as a fine grained material for heavy metal ions removal in tropical biofiltration systems. Appl. Clay Sci., 160: 106-115.

[37] Sileshi R, Pitt R, Clark S. (2016) Prediction of flow rates through various stormwater biofilter media mixtures. Proceedings of the World Environmental and Water Resources Congress, West Palm Beach.