Particle Swarm Optimization (PSO) Model for Hydroponics pH Control System

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

[1] *School of Electrical Engineering, College of Engineering, 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 optimise plant growth. However, adding fertiliser changes the pH of the nutrition solution. This would impact plant growth as each plant types requires a specific pH range to thrive. Due to the nonlinearity characteristics, pH neutralisation adjustment is difficult but essential. In addition, alkaline solutions are not completely dissociated due to the presence of acid. For these reasons, a mathematical model to estimate the solution's pH would help improve the alkaline and acidic delivery accuracy. This study represents a pH water neutralisation behaviour using Particle Swarm Optimisation algorithm (PSO). The project begins with input and output data acquisition leading to the development of the PSO model. The model fit and residual distribution have also been analysed for this model. The model's performance was accepted based on a correlation test because the lag signal exceeded 95% of the confidence interval. The model also recorded a very minimal error, and this proved that a good agreement is established between the predicted and actual pH values.**

Keywords – **PSO model, pH level, hydroponics, correlation test, acid and alkaline solution.**

1. Introduction

Many techniques can be used to grow plants. An alternative technique to grow plants is the hydroponics technique. Hydroponics is a technique to grow plants without using soil as a growing medium [1], [2]. The hydroponics method used mineral nutrient solutions in a water solvent to ensure the plant gets all the nutrients required from the water solution [3]. The absorption of water and nutrients takes place behind the root tip through modest root hairs. Temperature, water pH, light, and humidity are also necessary to enhance the plants healthy. Each plant needs a specific pH water level because the plant's growth and health will be risked if the pH level of the water is not in the required pH range.

A pH value changes according to chemicals mixed into them. In hydroponics applications, this can be the result of adding nutrients (fertiliser) to the water. Water pH range adjustment is an essential consideration for optimal growth for different plants [4]. This is because it maximises the plants' ability to absorb nutrients [5]. Therefore, hydroponics farmers typically adjust the water pH value using acid or alkaline solutions. Given that the optimal plant pH range is small (between 0.5 and 1), the pH level control is significant and needs to be monitored and controlled appropriately [6].

A computerised approach would increase the monitoring and control of this parameter. The initial step for this is to construct a mathematical model to represent the pH change in the water – leading to the development of the pH control system. The initial step for this is to construct a mathematical model to represent the pH change in the water, leading to the development of the pH control system. Thus, this paper was to discover optimal parameters for the NARX model using the Particle Swarm Optimisation algorithm (PSO). Model optimization is essential to search for an optimal ANN model structure. The optimization had a significant advantage, such as increased prediction accuracy [7], model stability [8], to perform in complex engineering structures [9], etc.

There were two categories of model optimization techniques; local and global techniques [10]. Local optimization techniques continue from an initial solution until a minimum is found. In contrast, global optimisation methods try to search the whole space of the problem to discover the absolute minimum [11]. Among the intensive study using global optimization techniques were the Genetic Algorithm and Particle Swarm Optimization [12].

Particle Swarm Optimization (PSO) technique introduced by Kennedy and Eberhart was inspired by a flock of birds motion [13]. Among the advantages of the PSO model are; good performance in detecting small targets [14], solved complex optimization problems [15], a consistent prediction and a high degree of accuracy [16]. Various researchers used the PSO technique in their research area, such as in image quality [17], robotic automation [18], medical [19], energy [20], etc. Besides that, the PSO model was prevalent in agricultural research [21]. But, there was a very limited study in the PSO model for hydroponics.

2. Research Method

The input and output data of the system were collected using the previously developed system [22]. Kennedy et al. [13] proposed a Particle swarm optimisation (PSO) algorithm to solve ongoing optimisation problems. The PSO was inspired by the movement of living things such as bird flocking and fish schooling that were generating many different movements. Birds or fish can move through a specific dimensional space without colliding with one another due to the best structure in their positioning and velocity. Since this organisational behaviour, it is suitable to solve unconstrained as well as constrained optimisation problems.

The PSO algorithm for particle velocity, V_{id} and position, X_{id} equation is given by [23]:

$$V_{id} = V_{id} + [C_1 \; x \; rand_1 \; (P_{best} - X_{id})] + [C_2 \; x \; rand_2 \; (G_{best} - X_{id})] \qquad (1)$$

$$X_{id} = X_{id} + V_{id} \qquad (2)$$

Where:

V_{id} = particle velocity
X_{id} = particle position
P_{best} = particle's best fitness
G_{best} = swarm's best solution
C_1 = cognition learning rate
C_2 = social learning rate
$rand_1$ and $rand_2$ = random numbers

Through the combination of every particle, the best data from the search space is recorded as P_{best}. The remaining best data in the past was recorded as G_{best}. The acceleration or learning rate of the particle is C_1 C_2 and C_3. The C_1 and C_2 are leaning if the learning followed a swarm algorithm or more on personal movement.

The parameters for the PSO experiments are shown in Table 1.

Table 1. PSO model experiment parameters

Parameter	value
Particles	5, 10, 15, 20, 25,
Random Seed	0, 20000, 30000
Max Iterations	50. 100, 150, 200, 250, 300, 350, 400, 450, 500
Cognition learning rate, c1	2
Social learning rate, c2	2
Particle velocity bounding parameter; Vmax, Vmin	1, -1
Particle position bounding parameters; Xmax, Xmin	1, 0
dims	5

The values of particles, random seed and maximum iterations were selected to test PSO convergence at different conditions and initial starting points. This is important because the PSO is a stochastic optimization algorithm that converges differently depending on its various parameters. In total, 150 parameter combinations were tested. The top combinations with the best results were selected for further analysis.

The cognition learning rate, c_1, and social learning rate, c_2 was both set to two to balance between the swarm and individual influences when updating particle positions. The values of x_{min} and x_{max} were set to 0 to 1, respectively, which means that the solutions will always be within this value. This avoids searches for x_{id} solutions beyond the acceptable range. The values will then be rescaled into their desired ranges in the fitness function. The values of v_{min} and v_{max} were also set accordingly to the maximum transition between x_{min} to x_{max} and x_{max} to x_{min}, respectively.

Figure 1 shows a flowchart of Particle Swarm Optimization (PSO) used in this research. The Mersenne-Twister pseudorandom number generator algorithm was used to initialize the initial particle positions randomly. The number stream is controlled by the seed, which allows the pseudorandom sequence to be duplicated if desired. With the same pseudorandom sequence, the experiments are to be repeatable if necessary.
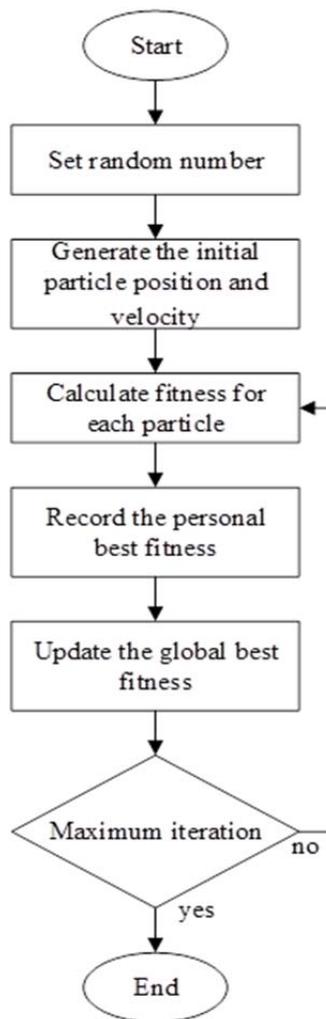
Figure 1. Particle Swarm Optimization (PSO) flowchart

Next, the initial particle velocities, v_{id}, and particle position, x_{id}, need to be generated, and the resulting fitness calculated. The problem dimension was set to five since there were five variables to be optimized (n_{u1}, n_{u2}, n_y, h, s). The personal best fitness, p_{best}, for each particle is then recorded, and global best fitness g_{best} is updated. The fitness values used by p_{best} and g_{best} were calculated based on the magnitude of correlation violations in training, validation, and testing sets. The search process was repeated until either the termination condition was met or the maximum number of iterations were met.

After the optimization has been completed, model fitness and residuals tests were used to measure the accuracy and validity of the model. For model accuracy tests, the One-Step-Ahead (OSA), MSE and R-squared tests were used. For residual tests, the correlation test and histogram analysis were used to analyze the randomness of the residuals.

3. PSO test Results

In this section, the results for PSO model validation have been discussed. The best combination of lags and hidden nodes were selected based on the lowest total correlation test and MSE value.

1) Testing for optimal lagging and node results

Table 2 shows the top twenty finest optimization results based on the magnitude of correlation violations. The best fitness achieved was 0.01503, with an MSE of 0.01213 at lags and hidden nodes of 29:30:30:49 (n_{u1}: n_{u2}: n_y: h). Based on the top three results, the least correlation violations occurred when the values of parameters were near their maximum. For example, the values of n_{u1}, n_{u2}. n_y and h were set to 30, 30, 30 and 50, respectively, and the three best solutions were consistently near these maximum values.

In terms of lag values, the models appear to require significantly delayed inputs and output to predict their future values, suggesting that the system has a long memory – its distant past values still have an influence on its future output. This observation is consistent with the nature of the pH control system, where pH values in the mixing tank require some time to change in response to introducing the acid and alkaline solution.

The high number of hidden nodes indicates that more nodes are required to learn the complex inter-relationships between the data. Potentially better results could be achieved if the maximum number of hidden nodes is increased. However, a search beyond this range is unfortunately beyond the computational capabilities of the hardware for this project.

The MSE values versus the correlation results are shown in Figure 2. The plot shows no significant relationship between them. This suggests that although both the MSE and the correlation values are affected by changes in the lags and hidden units, they are independent.
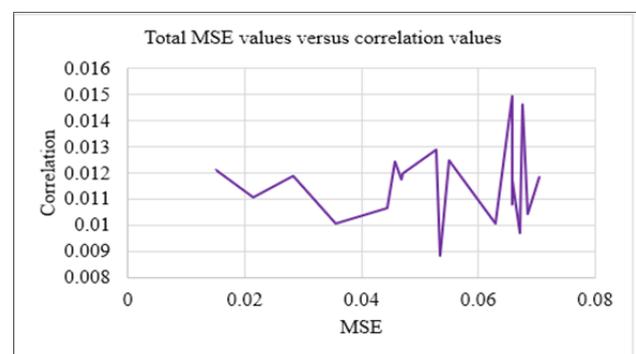

Figure 2. Relationship between correlation violations and MSE for top 20 PSO solutions

Table 2. Top twenty finest lags space and hidden nodes combination based on total correlation value for PSO

	Acid (nu1)	Alkaline (nu2)	pH (ny)	Hidden nodes (h)	Total R^2	Total Corr.	Total MSE	Seed
1	29	30	30	49	0.98035	0.01503	0.01213	98841
2	29	26	30	45	0.9800467	0.02148	0.01110	100000
3	30	29	30	50	0.9806067	0.02836	0.01189	27306
4	30	27	30	39	0.9830533	0.03560	0.01010	14577
5	30	5	30	50	0.98294	0.04447	0.01069	79041
6	5	30	30	43	0.9761567	0.04565	0.01244	45595
7	5	13	27	35	0.9814967	0.04675	0.01178	40282
8	16	5	27	50	0.9805467	0.04706	0.01200	43433
9	20	30	30	47	0.9773233	0.05272	0.01292	37338
10	19	5	20	48	0.9859267	0.05348	0.00887	40154
11	19	26	30	48	0.97749	0.05504	0.01248	100000
12	5	5	7	50	0.9838267	0.06301	0.01007	99342
13	5	30	30	50	0.9773533	0.06577	0.01497	79370
14	30	17	30	50	0.98107	0.06578	0.01080	44862
15	5	30	30	45	0.9779733	0.06579	0.01312	89658
16	30	16	30	45	0.9807733	0.06581	0.01175	1635
17	30	5	30	50	0.9842367	0.06712	0.00971	68006
18	30	29	30	50	0.97269	0.06755	0.01461	14084
19	30	5	30	50	0.9826133	0.06844	0.01044	73813
20	6	5	28	50	0.9755367	0.07042	0.01184	78714

2) One Step Ahead and MSE

Figure 3 shows the results of the one-step-ahead (OSA) tests for the top three solutions. The predicted results were able to follow the actual data's patterns generally. This observation is consistent with the low MSE values and R^2 values near one. Although this appears promising, another important test to validate the models is the residual test, which is described in next section.
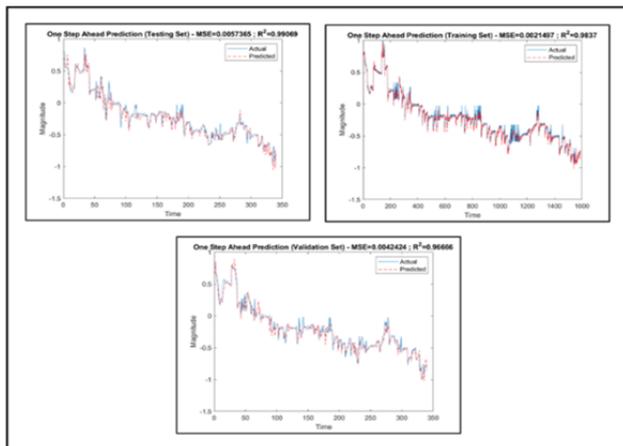

Figure 3. One step ahead test for PSO model

3) Residual Tests

Figure 4 shows the residuals of training, testing, and validation sets for the top three model optimization results. The residuals appear to be small (near 0) in most cases, which indicates good agreement between the model and actual output.

However, some issues with high residuals, particularly during sudden changes in the output, need to readjust themselves to compensate for the variation.
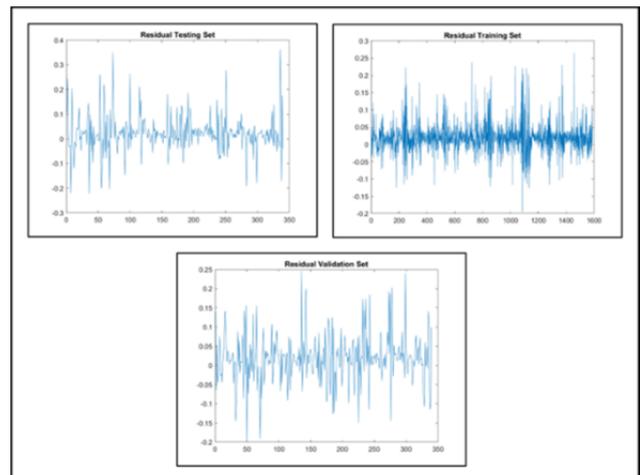

Figure 4. Residual test for PSO model

4) Correlation Tests

Figure 5 shows the autocorrelation test results for the training, testing, and validation sets of the top three models. Autocorrelation measures the similarity between a signal and itself over multiple lags. The residuals appear to be uncorrelated except for lag 0 (where the residual signals are aligned); a majority of them were within the 95% confidence limit.

Subsequently, the cross-correlation results of the best three models are shown in

Figure 6. In contrast to autocorrelation, cross-correlation measures the similarity between two different signals (in this case, the residuals and the output). Most of the correlation coefficients lie between the 95% confidence intervals, indicating that the output and residuals were uncorrelated with each other. When violations do occur, they tend to be small and isolated. In combination with the autocorrelation results, the observation indicates that the residuals are consistent with the profile of random noise and that the model is therefore unbiased and acceptable as the models were able to represent the most important information from the experimental data.
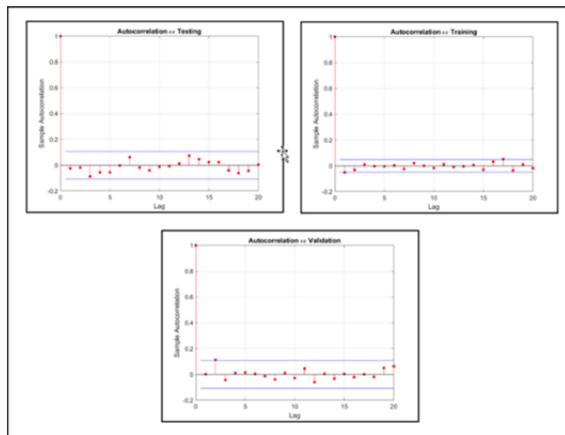

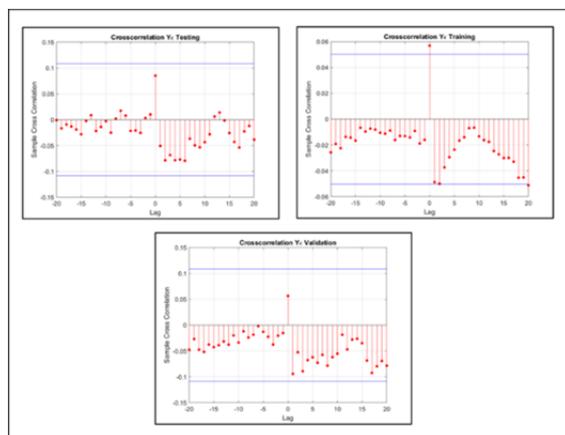*Figure 5. Autocorrelation test for PSO model*

Figure 6. Cross-correlation test for PSO model

5) Histogram Analysis

Figure 7 shows the results of the histogram test for lags and hidden nodes 29:30:30:49, 29:26:30:45, and 30:29:30:50. All graphs for this histogram test demonstrated normal distribution data sets of the residuals. In conclusion, the PSO model created passed all correlation and histogram tests with minimum violations.
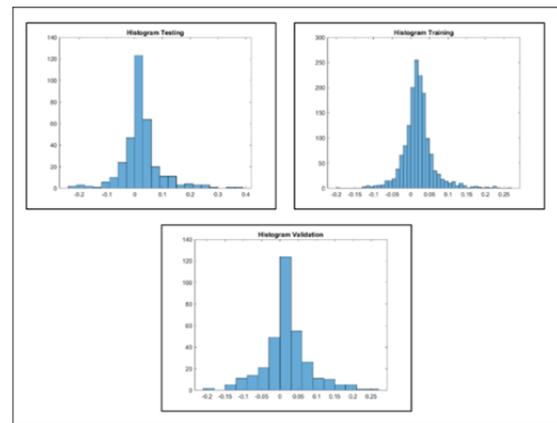

Figure 7. Histogram test for PSO model

In this section, the accuracy and validity of the PSO model have been tested. In a stage for optimal lagging, and node test generates a 150 data set. In this data, the parameter of swarm size, maximum iteration and an initial seed was identified and used for the second stage of the process. From that stage, the best optimal lags space and hidden nodes were at lag 29:30:30:49 for input1/acid (n_{u1})): input2/alkaline (n_{u2}) : output/pH (n_y): hidden nodes (h), based on the minimal total correlation value of 0.01503, with an MSE of 0.01213. Based on the OSA, correlation, and histogram test, the PSO model performance was accepted with the highest accuracy.

4. Conclusion

Then, several parameters are optimised using PSO that include the values of particles, random seed, maximum iterations, cognition, social learning rate, particle velocity and position. In a stage for optimal lagging, and node test generates a 150 data set. In addition, the model fit and residual distribution are analysed too for this model. Results showed that the new proposed model could record minimal error, proving that a good agreement is established between the predicted and the actual pH values.

Overall, the Particle Swarm Optimization (PSO) model has proven its dominance as a mathematical model to estimate the solution's pH. It would help to improve the delivery accuracy of the alkaline and acidic mixtures for the hydroponics pH control system.

References

[1]. Joshitha, C., Kanakaraja, P., Kumar, K. S., Akanksha, P., & Satish, G. (2021, February). An eye on hydroponics: The IoT initiative. In *2021 7th International Conference on Electrical Energy Systems (ICEES)* (pp. 553-557). IEEE.

[2]. Sarala, S. M. (2020, October). Robust Smart Irrigation System using Hydroponic Farming based on Data Science and IoT. In *2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC)* (pp. 1-4). IEEE.

[3]. Takeuchi, Y. (2019). 3d printable hydroponics: A digital fabrication pipeline for soilless plant cultivation. *IEEE Access*, 7, 35863-35873.

[4]. Johansen, J. L., Nielsen, M. L., Vestergård, M., Mortensen, L. H., Cruz-Paredes, C., Rønn, R., ... & Ekelund, F. (2021). The complexity of wood ash fertilization disentangled: Effects on soil pH, nutrient status, plant growth and cadmium accumulation. *Environmental and Experimental Botany*, 185, 104424.

[5]. Ávila, P. F., Da Silva, E. F., & 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.

[6]. Ravi, P. P., Lindner, J., Oechsner, H., & Lemmer, A. (2018). Effects of target pH-value on organic acids and methane production in two-stage anaerobic digestion of vegetable waste. *Bioresource technology*, 247, 96-102.

[7]. Fallah, B., Ng, K. T. W., Vu, H. L., & Torabi, F. (2020). Application of a multi-stage neural network approach for time-series landfill gas modeling with missing data imputation. *Waste Management*, 116, 66-78.

[8]. Hassan, M. A., Bailek, N., Bouchouicha, K., & Nwokolo, S. C. (2021). Ultra-short-term exogenous forecasting of photovoltaic power production using genetically optimized non-linear auto-regressive recurrent neural networks. *Renewable Energy*, 171, 191-209.

[9]. Momeni, E., Armaghani, D. J., Hajihassani, M., & Amin, M. F. M. (2015). Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks. *Measurement*, 60, 50-63.

[10]. Harrison, M. T., Roggero, P. P., & Zavattaro, L. (2019). Simple, efficient and robust techniques for automatic multi-objective function parameterisation: Case studies of local and global optimisation using APSIM. *Environmental Modelling & Software*, 117, 109-133.

[11]. Lloyd, J. J., Taylor, C. J., Lawson, R. S., & Shields, R. A. (1997). Comparison of local and global optimisation techniques for diffusion battery data analysis. *Journal of aerosol science*, 28(5), 821-831.

[12]. Nezhad, N. H. M., Niasar, M. G., Hagen, C. W., & Kruit, P. (2020, April). Local versus global optimization of electron lens system design. In *2020 IEEE 6th International Conference on Optimization and Applications (ICOA)* (pp. 1-6). IEEE.

[13]. Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks* (Vol. 4, pp. 1942-1948). IEEE.

[14]. Shahraki, H., Aalaei, S., & Moradi, S. (2021). Infrared small target detection based on the dynamic particle swarm optimization. *Infrared Physics & Technology*, 117, 103837.

[15]. Singh, G., & Singh, A. (2021). Extension of particle swarm optimization algorithm for solving transportation problem in fuzzy environment. *Applied Soft Computing*, 107619.

[16]. Hason, M. M., Hanoon, A. N., & Abdulhameed, A. A. (2021). Particle swarm optimization technique based prediction of peak ground acceleration of Iraq's tectonic regions. *Journal of King Saud University-Engineering Sciences*.

[17]. Kaur, M., & Dutta, M. K. (2021). Restoration and quality improvement of distorted tribal artworks using Particle Swarm Optimization (PSO) technique along with nonlinear filtering. *Optik*, 245, 167709.

[18]. Song, B., Wang, Z., & Zou, L. (2021). An improved PSO algorithm for smooth path planning of mobile robots using continuous high-degree Bezier curve. *Applied Soft Computing*, 100, 106960.

[19]. Júnior, D. A. D., da Cruz, L. B., Diniz, J. O. B., da Silva, G. L. F., Junior, G. B., Silva, A. C., ... & Gattass, M. (2021). Automatic method for classifying COVID-19 patients based on chest X-ray images, using deep features and PSO-optimized XGBoost. *Expert Systems with Applications*, 183, 115452.

[20]. Suman, G. K., Guerrero, J. M., & Roy, O. P. (2021). Optimisation of solar/wind/bio-generator/diesel/battery based microgrids for rural areas: A PSO-GWO approach. *Sustainable Cities and Society*, 67, 102723.

[21]. Sabzzadeh, I., & Shourian, M. (2020). Maximizing crops yield net benefit in a groundwater-irrigated plain constrained to aquifer stable depletion using a coupled PSO-SWAT-MODFLOW hydro-agronomic model. *Journal of Cleaner Production*, 262, 121349.

[22]. Nordin, M. K., Mohd Yassin, A. I., Md Tahir, N., & Saaid, M. F. (2021). Automated fertilizer mixer system for fertigation farming. *Journal of Electrical and Electronic Systems Research (JEESR)*, 18, 18-23.

[23]. Ma, T., Liu, S., & Xiao, H. (2020). Location of natural gas leakage sources on offshore platform by a multi-robot system using particle swarm optimization algorithm. *Journal of Natural Gas Science and Engineering*, 84, 103636.