ANFIS Controller Design Using PSO-Tuned PID Data for pH Regulation in Industrial Cooling Towers

*Basim Al-Najari[1], Chong Kok Hen[1], Johnny Koh Siaw Paw[1], Ali Fadhil Marhoon[2]
[1]Department of Electrical & Electronics Engineering, Universiti of Tenaga Nasional (UNITEN), Putrajaya Campus, Jalan IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia
[2]Department of Cyber Security Engineering, College of Engineering, Al-Maaqal University, Dinar Street, Al-Maqal, Al-Basrah, Iraq
E-mail: alnajaribasim@gmail.com
*Corresponding author

Keywords: ANFIS Controller, PID Controller, PSO Algorithm, pH Regulation, Instructions for design.

Received: March 13, 2025

The Adaptive Neuro-Fuzzy Inference System (ANFIS) controller is a modern alternative to the conventional PID controller. This paper presents the design of the ANFIS controller for pH regulation in cooling towers based on the dataset from the PID controller. This paper aims to design the ANFIS controller to achieve a lower RMSE than benchmark models, with improved transient response specifications such as rise time, settling time, and overshoot to optimize pH regulation characteristics. One of the fundamental requirements for designing the ANFIS controller is the availability of a dataset. The challenge in the design process is how to prepare the dataset. This issue was addressed by recording the PID controller dataset, which includes the error (e), the change in error (Δe), and the output. The methodology consists of the following steps: (1) modeling the pH loop of the cooling tower; (2) tuning a PID controller using the Particle Swarm Optimization (PSO) algorithm; (3) recording the PID controller's dataset, including error (e), change of error (Δe), and output; (4) training the ANFIS model using the MATLAB ANFIS Toolbox; and (5) enhancing the transient response by modifying the dataset. Following Instructions for design, the simulated results showed that the ANFIS controller achieved a root mean square error (RMSE) of 0.0081. The transient response characteristics of the best-performing ANFIS controller (Modified ANFIS_4) include: rise time = 0.5863 s, settling time = 1.4867 s, overshoot = 2.7958%, and peak = 7.6548. In comparison, the baseline PSO-tuned PID controller yielded a rise time of 0.6046 seconds, a settling time of 2.3155 seconds, an overshoot of 8.6770%, and a peak value of 8.0916. The results confirm that the ANFIS controller outperforms the PID controller in all key transient response parameters, offering improved accuracy, faster stabilization, and reduced overshoot. These findings demonstrate the effectiveness of the ANFIS design based on real PID controller data, supported by Instructions for design for reliable implementation in nonlinear industrial control systems.

Povzetek: Članek uvaja nov ANFIS regulator za regulacijo pH v industrijskih hladilnih stolpih na osnovi podatkov PID, nastavljenega s PSO. Znanstveni prispevek je metoda simetričnega zajema podatkov in podatkovne augmentacije.

1 Introduction

pH control is critically important in industrial applications. Accurate pH control prevents corrosion and scaling in mechanical equipment such as heat exchangers and pipelines. Extreme pH levels can lead to rapid corrosion or scale buildup, resulting in costly repairs, downtime, and safety hazards. Proper pH regulation enhances the efficiency of chemical reactions, reducing chemical usage and environmental impact [1]. The task of regulating pH is of high importance. Additionally, it is utilized in various industrial applications, with a particular emphasis on cooling towers. The neutralization of industrial cooling towers is contingent upon the correct pH control, which will facilitate the biological treatment of cooling water and significantly influence the utilization of chemical resources [2-4]. The pH neutralization process is exceedingly difficult to regulate. It is characterized by high nonlinearities and dynamic process dynamics, which are typically challenging to model [5]. Due to the pH regulation system's nonlinearity and time latency, the conventional PID controller finds it challenging to accomplish precise pH control [6]. According to the Mitsubishi operation limit, the pH should be maintained between 7.1 and 7.8. A PID controller maintains a pH of 7.45 [7]. The performance analysis of PID controllers in the Acid, Neutral, and Base regions is conducted with a set point of 7.45 [8]. Control methods that have been frequently employed in research include the use of PID [9]. Proportional-integral-derivative (PID) controllers have been extensively employed to regulate the speed and position of DC motors [10]. Due to their ability to model and control highly nonlinear processes, ANFIS controllers offer several potential benefits for pH control in industrial applications. pH control is inherently complex because of process variability, time delays, and sensitivity to disturbances, making traditional control methods less effective. The adaptive nature of ANFIS allows it to adapt to changing process dynamics and provide accurate, real-time control. This results in improved stability, faster response times, and reduced overshoot compared to conventional PID controllers [11]. Artificial intelligence (AI) models have substantially contributed in the past few decades by offering cost-effective and more precise solutions for simulating physical flood practices [12]. Artificial neural networks (ANNs) have garnered significant attention in control applications in recent years due to their independence from human intervention and expert-level expertise [13]. Expert behavior has been extensively modeled using artificial intelligence techniques, such as fuzzy inference and neural networks [14]. ANFIS controller is capable of autonomously learning and adapting to the state of a plant [15]. Conversely, ANFIS has been recognized for its superior features derived from neural networks and fuzzy logic. In 1997, Jang et al. developed ANFIS. It is a category of adaptive networks functionally equivalent to fuzzy inference systems. In fuzzy inference systems, neural networks revise the parameters from a training data set. The linguistic interpretability of Fuzzy Inference Systems (FIS) and the advantages asserted by neural networks (NNs) are present in ANFIS, which actively participates in pursuing specific objectives. Due to its adaptive capability, ANFIS is nearly immediately applicable to adaptive control and learning control. The primary advantages of ANFIS over classical linear approaches, such as linear control systems, are the structural knowledge representation and nonlinearity [16]. The ANFIS structure is developed using the PID configuration as a reference. The primary objective is to minimize transient response specifications, including overshoot, settling time, and rise time, to achieve optimal pH characteristics in the cooling water [17]. To enhance pH efficacy, the Adaptive Neuro-Fuzzy Inference System (ANFIS) has been suggested as an alternative to the traditional PID controller. The PID controller's advantages are strictly limited, which prevents the attainment of the necessary augmented response for specific applications. The most effective solution to these constraints is an ANFIS-based speed controller [18]. The ANFIS control algorithm's superior performance is due to its robustness in nonlinear systems. ANFIS also produces intelligent self-learning by integrating fuzzy logic with neural networks, resulting in numerous applications in the past. The ANFIS system is based on the Sugeno type of inference system and features a unique architecture that enables a hybrid learning algorithm. This paper aims to create an adaptive controller that outperforms the PID controller [19]. Using a novel optimization technique, the ANFIS controller's output stability is significantly enhanced compared to that of the PID controller for the designated data. In the context of stability control, the optimization of output is a complex module. Consequently, to enhance stability and power balancing, a novel ANFIS/PID optimization technique is implemented and analyzed using the Artificial Intelligence tool [20]. In general, the ANFIS controller outperforms other controllers. The ANFIS method is a hybrid of fuzzy logic and neural networks that capitalize on the capabilities of fuzzy logic to reason based on existing principles and learn and predict [21]. The ANFIS is a hybrid technique that combines artificial neural network algorithms and fuzzy logic theory. The ANFIS model was trained using datasets acquired by implementing a fuzzy logic controller in the MATLAB Simulink environment [22]. The efficacy of the controller, which features an artificial neural network-based fuzzy logic (ANFIS) control system, is compared to a conventional fuzzy logic system that is not based on an artificial neural net-work. Data necessary to model the fuzzy inference system based on an artificial neural network is derived from the induction motor system controlled by PI. Under all dynamic conditions, the ANFIS controller outperforms the controller that is exclusively implemented using fuzzy logic, as demonstrated by the results of the MATLAB-SIMULINK simulation [23].

This study presents an ANFIS controller design for pH regulation of cooling towers based on a PID controller. Instructions for design will be given to the design in this study. This study investigates whether a dataset generated from a PSO-tuned PID controller can be effectively used to train an ANFIS controller that achieves superior transient response performance—such as lower rise time, shorter settling time, and reduced overshoot—compared to the original PID controller and previously published ANFIS-based approaches. Additionally, the study explores whether symmetrical recording of both positive and negative dataset segments enhances the robustness of the ANFIS controller across varying setpoint directions and whether strategic augmentation of the training dataset can significantly improve the transient response of the modified ANFIS design.

Although DNNs (Deep Neural Networks), GP (Genetic Programming), and reinforcement learning are state-of-the-art (SOTA), the motivation for choosing ANFIS over these AI techniques is that it performs better in real-time control systems. This contribution is significant for researchers because it will help them design the ANFIS easily, accurately, and reliably.

2 ANFIS Controller design

2.1 Methodology

This section provides a practical guide on building an ANFIS pH controller for cooling towers based on the PID controller. To accomplish this, the steps below must be followed:

- pH loop of cooling towers modeling.
- PID controller tuning based on particle swarm optimization PSO.
- PID controller design.
- ANFIS Structure.

- PID controller dataset recording (error e, error change Δe, and output).
- ANFIS controller design based on PID controller dataset using ANFIS toolbox of MATLAB.
- ANFIS controller transient response enhancement.
- Instructions for design in ANFIS design and enhancement.

The methodology described above is illustrated in the flowchart shown in figure 1, which contains sufficient detail to enable replication of the process.

2.2 pH loop of cooling towers modeling

The cooling tower is a department of the State Company of Fertilizers (SCF) in Iraq. To design the controller for the pH loop of the cooling tower, the transfer function of the loop must be obtained. The cooling tower model was obtained using the method described in [24]. The mathematical model of pH response is described using the first-order system transfer function plus delay time FOBDT [25]. Figure 2 shows the cooling tower pH response. The final transfer function G(s) was calculated using the MATLAB system identification toolbox. The pH loop transfer function after pade approximation, curve fitting, and system identification toolbox of MATLAB is shown in the equation (1). The coefficients of equation (1) are unitless.

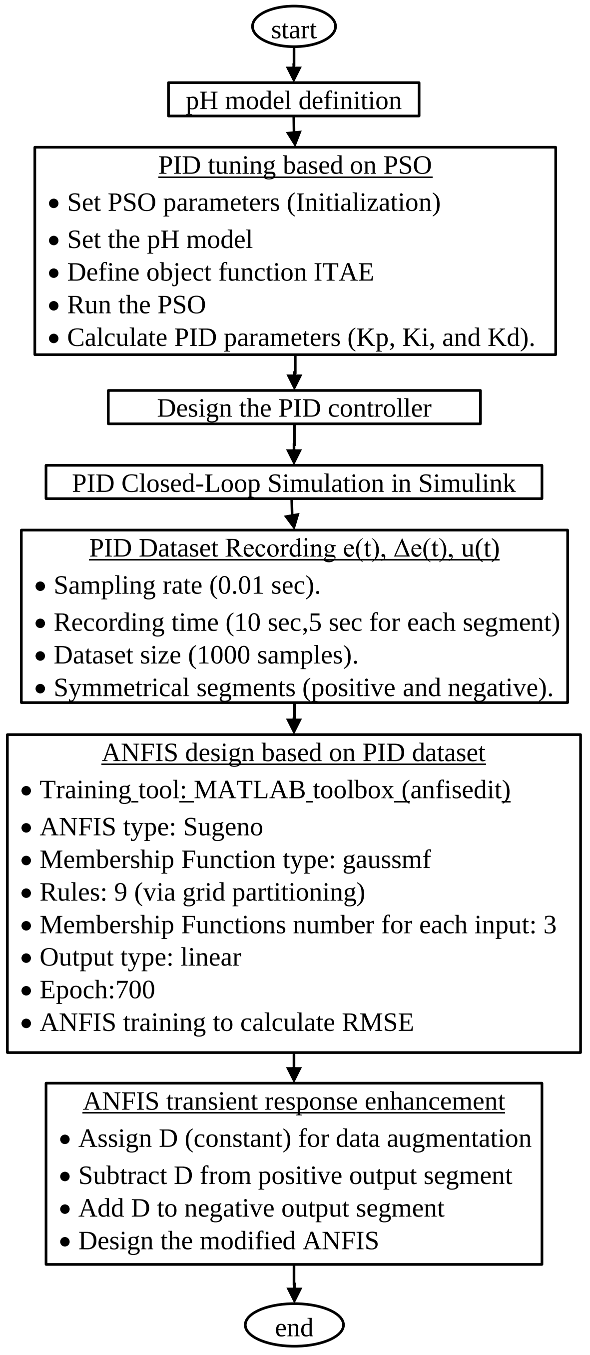

Figure 1: The methodology flowchart

$$G(s)=\frac{9.56e^{-5}s^2+2.32e^{-4}s+7.3486e^{-6}}{s^3+1.17s^2+0.036s+0.001122689} \quad (1)$$

The transfer function G(s) represents the dynamic response of the cooling tower's pH loop and is expressed in terms of the Laplace variable (s). The coefficients of the numerator and denominator are unitless and represent the system gain and time constants, respectively. The delay component introduced via the Padé approximation models the time lag in the system's pH response.

2.3 PID controller tuning based on PSO

2.3.1 PID

PID, or Proportional-Integral-Derivative, is a linear control mechanism currently the most often utilized control strategy in practical engineering applications. The PID control method is simple and practical [26]. The PID controller is frequently utilized as a feedback controller in process industries. Notwithstanding the process plant's dynamic characteristics, the PID controller provides outstanding control performance. It comprises three essential components: Proportional, Integral, and Derivative modes. Three fundamental parameters must be determined to create a PID controller. The three elements of a control system are proportional gain (Kp), integral gain (Ki), and derivative gain (Kd). The PID controller's output can be determined using equation (2). The PID controller's output signal is denoted as u(t), while the error signal is denoted as e(t) [27-28].

$$u(t) = Kp.e(t) + Ki\int_0^t e(t).dt + Kd.\frac{de(t)}{dt} \quad (2)$$

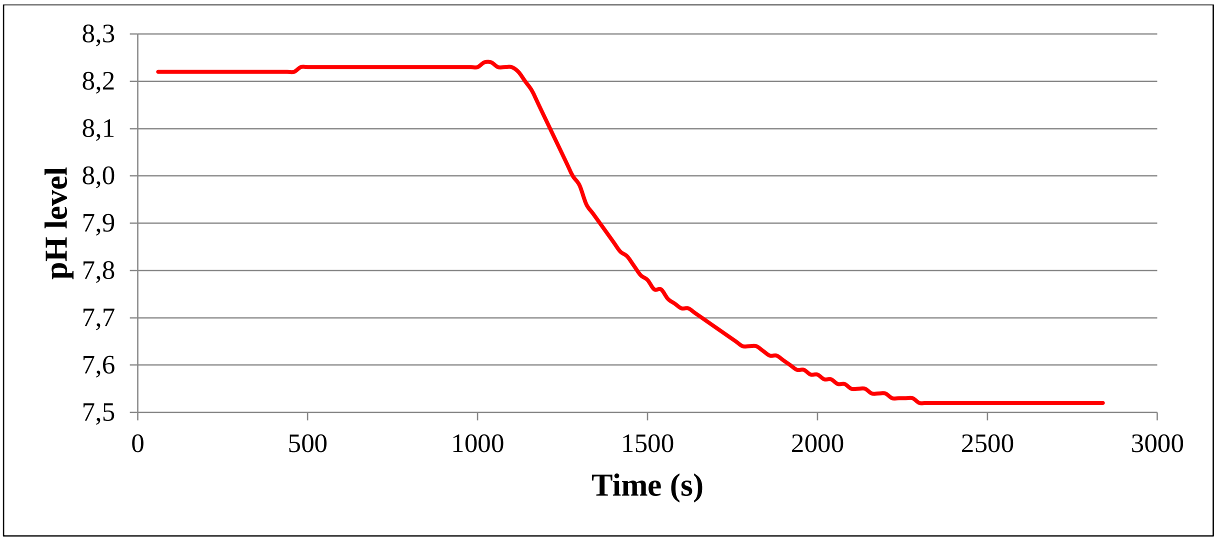

Figure 2: The pH response of the cooling towers at the State Company of Fertilizers (SCF).

2.3.2 PSO

The particle swarm optimization (PSO) algorithm, developed by Kennedy and Eberhart, is an evolutionary algorithm that functions according to swarm behavior [29]. The algorithm depends on the coordinated movement of the flock, dictated by the bird's location closest to the food source. The particles' position and velocity update equations represent the flock's movements. The equations for velocity (3) and position (4) are presented below:

$$V_i^{k+1} = w^k V_i^k + c_1 r_1 \left(P_{Best}^k - X_i^k\right) + c_2 r_2 \left(G_{Best}^k - X_i^k\right) \quad (3)$$

$$X_i^{k+1} = X_i^k + V_i^{k+1} \quad (4)$$

The variables in equations (3) and (4) are as follows: k denotes the number of repetitions, i represents the particle index, and w indicates the inertia weight, directly influencing velocity. The variables c1 and c2 denote the acceleration factors referred to as cognition and social constants, respectively. The variables r1 and r2 are stochastic values that lie within the interval of 0 to 1. Pbest denotes the ideal local solution, while Gbest signifies the optimal global solution. Vi and Xi represent the velocity and position of particle i, respectively [30].

The particle swarm optimization (PSO) algorithm seeks to identify the ideal solution inside a designated area to minimize the value of the objective function. This study designates the integrated time absolute error (ITAE), as illustrated in equation (5), as the system's objective function.

$$ITAE = \int_0^\infty t|e(t)|\, dt \quad (5)$$

Where e(t) is the error signal and (t) is the time [26].

Integral Time Absolute Error (ITAE) is a performance index used in control system design, particularly for tuning PID controllers. It quantifies the error over time, emphasizing the importance of minimizing the error signal's magnitude and duration. This approach is particularly beneficial in robust control applications where uncertainties in system parameters are prevalent. The ITAE index is instrumental in developing robust PID controllers, as demonstrated in various studies focusing on its application in different control scenarios [31].

The choice of ITAE as the objective function in this study is driven by its ability to optimize the PID controller parameters, its focus on error dynamics, robustness to uncertainties, optimization capabilities, and its applicability to multivariable systems. These factors collectively contribute to the effectiveness of the proposed control strategies in managing unstable multivariable systems, such as pH systems [32].

ITAE is adapted in this study to tune PID controller parameters using the PSO algorithm. This criterion is applied directly in the fitness evaluation stage of the PSO process. The PID controller's parameters—proportional gain (Kp), integral gain (Ki), and derivative gain (Kd)—are iteratively updated using PSO, and the ITAE value computed for each candidate solution guides the convergence toward optimal tuning. The optimized PID parameters improve response time, reduce overshoot, and enhance system stability. The PID controller will then generate the dataset (error, change in error, and controller output) required for the ANFIS controller design [33].

2.3.3 Tuning activity

The PID controller was tuned using the PSO algorithm; table 1 shows the PSO parameters while table 2 shows the tuning results. Figure 3 shows the PSO convergence characteristic. The PSO algorithm was configured with 700 iterations to ensure adequate objective function convergence (ITAE). This value was determined through preliminary testing, where lower iteration counts often resulted in premature convergence. A search range of –500 to 20,000 was selected to accommodate the significant

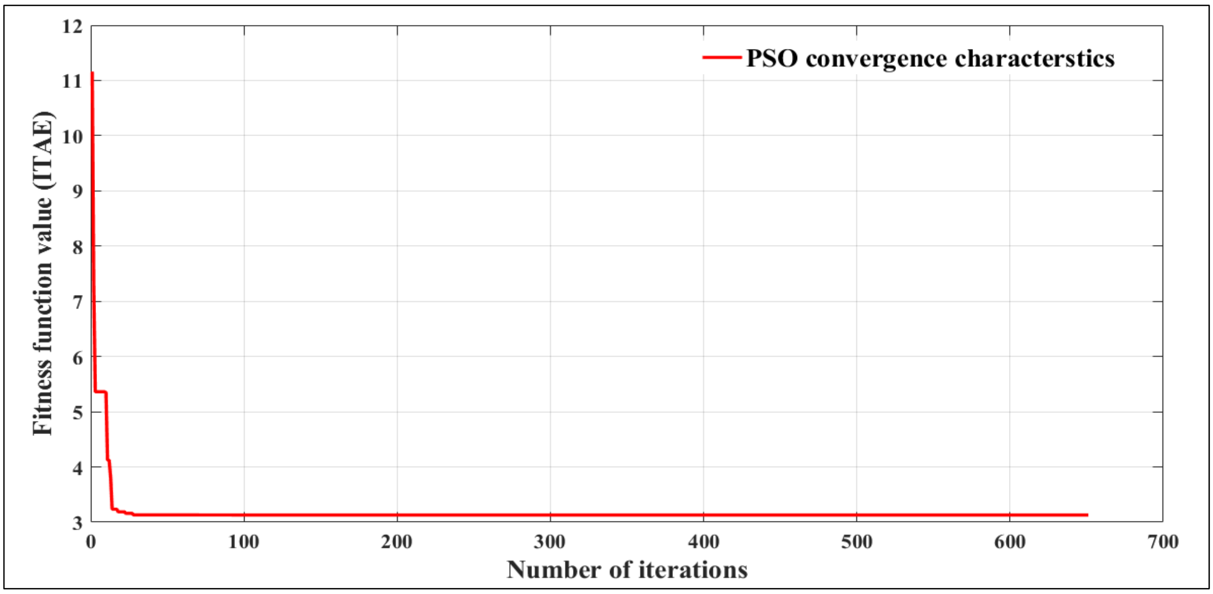

Figure 3: PSO convergence characteristics showing the decrease in the best objective function value (ITAE)

Table 1: PSO tuning parameters

Parameter name	Prefix	value
Number of variables	m	3
Population size	n	100
Maximum inertia weight	wmax	0.9
Minimum inertia weight	wmin	0.4
Acceleration factor 1 & 2	c1 & c2	2
Random factor 1 & 2	r1 & r2	1
Lower boundary	LB	-500
Upper boundary	UB	20000
Maximum iteration number	maxiter	1000

Table 2: PID tuning parameters based PSO

Parameter name	Prefix	value
Proportional gain	Kp	20000
Integral gain	Ki	100.2575
Derivative gain	Kd	14345.3379
Filter	N	0.1

controller gains typically required in nonlinear pH regulation systems. The PSO tuning process was executed over 10 independent runs with varying random seeds to assess repeatability. The convergence curve shown in Figure 3 represents a typical run, while the final PID parameters correspond to the run that achieved the lowest ITAE. Across trials, the convergence behavior remained consistent, with only minor variations in the final fitness values, confirming the method's reliability. The PSO algorithm was considered to have converged when the global best ITAE value improved by less than 1e-6 over 100 consecutive iterations.

2.3.4 The Robustness of the Cooling Towers pH Model

The stability of the system is an indication of its robustness. The pH model (equation (1)) underwent a stability test to assess the system's robustness using a Bode plot. The gain margin (GM) and phase margin (PM) must be derived from the open-loop transfer function to determine system stability using a Bode plot. GM is measured at the phase crossover frequency (where phase = -180°), while PM is measured at the gain crossover frequency (where gain = 0 dB). The stability criteria are that GM in dB and PM in degrees must be positive. Positive GM and PM indicate system robustness. The Bode plots have been included below to verify that the control system remains stable and robust, as shown in Figure bp.

Figure 4 shows GM = ∞ dB and PM = 70.5081 degrees at 2.43 rad/s. These plots confirm that the Gain Margin and Phase Margin are within acceptable bounds, indicating robustness to model uncertainties. These results confirm that although the controller gains are large, the closed-loop system remains stable and robust and does not destabilize the system. Therefore, the system is considered robust against gain variations and modeling uncertainties. The magnitude may stay below 0 dB across the frequency range, implying no phase margin violation. The phase may never reach -180° while the gain is above 0 dB, implying no gain margin violation. Thus, even with high PID gains, the system is stable, and the design is well-controlled and robust.

2.4 PID controller design

Figure 5 shows the closed-loop control system block diagram while Figure 6 shows the step (pH=7.45) transient response. The set point of pH 7.45 was selected based on industrial requirements for cooling tower operations, particularly following the Mitsubishi operational guidelines. These guidelines specify an acceptable pH range from 7.1 to 7.8, with 7.45 commonly used as a nominal set point. This is the optimal set point to ensure safe operation and optimal biological treatment of cooling water. Table 3 shows the pH transient response characteristics.

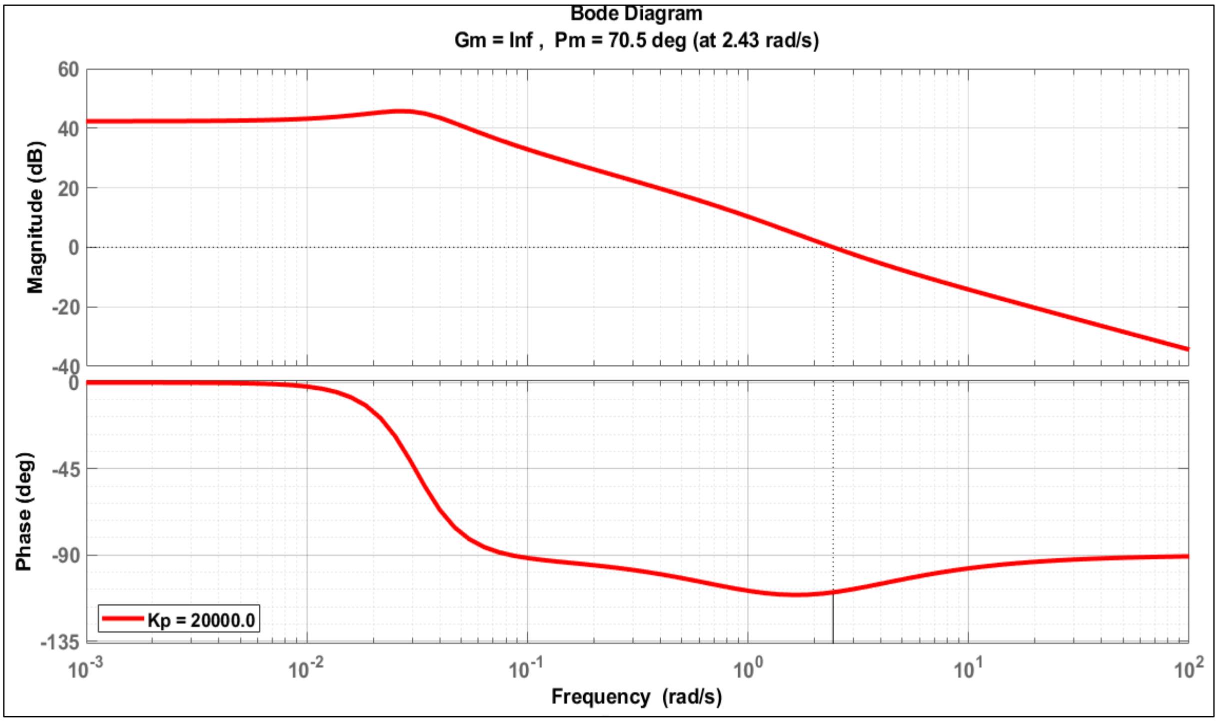

Figure 4: bode plot

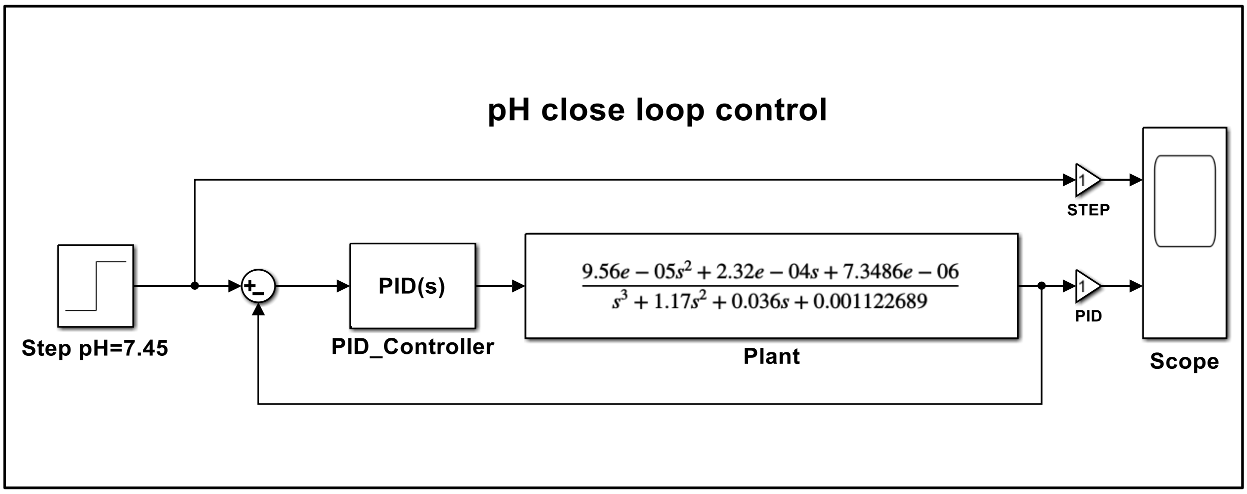

Figure 5: pH close loop control

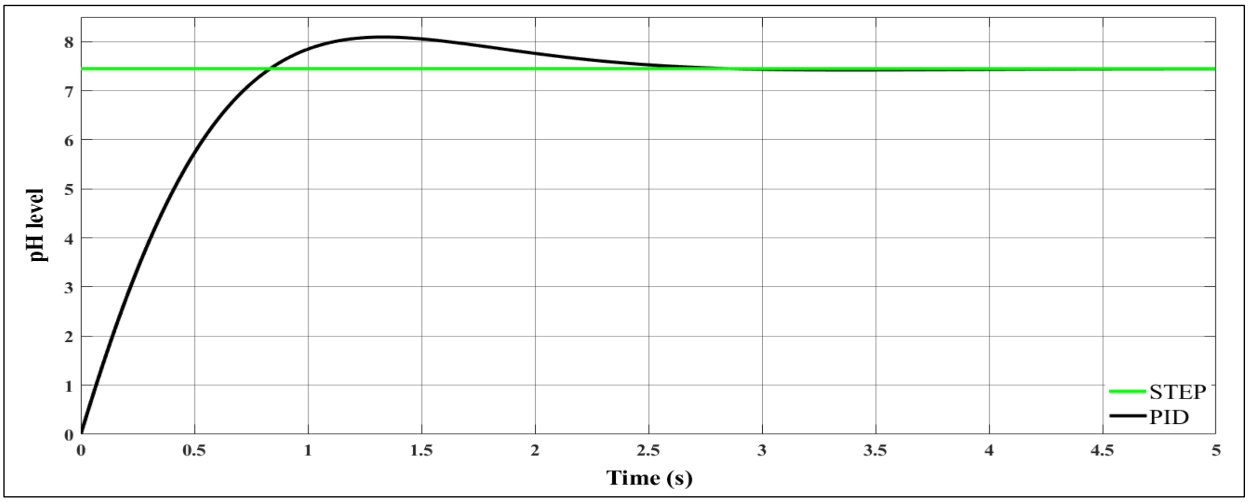

Figure 6: pH close loop response

Table 3: pH response characteristics

Parameter name	Value	Unit
Rise time	0.6046	sec
Settling time	2.3155	sec
overshoot	8.6770	%
peak	8.0916	pH level

2.5 ANFIS structure

The Adaptive Neuro-Fuzzy Inference System (ANFIS) is a hybrid model integrating Neural Networks (NN) and Fuzzy Inference Systems (FIS). It is capable of solving optimization challenges [34]. The standard FLC generally utilizes input signals of the system error (e) and the rate of change (Δe) of the error. The system error is the difference between the set point r(t) and the plant output u(t) at time t, whereas the rate of change in the error Δe is the difference between the current error e(t) and the preceding error e(t-1) at time t. The error and its variation are presented in equations (6) and (7) [35]:

$$e\ (t) = r(t) - u\ (t) \tag{6}$$

$$\Delta e\ (t) = e\ (t) - e\ (t-1) \tag{7}$$

Figure 7 illustrates the typical ANFIS model. In this diagram, a circle denotes a fixed node, whereas a square signifies an adaptable node. The Sugeno fuzzy model is the most prevalent among other FIS models owing to its superior interpretability, computing efficiency, and incorporation of optimal and adaptive techniques. For each model, a standardized rule set of two fuzzy if-then rules (as illustrated in equations (8) and (9)) can be expressed as:

$$Rule1: if\ e\ is\ A1\ and\ \Delta e\ is\ B1, then\ f1 = p1 * e\ +\ q1 * \Delta e\ +\ r1 \tag{8}$$

$$Rule2: if\ e\ is\ A2\ and\ \Delta e\ is\ B2, then\ f2 = p2 * e\ +\ q2 * \Delta e\ +\ r2 \tag{9}$$

Where A1, A2, B1, and B2 are matching fuzzy sets, while p1, p2, q1, q2, r1, and r2 denote constraints on the output function, and z = f(e, Δe) is a crisp function in the consequent [36].

Figure 7 shows that the ANFIS is comprised of five layers. The five layers are the fuzzification layer, product layer, normalized layer, defuzzification layer, and output layer [37], which are detailed as follows:

Layer 1 (fuzzification): This layer uses square nodes to represent an adaptive functional element. Each entry into node i invokes an adaptable membership function to create the degree of membership for linguistic elements.

Membership functions can take on various forms, such as Gaussian, trapezoidal, tri-angular, or extended Bell functions. The layer outputs are shown in the Equation (10):

O1,i = μAi (e), i = 1,2 or O1, i = μBi-2(Δe), i = 3,4. (10)

The two inputs in this context are designated as e and Δe. The input specifications, μAi and μBi, evaluated as Gaussian membership functions (Equation (11)), require two variables called premise variables, consisting of the center c and the width σ.

$$Gaussian\ (x: c, \sigma)\ =\ e^{-\frac{1}{2}(\frac{x-c}{\sigma})^2} \tag{11}$$

where O1, i denotes the output of layer 1 at the ith node.

Layer 2 (product): These fixed nodes denote the product Π for the computation of a rule's firing strength. This layer receives input values from the first layer and transforms them into a membership function representing fuzzy input variable sets. The output of each node is the multiplication of all its received signals. The outputs of this layer are shown in Equation (12):

O2,i = wi = μAi (x) μBi(y), i = 1, 2 (12)

Using grid partitioning, the autogenerated rules are mn, where m is the number of MFs in each input and n is the total number of inputs.

Layer 3 (normalization): Layer 3 nodes are also fixed nodes. Each node in this layer is marked as N. Every node normalizes the firing strength of a rule from the preceding layer by computing the ratio of the ith rule's firing strength to the aggregate firing strength of all rules. The outputs of this layer can be seen in Equation (13):

$$o_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2}\ i = 1, 2 \tag{13}$$

Where $\overline{w}$ is defined as the normalized firing strength of a rule.

Layer 4 (defuzzification): The nodes in this layer are adaptive, with the node function defined by Equation (14):

$$o_{4,i} = \overline{w}_i\ f_i = \overline{w}_i \left(p_i x + q_i y + r_i\right)\ i = 1, 2 \tag{14}$$

where $\overline{w}$ is the rule's normalized firing strength and {pi, qi, ri} is a first-order polynomial. O4, i represents the output of layer 4. Parameters in this layer are linear and are well known as consequent parameters. These parameters are identified during the training process of the ANFIS.

Layer 5 (overall output): This single node is called the output layer, which is labeled as (Σ). This layer only sums up the outputs of all rules in the previous layer and converts fuzzy results into crisp outputs, as shown in Equation (15):

$$o_{5,i} = \sum_{i=1}^{2} \overline{w}_i f_i\ = \frac{\sum_{i=1}^{2} \overline{w}_i f_i}{w_1 + w_2} \tag{15}$$

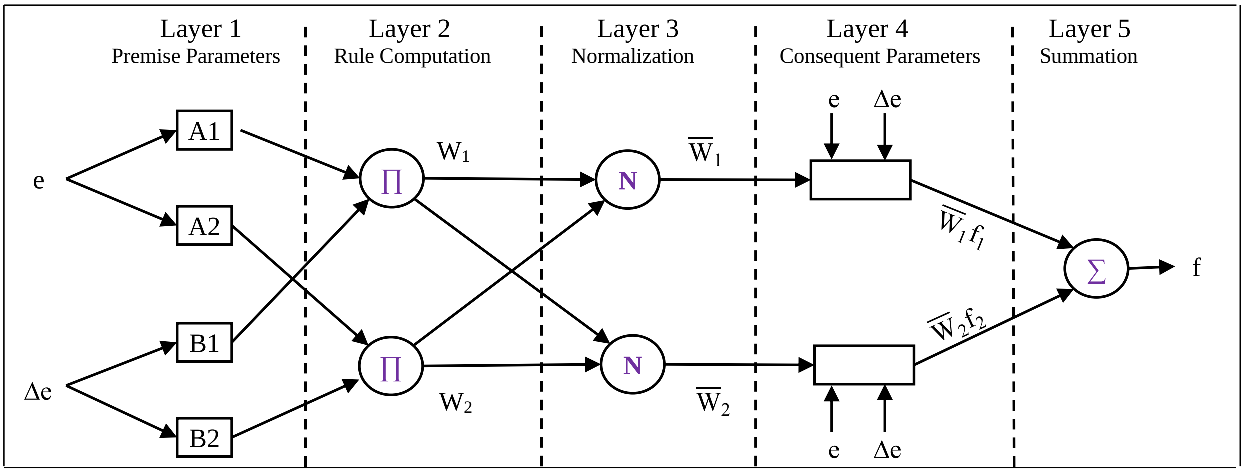

Figure 7: Typical ANFIS structure

The error is the difference between the actual and predicted outputs of the ANFIS. The fewer the errors, the more successful the ANFIS [38-40].

2.6 PID controller datasets recording (error e, change of error Δe, and output)

For ANFIS design, the input and output dataset must be available. There are two input signals: error (e), significant rate of change of error (Δe), and one output signal. The PID datasets were obtained through simulation using a closed-loop control model developed in MATLAB/Simulink. The input and output dataset has been derived from a designed PID controller tuned using particle swarm optimization PSO algorithm. Figure 8 shows the system close loop control with the recording circuit, figure 9 shows the system transient response, table 4 shows the recorded dataset information of figure 9, and Figure 10 shows the recording circuit where the recording circuit was built based on equations (6) and (7). The recorded data's resolution is 0.01 seconds over a total period of 10 seconds, resulting in 1,000 data samples (500 for the positive part and 500 for the negative part).

The terms "positive part" and "negative part" refer to the direction of the setpoint change and the specific segments of the dataset used to train the ANFIS controller. The positive part corresponds to the dataset recorded during the increase in setpoint from 0 to 7.45 (pH units) over the first 5 seconds of the simulation. The negative part corresponds to the dataset recorded during the decrease in setpoint from 7.45 to 0 over the subsequent 5 seconds (5 10-second intervals). This bidirectional data acquisition ensures that the ANFIS controller learns to respond to both upward and downward changes in setpoint, enhancing its robustness and generalization across typical operating conditions.

The recorded datasets (the PID controller inputs and output datasets) were sent to the MATLAB workspace for ANFIS design activities.

Table 4: Recorded dataset information

Parameter name	Value	Unit
Recorded dataset size	1000	points
Each part dataset size	500	points
Recording period	10	sec
Positive part pH=7.45	0~5	sec
Negative part pH=0	5~10	sec
Recording time	1e-2	sec

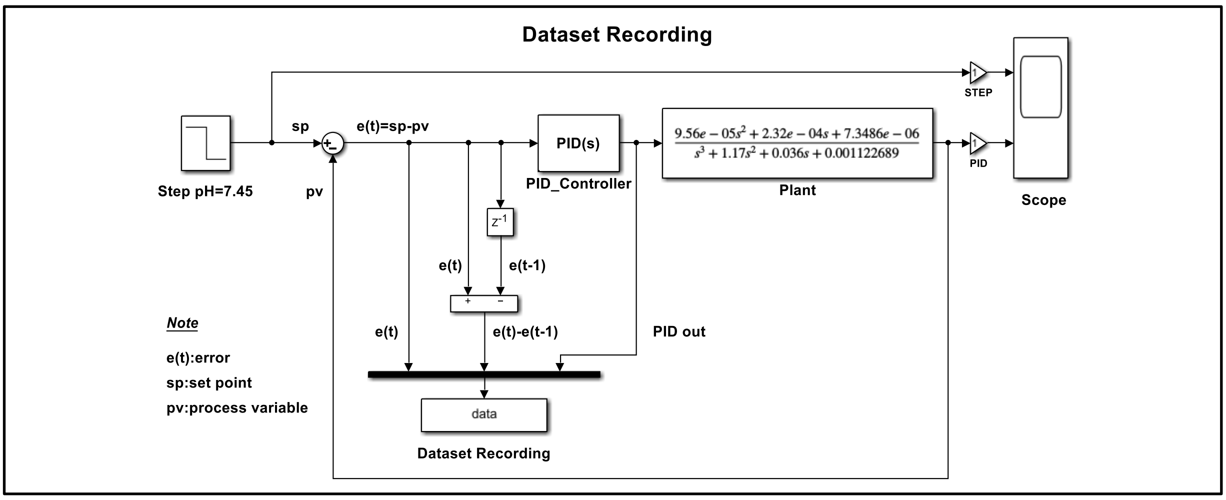

Figure 8: Dataset recording circuit

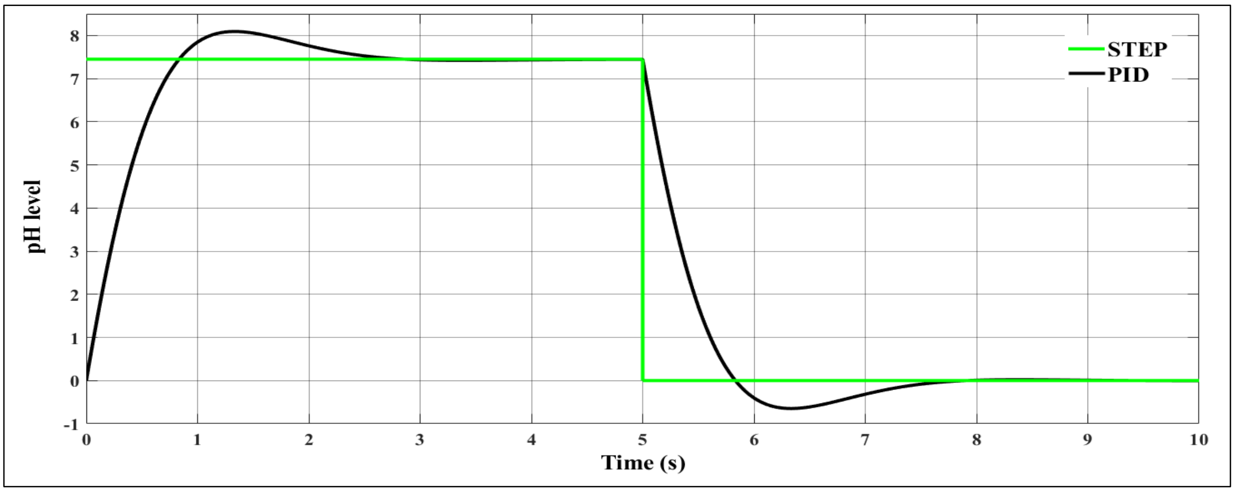

Figure 9: System transient response

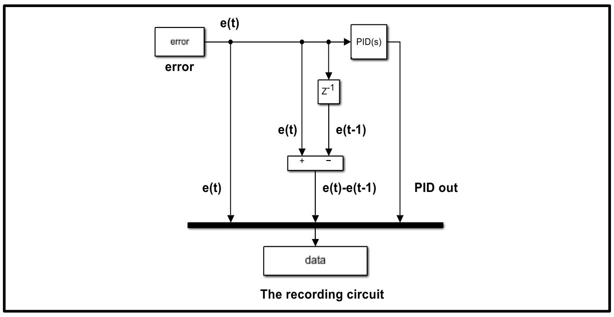

Figure 10: The recording circuit

Table 5: The noise metric for inputs-output variable

variable	error	Error change	output
Standard deviation	1.5018	0.3355	32177.22
variance	2.2555	0.1126	1035373000
mean	0.0034	0.0000046	100.75
median	0.0045	0.0000088	135.21
min	-7.45	-7.45	-159564.88
max	7.45	7.45	159687.03
range	14.9	14.9	319251.91

2.6.1 Quantifying the noise and variability of the dataset

The dataset used for ANFIS training was collected from a MATLAB-Simulink model of a PID-controlled pH regulation loop. The recorded data was generated under controlled simulation conditions; therefore, it was not subject to environmental or sensor-induced noise as in real-world data. The data is free from errors and consistent. Table 5 below presents the noise metrics calculated for each numeric column.

On the other hand, the cleaned data means that it is free from missing, duplicate, and null values. The training data consists of three columns: error, error change, and output. The dataset appears to be numeric and continuous, with no missing values visible in the sample rows and no data type consistency (all numeric).

2.6.2 Dataset statistical measures and their size impact on controllers

This section will present a statistical measure of the dataset and its effect on controller.

- The impact of dataset size on the controller's performance and generalization capabilities: The dataset's size significantly impacts the performance and generalization capabilities of controllers like ANFIS (Adaptive Neuro-Fuzzy Inference Systems). Table 6 shows the impact of the data size on controller performance
- Statistical Measures Characterizing the Dataset: The dataset used for designing the ANFIS controller comprises three key variables: error (e), change of error (Δe), and output. Table 7 shows the primary statistical measures calculated for each variable, providing a comprehensive dataset characterization.

These statistical measures provide a detailed summary of the dataset's central tendency, dispersion, and range, which are crucial for understanding the behavior of the control system and for training robust ANFIS models.

Table 6: the impact of the data size on controller performance

Data size	Performance	Generalization	Risk
Small	Low accuracy	Poor adaptability	Underfitting or overfitting
Large	High accuracy	Strong adaptability	Computational cost
Unbalance	Inconsistent output	Unreliable control	Loss of control
Balance	Optimized accuracy	Robust response	Best configuration

Table 7: The primary statistical measures

Statistic	error	error change	output
Count	1001	1001	1001
Mean	0.0034	0.000005	100.75
Standard deviation	1.5018	0.3355	32,177.22
Minimum	-7.45	-7.45	-159,564.88
25th percentile	-0.0642	-0.0023	-1,388.80
Median	0.0045	0.0000088	135.21
75th percentile	0.0711	0.0023	1,597.33
Maximum	7.45	7.45	159,687.03
Variance	2.2555	0.1126	1035373000

2.7 ANFIS controller design using ANFIS editor GUI

This section presents two subjects: the description of the ANFIS Editor GUI and the design of ANFIS using the ANFIS Editor GUI based on the recorded dataset from the PID controller, as outlined below:

2.7.1 ANFIS editor GUI

The ANFIS Editor GUI is a graphical user interface in MATLAB's Fuzzy Logic Toolbox that allows users to interactively create, train, and test Adaptive Neuro-Fuzzy Inference Systems (ANFIS). ANFIS is a hybrid modeling approach that combines neural networks and fuzzy logic to automatically tune fuzzy inference systems using input/output data sets. The ANFIS Editor GUI is opened using "anfisedit" instruction. The "anfisedit" instruction is a MATLAB command that opens ANFIS Editor GUI, enabling interactive design, training, and evaluation of adaptive neuro-fuzzy inference systems.

The ANFIS Editor GUI allows users to:

- Load training, testing, and checking data sets from the MATLAB workspace.
- Generate or load an initial fuzzy inference system (FIS) model of the Sugeno type.
- Visualize and edit the structure of the FIS, including membership functions and rules.
- Select training parameters, such as the optimization method (backpropagation or hybrid), number of epochs, and error tolerance.
- Train the FIS model on the provided data, adjusting membership function parameters to fit the input/output relationships best.
- Test and validate the trained model using separate data sets to check for overfitting and generalization.
- Save and export the trained FIS for further use in simulations or control systems.

2.7.2 ANFIS Controller Design Based on PID Recorded Dataset

Based on the PID controller dataset that was recorded in section (2.6), the ANFIS controller was designed using the ANFIS Editor GUI (ANFIS Toolbox of MATLAB). The design activities are:

- Opening the ANFIS Editor GUI (ANFIS Toolbox of MATLAB for ANFIS controller design using ("anfisedit") instruction.
- Loading the workspace's recorded datasets to the ANFIS Toolbox. Figure 11 shows the loaded dataset.
- Generation of ANFIS for the loaded dataset is based on the data mentioned in Table 8. Figure 12 shows the ANFIS input-output variables: error (e), change of error (Δe), and output. Table 9 shows the dataset ranges with set point=7.45.
- Training the loaded datasets with 700 Epoch. The training result is RMSE=24.2599.

A symmetrical dataset validation strategy was used instead of adopting a traditional train/test data split or cross-validation. It was employed using both positive and negative dataset segments. Thus, the positive and negative dataset design serves as a domain-specific validation approach, ensuring that the ANFIS controller maintains stability and performance across the full operational range.

Table 8: ANFIS design parameters

Item	Description
Membership functions type	gaussmf
Number of memebership functions	3
Number of fuzzy rules	9
Epoch number	700
Membership functions output type	linear
No. inputs	2
No. outputs	1
Type of partition	Grid partition

Table 9: The dataset ranges

Variable	Min	Max
e	-7.4455	7.45
Δe	-7.45	7.45
out	-1.98e+5	1.948e+5

Figure 13 shows the trained datasets. Figure 14 shows the Epoch error. Figure 15 shows the ANFIS controller structure. Figure 16 shows the details of the fuzzy rule bases. Figure 17 shows the close-loop control using the PID and ANFIS controllers. Figure 18 shows the system's transient response using ANFIS and PID controllers, and Table 10 shows the transient response performance parameters for the two controllers.

The ANFIS (Adaptive Neuro-Fuzzy Inference System) controller interacts with the cooling tower system in a closed-loop control framework to regulate the pH level. The interaction involves several key aspects, such as sampling rate, data acquisition, error computation, and control output generation. The pH sensor measures the cooling water's pH level. The controller compares the current pH reading with the setpoint (7.45). The error signal and its rate of change are computed.

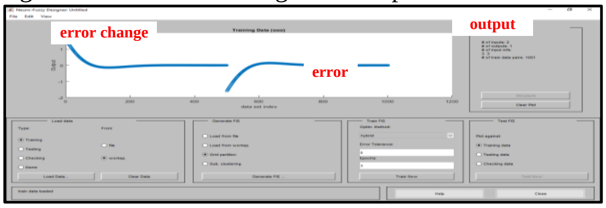

Figure 11: the loaded dataset

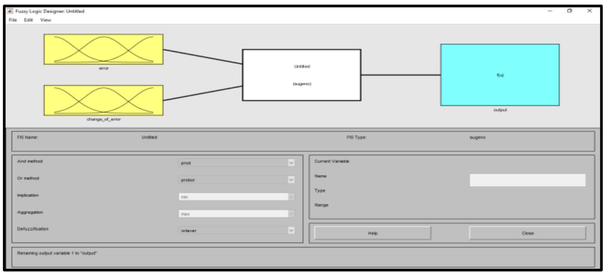

Figure 12: The ANFIS input-output signals

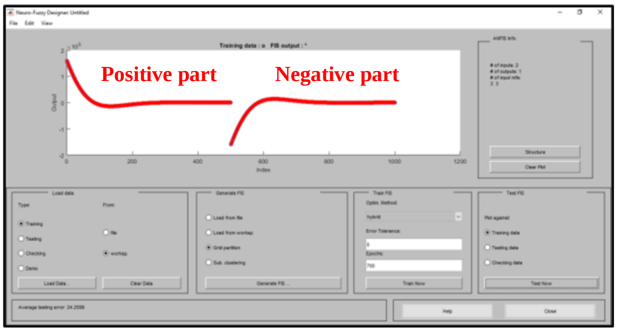

Figure 13: The trained dataset

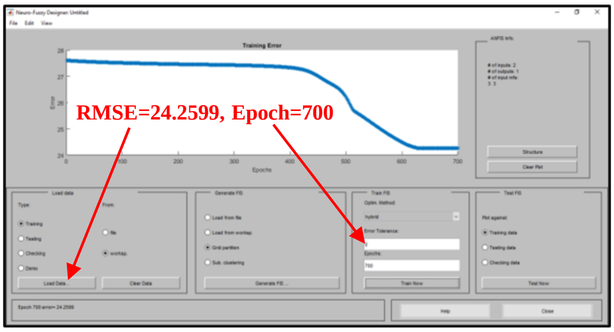

Figure 14: The Epoch error

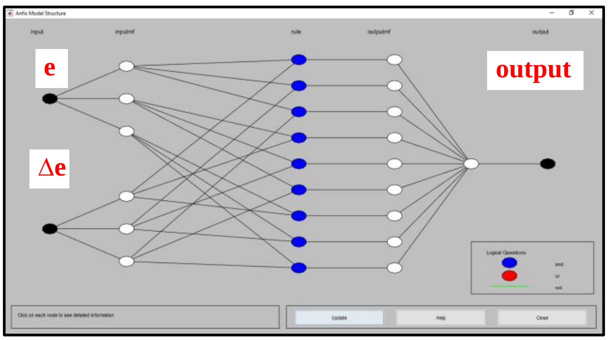

Figure 15: The ANFIS controller structure

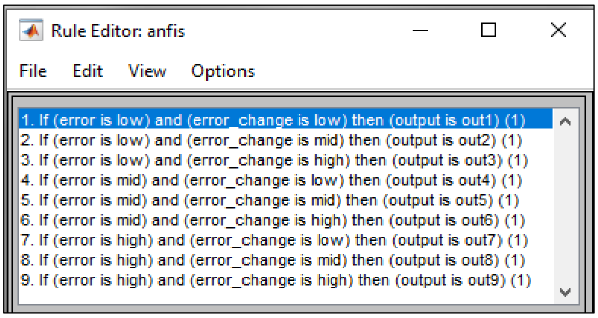

Figure 16: Fuzzy rule bases

Table 10: Response performance parameters

Parameter	PID	ANFIS	Unit	Best
Rise time	0.6046	0.6047	sec	PID
Settling time	2.3155	2.3149	sec	PID
overshoot	8.6770	8.6760	%	ANFIS
peak	8.0916	8.0922	pH	PID
Peak time	1.335	1.330	sec	ANFIS

The ANFIS controller outputs a control signal to adjust the flow rate of chemical dosing (sulfuric acid injection) to regulate pH. Total number of recorded values: 1000. The recording period used was 10 seconds total, with a time step of 0.01 seconds, so the sampling rate is 100 Hz, meaning the controller reads sensor data and updates its output every 10 milliseconds. The ANFIS controller successfully replaces the conventional PID controller in this application by leveraging the strengths of fuzzy logic and neural networks. Its ability to learn from real-world PID-generated data and adapt dynamically to changing conditions makes it ideal for complex, nonlinear processes like pH regulation in cooling towers.

2.8 Data augmentation and statistical test

This section presents two subjects as follows:

2.8.1 ANFIS controller transient response enhancement

The ANFIS controller that was designed based on the PID controller dataset in section (2.7) took the same transient response as the PID controller. The total number of the dataset is 1000 values. The dataset was divided into two parts: a positive part and a negative part where for each part 500 values. The recording information is shown in table 4. Figure 19 shows the ANFIS controller's transient response enhancement based on data augmentation.

A targeted augmentation technique was employed to improve the transient response of the original ANFIS controller. The inputs of the original ANFIS (e and Δe) remain unchanged. Specifically, for the positive segment of the dataset (setpoint increase from 0 to 7.45), the modification involved subtracting a constant value D of the original ANFIS output, represented by equation (16):

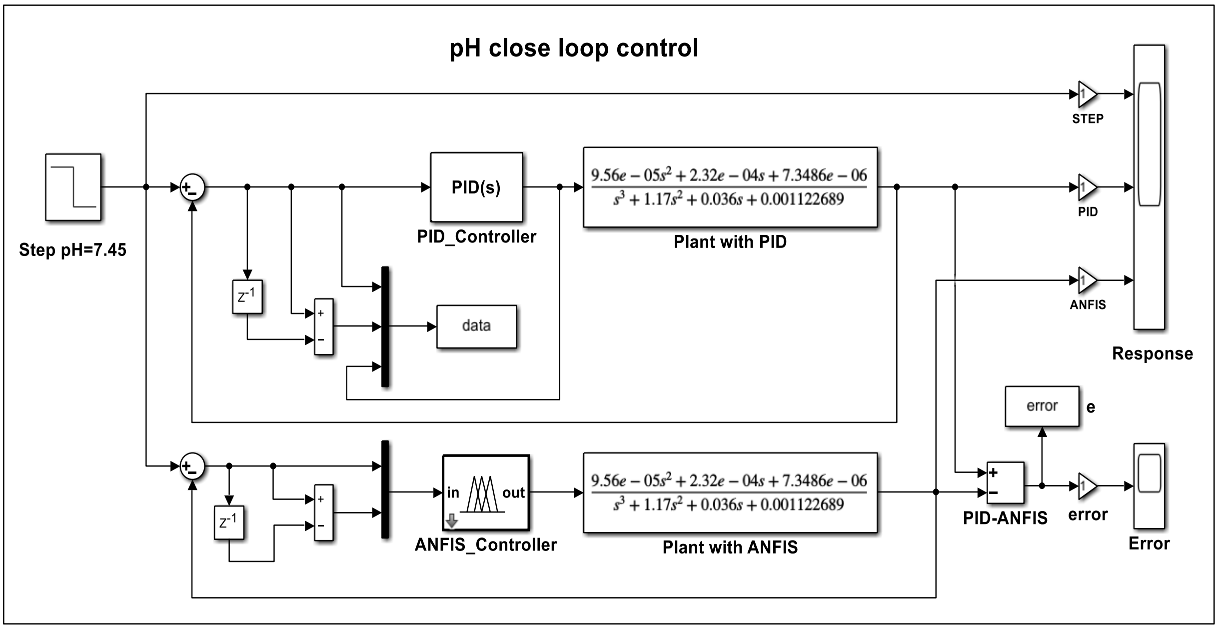

Figure 17: The close loop using the PID and ANFIS controllers

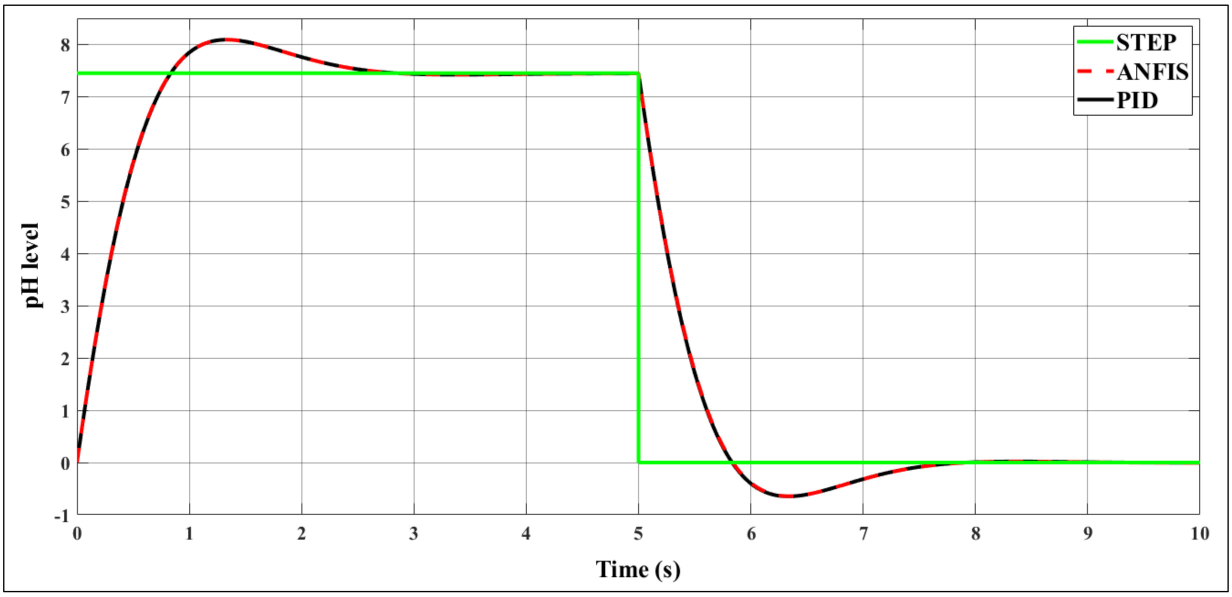

Figure 18: The transient response of PID and ANFIS

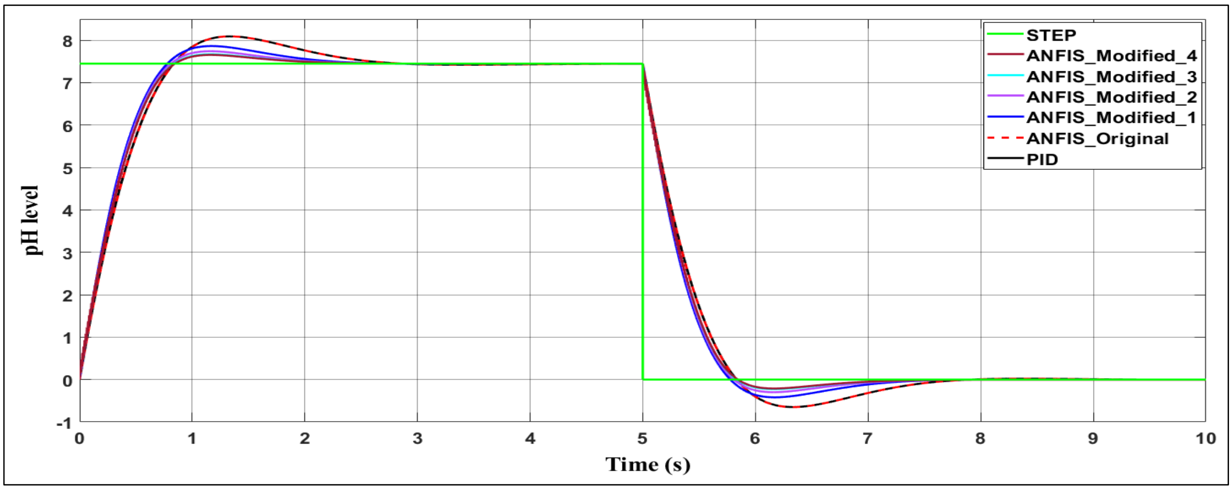

Figure 19: The transient response enhancement of multi-controllers

$$M_i = O_i - D \quad where\ i = 1{\sim}500 \tag{16}$$

Conversely, for the negative segment of the dataset (set point decrease from 7.45 to 0), the modification involved adding the same constant value D, as described by equation (17):

$$M_i = O_i + D \quad where\ i = 501{\sim}1000 \tag{17}$$

In these expressions, M_i denotes the modified ANFIS output, O_i refers to the original ANFIS output, and D is the augmentation constant introduced to enhance the transient response characteristics. Table 11 presents the transient response performance parameters of the multi-controllers.

The augmented dataset helps the ANFIS model to:

- Enhance generalization beyond the originally trained setpoints.
- Provide smoother transitions and better transient performance under variable conditions.
- Improve key performance indicators such as rise time, settling time, and overshoot.

This strategy, referred to as target output-shift augmentation or data reflection and scaling augmentation, provided the best results in the ANFIS_Modified_4, as reported in Table 11.

2.8.2 Statistical Validation of the ANFIS Data Augmentation Technique"

A statistical significance test (one-tailed t-test) was conducted to evaluate the effectiveness of the data augmentation technique applied to the ANFIS controller. Ten independent ANFIS models were trained, each using a modified output dataset generated via the augmentation method, while the input datasets (error e and change of error Δe) remained unchanged. This modification procedure is described by equations (16) and (17) in Section 2.8.1. All ANFIS models were designed following the methodology outlined in Section 2.7.

The transient response performance metrics rise time, settling time, overshoot, peak, and peak time were recorded for each model run. Table 12 shows that the resulting p-values for all metrics were less than 0.05, and the corresponding t-statistics exceeded the critical t-value. These results confirm that the improvements observed due to data augmentation are statistically significant and not due to random chance.

2.9 Instructions for design

In this section, several subjects will be presented as follows:

2.9.1 ANFIS design and enhancement

- The ANFIS controller must be designed based on the PID controller dataset.
- The PID controller dataset is (error e (t), change of error Δe (t), and output.
- Several membership functions and various types of membership functions must be tested as performance metrics to determine the optimal number and type of membership functions for the ANFIS design.
- The dataset recording circuit must be built based on equations (6) and (7).
- The delay function (Z^{-1}) and the subtract function must be used in the dataset recording circuit instead of the function ($\frac{de}{dt}$) used in previous papers.
- The controller set point may be changed (increasing or decreasing) according to the need, therefore the positive and negative parts of the dataset must be recorded. If only the positive part is recorded, the ANFIS controller may lose control in case of set point decreasing.
- The positive and negative parts of the recorded dataset must be symmetrical; otherwise, the ANFIS controller may lose control.
- The ANFIS controller follows the set point within its range; otherwise, the ANFIS controller may lose control, unlike the PID controller which remains in control even if the given set point is out of its range.
- The root mean square error (RMSE) can be a reasonable value when designing the ANFIS controller using MATLAB's ANFIS toolbox (ANFIS Editor GUI).
- To improve the response of the ANFIS controller, negative data must be added to the positive part and positive data must be added to the negative part.
- The data must be added for all values of the dataset; otherwise, the controller may lose control.
- In the case of multiple set points, the modified ANFIS transition must be from and to the recorded datasets; otherwise, it may encounter issues with its response.

Table 11: The transient response performance parameters of Multi controllers

Controller	Addition	Rise time (s)	Settling time (s)	Overshoot (%)	Peak pH level	Peak time (s)	Integral Absolute Error (IAE)
PID	---	0.6046	2.3155	8.6770	8.0916	1.335	3.072
ANFIS_Original	0	0.6047	2.3149	8.6760	8.0922	1.330	3.075
ANFIS_Modified_1	10000	0.5485	1.8928	5.6011	7.8633	1.170	2.495
ANFIS_Modified_2	20000	0.5633	1.6874	3.9657	7.7418	1.160	2.428
ANFIS_Modified_3	30000	0.5745	1.5302	3.0102	7.6708	1.160	2.391
ANFIS_Modified_4	33000	0.5863	1.4867	2.7958	7.6548	1.165	2.380

Table 12: Performance metric statistical t-test

run	addition	Rise time	Settling time	Overshoot	Peak	Peak time
1	0	0.6047	2.3149	8.6760	8.0922	1.33
2	5000	0.5471	2.0425	6.9318	7.9623	1.2
3	10000	0.5485	1.8928	5.6011	7.8633	1.17
4	15000	0.5559	1.7805	4.6605	7.7934	1.16
5	20000	0.5633	1.6874	3.9657	7.7418	1.16
6	25000	0.5695	1.6055	3.4325	7.7021	1.16
7	30000	0.5745	1.5302	3.0102	7.6708	1.16
8	31000	0.5837	1.5156	2.9359	7.6652	1.16
9	32000	0.5852	1.5012	2.8645	7.6599	1.165
10	33000	0.5863	1.4867	2.7958	7.6548	1.165
p-value		3.65e-15	4.72e-09	2.90e-05	2.80e-17	7.59e-14
t. statistical		96.022	19.911	7.076	165.028	68.516
t. critical		1.833	1.833	1.833	1.833	1.833
Statistically significance		yes	yes	yes	yes	yes

2.9.2 The challenges in ANFIS physical implementation

The ANFIS designed in section 2.9.1 can be physically implemented but will encounter several challenges. These include hardware, integration, and operational limitations that differ from the controller simulation environment. The challenges are outlined below:

- Real-Time Computation Constraints: ANFIS involves multiple fuzzy rules, Gaussian membership functions, and continuous parameter evaluation. This can be computationally demanding, especially if implemented on low-power PLCs or microcontrollers. Ensuring sufficient processing speed for a 100 Hz sampling rate may require more powerful hardware or optimization of the rule base.
- Sensor Noise and Disturbances: The current model was trained on a clean or mildly noisy dataset. In practice, sensor drift, industrial electrical noise, and unmodeled disturbances may introduce unexpected behavior. Real-time filtering or retraining with noise-augmented data would be necessary to preserve accuracy and stability.
- Limited Generalization Outside Training Range: The ANFIS controller is trained for setpoints within a specific range ([0, 7.45]). Physical variations, such as changes in chemical composition, unexpected pH spikes, or controller saturation outside this range, may cause the controller to lose stability or output extreme values. This limitation necessitates careful consideration of the operational parameters and the design of the ANFIS controller to ensure it can handle the expected conditions.
- Dataset collection on-site: The design of the ANFIS controller relies heavily on the availability of a well-prepared dataset from the PID controller, which includes error, change of error, and output values. If the dataset is not comprehensive or accurately recorded, it can lead to poor controller performance. The methodology for preparing this dataset must be clearly defined and followed to avoid issues during implementation.
- Integration and Interfacing: Integrating the ANFIS controller into an existing industrial control system, such as a PLC, requires robust interfacing with legacy hardware and software. Compatibility issues, communication delays, and synchronization problems may arise, especially if the new controller must coexist with other plant automation systems or safety interlocks.
- Robustness to Disturbances and Nonlinearities: The cooling tower pH process is highly nonlinear and sensitive to disturbances. While ANFIS can effectively model nonlinearities, any sudden or unmodeled disturbance could lead to control failure unless the system has been trained with a comprehensive dataset that includes such anomalies.

2.9.3 The robustness of the proposed ANFIS design

The robustness of the designed ANFIS controller concerning possible variations in the cooling tower system is addressed through several aspects as follows:

- Use of Real PID Controller Data: The ANFIS model is trained using a dataset derived directly from a PID controller tuned by Particle Swarm Optimization (PSO). This approach ensures that the ANFIS controller inherits practical, system-specific characteristics, making it more robust to real-world disturbances and nonlinearities.
- Dataset Symmetry and Coverage: Ensure the training dataset includes positive and negative parts (i.e., scenarios where the set point increases or decreases) to significantly enhance the robustness. This bidirectional coverage prevents the ANFIS controller from losing control when faced with reverse dynamics, a risk observed when only a unidirectional dataset is used.
- Enhanced ANFIS with Data Augmentation: A modified ANFIS version was developed by adding

supplementary data to both the positive and negative parts of the dataset. This augmentation led to improved transient response specifications.

- Validation Under Multiple Set Points: The system was tested with multiple set points (2, 5, 7, 4, 0), simulating real-world dynamic changes. The ANFIS controller successfully tracked these changes as long as they were within the trained dataset's range, proving its robustness in bounded dynamic environments.
- RMSE and Performance Comparison: The ANFIS controller achieved a very low RMSE of 0.0081, indicating minimal deviation from the expected output. Compared to other studies, this RMSE is significantly better, reinforcing the robustness and precision of the proposed design.

2.9.4 Guidelines for tuning and maintaining the ANFIS controller in a real-world environment

Tuning and maintaining an ANFIS (Adaptive Neuro-Fuzzy Inference System) controller in a real-world environment involves several key steps to ensure optimal performance. Here are some practical guidelines:

- Data Collection for Training: Use historical data from existing controllers, such as PID controllers, to train the ANFIS model. This data should include input variables like error and change in error, as well as the output variable, such as dosing pump speed.
- Defining Membership Functions: Carefully select the number and type of membership functions for the fuzzy inference system (FIS), as this can significantly impact the performance of the ANFIS controller. Experiment with different configurations to find the most effective setup.
- Training the ANFIS Model: Utilize MATLAB and its "anfisedit" command to train the ANFIS model. Ensure the training process is thorough, adjusting parameters to minimize errors and improve response times.
- Performance Evaluation: After training, evaluate the ANFIS controller's performance by testing it with various pH set points. Monitor key performance metrics such as rise time, settling time, overshoot, and steady-state error.
- Continuous Monitoring and Adjustment: Implement a system for continuously monitoring the ANFIS controller's real-time performance. This allows for timely adjustments if the system's performance deviates from expected results. Regularly check for changes in the process dynamics that may require retraining of the ANFIS model.
- Integration with PLC: Ensure that the ANFIS controller is correctly integrated with the PLC (Programmable Logic Controller) using OPC (OLE for Process Control) for effective communication. This integration is crucial for controlling the dosing pump speed in real-time based on the desired sulfuric acid flow rate.
- Documentation and Feedback Loop: Maintain thorough documentation of the tuning process, performance metrics, and any adjustments made. Establish a feedback loop to incorporate insights from operational data into the tuning process for continuous improvement.

3 Results and discussion

In this section, the results obtained from the ANFIS design based on the PID controller will be presented and described in detail. There are two types of ANFIS controllers: original and modified. The original ANFIS controller is designed based on the PID controller row dataset. In contrast, the Modified ANFIS controller is designed based on data addition to the PID controller row dataset for performance enhancement. This section is divided into four items as below:

3.1 Positive part of the original ANFIS controller

If the ANFIS controller is designed based on the positive part of the PID controller dataset, the ANFIS controller will work normally with set point increasing or decreasing. However, in some cases, it may lose control in case of set point decreases and visa-versa. Figure 20 shows the ANFIS controller response.

In Figure 20, the ANFIS controller was given multi-set points (set points: 2, 5, 7, 4, 0). It operated normally in cases of set point increases or decreases, even though the dataset is only for the positive part. However, in some cases, the ANFIS may lose control when the set point decreases. For safety, the ANFIS dataset must contain both positive and negative values.

3.2 Positive and negative parts of the original ANFIS controller

Figure 11 shows the positive and negative parts of the PID controller dataset. The ANFIS controller was designed based on the PID controller dataset. The root mean square error (RMSE) is 0.0081. Table 10 shows the response performance parameters for the PID controller and ANFIS controller. The parameters are identical. The lower the error rate (RMSE), the more identical the performance.

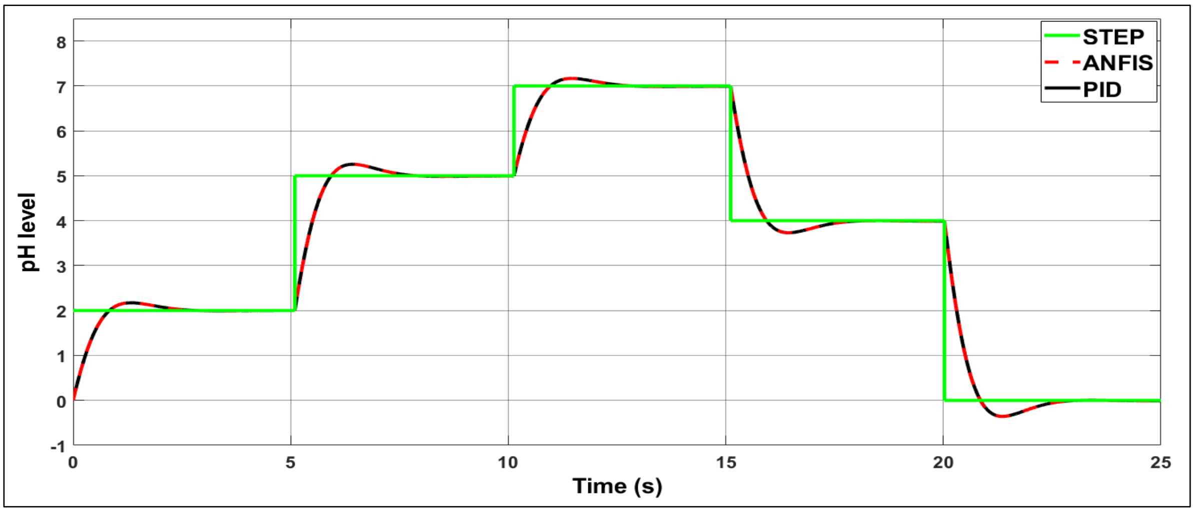

Figure 20: ANFIS controller response (Positive part).

Sometimes, the operator in the central control room CCR of the plants needs to change the value of the controller set point according to the industrial process. The controller's output is assumed to follow the set point to reach a stable state. Multi-set points (increasing and decreasing: 2, 5, 7, 4, and 0) were given to the controllers to see their response, as shown in Figure 21.

From Figure 21, the ANFIS controller works normally. It follows the set points, increasing and decreasing them, provided the set points are within the recorded range.

3.3 Modified ANFIS controller

From Figure 19 and Table 11, it can be seen that the best transient response characteristics belong to ANFIS_Modified_4, with a rise time of 0.5863 seconds, a settling time of 1.4867 seconds, an overshoot of 2.7958%, and a peak of 7.6548, outperforming the original PID controller in all key criteria. The enhancement activity produced improved transient response metrics well-suited for our system's performance requirements. These characteristics meet typical performance targets for chemical process control, where low overshoot (<5%), fast settling (<2 seconds), and minimal rise time is critical to ensure the safe, fast, accurate, efficient, and stable operation of pH regulation systems in industrial cooling towers. Multi set points (increasing and decreasing: 2, 5, 7, 4, and 0) were given to the controllers to see their response, as shown in Figure 22.

Two set points (7.45 and 0) were given during the PID controller dataset recording activity. From Figure 22, it is clear that multiple set points were assigned to the original and modified ANFIS controllers (set points: 2, 5, 7, 4, and 0), and they operate normally with the set points increasing or decreasing.

Changing the setpoint does not affect control performance; however, the goal of providing multiple setpoints is to determine whether the ANFIS controller remains under control and does not break out of control. The test shows that the ANFIS controller, trained using both positive and negative dataset segments, successfully tracked all setpoint transitions within the tested range without losing control, demonstrating robust generalization. This analysis is visually supported by

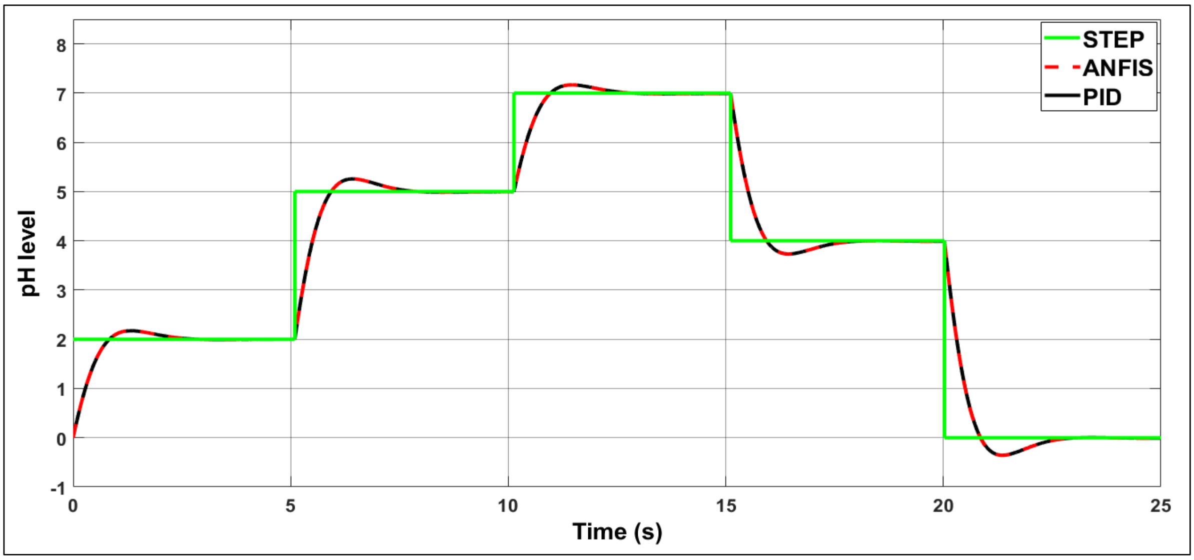

Figure 21: ANFIS controller response (Positive and negative parts).

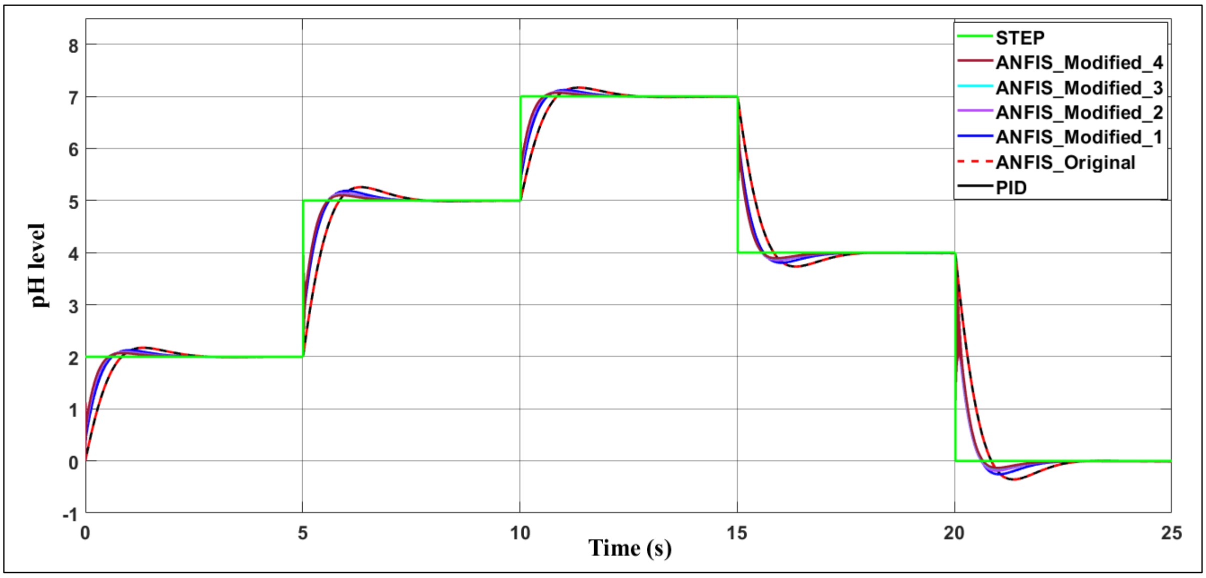

Figure 22: ANFIS models response for Multi set points

Figures 20 through 22 in the manuscript and numerically in Table 10, which presents the transient performance of both the original and modified ANFIS controllers under these conditions. The modified ANFIS (ANFIS_Modified_4) exhibited the best transient response characteristics. This confirms the importance of including a symmetrical dataset when designing an ANFIS controller and validates the modified controller's ability to maintain performance across dynamic operating conditions.

3.4 Comparative data from previous research.

In this study, the ANFIS controller was designed based on the dataset obtained from the PID controller response. Then, the PID controller was replaced with the ANFIS controller. Figure 18 shows that the ANFIS response is identical to the PID controller response. This concern was addressed through analysis under multiple operating points (set points 2, 5, 7, 4, and 0), which better reflects real-world fluctuations in industrial processes. The comparative transient responses are illustrated in Figures 17, 18, and 19, and the quantitative results are summarized in Table 11. Therefore, the proposed ANFIS-based solution is not merely a direct replacement for the PID under the same conditions but a future-ready control strategy designed to handle variable, nonlinear, and dynamic environments more reliably and efficiently. Figure 23 shows the error between the actual values represented by the PID controller response and the predicted values represented by the ANFIS controller response. The error range is between -02731 and 0.00213, and the RMSE is 0.0081.

This section presents the ANFIS controller design in the neuro-fuzzy technique based on different data from previous research. Table 13 shows the comparison between the current study and similar previous research. From Table 13, [41] designed an ANFIS controller for pH regulation based on data obtained from a CSTR simulator. The RMSE was 0.130348. [42] Designed an ANFIS controller for pH regulation based on data obtained from a bench-scale (pilot) plant. The RMSE was 0.0602 and the settling time was 350 seconds. [43] Designed an ANFIS controller for pH regulation based on data obtained from a wastewater treatment plant. The RMSE was 0.1825. [44] Designed an ANFIS controller for pH regulation based on data obtained from the sugar industry. The RMSE was 0.265. The overshoot was 22.7 %, and the settling time was 84.1 seconds. These comparisons demonstrate that the ANFIS controller designed in this paper, based on the PID controller dataset, achieves better performance metrics, such as a lower RMSE of 0.0081, a rise time of 0.6046 seconds, a lower overshoot of 8.6760, and a lower settling time of 2.3149 seconds, compared to ANFIS and other control approaches reported in previous related research.

It is important to note that the performance metrics presented in Table 13 should be interpreted in the context of differing experimental conditions across studies. These include variations in dataset size, dataset origin (simulated vs. real-world), system complexity, and process dynamics. For instance, while the current study utilized 1,000 real-world data samples from an industrial cooling tower, some referenced studies relied on synthetic or lab-scale datasets. Furthermore, system characteristics such as time delays, nonlinearity levels, and control objectives vary significantly. Therefore, the proposed ANFIS controller demonstrates lower RMSE and faster response metrics.

In addition, this study has extra novel contributions, such as a Practical Dataset Generation Approach, a

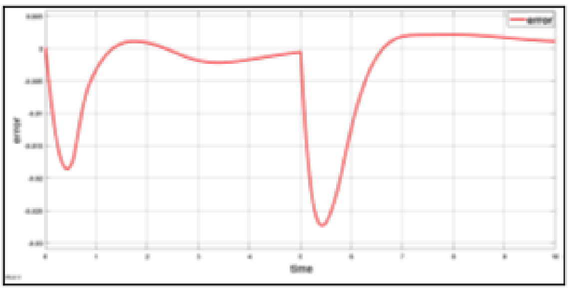

Figure 23: Error shape

Table 13: ANFIS controller performance comparison with prior studies

Reference	This study	[41]	[42]	[43]	[44]
System type	Industrial cooling towers	CSTR Simulator	Bench-scale (Pilot) plant	Wastewater treatment plant	Sugar industry
Dataset source	PID controller Dataset	Simulated CSTR model	Small-scale physical system	Real-world plant data	Operational data
Dataset type	Real-world, process based	Synthetic/simulated	Lab data	Operational data	Industrial data
Dataset size	1000	499	500	140	200
Methodology	PSO+ANFIS	ANFIS	ANFIS	ANFIS	ANFIS
RMSE	0.0081	0.130348	0.0602	0.18250	0.265
Overshoot	2.7958 %	-------------	---------------	------------	22.7 %
Rise time	0.5863 s	-------------	---------------	------------	------------
Settling time	1.4867 s	-------------	350 s	------------	84.1 s

Symmetrical Dataset Strategy for Bidirectional Control, Response Enhancement via Data Augmentation, Comprehensive Practical Guidelines and Considerations, and a Comprehensive Practical Implementation Framework. Section 3.5 describes the novel contributions of this study in detail.

3.5 The novel contributions

This section presents the novel contributions of this study that distinguish it from prior research, as listed below.

- Practical Dataset Generation Approach: This study presents a structured method for recording a real-world PID controller dataset (error, change in error, and output) from a closed-loop pH control system. Unlike previous studies that rely on synthetic or simulator-generated data, this method captures practical dynamics that enhance training realism.
- Symmetrical Dataset Strategy for Bidirectional Control: A novel contribution is the recording and use of symmetrical positive and negative dataset segments, enabling the ANFIS controller to respond accurately to increasing and decreasing setpoints—an aspect often overlooked in existing ANFIS-based control implementations.
- Response Enhancement via Data Augmentation: A data augmentation strategy involving adding opposing polarity samples was introduced that significantly improves the transient response (e.g., reducing overshoot to 2.79% and settling time to 1.49 seconds), achieving better performance than both the original ANFIS and previous benchmark studies.
- Comprehensive Practical Guidelines and Considerations: The paper includes detailed practical implementation guidelines, robustness analysis, and sensitivity testing that are not commonly addressed in ANFIS literature. This bridges the gap between simulation studies and deployable industrial control solutions.
- Comparative Performance Benchmarking: This study compares performance with existing research, using consistent metrics (e.g., RMSE, rise time) and clearly noting the differences in dataset sources and system types—something that is rarely highlighted in other publications.
- Comprehensive Practical Implementation Framework: The study includes detailed implementation guidelines, covering dataset generation, training configuration, controller tuning, deployment in Simulink, and expected challenges in real-world industrial integration—bridging the gap between theoretical ANFIS designs and field applications.

To clarify the novel contributions, Table 14 summarizes the differences between this study and other studies, as well as the reasons for those differences.

4 Conclusions

This study proposed an ANFIS-based controller for pH regulation in industrial cooling towers, designed using a dataset derived from a PSO-tuned PID controller to achieve a lower RMSE compared to benchmark models, along with enhanced transient response parameters to optimize pH regulation characteristics. The methodology involved generating a PID dataset, training ANFIS using MATLAB, and systematically enhancing performance through targeted data augmentation. The highlighted points in this study are listed below:

4.1 Key Findings

- The original ANFIS controller achieved an RMSE of 0.0081, closely matching the performance of the baseline PID controller.
- The modified ANFIS controller (ANFIS_Modified_4) further improved transient response metrics, including a 2.8% overshoot, 1.49 s settling time, 0.586 s rise time, and IAE 2.380.
- A symmetrical dataset design ensured robustness during both setpoint increases and decreases, enhancing controller generalization.

Table 14: The differences between this study and other studies and “why”

Aspect	This Study	Other Studies	Why This Matters (The 'Why')
Dataset Source	Real-world PID controller data (Section 2.6)	Simulator-based, lab-scale, or black-box data	Real-world PID-based data ensures that the trained ANFIS inherits practical system characteristics, leading to better generalization and reliability.
Symmetrical Dataset Strategy	Both positive and negative setpoint segments used	Many prior works use only unidirectional or non-symmetrical data	Using symmetrical data ensures that the ANFIS remains under control during both increasing and decreasing setpoints, improving safety and robustness.
Data Augmentation Technique	Introduces a novel output shift method to modify dataset (Section 2.8)	No such targeted augmentation in prior works	This enhancement significantly reduces overshoot and settling time. It’s a practical method to refine control performance without altering model structure.
RMSE and Performance	RMSE = 0.0081; Overshoot = 2.8%; Settling Time = 1.49s	Higher RMSEs (0.0602–0.265); Overshoot = up to 22.7%; Settling time = up to 84.1s	Your method clearly outperforms prior studies in quantitative metrics, showing the effectiveness of your data preparation and ANFIS tuning strategy.
Implementation Focus	Discusses simulation-to-physical deployment and practical limitations (Section 3.6, 3.7)	Focuses mainly on simulation or academic benchmarking	You offer practical design guidance (e.g., dataset symmetry, augmentation strategy) that can be used by engineers in real systems.
Statistical Validation	One-tailed t-test to confirm significance of improvements (Table 12)	No statistical validation of performance	Including statistical tests demonstrates rigor and makes your conclusions more credible and reproducible.

Table 15: Assumptions and limitations of ANFIS design

Aspect	Assumption	Limitation
Dataset	PID-generated data is sufficient	May not capture all dynamics or disturbances
Process Model	First-order + delay is adequate	Real system may be more nonlinear/complex
Simulation Environment	MATLAB is representative of reality	Ignores hardware, noise, delays, faults
ANFIS Structure	Standard structure is suitable	Suboptimal parameterization can harm performance
Disturbances	Not fully modeled	Real-world disturbances can degrade performance
Fault Handling	No explicit consideration	Vulnerable to unexpected events
Data Dependency	PID dataset captures all relevant dynamics	Poor generalization beyond training data
Setpoint Range	Controller operates within training range	Fails or becomes unstable outside this range
Dataset Symmetry	Positive/negative parts are balanced	Unbalanced data leads to instability
Extrapolation	Behavior predictable within known ranges	Cannot handle unknown/unseen scenarios
Online Adaptability	System dynamics are static	No real-time adaptation capability

4.2 Novel contributions compared to prior work

- Real-world dataset generation from a PSO-tuned PID controller rather than simulated systems.
- A bidirectional (positive and negative) dataset structure is introduced for improved control stability.
- Application of target output-shift data augmentation to enhance transient response.
- Practical design guidelines bridging simulation and real-world deployment, including robustness testing and statistical validation.

Table 14 summarizes the differences between this study and other studies to clarify the novel contributions.

As shown in Table 14, this study outperforms previous works by using real-world PID data, a symmetrical dataset strategy, and a novel data augmentation method, resulting in significantly better control performance. It achieves a lower RMSE, minimal overshoot, and a faster settling time, and its findings are statistically validated, making it more practical and credible than earlier studies.

4.3 Limitations

- The controller is trained within a fixed setpoint range; extrapolation beyond this range may lead to instability.
- Real-time deployment challenges such as hardware limitations and sensor noise were not modeled.
- The controller lacks online learning or adaptation capabilities.
- Table 15 presents the assumptions and limitations of the ANFIS controller design.

4.4 Future work

- The methodology presented in this study is broadly applicable to various systems. It can be utilized to design an Adaptive Neuro-Fuzzy Inference System (ANFIS) controller using datasets derived from PID controllers. Moreover, the performance of the ANFIS controller can be further improved through dataset augmentation techniques, thereby enhancing its overall response.
- Future studies will focus on implementing the ANFIS controller in a physical industrial setting, introducing online adaptation, and expanding the training dataset to include a broader range of operational conditions and disturbance scenarios.

Acknowledgments

The authors would like to acknowledge the staff of the State Company of Fertilizers SCF in Iraq, particularly the cooling towers department staff, who provided us with the necessary practical information for the cooling tower system, and the instrumentation department staff, who installed the Yokogawa recorder locally to calculate the cooling tower system's pH loop transfer function.

References

[1] R. Tiwari and G. S. Mahalpure, “A Detailed Review of pH and its applications,” *Journal of Pharmaceutical and Biopharmaceutical Research*, vol. 6, no. 2, pp. 492–505, Apr. 2025, doi: 10.25082/JPBR.2024.02.001.

[2] S. S. Ram, D. D. Kumar, B. Meenakshipriya, and K. Sundaravadivu, “Designing and comparison of controllers based on optimization techniques for pH neutralization process,” no. Icices, pp. 1–5, 2016, doi: 10.1109/icices.2016.7518847.

[3] A. Fadzlullah*et al., “Design and Simulation of PID Controller for PH Neutralization Process,” Int. J. Innov. Technol. Explor. Eng., vol. 9, no. 3, pp. 2740–2743, 2020, doi: 10.35940/ijitee.c9236.019320.

[4] Y. Xu, X. Zeng, S. Bernard, and Z. He, “Data-driven prediction of neutralizer pH and valve position towards precise control of chemical dosage in a wastewater treatment plant,” J. Clean. Prod., vol. 348, no. February, p. 131360, 2022, doi: 10.1016/j.jclepro.2022.131360.

[5] K. Bingi, R. Ibrahim, M. N. Karsiti, T. D. Chung, and S. M. Hassan, “Optimal PID control of pH neutralization plant,” 2016 2nd IEEE Int. Symp. Robot. Manuf. Autom. ROMA 2016, no. September, 2017, doi: 10.1109/ROMA.2016.7847812.

[6] Z. Meng et al., “Design and Application of Liquid Fertilizer pH Regulation Controller Based on BP-PID-Smith Predictive Compensation Algorithm,” Appl. Sci., vol. 12, no. 12, 2022, doi: 10.3390/app12126162.

[7] T. Kalavathi Devi et al., “Sensor Technology and Regulation method for Sustaining the pH value in Sugar Mechanized Process,” 2021 4th Int. Conf. Electr. Comput. Commun. Technol. ICECCT 2021, 2021, doi: 10.1109/ICECCT52121.2021.9616891.

[8] S. Sakthiya Ram, D. Dinesh Kumar, and B. Meenakshipriya, “Designing of PID controllers for pH neutralization process,” Indian J. Sci. Technol., vol. 9, no. 12, 2016, doi: 10.17485/ijst/2016/v9i12/89940.

[9] Muhlasin, Budiman, M. Ali, A. Parwanti, A. A. Firdaus, and Iswinarti, “Optimization of Water Level Control Systems Using ANFIS and Fuzzy-PID Model,” Proceeding - 2020 3rd Int. Conf. Vocat. Educ. Electr. Eng. Strength. Framew. Soc. 5.0 through Innov. Educ. Electr. Eng. Informatics Eng. ICVEE 2020, 2020, doi: 10.1109/ICVEE50212.2020.9243229.

[10] W. M. Elsrogy, M. A. Fkirin, and M. A. M. Hassan, “Speed control of DC motor using PID controller based on artificial intelligence techniques,” 2013 Int. Conf. Control. Decis. Inf. Technol. CoDIT

2013, pp. 196–201, 2013, doi: 10.1109/CoDIT.2013.6689543.

[11] Y. Wang, N. Zhang, C. Chen, Y. Jiang, and T. Liu, "Nonlinear Adaptive Generalized Predictive Control for PH Model of Nutrient Solution in Plant Factory Based on ANFIS," *Processes*, vol. 11, no. 8, Aug. 2023, doi: 10.3390/pr11082317.

[12] D. K. Ghose, K. Tanaya, A. Sahoo, and U. Kumar, "Performance Evaluation of hybrid ANFIS model for Flood Prediction," 8th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2022, pp. 772–777, 2022, doi: 10.1109/ICACCS54159.2022.9785002.

[13] C. Navaneethakkannan and M. Sudha, "Analysis and implementation of ANFIS-based rotor position controller for BLDC motors," J. Power Electron., vol. 16, no. 2, pp. 564–571, 2016, doi: 10.6113/JPE.2016.16.2.564.

[14] C. eddine, K. MANSOURI, M. mourad, and A. BELMEGUENAI, "Adaptive Neuro-Fuzzy Inference Systems for Modeling Greenhouse Climate," Int. J. Adv. Comput. Sci. Appl., vol. 7, no. 1, pp. 96–100, 2016, doi: 10.14569/ijacsa.2016.070114.

[15] Hidayat, S. Pramonohadi, Sarjiya, and Suharyanto, "A comparative study of PID, ANFIS and hybrid PID-ANFIS controllers for speed control of Brushless DC Motor drive," Proceeding - 2013 Int. Conf. Comput. Control. Informatics Its Appl. "Recent Challenges Comput. Control Informatics", IC3INA 2013, pp. 117–122, 2013, doi: 10.1109/IC3INA.2013.6819159.

[16] E. Joelianto and D. Candra Anura, "Transient response improvement of PID controller using ANFIS-hybrid reference control," Proc. 2011 2nd Int. Conf. Instrum. Control Autom. ICA 2011, no. November, pp. 41–46, 2011, doi: 10.1109/ICA.2011.6130127.

[17] H. Chaudhary, S. Khatoon, and R. Singh, "ANFIS based speed control of DC motor," 2nd IEEE Int. Conf. Innov. Appl. Comput. Intell. Power, Energy Control. with their Impact Humanit. CIPECH 2016, pp. 63–67, 2017, doi: 10.1109/CIPECH.2016.7918738.

[18] D. Yadav and A. Verma, "Behaviour Analysis of PMSM Drive using ANFIS Based PID Speed Controller," 2018 5th IEEE Uttar Pradesh Sect. Int. Conf. Electr. Electron. Comput. Eng. UPCON 2018, pp. 1–5, 2018, doi: 10.1109/UPCON.2018.8596845.

[19] M. H. Jali, N. E. S. Mustafa, T. A. Izzuddin, R. Ghazali, and H. I. Jaafar, "ANFIS-PID controller for arm rehabilitation device," Int. J. Eng. Technol., vol. 7, no. 5, pp. 1589–1597, 2015.

[20] B. Pramod Raja and N. Anbuselvan, "Stability Improvement of Grid Connected MLI using the Controllers (PID & ANFIS) by Limiting Voltage Level," Proc. 3rd Int. Conf. Intell. Eng. Manag. ICIEM 2022, pp. 582–587, 2022, doi: 10.1109/ICIEM54221.2022.9853123.

[21] Y. Z. Maulana, S. Hadisupadmo, and E. Leksono, "Performance analysis of PID controller, fuzzy and ANFIS in pasteurization process," Proc. 2016 Int. Conf. Instrumentation, Control. Autom. ICA 2016, pp. 171–177, 2017, doi: 10.1109/ICA.2016.7811496.

[22] H. Oubehar, A. Selmani, A. Ed-Dahhak, A. Lachhab, M. E. H. Archidi, and B. Bouchikhi, "ANFIS-based climate controller for computerized greenhouse system," Adv. Sci. Technol. Eng. Syst., vol. 5, no. 1, pp. 8–12, 2020, doi: 10.25046/aj050102.

[23] G. Joshi and A. J. Pinto Pius, "ANFIS controller for vector control of three phase induction motor," Indones. J. Electr. Eng. Comput. Sci., vol. 19, no. 3, pp. 1177–1185, 2020, doi: 10.11591/ijeecs.v19.i3.pp1177-1185.

[24] B. Al-Najari, C. Kok Hen, J. Koh Siaw Paw, and A. Fadhil Marhoon, "Design and Implementation of PID Controller for the Cooling Tower's pH Regulation Based on Particle Swarm Optimization PSO Algorithm," Iraqi J. Electr. Electron. Eng., vol. 20, no. 2, pp. 59–67, 2024, doi: 10.37917/ijeee.20.2.5.

[25] R. Zhou et al., "Fuzzy Neural Network PID Strategy Based on PSO Optimization for pH Control of Water and Fertilizer Integration," Appl. Sci., vol. 12, no. 15, 2022, doi: 10.3390/app12157383.

[26] Y. Li, Y. Zhang, S. Cui, Y. Liu, and X. Lv, "Application of optimizing PID Parameters based on PSO in the Temperature Control System of Haematococcus Pluvialis," vol. 2020, no. 2, pp. 1850–1853, 2020, doi: 10.1109/ITAIC49862.2020.9338777.

[27] N. K. Ray, S. K. Mohapatra, and S. S. Dash, "Gravitational Search Algorithm for Optimal Tunning of controller parameters in AVR system," Int. Conf. Comput. Intell. Smart Power Syst. Sustain. Energy, CISPSSE 2020, no. iv, 2020, doi: 10.1109/CISPSSE49931.2020.9212197.

[28] A. A. Ali and M. T. Rashid, "Design PI Controller for Tank Level in Industrial Process," Iraqi J. Electr. Electron. Eng., vol. 18, no. 1, pp. 82–92, 2022, doi: 10.37917/ijeee.18.1.10.

[29] James Kennedy, Russell Eberhart, "Particle Swarm Optimization," Proceedings of ICNN'95 - International Conference on Neural Networks, Perth, WA, Austuralia, 1995, pp. 1942-1948, Vol. 4.

[30] B. Ozgenc, M. Sinasi Ayas, and I. H. Altas, "A Hybrid Optimization Approach to Design Optimally Tuned PID Controller for an AVR System," HORA 2020 - 2nd Int. Congr. Human-Computer Interact. Optim. Robot. Appl. Proc., pp. 0–4, 2020, doi: 10.1109/HORA49412.2020.9152898.

[31] S. P. Simon, L. Dewan, and M. P. R. Prasad, "Design and Analysis of ITAE Tuned Robust PID Controller for Brushed DC Motor," in Proceedings - 2022 IEEE Silchar Subsection Conference, SILCON 2022, Institute of Electrical and

Electronics Engineers Inc., 2022. doi: 10.1109/SILCON55242.2022.10028938.

[32] K. P. Sanoj and V. Dhanya Ram, “Optimal PID Controller Tuning for Multivariable Unstable Systems by Minimizing ITAE Criteria,” in IFAC-PapersOnLine, Elsevier B.V., Mar. 2024, pp. 391–396. doi: 10.1016/j.ifacol.2024.05.067.

[33] A. Dubravic, D. Demirovic, and A. Serifovic-Trbalic, “Optimization of PID Controller Using PSO Algorithm for a First Order Plus Dead Time (FOPDT) Process -A Simulation Study,” in International Conference on Electrical, Computer, and Energy Technologies, ICECET 2022, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICECET55527.2022.9872631.

[34] U. Subramaniam, K. S. Reddy, D. Kaliyaperumal, V. Sailaja, P. Bhargavi, and S. Likhith, “A MIMO–ANFIS-Controlled Solar-Fuel-Cell-Based Switched Capacitor Z-Source Converter for an Off-Board EV Charger,” Energies, vol. 16, no. 4, 2023, doi: 10.3390/en16041693.

[35] M. Petrov, I. Ganchev, and A. Taneva, “Fuzzy PID control of nonlinear plants,” 2002 1st Int. IEEE Symp. , vol. 1, no. September, pp. 30–35, 2002, doi: 10.1109/IS.2002.1044224.

[36] E. Klir, J.R., Sun, C.T., Mizutami, “Neuro Fuzzy and Soft Computing.” 1997, Prentice-Hall Inc., Upper Saddle River, NJ 07458, USA.

[37] M. N. M. Salleh, N. Talpur, and K. H. Talpur, “A Modified Neuro-Fuzzy System Using Metaheuristic Approaches for Data Classification,” in Artificial Intelligence - Emerging Trends and Applications, InTech, 2018. doi: 10.5772/intechopen.75575.

[38] M. N. M. Salleh and K. Hussain, “A Review of Training Methods of ANFIS for Applications in Business and Economics,” International Journal of u- and e- Service, Science and Technology, vol. 9, no. 7, pp. 165–172, Jul. 2016, doi: 10.14257/ijunesst.2016.9.7.17.

[39] A. Yonar and H. Yonar, “Modeling air pollution by integrating ANFIS and metaheuristic algorithms,” Model Earth Syst Environ, vol. 9, no. 2, pp. 1621–1631, Jun. 2023, doi: 10.1007/s40808-022-01573-6.

[40] V. S. Ghomsheh, M. A. Shoorehdeli, and M. Teshnehlab, "Training ANFIS structure with modified PSO algorithm," in Proceedings of the 2007 Mediterranean Conference on Control & Automation (MED'07), Athens, Greece, pp. 1–6, July 2007, doi: 10.1109/MED.2007.4433927.

[41] S. S. . Navghare, G. G. . Bodhe, and Shruti, “Design of Adaptive pH Controller using ANFIS,” Int. J. Comput. Appl., vol. 33, no. 6, pp. 41–48, 2011, [Online]. Available: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Design+of+Adaptive+pH+Controller+using+ANFIS#0

[42] A. S. Mota, M. R. Menezes, J. E. Schmitz, T. V. Da Costa, F. V. Da Silva, and I. C. Franco, “Identification and online validation of a pH neutralization process using an adaptive network-based fuzzy inference system,” Chem. Eng. Commun., vol. 203, no. 4, pp. 516–526, 2016, doi: 10.1080/00986445.2015.1048799.

[43] M. S. Gaya, N. A. Wahab, Y. M. Sam, and S. I. Samsuddin, “ANFIS based effluent pH quality prediction model for an activated sludge process,” Adv. Mater. Res., vol. 845, pp. 538–542, 2014, doi: 10.4028/www.scientific.net/AMR.845.538.

[44] S. K. Sunori et al., “Neuro-fuzzy Controller Design for pH Control in Sugar Refineries,” 8th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2022, pp. 197–202, 2022, doi: 10.1109/ICACCS54159.2022.9785146.