Nonlinear Auto-Regressive with Exogenous Inputs (NARX) Modelling of Hydroponics Water pH Level in response to Acid and Alkaline Solutions

Mohammad Farid Saaid [1], Mohd Khairi Nordin [1], Ihsan Mohd Yassin [2], Nooritawati Md Tahir [1,3]

[1] *School of Electrical Engineering, Universiti Teknologi MARA (UiTM), Shah Alam, Malaysia*
[2] *Microwave Research Institute, UiTM, Shah Alam, Malaysia*
[3] *Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), UiTM, Shah Alam, Malaysia*

Abstract **– Nutrients are essential to optimising plant growth. However, the introduction of fertiliser in a hydroponics setup influences the pH level of the nutrient solution. This, in turn, could affect plants' growth as many types of plants require a specific pH range to grow optimally. Conventional hydroponics cultivation performs pH adjustment manually – a meticulous and error-prone process. Manual adjustment of pH solutions is prone to estimation errors, particularly when the pH levels change drastically due to the slow response of the solution to the addition of alkaline or acidic mixtures and sensitivity to minute errors in mixture delivery. For these reasons, a model to estimate the solution's pH would help improve the delivery accuracy of the alkaline and acidic mixtures. Past research offers minimal study to optimally construct the model from a System Identification (SI) perspective. This study represents a pH water neutralisation behaviour using the Nonlinear Autoregressive model with Exogeneous Inputs (NARX). The project begins with input and output data acquisition, leading to the development of the NARX model. Model performance was then evaluated by analysing the model fit and residual distribution.**

Keywords **– NARX, pH level, acid and alkaline, hydroponic, system identification.**

DOI: 10.18421/TEM123-04
https://doi.org/10.18421/TEM123-04

Corresponding author: Mohammad Farid Saaid,
School of Electrical Engineering, College of Engineering, UiTM Shah Alam, Malaysia
Email: mfarids@uitm.edu.my

Received: 05 April 2023.
Revised: 21 June 2023.
Accepted: 04 July 2023.
Published: 28 August 2023.

1. Introduction

Hydroponics is an innovative agricultural technique that utilizes nutrient-rich water to cultivate plants. This technology is gaining traction in the agricultural industry due to its cost-effectiveness, sustainability, and efficiency. Hydroponics is more cost-effective than traditional farming because it requires less land, fewer resources, and involves less labour and maintenance. Additionally, hydroponic farming can be done in a variety of places, from greenhouses to warehouses, making it a very versatile and efficient way to grow produce [1].

As more people become aware of the benefits of hydroponic farming, the demand for hydroponically grown produce is expected to grow. Additionally, advances in hydroponic technology are making the process more efficient and cost-effective. This, combined with the growing demand for locally grown produce, is likely to result in increased investment in hydroponic farming.

Nutrients are required for plant growth in addition to temperature, pH, water, light, and humidity. To thrive, plants require a specific pH range [1]. Acidic solutions contain more hydrogen ions because of water molecules dissociating to release hydrogen ions during nitrification. Alkaline solutions, on the other hand, contain a greater concentration of hydroxyl ions [2]. The pH scale ranges from zero to 14, with seven indicating neutral pH such as in distilled water, above seven indicates alkaline, and pH below seven as acidic [3].

In a hydroponics setup, the pH value of the solution varies according to the chemical composition present in the solution. This composition can change because of the addition of nutrients (fertiliser) to the water. Adjusting the pH range of water is critical for optimal plant growth as it increases the plants' capacity for nutrient absorption [4]. Additionally, it is critical to monitor and precisely control the pH level because the optimal range for plants is very narrow (between 0.5 and 1). As a result, hydroponic farmers frequently adjust the pH of the water using acid or alkaline solutions.

Controlling pH neutralisation is difficult but necessary due to the process's highly nonlinear nature. The requirement to maintain a specific pH value is based on environmental, legislative, and quality standards [5]. Traditionally, pH levels are manually measured using litmus paper and a pH meter. Manual pH solution adjustment is prone to estimation errors, resulting in dramatic changes in pH levels as the acid and alkaline mixture is gradual (resulting in a nonlinear behaviour with a slow response time). A computerised approach would allow farmers to effectively improve monitoring and control of plant growth. To achieve this, the first step is to develop a mathematical model representing the pH change in water, which will eventually result in the development of the automated pH control system.

As a result of these considerations, an artificial intelligence-based model was developed to estimate the pH of the solution to improve the accuracy of the alkaline and acidic mixtures' delivery. Although some research has been conducted previously, relatively few studies have been conducted to optimise the model's construction from a System Identification (SI) perspective. Therefore, the novelty of this study attempts to establish a pH water neutralisation behaviour using the Nonlinear Autoregressive model with Exogenous Inputs (NARX) model, a popular and robust control engineering method for identification of systems based on the input and output response(s) [6].

In SI, input and output pairs are collected at regular intervals, and a predictive model is built to represent the system's behaviour base on these pairs. Artificial Intelligence (AI)-based models created using SI are adaptive to various situations and able to readjust control measures to change within the modified environment [7]. NARX, among the many types of SI models, is a highly robust and practical regression model for time series prediction that uses lagged values of the output (target) and the input (exogenous) variables as inputs to predict the future values of the target variable [8].

NARX is well suited for nonlinear time series with long-term dependencies and multiple inputs, making them a useful tool for complex forecasting tasks. Its benefits for time series modelling and nonlinear systems have been widely studied [9], [10], [11], [12], [13].

Applications of NARX in hydroponics and pH modelling are reviewed next. Reference [13] used a NARX model to predict the response of chilli pepper plants to changes in root zone temperature in a hydroponics setup. The model was comprised of a Single Input Singe Output (SISO) system tested on five datasets of plant growth responses. The model was able to forecast the root zone temperature accurately, albeit limited robustness due to lack of data on plants' responses to extreme conditions.

Rattanawaorahirunkul et al. [14] built a pH model using several non-linear SI models. The Hammerstein-Wiener, NARX, Autoregressive model with Exogenous Inputs (ARX), and state-space model. Their study validated the model using conventional model fit tests (fit percentage, Mean Absolute Error (MAE), and Mean Squared Error (MSE). However, the work did not investigate the residual distribution of the models, which is a critical part of the modelling process to disprove model bias.

A similar study by Ferentinos [15] developed a pH model without a strong validation of the residuals, but instead focused on discovering the optimal solution to the parameter-dependent problem by minimising a cost function related to the problem [16]. Xia [17] developed a new non-linear optimisation algorithm for Model Predictive Control (MPC). The model was used for pH set-point tracking optimization by combining Chaotic Local Search (CLS) and Roulette Wheel Mechanism (RWM). Testing related to the residuals were not extensively deliberated.

The uniqueness of our study compared to previous works is derived from the extensive validation of the NARX model for pH forecasting in a hydroponics setup. Along with standard model fitting tests (such as the One Step Ahead (OSA) prediction, we used a variety of methods to validate the randomness of the residuals. Random residuals indicate that the model is properly capturing the underlying relationships in the data and is not overfitting. Validation tests such as correlation tests and histogram analysis are an important diagnostic tool for testing the validity of a predictive model. Random residuals, as any non-random patterns in the residuals may indicate that the model is not accurately capturing the underlying relationships in the data. These tests are critical for practitioners to use as a basis and reference when selecting a valid and unbiased model to represent the system.

Although this issue may appear trivial to some, SI theory states that if the model's residuals are not random, this indicates the presence of unmodeled dynamics. This could result in a variety of issues regarding the validity of model predictions, most notably model bias.

The rest of this article is organised as follows: Section 2 presents the research methodology, followed by the results in Section 3. Finally, concluding remarks are presented in Section 4.

2. Research Method

This section summaries the research methodology. The model-building process began with data collection and was followed by preprocessing to enhance the model's learning ability. Following that, the model construction process is described, followed by information about how the validation tests were conducted.

a) Data Collection

The input and output data was collected from a previously-developed hydroponics apparatus [18]. As illustrated in Figure 1, the developed apparatus was composed of three major components: an input, a microcontroller, and an output. Two syringe pumps were used to drip the acid and alkaline solution randomly. The main container was dripped with either acidic (pH down) or alkaline (pH up) solution at random intervals via a two-syringe pump mechanism. Subsequently, the stepper motor was connected to the syringe via a lead screw rod. The stepper motor rotates the lead screw rod, which advances the syringe pump linearly.

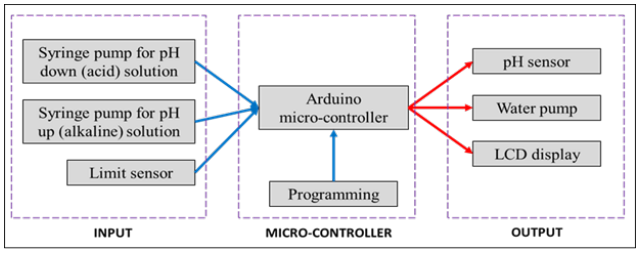

Figure 1. Block diagram of the automated random drip and alkaline system

By activating both syringes in random order, the effect of acid and alkaline solutions on pH was determined. Random activations result in a change in the pH value, which serves as the study's output. For the duration of the experiment, the syringes were activated every minute, and pH samples were taken every second. The dataset used in this study comprised of 2,963 data points.

b) Data Pre-processing

Preprocessing is a critical step in machine learning research. Properly preprocessed data enables the models to learn more precisely, resulting in a better result [19].

Tangent-sigmoid activation functions are used in the MLP-NARX model because they provide a non-linear activation function which helps to capture complex non-linear relationships in data. This non-linear activation helps to increase the accuracy of the network and enables it to solve complex problems. In addition, the tangent-sigmoid activation function helps to ensure that the outputs of the neurons remain in a range of -1 to 1 which makes the learning process more efficient. It is critical to avoid having the trained network weights be extremely large or extremely small, as this can result in a reduction in the resulting performance. During data collection, pH values ranged from 3.89 to 9.9. These values are outside the activation range of the MLP's tangent-sigmoid activation function, and thus need to be scaled to between -1 and +1.

The MLP-NARX model requires training, testing and validation sets because they are complex algorithms that require a lot of data to train properly. The training set is used to train the model so that it can learn the patterns and relationships in the data. The testing set is used to measure the performance of the model and see how well it generalizes to new data. Finally, the validation set is used to fine-tune the model and make sure it is not overfitting or underfitting the data. By splitting the data into these three sets, the model can be trained and tested more effectively. For this experiment, the dataset was divided into three groups, with a 70% split for training, 15% for testing, and 15% for validation. This ratio is commonly used by many researchers who work in similar fields.

c) Model Construction

The NARX model can be represented using various methods. However, for this experiment, the Multi-Layer Perceptron (MLP) model was selected to represent the NARX structure because of its ability to automatically learn the relationships between the input and output. Additional evidence has been provided that the MLP is both a robust and accurate estimator in a variety of SI applications [20].

The Levenberg-Marquardt (LM) algorithm was used to train the MLP-NARX model. It is a numerical optimization technique used to adjust parameters of a model to minimize the sum of square errors between observed and model-predicted values.

LM is a combination of the Gauss-Newton and gradient descent algorithms and is used for problems involving nonlinear least squares. The algorithm works by first approximating the local curvature of the error surface and then using this to determine an update direction for the parameters. It then adjusts the step size and direction of the update to find an optimal solution. LM has been shown to be effective at approximating functions in a test by [21].

The model structure of the MLP-NARX model is shown in Figure 2, described in Eq. (1):

$$y(t) = f\big[y(t-1), y(t-2), \ldots\ldots, y\big(t-n_y\big), u(t-n_k-1), u(t-n_k-2), \ldots, u(t-n_k-n_u)\big] + \varepsilon(t) \quad (1)$$

where $y(t)$ is the MLP-NARX model output, f is the MLP-NARX model, and $\varepsilon(t)$ are the residuals. The model structure that utilized past inputs and outputs of the actual system to predict its future values. Four variables can be a
djusted for the MLP-NARX model, namely the alkali and acidic syringe control inputs, the pH value, and the number of hidden units. A brute-force trial and error method were used to determine the optimal parameters for the model. The trial-and-error method process used a sequence combination beginning from 5:5:5:5 to 30:30:30:50 (for input1/acid (nu_1): input2/alkaline (nu_2): output/pH (ny): and hidden nodes (h), generating a total of 175,760 different candidate solutions. The best optimal lag space and hidden nodes have been discussed in Section III.

d) Model Validation

The test to measure the ability of a model to predict the one step ahead output is referred to as One Step Ahead (OSA). The differences between the actual data and OSA predictions are referred to as residuals, and they can be calculated using a variety of methods, including MSE and r-squared analysis.

MSE measures the quality of the model for data prediction. The error between the actual and predicted output can be calculated according to Eq. (2):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\big(y(t) - \hat{y}(t)\big)^2 \quad (2)$$

where $\hat{y}(t)$ and $y(t)$length n is the number of data, $\hat{y}(t)$ is the predicted value and $y(t)$ is the observed value at point i.

The r-squared test [22], also known as the coefficient of determination, represents the statistical measure known as the goodness of a model. The quality of linear regression will produce a value of R-Squared approaching one (3):

$$r^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \acute{y}_i)^2}{\sum_{i=1}^{n}(y_i - \acute{y}_i)^2} \quad (3)$$

where length n is the number of data, $\acute{y}_i$ is the predicted value and y_i is the observed value at point i.

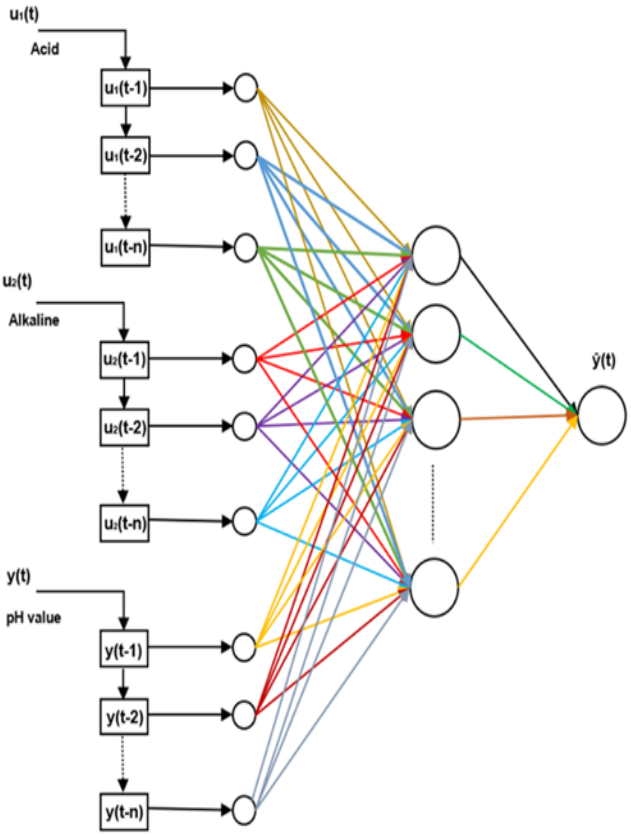

Figure 2. Time series MLP-NARX

Finally, correlation tests are important to measure the correlation between the residuals and other data in the time series. The result from this test indicates the time-series dependencies and the degree of correlation of the sequence to the residuals. The common version of the correlation tests are the autocorrelation and cross-correlation, as described in Eq. (4) and Eq. (5):

$$\theta_{\varepsilon\varepsilon}(\tau) = [\varepsilon(t-\tau)\varepsilon(t)] = \delta(\tau) \quad (4)$$

$$\theta_{y\varepsilon}(\tau) = E \quad (5)$$

where $\boldsymbol{\theta_{\varepsilon\varepsilon}(\tau)}$ and $\boldsymbol{\theta_{y\varepsilon}(\tau)}$ are the correlation coefficients with respect to the lag, $\boldsymbol{\tau}$. The Kronecker Delta, $\boldsymbol{\delta(\tau)}$, is defined as

$$\delta(\tau) = \begin{cases} 1, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases}$$

The Kronecker delta is a generalization of the Dirac delta function, a mathematical operator used to express the discrete delta function in terms of its components. It is defined as a function that takes two integers as arguments and returns one if they are equal, and zero otherwise.

If the correlation coefficient falls within the 95 percent confidence interval, the model is recommended as acceptable based on these correlation tests.

Additionally, in this study, a histogram of the residuals was used to analyse the distribution of the residuals. A bell-shaped curve indicates that the residuals are normally distributed (random).

3. Results and Discussion

This section discusses the results of modelling using NARX structure selection. The lags and number of hidden units in the MLP-NARX are adjustable parameters. We tested many combinations of these parameters to obtain the best-performing one. This was done by first shifting the input signals (acid (n_{u1}), alkaline (n_{u2}), pH (n_y)) by five lags, then gradually increasing the lag value by five until 30. The input signals are further used to determine the optimum hidden node (h) from five to 50 nodes. The selection criteria for the tests were the MSE and magnitude of correlation test violations (the magnitude of correlation coefficients outside the 95% confidence limits in the autocorrelation and cross-correlation tests) in the model. The number of parameter combinations was 175,760. The top ten best-performing parameter combinations are shown in Table 1, sorted in ascending order based on the magnitude of correlation coefficient violations during the auto and cross-correlation tests.

The lowest correlation violation value was found in the first row of Table 1. However, due to its relatively high MSE value when compared to the other parameter combinations, it was not selected as the optimal parameter combination. The optimal parameter combination was chosen from the sixth row of Table 1 ($n_{u1}:n_{u2}:n_y:h = 13:5:14:15$). Although this combination was ranked sixth in terms of correlation, the low total MSE value makes it a more appealing combination compared to the rest. The optimal parameter combination for second and third positions, respectively, were ($n_{u1}:n_{u2}:n_y:h = 19:18:24:15$) and ($n_{u1}:n_{u2}:n_y:h = 29:25:30:50$). The MLP-NARX with the optimal parameter combination of ($n_{u1}:n_{u2}:n_y:h = 13:5:14:15$) will be discussed in greater detail beginning from Section III-A.

a) Data Collection

Figure 3 shows the raw input-output data used for training, testing, and validation. The input data ranged from 0 to 1, representing the dripping activity and volume of acidic and basic solutions. According to the total quantity of collected data (2,963), the automatic random number was generated for a greater proportion of acid solution syringe pump activation cases than alkaline cases, as the pH value decreased with time.

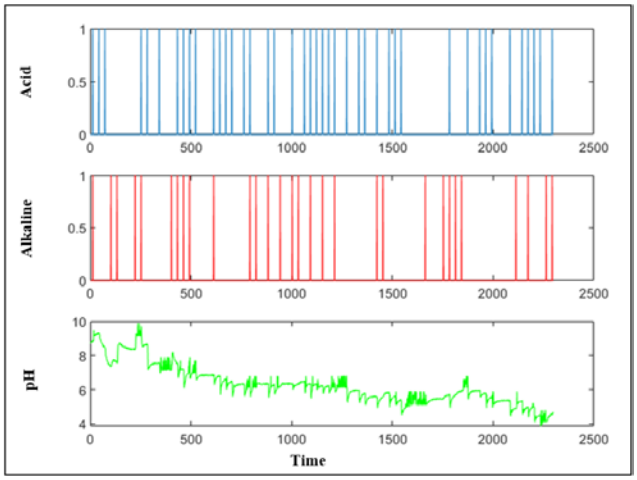

Figure 3. Raw input-output data (acid, alkaline and pH)

Figure 4 demonstrates that data pre-processing was necessary for the model to accurately train and generate better results. The data magnitudes were scaled to a range between -1 and 1, which avoids network weights from being excessively large or tiny during training, which would lead to a decrease in performance.

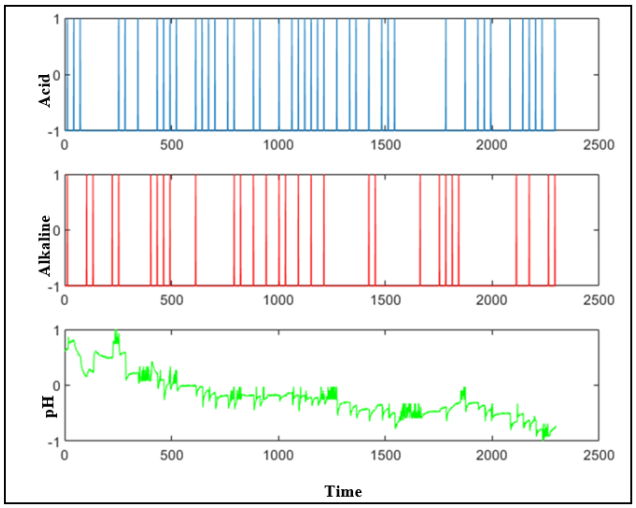

Figure 4. Pre-processed input-output data (acid, alkaline and pH)

Table 1. Top ten best performing lags space and hidden nodes combination based on the total magnitude of correlation violations.

	Acid (nu1)	Alkaline (nu2)	pH (ny)	Hidden nodes (h)	R^2 train	R^2 test	R^2 val.	Total Corr.	Total MSE
1	27	23	27	15	0.9880	0.9871	0.9357	0.0190	1.83×10^{-2}
2	19	18	24	15	0.9853	0.9914	0.9601	0.0381	1.26×10^{-2}
3	17	11	20	35	0.9885	0.9880	0.9435	0.0445	1.67×10^{-2}
4	29	25	30	50	0.9837	0.9917	0.9649	0.0463	1.17×10^{-2}
5	14	29	29	10	0.9859	0.9912	0.9556	0.0480	1.27×10^{-2}
6	13	5	14	15	0.9839	0.9951	0.9783	0.0481	8.30×10^{-3}
7	29	23	30	35	0.9887	0.9929	0.9668	0.0505	9.72×10^{-3}
8	21	27	29	50	0.9821	0.9932	0.9660	0.0530	1.13×10^{-2}
9	6	25	30	50	0.9819	0.9906	0.9614	0.0551	1.32×10^{-2}
10	26	26	30	35	0.9820	0.9919	0.9654	0.0556	1.14×10^{-2}

b) One Step Ahead (OSA) Test

As depicted in Figure 5, the MLP-NARX model's training, testing, and validation of OSA results all occur simultaneously. It was found that the predicted pH values were consistent with the measured pH values. The lag values from the best model indicate that it takes a significant amount of time (more than one minute) for the pH mixture to be realised before any changes are detected by the pH sensing device. Additionally, the presence of previous pH values (lagged y terms) in the input implies that there is a relationship between past and future pH values.

Following the validation of the OSA graph using the results of the MLP-NARX model, it was discovered that there were limitations in prediction when sharp changes occurred in the pH values. These can be observed particularly at the 50, 120, 240, and 270 in Figure 5. There were some improvements in model fit for most of the predictions when using OSA training data, but the results were not statistically significant. This is to be expected, given that the network was trained using the same training data as the previous networks. This would have led to a network outperforming previously unseen validation and testing sets on this dataset, if the network had been used.

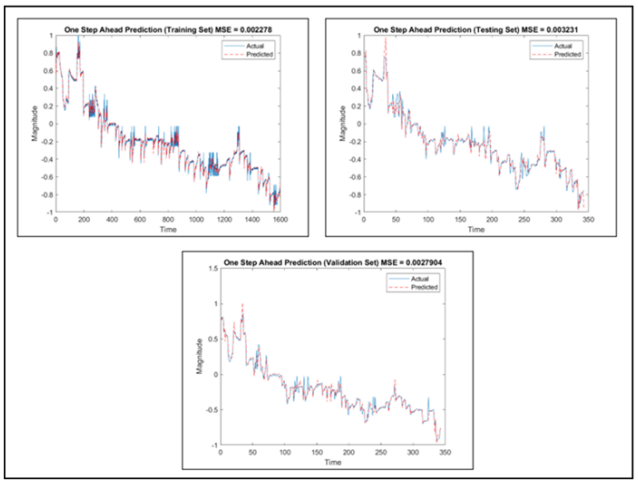

Figure 5. One Step Ahead test for MLP-NARX model

c) Correlation Test

As mentioned previously, correlation tests help gauge the randomness of the residuals, which is an important criterion for validating the model. The autocorrelation test measures the correlation between the residuals and themselves, while the cross-correlation test measures the residuals' correlation against the output.

Autocorrelation analysis of the MLP-NARX model is illustrated in Figure 6. Except for lag 0, all correlation values remained within the 95% confidence intervals. Next, the MLP-NARX model cross-correlation with the output was tested (Figure 7). There were several correlation violations (or near violations) at certain lags (zero, three, four, six and nine). However, the magnitudes were relatively small. As shown in Table 1, the total correlation coefficient for the MLP-NARX model was only 0.0481 outside of the confidence interval region.

Based on these results, the model was deemed valid and acceptable.

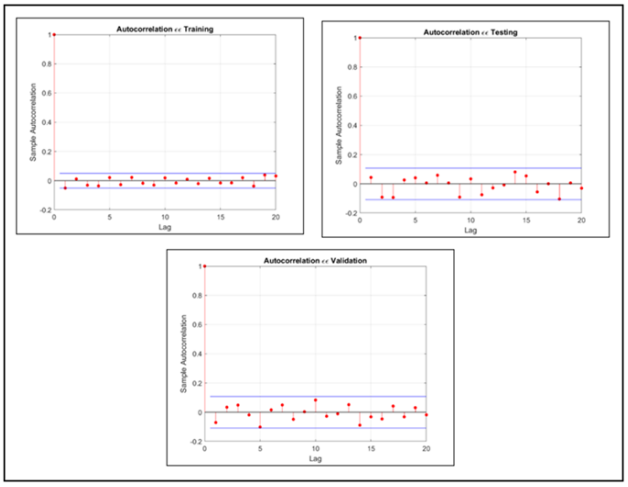

Figure 6. Autocorrelation test (ℰℰ) for MLP-NARX model

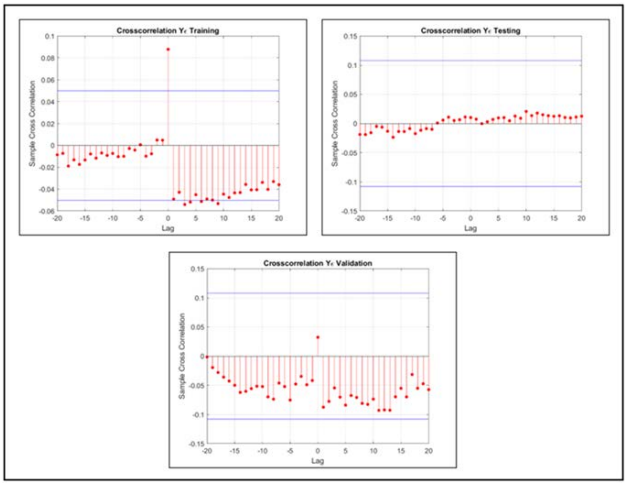

Figure 7. Cross-correlation test (Yℰ) for MLP-NARX model

4. Conclusion

An MLP-NARX model was developed to predict the non-linear behaviour of pH levels in a hydroponics nutrient solution in response to acid and alkaline solutions. Data collection and preprocessing was deliberated, followed by the model construction process. After 175,760 parameter combinations were tested, the best-performing MLP-NARX model was found to be $(n_{u1}{:}\,n_{u2}{:}\,n_y{:}\,h = 13{:}\,5{:}\,14{:}\,15)$. Further tests on the model fit, including correlation, MSE, R-square, and histogram, confirmed the model's performance and validity. The correlation test showed several small correlation violations, with minimal error, showcasing a good agreement between predicted and actual pH values. Findings from this thesis have proven that the developed model could achieve the objectives.

However, several recommendations for future work are also suggested. Firstly, the number of data samples can be increased to improve the model robustness as the MLP-NARX model may learn new relationships over a longer interval. Based on the experiments performed, the sampling rate was 1 Hz. This rate was considered adequate as any dripping of the alkaline or acidic solutions appears to take 60 seconds before the pH reading stabilizes. The MLP-NARX model's training process could learn more about the relationship between acid and alkaline and the influence of acid and alkaline on the pH value.

Acknowledgements

The authors acknowledge the Universiti Teknologi MARA for funding under the MyRA Research Grant (600-RMC/MYRA 5/3/LESTARI (005/2020).

The authors would also like to thank the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, for conducting this research

References:

[1]. Belmehdi, O., El Harsal, A., Benmoussi, M., Laghmouchi, Y., Skali Senhaji, N., & Abrini, J. (2018). Effect of light, temperature, salt stress and pH on seed germination of medicinal plant Origanum elongatum (Bonnet) Emb. & Maire. *Biocatalysis and Agricultural Biotechnology*, *16*, 126-131.

[2]. Alves, L. A., Denardin, L. G. d. O., Martins, A. P., Anghinoni, I., Carvalho, P. C. d. F., & Tiecher, T. (2019). Soil acidification and P, K, Ca and Mg budget as affected by sheep grazing and crop rotation in a long-term integrated crop-livestock system in southern Brazil. *Geoderma*, *351*, 197-208.

[3]. Zhu, M., Zhang, D., Lu, Q., Yan, Y., Zhu, K., Ye, K., Yan, J., Cao, D., Wang, Q., Huang, X., & Wang, G. (2021). Hollow hexagonal NiSe–Ni3Se2 anchored onto reduced graphene oxide as efficient electrocatalysts for hydrogen evolution in wide-pH range. *International Journal of Hydrogen Energy*, *46*(39), 20524-20533.

[4]. Ávila, P. F., Ferreira da Silva, E., & Candeias, C. (2017). Health risk assessment through consumption of vegetables rich in heavy metals: the case study of the surrounding villages from Panasqueira mine, Central Portugal. *Environmental Geochemistry and Health*, *39*(3), 565-589.

[5]. Abdullah, N. H. S., Karsiti, M. N., & Ibrahim, R. (2012). A review of pH neutralization process control. In *2012 4th International Conference on Intelligent and Advanced Systems (ICIAS2012)*, 594-598.

[6]. Quintana, R., Qiu, X., Benjamin, A., Panicker, P., Schmidt, G., & Hübler, A. C. (2021). The printed piezoelectric loudspeaker model: A modal perspective from system identification. *Applied Acoustics*, *174*, 107722.

[7]. Chakravorti, T., & Satyanarayana, P. (2020). Non linear system identification using kernel based exponentially extended random vector functional link network. *Applied Soft Computing*, *89*, 106-117.

[8]. Laadissi, E. M., El Filali, A., & Zazi, M. (2018). A nonlinear TSNN based model of a lead acid battery. *Bulletin of Electrical Engineering and Informatics*, *7*(2), 169-175.

[9]. Lee, W. J., Na, J., Kim, K., Lee, C.J., Lee, Y., & Lee, J. M. (2018). NARX modeling for real-time optimization of air and gas compression systems in chemical processes. *Computers & Chemical Engineering*, *115*, 262-274.

[10]. Amirkhani, S., Tootchi, A., & Chaibakhsh, A. (2021). Fault detection and isolation of gas turbine using series–parallel NARX model. *ISA Transactions*, *120*, 205-211.

[11]. Cheng, J., Ji, Z., Li, M., & Dai, J. (2020). Study of a noninvasive blood glucose detection model using the near-infrared light based on SA-NARX. *Biomedical Signal Processing and Control*, *56*, 101694.

[12]. Zhu, Y.-P., Lang, Z. Q., Mao, H.-L., & Laalej, H. (2022). Nonlinear output frequency response functions: A new evaluation approach and applications to railway and manufacturing systems' condition monitoring. *Mechanical Systems and Signal Processing*, *163*, 108179.

[13]. Aji, G. K., Hatou, K., & Morimoto, T. (2020). Modeling the Dynamic Response of Plant Growth to Root Zone Temperature in Hydroponic Chili Pepper Plant Using Neural Networks. *Agriculture*, *10*(6), 234.

[14]. Rattanawaorahirunkul, R., Sanposh, P., & Panjapornpon, C. (2016). Nonlinear system identification of pH process using Hammerstein-Wiener model. In *2016 International Conference on Electronics, Information, and Communications (ICEIC)*, 1-4.

[15]. Ferentinos, K. P., Albright, L. D., & Scott, N. R. (2000). Modeling pH and Electrical Conductivity in Hydroponics using Artificial Neural Networks. *IFAC Proceedings Volumes*, *33*(19), 173-178.

[16]. Mo, Y. (2012). Particle swarm optimization for cylinder helical gear multi-objective design problems. In *Applied Mechanics and Materials*, *109*, 216-221.

[17]. Xia, X. (2012). Nonlinear predictive control based on particle swarm optimization applied to pH neutralization reaction. *Proceedings 2012 International Conference on System Science and Engineering, ICSSE 2012*, 445-448.

[18]. Saaid, M. F., Yassin, A. I. M., & Tahir, N. M. (2020). Automated monitoring and controlling pH levels for hydroponics cultivation technique. *Indonesian Journal of Electrical Engineering and Computer Science*, *18*(3), 1236-1243.

[19]. Felix, E. A., & Lee, S. P. (2019). Systematic literature review of preprocessing techniques for imbalanced data. *IET Software*, *13*(6), 479-496.

[20]. Rachburee, N., & Punlumjeak, W. (2021). Oversampling technique in student performance classification from engineering course. *International Journal of Electrical and Computer Engineering (IJECE)*, *11*, 3567.

[21]. Khalfi, J., Boumaaz, N., Soulmani, A., & El Mehdi, L. (2021). An electric circuit model for a lithium-ion battery cell based on automotive drive cycles measurements. *International Journal of Electrical and Computer Engineering (IJECE)*, *11*, 2798.

[22]. Chum, P., Park, S. M., Ko, K. E., & Sim, K. B. (2012). Optimal EEG feature extraction based on R-square coefficients for motor imagery BCI system. In *2012 12th International Conference on Control, Automation and Systems*, 754-758.