Article

Adaptive Neural Network Q-Learning-Based Full Recurrent Adaptive NeuroFuzzy Nonlinear Control Paradigms for Bidirectional-Interlinking Converter in a Grid-Connected Hybrid AC-DC Microgrid

Muhammad Awais [1,†] , Laiq Khan [2,†] , Said Ghani Khan [3] , Qasim Awais [4] and Mohsin Jamil [5,*]

1 Department of Electrical and Computer Engineering, COMSATS University Islamabad, Abbottabad Campus, Abbottabad 22060, Pakistan
2 Department of Electrical and Computer Engineering, COMSATS University Islamabad, Islamabad 45550, Pakistan
3 Department of Mechanical Engineering, College of Engineering, University of Bahrain, Isa Town 32038, Bahrain
4 Department of Electronics and Computer Science, Fatima Jinnah Women University, Rawalpindi, Old Presidency, Rawalpindi 46000, Pakistan
5 Department of Electrical and Computer Engineering, Faculty of Engineering and Applied Sciences, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada
* Correspondence: mjamil@mun.ca; Tel.: +1-709-864-2751
† These authors contributed equally to this work.

Abstract: The stability of a hybrid AC-DC microgrid depends mainly upon the bidirectional interlinking converter (BIC), which is responsible for power transfer, power balance, voltage solidity, frequency and transients sanity. The varying generation from renewable resources, fluctuating loads, and bidirectional power flow from the utility grid, charging station, super-capacitor, and batteries produce various stability issues on hybrid microgrids, like net active-reactive power flow on the AC-bus, frequency oscillations, total harmonic distortion (THD), and voltage variations. Therefore, the control of BIC between AC and DC buses in grid-connected hybrid microgrid power systems is of great importance for the quality/smooth operation of power flow, power sharing and stability of the whole power system. In literature, various control schemes are suggested, like conventional droop control, communication-based control, model predictive control, etc., each addressing different stability issues of hybrid AC-DC microgrids. However, model dependence, single-point-failure (SPF), communication vulnerability, complex computations, and complicated multilayer structures motivated the authors to develop online adaptive neural network (NN) Q-learning-based full recurrent adaptive neurofuzzy nonlinear control paradigms for BIC in a grid-connected hybrid AC-DC microgrid. The proposed strategies successfully ensure the following: (i) frequency stabilization, (ii) THD reduction, (iii) voltage normalization and (iv) negligible net active-reactive power flow on the AC-bus. Three novel adaptive NN Q-learning-based full recurrent adaptive neurofuzzy nonlinear control paradigms are proposed for PQ-control of BIC in a grid-connected hybrid AC-DC microgrid. The control schemes are based on NN Q-learning and full recurrent adaptive neurofuzzy identifiers. Hybrid adaptive full recurrent Legendre wavelet-based Neural Network Q-learning-based full recurrent adaptive NeuroFuzzy control, Hybrid adaptive full recurrent Mexican hat wavelet-based Neural Network Q-learning-based full recurrent adaptive NeuroFuzzy control, and Hybrid adaptive full recurrent Morlet wavelet-based Neural Network Q-learning-based full recurrent adaptive NeuroFuzzy control are modeled and tested for the control of BIC. The controllers differ from each other, based on variants used in the antecedent part (Gaussian membership function and B-Spline membership function), and consequent part (Legendre wavelet, Mexican hat wavelet, and Morlet wavelet) of the full recurrent adaptive neurofuzzy identifiers. The performance of the proposed control schemes was validated for various quality and stability parameters, using a simulation testbench in MATLAB/Simulink. The simulation results were bench-marked against an aPID controller, and each proposed control scheme, for a simulation time of a complete solar day.

1. Introduction

Renewable resource-based distributed generators (DGs), like wind turbines (WTs), photovoltaic (PV) generators, micro-turbines (MTs) etc., are in fashion nowadays to reduce greenhouse emissions and to improve the efficiency of the electric power system. A large power system is divided into many small microgrids by the integration of DGs, energy storage system (ESS), and loads, and, thus, making it more controllable, flexible, and effective [1–4].

The DC microgrid has no frequency issues, produces low power loss, due to the absence of reactive power, and has simple dynamic variations and easy power conversions. These advantages of the DC microgrid make it superior to the AC microgrid. However, hybrid AC-DC microgrids are gaining huge attention, due to their capability of combining the benefits of both AC and DC networks [1,2,5–8]. The AC and DC microgrids are connected through a bidirectional AC-DC interlinking converter, thus, making a hybrid AC-DC microgrid that is capable of utilizing the advantages of both sides and which can entertain multiple types of loads. However, new technological challenges arise for the smooth and reliable operations of hybrid AC-DC microgrids [1,2,9–13].

AC and DC subgrids are interconnected via bidirectional interlinking converters (BIC). The BIC is responsible for smooth bidirectional power flow, power balance, microgrid stability, voltage solidity, frequency and transients firmness. Failure in the right performance of BIC leads to instability of the microgrid and causes power loss. The unpredictable nature of renewable DGs, nonlinear uncertain loads, switching harmonics, and bidirectional power flow between the utility grid and storage devices produces a variety of stability issues, like net active-reactive power flow on the AC-bus, increased THD level, voltage fluctuation, and frequency instability. Therefore, the key challenge in hybrid AC-DC microgrid is the control problem [2,14–19].

Generally, the whole network and, particularly, the BIC requires intelligent adaptive control to ensure the coordination of DC and AC subsystems and appropriate power flow across the hybrid microgrid [1,2,20,21]. The problem is non-trivial and becomes more difficult, due to the difference in response time of AC and DC networks. The faster response of the DC network over the AC network, and interactions between corresponding sides, may lead to destabilization of the power system [22,23]. Thus, the key issue in the hybrid AC-DC microgrid is the control of BIC.

Several control schemes are reported in the literature for the control of BIC in a hybrid AC-DC microgrid. Reference [24] presented coordination control and dual closed-loop control for the interlinking converter. However, the study lacked real time simulation results and presented two cases with specific variation in the generation and loads. Reference [2] designed a uniform control scheme for the bidirectional interlinking converter for economical and resilient objectives of hybrid AC-DC microgrids. However, the study lacked a real time simulation testbed and was studied for specific power flow cases. Reference [25] presented the voltage regulation of the DC-bus by designing a model-based controller for the bi-directional interlinking converter. However, the study required proper modeling of the system and the PID control used might not work for large complex power systems. Reference [26] presented the application of a passivity framework to AC-DC grids. However, a lossless line was assumed in the study and no real time simulation results were presented. Reference [27] presented virtual inertia support for BIC in the hybrid AC-DC microgrid. However, use of PI control and only the regulation of active power reduced its impact for large power systems with various operating points. Many other articles are found in the literature but the lack of real time systems and a small number of DGs reduce their productivity.

This article presents three novel control schemes for a bidirectional interlinking converter for a hybrid AC-DC microgrid. The control schemes are hybrid adaptive NN Q-learning-based full recurrent adaptive NeuroFuzzy (FRNF) control schemes. The difference in the three proposed control schemes is based on the different variants in the antecedent and consequent parts of adaptive NeuroFuzzy architecture. The variants in the antecedent part are Gaussian membership function and B-Spline membership function, while the variants used in the consequent part are Mexican hat wavelet, Legendre wavelet, and Morlet wavelet. The real time simulation testbed consisted of multiple DGs and an utility grid connected to DC- and AC-buses via converters and transformers. The real time solar and wind profiles, along with maximum power point tracking, used in this article were obtained from our previously published articles [28–32]. The use of multiple DGs, real time solar and wind profiles, and one day complete simulation are the key features of this study. The performance of the proposed control scheme was verified through various quality and stability parameters, using MATLAB/SIMULINK software, 2015a. The various results obtained from the simulation were compared with an aPID controller and against each other. Table 1 gives the features and contribution of the proposed solution, compared to previously reported work.

The article is divided into sections for understanding. Section 2 describes the testbed, Sections 4–13 discusses the detailed modeling of the proposed control schemes, Section 14 discusses the various results of the simulation, Section 15 concludes the article, and Section 16 shows suggested future work.

Remark

This work is the extension of the authors' previous work reported in [28–32], which focused on maximum power point tracking of the variable speed wind energy conversion system, PV-farm, swift response of solid oxide fuel cell, and charging/discharging scheduling of the plug in hybrid vehicles (PHEVs) in a grid-connected hybrid AC-DC microgrid. The voltage, frequency and THD regulations were considered in the above studies. However, the control of the interlinking converter in this study regulates the concerns (voltage, frequency, THD) at a higher efficiency level. Hence, the new control typologies were developed to increase the stability and reliability of the hybrid microgrid.

Table 1. Main contributions, features and limitations of the existing and proposed work.

Ref. No.	Year	Frequency Regulation	Voltage Regulation	THD Reduction		Limitations/Features
[33]	2016	✗	✓	✓	☞	Independent of communication and model parameters. However, general hybrid AC-DC model is assumed with negligence of frequency regulation.
[34]	2017	✗	✓	✗	☞	Applicable only on low voltage hybrid grid. Frequency regulation and THD reduction are neglected.
[35]	2018	✗	✓	✗	☞	Leader follower consensus and particle swarm optimization are prone to SPF [1]. Only voltage regulation is considered.
[36]	2019	✗	✗	✗	☞	Synchronization control increases the risk of SPF. Only voltage regulation is considered.
[37]	2020	✓	✓	✗	☞	Hierarchical control strategy has drawback of SPF. THD reduction is neglected.
[18]	2021	✗	✓	✓	☞	Model predictive control algorithm requires model and parameters. It is complex and consumes more time. Moreover, frequency regulation is neglected.
[38]	2021	✗	✗	✓	☞	Plug-in-hybrid vehicles and other renewable resources are not taken into account. THD reduction is neglected.
[39]	2021	✗	✓	✗	☞	PI controller cannot operate under highly nonlinear conditions and prone to SPF. Only voltage regulation is considered.
[40]	2022	✓	✓	✗	☞	Coordinated control scheme requires communication network. THD regulation is neglected.
[41]	2022	✗	✓	✓	☞	Multiple control schemes are integrated. SPF of any control scheme results in power instability that may lead to blackout. Only wind power, PV and BSS are considered. Frequency regulation is neglected.
Proposed	2022	✓	✓	✓	☞	Model free control strategy is used. No communication required. Advantages of Q-learning and full recurrent neurfuzzy are combined to avoid SPF. Complete microgrid with multiple DGs is considered. Real time solar irradiance, temperature and wind speed is used. Optimal active-reactive power flow, voltage-frequency regulation, and THD reduction are considered.

[1] SPF = Single-Point-Failure.

2. System Overview and Model Description

Figure 1 sketches the proposed microgrid. It consists of AC and DC-buses which are connected to each other via a main interlinking inverter. The DC-bus has multiple renewable energy resources connected to it through various AC-DC converters. The 165 F super-capacitor bank, 260 kW Photovoltaic (PV) array, 260 kW solid oxide fuel cell (SOFC), 150 kW electrolyzer, 200 Ah batteries, and 100 kW wind turbine (WT) are connected to the DC-bus through AC-DC and DC-DC converters. The AC-bus is connected with 200 kVA micro-turbine, 11 kV grid, bidirectional smart charging station (CS), and the AC-load through transformers, as well as AC-DC and AC-DC-AC converters. The DC-bus is interlinked with the AC-bus through the main inverter. The technical details of the proposed microgrid and its components are given in Appendix A [28–32].
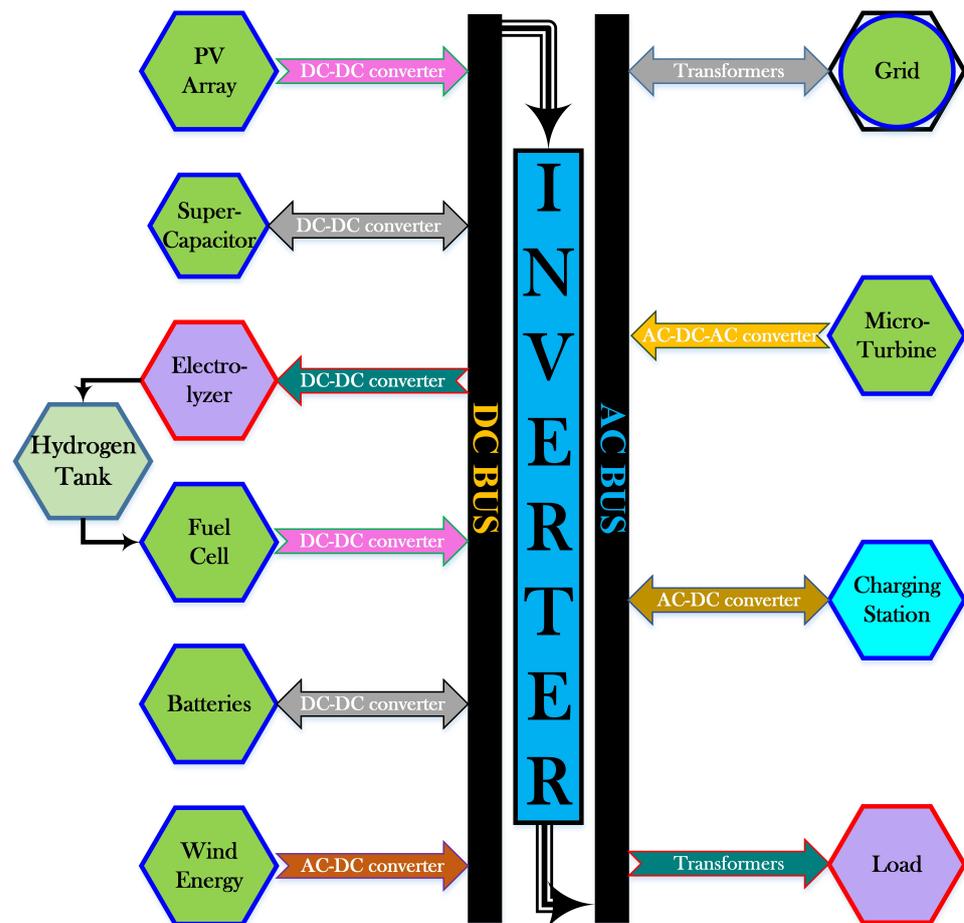
Figure 1. General sketch of the proposed microgrid.

Modeling and Description of Interlinking Inverter

AC and DC sub-grids are linked with each other via an interlinking power converter in the smart microgrid hybrid power system (SMG-HPS). The role of the interlinking converter is important in terms of bidirectional power transfer and stability of the whole SMG-HPS. The uncertain load and generation conditions enforces the interlinking converter to operate in various modes. If no power transfer takes place to either side, the operational mode is known as stop mode. If power transfers from the DC subgrid side to the AC subgrid side, the operational mode is known as inverter mode. If power transfers from the AC subgrid side to the DC subgrid side, the operational mode is known as rectifier mode.

Smooth power transfer between AC and DC subgrids is guaranteed by the interlinking converter. During grid-connected mode, the power balance is ensured by the utility grid, while the DC-bus voltage stability is ensured by an interlinking converter. The

increasing use of electronic devices in the power sector, at the household and industrial levels, dramatically increases the complexity of the power system and reduces efficiency. The increase in THD level, and fluctuations in the frequency and voltage profiles, reduce the life of many electric components, as well as decrease the reliability and quality of power.

The specifications of interlining inverter used in this research work are given in Table A9, Appendix A.9.

The interlinking or main inverter acts as a bridge between AC and DC sub-grids. It controls the active and reactive power and, thus, ensures the voltage stability. The terminal voltages of this voltage source converter are as follows:

$$v_{an} = L_{11}\frac{di_a}{dt} + R_{11}i_a + v_{an-AC} \tag{1}$$

$$v_{bn} = L_{11}\frac{di_b}{dt} + R_{11}i_b + v_{bn-AC} \tag{2}$$

$$v_{cn} = L_{11}\frac{di_c}{dt} + R_{11}i_c + v_{cn-AC} \tag{3}$$

where, inductance and resistance are shown by L_{11} and R_{11}, respectively. The dq-reference frame is calculated by using Park's transformation as follows:

$$\frac{di_d}{dt} = \omega i_q - \frac{R_{11}i_d}{L_{11}} + \frac{1}{L_{11}}(v_d - v_d^*) \tag{4}$$

$$\frac{di_q}{dt} = \omega i_d - \frac{R_{11}i_q}{L_{11}} + \frac{1}{L_{11}}(v_q - v_q^*) \tag{5}$$

Voltages in the dq-reference frame are calculated using the following equations:

$$v_d = L_{11}\frac{di_d}{dt} - \omega L_{11}i_q + v_d^* + R_{11}i_d \tag{6}$$

$$v_q = L_{11}\frac{di_q}{dt} - \omega L_{11}i_d + v_q^* + R_{11}i_q \tag{7}$$

In the dq-reference frame the power is calculated using the following relation:

$$P_{dq} = \frac{3}{2}(v_d i_d + v_q i_q) \tag{8}$$

3. Supervisory Control of Microgrid and Operation Strategy

The principal purpose of supervisory control is to provide continuous and reliable power to all the connected loads. The loads used in this research work are residential load (P_L) and CS load (P_{CS}). The significant power is transmitted to these main loads. However, the excess power during off-peak hours is transmitted to BSS and SC for storing purposes and later use. The electrolyzer also acts as a load and obtains power from the DC-bus of SMG-HPS. The supervisory control enforces the use of renewable energy to entertain loads and store electrical power in various storing devices. However, in the case of a deficit power, the extra power is obtained from BSS, SC, and CS (PHEVs).

The modes of operation of supervisory control depend on the balance of available and required power.

3.1. Modes of Operation of Supervisory Control System

The modes are based on the deficit and excess power. The description of the modes used in this research work is given under.

3.1.1. Mode of Power Deficit

The operational modes of supervisory control systems with deficit power are described below.

Mode 1: WT, PV and BSS Fulfill the Load Demand

In this mode, only RES WT, PV, and BSS provide power to fulfil the load demand. The WT and PV provide their maximum output powers while BSS power is restricted to up to 20% of the state-of-charge (SOC). The mathematical representation of this mode is:

$$P_{Load} \geq P_{ren} + P_{BSS} \tag{9}$$

where, $P_{Load} = P_{CS} + P_L = V_L \times I_L$ is the load demand, $P_{ren} = P_{WT} + P_{PV}$ is the cumulative renewable power, and P_{BSS} is the BSS power for $SOC_{BSS} \geq 20\%$. The BSS discharges in this mode of operation.

Mode 2: WT, PV, BSS and SC Fulfill the Load Demand

If the load demand is higher than the power produced in mode 1, then deficit power is obtained from SC to fulfil the load demand. The mathematical representation of this mode is:

$$P_{Load} \geq P_{ren} + P_{BSS} + P_{SC} \tag{10}$$

where P_{SC} is the power obtained by SC for $SOC_{SC} \geq 20\%$. All other sources are kept OFF and power is obtained only through RES, BSS, and SC during this mode of operation.

Mode 3: WT, PV, BSS, SC and SOFC Fulfill the Load Demand

If the power demand of load is not fulfilled from RES and the discharging of BSS and SC, then the deficit power is obtained from SOFC to satisfy the load demand. The SOFC tracks a power reference that depends on the deficit power. The mathematical representation of this mode is:

$$P_{Load} \geq P_{ren} + P_{BSS} + P_{SC} + P_{SOFC} \tag{11}$$

where P_{SOFC} is the power obtained from SOFC.

Mode 4: WT, PV, BSS, SC, SOFC and Grid Fulfill the Load Demand

If the power required by the load increases from mode 3, then the extra power deficit is obtained from the utility grid to satisfy the load requirements. The mathematical expression for this mode is:

$$P_{Load} \geq P_{ren} + P_{BSS} + P_{SC} + P_{SOFC} + P_{Grid} \tag{12}$$

where P_{Grid} is the power obtained from the utility grid.

Mode 5: WT, PV, BSS, SC, SOFC, Grid and MT Fulfill the Load Demand

If the power delivered from mode 4 cannot satisfy the load demands during peak hours, then the deficit power is obtained from the MT. The mathematical relation for this mode is:

$$P_{Load} \geq P_{ren} + P_{BSS} + P_{SC} + P_{SOFC} + P_{Grid} + P_{MT} \tag{13}$$

where P_{MT} is the power obtained from MT to satisfy the load demand.

3.1.2. Modes of Excess Power

The modes of supervisory control systems with excess power are described below.

Mode 6: Excess Power Given to Electrolyzer

If the power obtained from RES and BSS is more than the power required by the load, then the extra energy is absorbed by the electrolyzer. This mode can be mathematically represented as:

$$P_{Load} \leq P_{ren} + P_{BSS} \tag{14}$$
$$P_{excess} = P_{Load} - P_{ren} - P_{BSS} \tag{15}$$
$$P_{elect} = P_{excess} \tag{16}$$

where, P_{excess} is the extra amount of power, P_{elect} is the electrolyzer power. The BSS is in discharge mode.

Mode 7: Excess Power Given to SC and Electrolyzer

If the power generated from RES and BSS is greater than the load demand, while SOC of SC is less than 90%, then the excess energy is utilized to charge SC. In the case of more surplus power, the electrolyzer is entertained as well. The mathematical representation of this mode is:

$$P_{Load} \leq P_{ren} + P_{BSS} \tag{17}$$
$$P_{excess} = P_{Load} - P_{ren} - P_{BSS} \tag{18}$$
$$P_{SC} = P_{excess} \tag{19}$$
$$P_{elect} = P_{excess} - P_{SC} \tag{20}$$

where P_{SC} shows the power absorbed by SC in charging mode. In this mode, RES and BSS are providing power, while load, SC and electrolyzer are absorbing power.

Mode 8: Excess Power Given to SC, Grid, and Electrolyzer

In this mode of operation, the power generated from the RES is greater than the load demand. Therefore, the excess energy is given to SC for its charging. If SMG-HPS still has surplus power generated from RES then the excess energy is delivered to the utility grid during its peak hours and the remaining surplus power is given to the electrolyzer for the production of hydrogen gas. The mathematical expression for this mode is:

$$P_{Load} \leq P_{ren} + P_{BSS} \tag{21}$$
$$P_{excess} = P_{Load} - P_{ren} - P_{BSS} \tag{22}$$
$$P_{SC} = P_{excess} \tag{23}$$
$$P_{Grid} = P_{excess} - P_{SC} \tag{24}$$
$$P_{elect} = P_{excess} - P_{SC} - P_{Grid} \tag{25}$$

where P_{Grid} shows the power absorbed by the utility grid during peak hours.

Mode 9: Excess Power Given to SC, Grid, and Electrolyzer, while BSS Is Disconnected

This mode of operation is similar to mode 8. However, in this mode, the BSS is disconnected, i.e., $P_{BAT} = 0$. The excess power generated from RES is delivered to SC for charging, for the utility grid during peak hours, and for the electrolyzer for production of hydrogen gas, respectively. The mathematical expression is:

$$P_{Load} \leq P_{ren} \tag{26}$$
$$P_{excess} = P_{Load} - P_{ren} \tag{27}$$
$$P_{SC} = P_{excess} \tag{28}$$
$$P_{Grid} = P_{excess} - P_{SC} \tag{29}$$
$$P_{elect} = P_{excess} - P_{SC} - P_{Grid} \tag{30}$$

The supervisory control flow chart is given in Figure 2.

$P_{ren} = P_W + P_{PV}$
$P_{ref} = $ reference signal for activation

Start

$P_{ren}, P_{grid}, P_{load}, P_{sofc}, P_{bat}, P_{mt}, P_{elect}, SOC_{bat}, SOC_{sc}$

From 1

$P_{load} = P_{ren} \pm P_{bat} \pm P_{sc} \pm P_{sofc} - P_{elect} \pm P_{grid} + P_{mt}$

$SOC_{sc} > 20\%$ — Yes / No

$P_{sc\text{-}d\text{-}ref} = 1$

$P_{load} > P_{ren} + P_{bat\text{-}d} + P_{sc\text{-}d}$ — No / Yes — Goto 1

$P_{sofc\text{-}ref} = 1$

$P^*_{load} = P_{load} - P_{pro}$

$P_{load} > P_{ren} + P_{bat\text{-}d} + (P_{sc\text{-}d}) + P_{sofc}$ — No — Goto 1 / Yes

$P^*_{load} > P_{ren} + P_{bat\text{-}d} + (P_{sc\text{-}d}) + P_{sofc} + P_{mt} + P_{grid}$ — No — Goto 1 / Yes

$P_{mt\text{-}ref} = 1$

$P^*_{load} = P_{load} - P_{dc}$

$P_{load} > P_{ren} + P_{bat\text{-}d} + (P_{sc\text{-}d}) + P_{sofc} + P_{mt}$ — Yes

$P^*_{load} > P_{ren} + P_{bat\text{-}d} + (P_{sc\text{-}d}) + P_{sofc} + P_{mt} + (P_{grid})$ — No — Goto 1 / Yes

$P^*_{load} = P_{load} - P_H$

Grid unavailable or peak hour — Yes / No

$P_{grid\text{-}ref} = 1$

$P_{load} > P_{ren} + P_{bat\text{-}d} + (P_{sc\text{-}d}) + P_{sofc} + P_{mt} + P_{grid}$ — Yes — Goto 1

$P_{load} > P_{ren} + P_{bat\text{-}d} + (P_{sc\text{-}d}) + P_{sofc} + P_{mt} + P_{grid}$ — No — Goto 1

$P_{load} > P_{ren}$ — Yes / No

$SOC_{bat} > 20\%$ — Yes / No — Goto 2

$SOC_{bat} > 90\%$ — Yes — Goto 3 / No

$P_{bat\text{-}d\text{-}ref} = 1$

$P_{bat\text{-}c\text{-}ref} = 1$

$P_{load} > P_{ren} + P_{bat\text{-}d}$ — Yes / No — Goto 1

$P_{load} - (P_{ren} + P_{bat\text{-}c}) < 0$ — No — Goto 1 / Yes

$SOC_{sc} > 90\%$ — Yes / No

$P_{sc\text{-}c\text{-}ref} = 1$

From 3

$P_{load} - (P_{ren} + P_{bat\text{-}c} + P_{sc\text{-}c}) < 0$ — No — Goto 1 / Yes

If H_2 tank is full — Yes / No

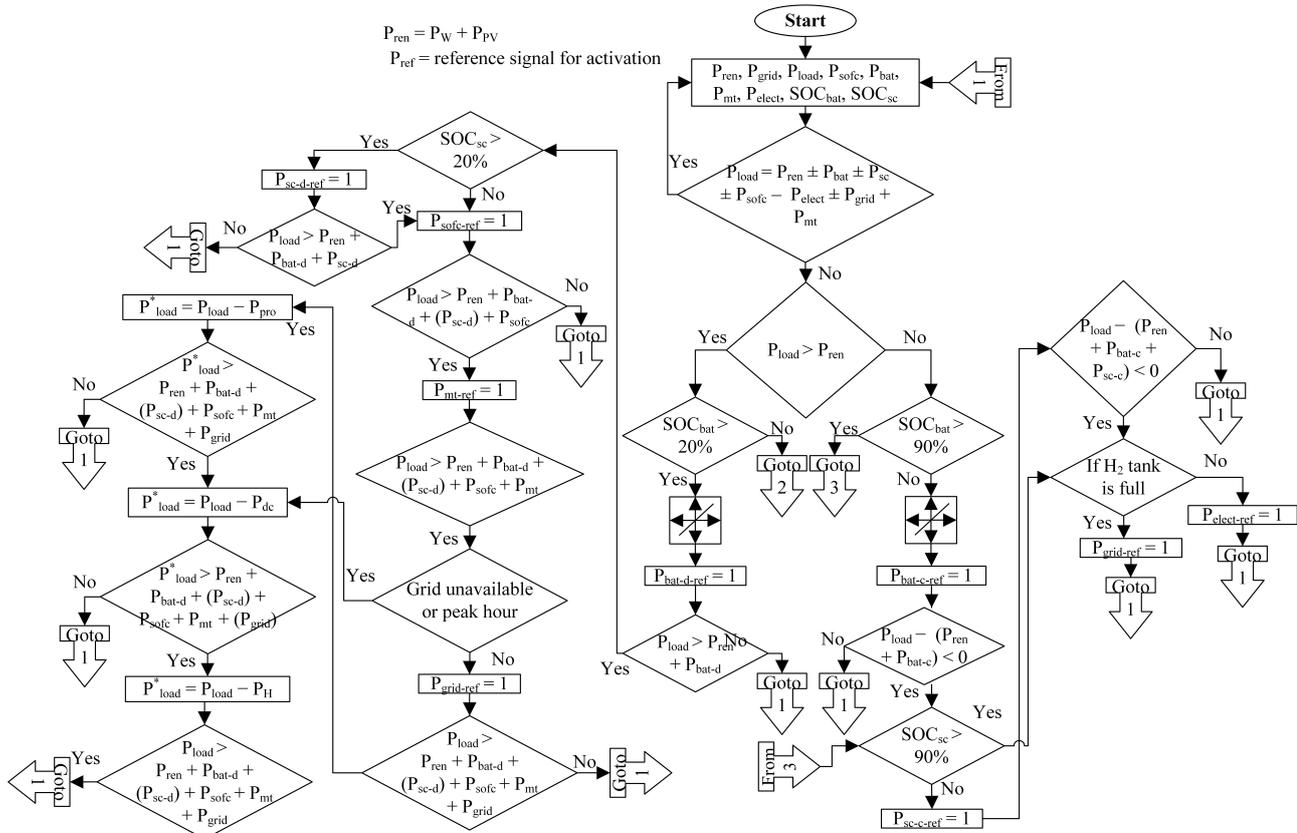$P_{grid\text{-}ref} = 1$

$P_{elect\text{-}ref} = 1$

Goto 1

Figure 2. Flow chart for supervisory control system. * shows the new/required power.

4. Description and Modeling of Proposed Control Schemes

This section discusses the mathematical modeling of, and describes, the three proposed hybrid adaptive NN Q-learning0based full recurrent adaptive NeuroFuzzy (NNQLNF) control schemes. The proposed control schemes consist of two parts: NeuroFuzzy parameter tuning (NFPT) and optimal action-value function $Q^*(x, u)$ estimator network (QEN). $Q^*(x, u)$ is estimated using a backpropagation (BP) neural network (NN). The NeuroFuzzy systems used are discussed in Sections 6 and 7. The NeuroFuzzy parameters are updated using $Q^*(x, u)$. The action exploration modifier (AEM) guarantees the trail of all possible actions. The NNQLNF control technique does not depend on prior information for future driving conditions, due to online adaptation of the learning algorithm and fuzzy parameters, which make it a prominent and advantageous control paradigm.

5. Neural Network Q-Learning-Based Full Recurrent Adaptive NeuroFuzzy Control

In control problems, Q-learning and actor-critic learning are two major types of reinforcement learning. The actor-critic learning estimates the state value function and chooses an optimal action for every state. However, in Q-learning, the system approximates the action-value function for all action-state sets and chooses the optimal control technique based on this [42,43].

The schematic of NNQLNF is given in Figure 3. It consists of two parts: the QEN network, that estimates $Q^*(x, u)$, and the FRNF architecture, that tunes the parameters. The QEN comprises a BP NN for estimation purpose, while the FRNF is an intelligent identifier discussed in Sections 6 and 7. Three different FRNF architectures are used as FPT, and, thus, three distinct hybrid adaptive FQL embedded FRNF control paradigms are derived, which are used in this research work.

5.1. Back Propagation NN for Estimating $Q^*(x, u)$

If the initial state is x and initial action is u, then the expected discounted sum of rewards gives the action-value function $Q(x, u)$, as given below [42]:

$$Q(x, u) = E(\sum_{k=1}^{\infty} \gamma^k r_{t+k+a} | x_t = x, u_t = u) \tag{31}$$

where, E is the expected value function, and u is the action value. The optimal action-value $Q^*(x, u)$ is given as:

$$Q^*(x, u) = E(r(x_{t+1}) + \gamma \max_{u'} Q^*(x_{t+1}, u') | x_t = x, u_t = u) \tag{32}$$

QEN approximates/predicts the ideal action-value function $Q^*(x, u)$ related to different inputs and control output states. The approximation property of BP NN is used for the estimation of $Q^*(x, u)$.

The internal architecture of QEN is given in Figure 4. The network consists of a three-layer structure: input node, hidden node, and output node. The inputs are state variable and control action. However, the output is the required QEN i.e., $Q(x, u)$.

The $Q(x, u)$ can be represented mathematically as:

$$Q(x, u) = f(V) \tag{33}$$

where, V gives the weighted sum input of the output node and is given as:

$$V = \sum_{k=1}^{10} w_{(40+k)} \times y_k \tag{34}$$

where, $w_{(40+k)}$ represents the weight between hidden and output node and y_k is the output of hidden layer and is given as:

$$y_k = f(a_k) \tag{35}$$

where, a_k gives the added input of kth hidden node and is given as:

$$a_k = \sum_{i=1}^{4} x_i \times w_{(i-1,j)} \tag{36}$$

where, $w_{(i-1,j)}$ shows the weight between input and hidden node, $x = x_i$ represents the QEN's input, and f shows the activation function of the node. The sigmoid function is used as the activation function of the node and is given as:

$$f(x) = \frac{1}{1 + exp^{-x}} \tag{37}$$

Generalized policy iteration is used for tuning the parameters of QEN. The optimal action-value function is approximated by reducing temporal difference (TD) error, δ_t continuously with the help of NN. The TD error is given as follows:

$$\delta_t = r_{t+1} + \gamma \max_{u'} Q^*(x_{t+1}, u') - Q(x_t, u_t) \tag{38}$$

The cost function is the mean square error and is given as:
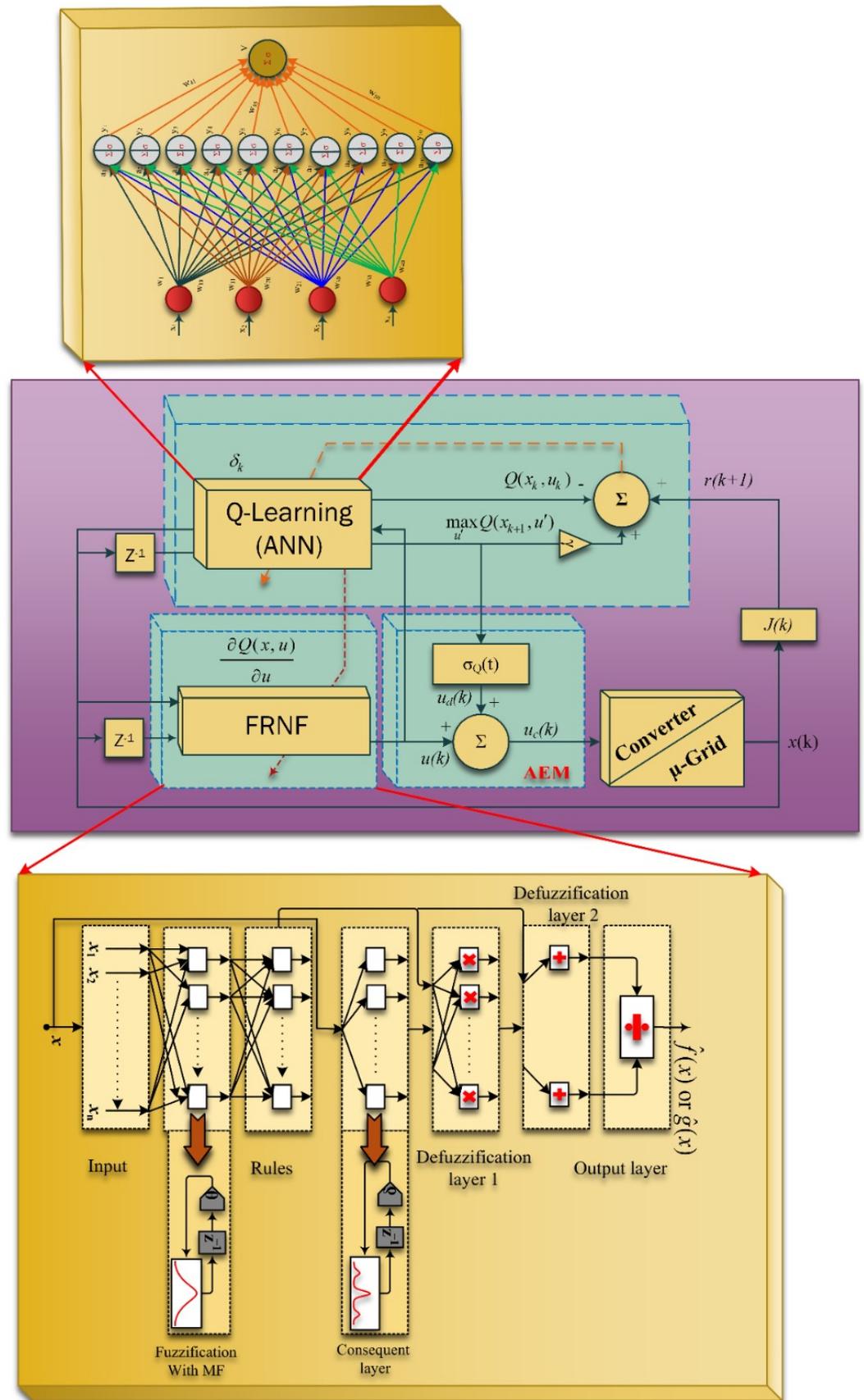
$$E = \frac{1}{2} \delta_t^2 \tag{39}$$

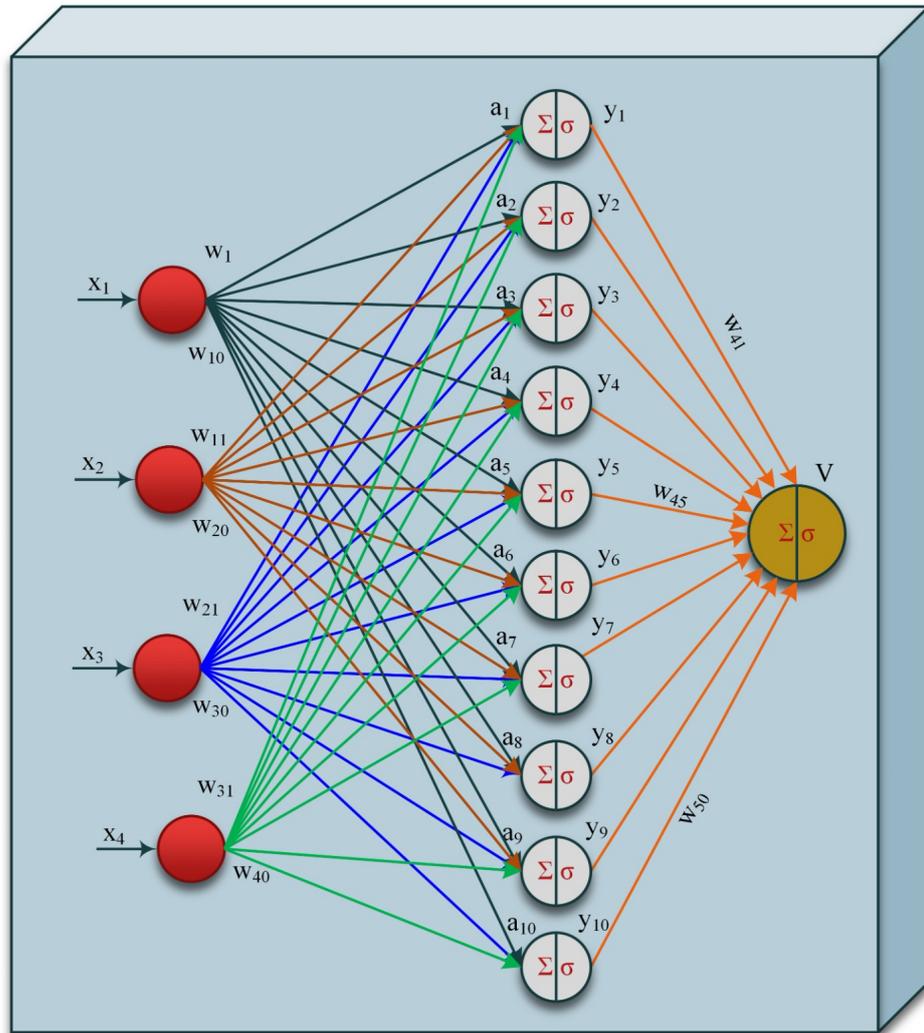**Figure 3.** Neuro Network Q-learning-based full recurrent adaptive NeuroFuzzy internal architecture.

Figure 4. QEN internal architecture.

The gradient descent method is used for fast convergence. The rule for weight update is given as:

$$w(t+1) \quad = \quad w(t) - \zeta \frac{\partial E}{\partial w} \tag{40}$$

$$\frac{\partial E}{\partial w} \quad = \quad \sigma_t \frac{\partial \delta_t}{\partial w} = -\delta_t \frac{\partial Q(x_t, u_t)}{\partial w} \tag{41}$$

$\frac{\partial Q(x_t, u_t)}{\partial w}$ can be obtained by using chain rule for $w_{(40+k)}$ and $w_{(i-1,j)}$ as shown below:

$$\frac{\partial Q(x_t, u_t)}{\partial w_{(40+k)}} = \frac{\partial Q(x_t, u_t)}{\partial V} \times \frac{\partial V}{\partial w_{(40+k)}}$$

$$= f'(V) \times y(k) \tag{42}$$

$$= y_k \times Q(x_t, u_t) \times (1 - Q(x_t, u_t))$$

$$k = 1, \ldots, 10 \tag{43}$$

$$
\begin{aligned}
\frac{\partial Q(x_t, u_t)}{\partial w_{(i-1,j)}} &= \frac{\partial Q(x_t, u_t)}{\partial V} \times \frac{\partial V}{\partial y_j} \times \frac{\partial y_j}{\partial a_j} \times \\
&\quad \frac{\partial a_j}{\partial w_{(i-1,j)}}
\end{aligned}
\tag{44}
$$

$$
\begin{aligned}
&= f'(V) \times y_j \times w_{(40+j)} \\
\times f'(a_j) \times u_i &
\end{aligned}
\tag{45}
$$

$$
\begin{aligned}
&= w_{(40+j)} \times u_i \times Q(x_t, u_t) \times \\
&\quad [1 - Q(x_t, u_t)] \times y_j \times [1 - y_j] \\
&\qquad i = 1, \ldots, 4 \quad j = 1, \ldots, 10
\end{aligned}
\tag{46}
$$

$\frac{\partial Q(x_t, u_t)}{\partial u}$ can also be obtained using chain rule as given below:

$$
\begin{aligned}
\frac{\partial Q(x_t, u_t)}{\partial u} &= \frac{\partial Q(x_t, u_t)}{\partial V} \times
\end{aligned}
\tag{47}
$$

$$
\sum_{j=1}^{10} \left(\frac{\partial V}{\partial y_j} \times \frac{\partial y_j}{\partial w_j} \times \frac{\partial w_j}{\partial u} \right)
$$

$$
= f'(V) \times \sum_{j=1}^{10} \left(w_j^{(1)} f'(w_i) \times w_{j,4}^{(2)} \right)
\tag{48}
$$

$$
\begin{aligned}
&= Q(x_t, u_t) \times (1 - Q(x_t, u_t)) \times \\
&\quad \sum_{j=1}^{10} \left(w_i^{(1)} \times w_{i,4}^{(2)} \times y_j \times [1 - y_j] \right)
\end{aligned}
\tag{49}
$$

Also;

$$
\begin{aligned}
\frac{\partial Q(x_t, u_t)}{\partial u} &= \frac{\partial Q(x_t, u_t)}{\partial V} \times
\end{aligned}
\tag{50}
$$

$$
\sum_{j=1}^{10} \left(\frac{\partial V}{\partial y_j} \times \frac{\partial y_j}{\partial a_j} \times \frac{\partial a_j}{\partial u} \right)
$$

$$
= f'(V) \times \sum_{j=1}^{10} \left(w_{(40+j)} f'(a_5) \right.
\tag{51}
$$

$$
\left. \times w_{(30+j)} \right)
$$

$$
\begin{aligned}
&= Q(x_t, u_t) \times (1 - Q(x_t, u_t)) \times \\
&\quad \sum_{j=1}^{10} \left(w_{(40+j)} \times w_{(30+j)} \times \right.
\end{aligned}
\tag{52}
$$

$$
\left. y_j \times [1 - y_j] \right)
$$

It must be noted that the control output of FRNF is an input to the NN too.

6. Full Recurrent Adpative NeuroFuzzy Architectures

A variety of FRNF identifiers are used in order to identify the nonlinear $\widehat{f}(x)$ and $\widehat{g}(x)$ functions for different sub-systems in SMG-HPS. The seven-layered FRNF system uses NeuroFuzzy concept for estimation.

Fuzzy logic uses IF-THEN rules for approximation of unknown functions using standard fuzzy model. The unknown functions, $\widehat{f}(x)$ and $\widehat{g}(x)$, can be identified by the standard fuzzy model using a set of rules:

R^m: IF x_1 is $A_1^j \ldots$ and x_n is A_n^j THEN y is β_l^j

Let fuzzy logic controller have q inputs, $\rho_1, \rho_2, \ldots, \rho_q$. The output of NeuroFuzzy system is given as:

$$Y = \frac{\sum_{l=1}^{m} \prod_{j=1}^{q} \mu_{F_j^l}(\rho_j) \beta_l}{\sum_{l=1}^{m} \prod_{j=1}^{q} \mu_{F_j^l}(\rho_j)} \tag{53}$$

where $\mu_{F_j^l}$ is the membership function, ρ_j and β_l are adjustable parameters. It is the point in R (set of rules) at which μ_{β_j} achieves its maximum value. m is the number of fuzzy rules used to construct the identifier, F_j^l is the jth fuzzy set corresponding to the lth fuzzy rule, and β_l is centroid of the lth fuzzy set corresponding to identifier output, $\widehat{f}(x)$ and $\widehat{g}(x)$. Equation (53) can be written for $\widehat{f}(x)$ and $\widehat{g}(x)$ using fuzzy-basis function vector $\xi(x)$, as:

$$\widehat{\mathbf{f}}(\mathbf{x}) = \xi_{\mathbf{f}}^{\mathbf{T}} \xi(\mathbf{x}) \tag{54}$$

and

$$\widehat{\mathbf{g}}(\mathbf{x}) = \xi_{\mathbf{g}}^{\mathbf{T}} \xi(\mathbf{x}) \tag{55}$$

where

$$\beta_f = [\beta_{f1} \ \beta_{f2} \ \cdots \ \beta_{fm}]^T \tag{56}$$

and

$$\beta_g = [\beta_{g1} \ \beta_{g2} \ \cdots \ \beta_{gm}]^T \tag{57}$$

and $\xi(x)$ is given as

$$\xi = [\xi_1 \ \xi_2 \ \cdots \ \xi_m]^T = \left[\frac{\prod_{j=1}^{q} \mu_{F_j^1}(\rho_j)}{\sum_{l=1}^{m} \prod_{j=1}^{q} \mu_{F_j^l}(\rho_j)} \right.$$
$$\left. \cdots \ \frac{\prod_{j=1}^{q} \mu_{F_j^l}(\rho_j)}{\sum_{l=1}^{m} \prod_{j=1}^{q} \mu_{F_j^l}(\rho_j)} \right] \tag{58}$$

Several mathematical relations and functions are available for designing a fast and robust NeuroFuzzy identifier. The following variants were used to design antecedent and the consequent part of the fuzzy logic system for this research work.

6.1. Variants of Antecedent Part

The transformation of continuous input variables into linguistic variables is fuzzification. A membership function is always required for the transformation. The importance of the membership function is based on its shape, that translates complete information of the plant (uncertainties and nonlinearities) in a fuzzy inference system.

6.1.1. Gaussian Membership Function

The Gaussian membership function has the following properties:
- local and nonlinear nature
- smooth output

Gradient-based techniques are highly suitable for use, due to the continuous differentiable nature of the Gaussian membership function. It is expressed as:

$$\mu_j^r(x_i) = exp\left[-\left(\frac{x_i(k) + m_{ij}}{\sigma_{ij}} \right)^2 \right] \tag{59}$$

where m_{ij} and σ_{ij} are the mean and variance of the *i*th input and *j*th membership function.

6.1.2. B-Spline Membership Function

The B-Spline membership function (locally controllable membership function formed by polynomial pieces) is used as a variant of the antecedent part. The B-Spline membership function is defined as [44]:

$$\varrho_{ij}(x_i) = O_i^{(2)} = \sum_{k=0}^{n} Y_k \Delta_{k,p}(x_i) \quad 1 \le p \le n. \tag{60}$$

where, ϱ_{ij} represents the degree of B-Spline membership function, Y_k is the control point with $k = 0, 1, 2, 3, \ldots, n$ with $n + 1$ total control points. p is the order of B-Spline basis function. $\Delta_{k,p}(x_i) = \Delta(x_i \backslash \wp_1, \wp_2, \ldots, \wp_{n+p})$ is the *k*th B-Spline basis function and is given by the following Cox-de Boor recursion formula [44]:

$$\Delta_{k,p}(x_i) = \begin{cases} 1 \text{ if } p = 1 \quad x \in [\wp_i \ \wp_{i+1}[\\ \left(\frac{x - \wp_i}{\wp_{i+p-1} - \wp_i}\right) \Delta_{i,p-1} + \left(\frac{\wp_{i+p} - x}{\wp_{i+p} - \wp_{i+1}}\right) \\ \Delta_{i+1,p-1} \text{ if } p > 1, x \in [\wp_i \ \wp_{i+p}[\\ 0 \text{ if } p = 1 \quad x \notin [\wp_i \ \wp_{i+1}[\end{cases} \tag{61}$$

where, $\wp = [\wp_1, \wp_2, \ldots, \wp_{n+p}] \in \mathbb{R}$ is the knot vector such that $\wp_{i+1} - \wp i \ge 0$.

This research work employed a second-order B-Spline membership function with nine control points and thirteen-knot vectors.

6.2. Variants of Consequent Part

The consequent part generates weights based on different mathematical functions, like the Fourier series function, wavelet networks, and polynomial NN. The operation of the consequent part takes place in parallel to the antecedent part and produces the final output of the identifier at the defuzzification layer. The variants of the consequent part used in this research work are given below.

6.2.1. Fuzzy Wavelet Neural Networks (NNs)

For a better estimation of nonlinear functions, wavelet NNs were proposed as a substitute to feedforward NNs. Due to numerous neurons, NNs may get struck in the local minima, that results in slower convergence of the network. To get rid of this, wavelet functions can be used in the structure. Wavelets are waves having a limited duration and zero mean value. The localization characteristics of wavelets, and the fast learning abilities of NNs, result in better outcomes for complex nonlinear system modeling. The schematic diagram of wavelet NN is given in Figure 5.
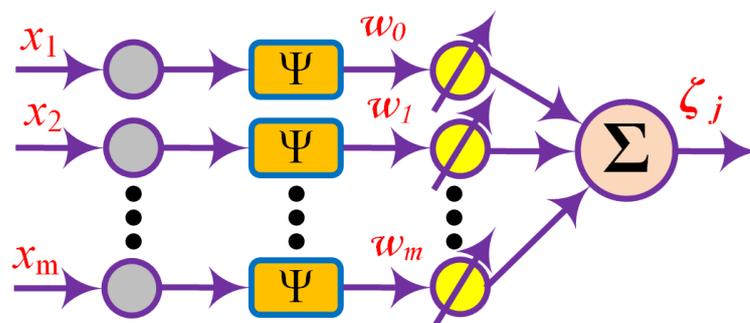

Figure 5. Wavelet Neural Network.

Following are the wavelet activation functions used in this research work.

- *Mexican hat wavelet* (MHW) is a negative normalized, non-orthogonal second derivative of Gaussian function. MHW function is expressed as;

$$\Psi_i(x_i) = |d_{ij}|^{-\frac{1}{2}} \psi(z_i) \tag{62}$$

where, $\Psi_i(x_i)$ is the family of wavelets obtained by single $\psi(x_i)$ function, $z_{ij} = \frac{x_i(k) + H_{ij}F_{ij} - t_{ij}}{d_{ij}}$, H_{ij} represents the output of the consequent part, F_{ij} is adaptive recurrent feedback weight, t_{ij} is the translation, and $d_{ij} \neq 0$, for $i = 1, 2, \ldots, n$ is dilation respectively.

- *Morlet wavelet* (Mor-W) is given as [44]:

$$\Psi_{ij}(z_{ij}) = exp\left[-0.5(z_{ij})^2 cos(5z_{ij})\right] \tag{63}$$

where;

$$z_{ij} = \frac{x_i(k) + H_{ij}F_{ij} - t_{ij}}{d_{ij}} \tag{64}$$

where H_{ij} represents the output of the consequent part, F_{ij} is adaptive recurrent feedback weight, t_{ij} is the translation, and d_{ij} is dilation of Mor-W.

- *Legendre wavelets* (Leg-W) are also known as spherical harmonic wavelets. They are based on Legendre polynomial, compactly supported, and orthonormal wavelets. They can be expressed as [45]:

$$\Psi_{pq}^k(x) = \begin{cases} 2^{p+1}\sqrt{2k+1} \times & \forall \frac{q}{2^p} \leq x \leq \frac{q+1}{2^p} \\ L_k(2^p x - q - 0.5), & \\ 0, & otherwise \end{cases} \tag{65}$$

where, $p = 1, 2, \ldots, m$ and $q = 0, 1, \ldots, 2^p - 1$ shows decomposition level and integer translation, respectively. L_k shows the Legendre polynomial as given in (66), with k being the degree of the polynomial [45].

$$L_p(x) = \frac{1}{2^p n!} \frac{d^n}{dx^n}\left(x^2 - 1\right)^p \tag{66}$$

where, $p \geq 0, x \in [-1 \ 1]$. The first four Legendre polynomials for $p = 0, \ldots, 3$ used in this work are given below [45]:

$$
\begin{aligned}
L_o(x) &= 1, \\
L_1(x) &= x, \\
L_2(x) &= 0.5\left(3x^2 - 1\right), \\
L_3(x) &= 0.5\left(5x^3 - 3x\right).
\end{aligned} \tag{67}
$$

Six Leg-W basis functions were used in this research work for $p = 1$ and $q = 0, \ldots, 2$ defined on $[0 \ 1]$.

7. Proposed Full Recurrent Adaptive NeuroFuzzy Identifier

The full recurrent adaptive NeuroFuzzy (FRNF) has seven layers as shown in Figure 6. The antecedent part consists of the first three layers, whereas the remaining four layers are consequent part layers. The n number of input signals in the first layer is equivalent to the m number of nodes and these nodes are used for input distribution.
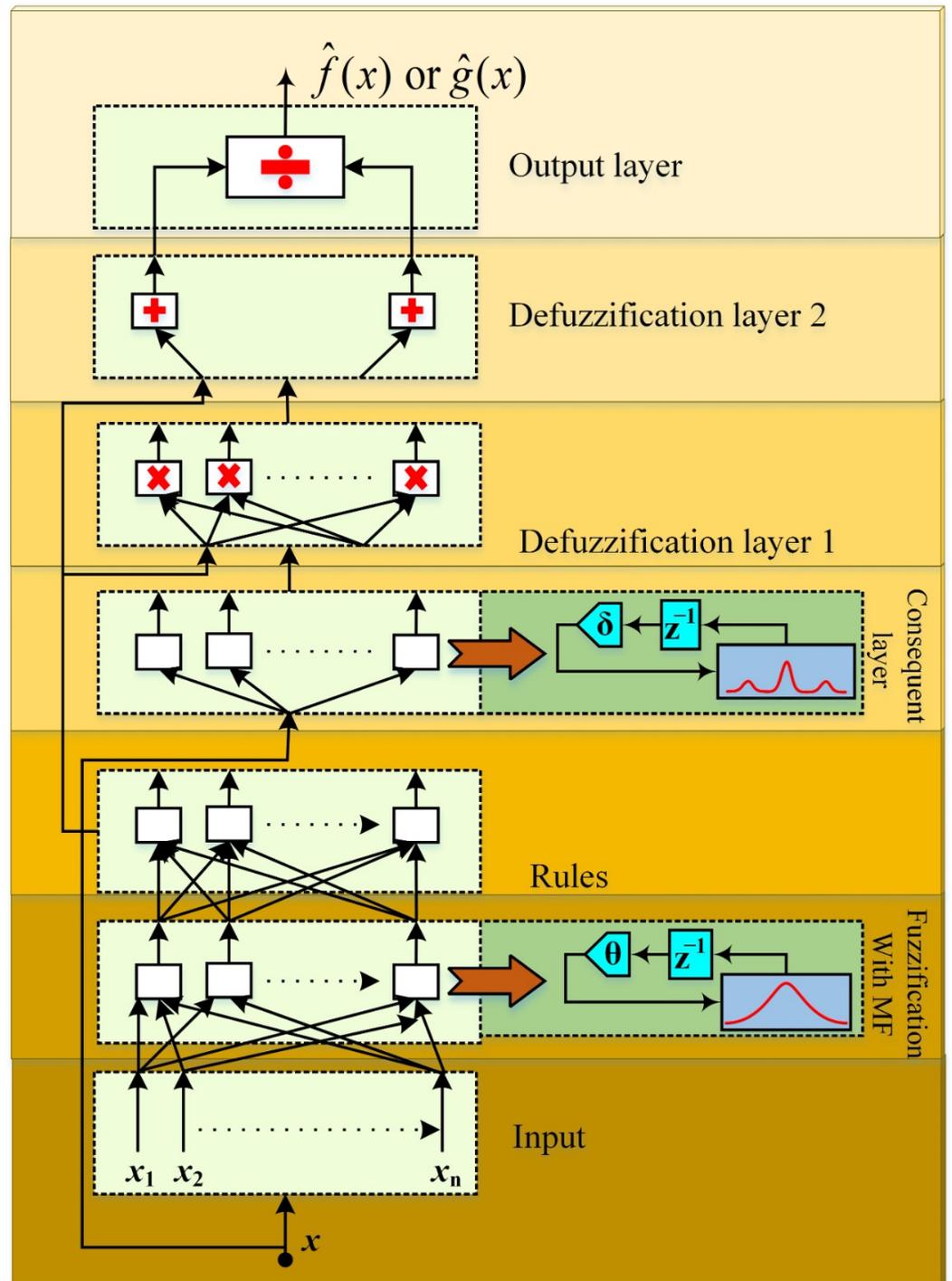
Figure 6. Architecture of Full Recurrent Adaptive NeuroFuzzy System.

Let I_i^k and O_i^k represents input and output of the ith node in the kth layer. The operation function of nodes and the signal propagation in each layer of FRNF is given below.

Layer 1

The input is received at this layer and is passed to the next layer. The temporal relationship in the network is shown by the feedback connection of this layer.

The output is: $O_i^{(1)} = x_i$.

Layer 2

One of the variants described in Section 6.1 is used as the membership function in this layer. This layer estimates the membership degree and fuzzy sets.

The output of this layer, in the case of the B-Spline membership function, is:

$$O_i^{(2)} = \varrho_{ij}(k) + v\varrho_{ij}(k-1) \tag{68}$$

where, ϱ represents one of the variants given in (60), v is the closed-loop adjustable feedback gain of the antecedent part in case of B-Spline membership function.

The output of this layer in the case of Gaussian membership function is:

$$O_i^{(2)} = exp\left[-\left(\frac{x_i(k) + O_i^{(2)}(k-1)\theta_{ij} + m_{ij}}{\sigma_{ij}}\right)^2\right] \tag{69}$$

where, O_i represents the output of the ith node, superscript (2) indicates layer number, ij subscript shows the jth term of the ith input, x_i is the input, and θ_{ij} is the recurrent weight, σ_{ij} and m_{ij} are the variance and mean of the ith input and jth membership function. It must be noted that the recurrent weight of the Gaussian membership function in antecedent part is adaptive.

Layer 3

In the rule layer, the product of the membership function is calculated. The number of rules in this layer determines the number of nodes.

The output, in the case of B-Spline membership function, is:

$$O_i^{(3)} = \prod_{i=1}^{n} O_i^{(2)} \tag{70}$$

For the Gaussian membership function, the output of this layer is:

$$O_i^{(3)} = \prod_{i=1}^{n} exp\left[-\left(\frac{x_i + O_i^{(2)}(k-1)\theta_{ij} - m_{ij}}{\sigma_{ij}}\right)^2\right] \tag{71}$$

Layer 4

This layer approximates the weighted firing strength and shows the THEN-part of fuzzy rules. The inputs to this layer are the error signal and own weighted feedback signal.

The output of this layer for Leg-W given in Section 6.2 is:

$$O_i^{(4)} = H_i(k) + \delta_i H_i(k-1) \tag{72}$$

where, for Leg-W using (65), $H_i = \sum\limits_{p=1}^{N} \sum\limits_{q=0}^{2^p-1} \sum\limits_{k=0}^{K} w_{nm}^k \Psi_{pq}^k$, w_{nm} represents Leg-W coefficients. The feedback weight is a closed-loop fixed gain for this layer.

In the case of MHW and Mor-W, the output of this layer is:

$$O_i^{(4)} = w_i \sum\limits_{i=1}^{n} \Psi_{ij}^k(Z_{ij}) \tag{73}$$

where, w_i is the weight and $\Psi_{ij}^k(Z_{ij})$ is given in Equations (62) and (63) for MHW and Mor-W, respectively.

It must be kept in mind that the feedback gain for MHW and Mor-W is adaptive, while the feedback gain for all other variants used is conventional feedback adjustable fixed gain.

Layer 5

In the first defuzzification layer, the sum of products of antecedent and consequent parts of (70) and (72) is calculated.

The output is:

$$O_i^{(5)} \quad = \quad \sum_{i=1}^{n} O_i^{(4)} O_i^{(3)} \tag{74}$$

Layer 6

In the second defuzzification layer, the sum of all the rules from (70) is calculated.

The output is:

$$O_i^{(6)} \quad = \quad \sum_{i=1}^{n} O_i^{(3)} \tag{75}$$

Layer 7

The required nonlinear functions are approximated in the output layer.

The output is:

$$\widehat{u_f} = O_f^{(7)} = \widehat{f}(x) = \frac{O_{fi}^{(5)}}{O_{fi}^{(6)}} \tag{76}$$

$$\widehat{u_g} = O_g^{(7)} = \widehat{g}(x) = \frac{O_{gi}^{(5)}}{O_{gi}^{(6)}} \tag{77}$$

7.1. Optimization Algorithm

The training of FRNF is for the adjustment of input-output pairs or a given function by fine-tuning network parameters. Mean square error is used as the cost function for the training purpose, given as [46–51];

$$E \quad = \quad \frac{1}{2}(\widehat{y} - y)^2, \tag{78}$$

where E is the identification error, $\widehat{y}$ is the approximated output of the subsystem, and y is the actual output of the subsystem, respectively. The gradient descent method is used for fast cost function reduction and convergence [46–51] . The general equation is given as follows:

$$\Omega(k+1) \quad = \quad \Omega(k) - \gamma g_k \tag{79}$$

where, g_k is the gradient of cost function at kth iteration, $\gamma > 0$ is the learning rate and k is the iteration index.

8. Full Recurrent Adaptive NeuroFuzzy Identifiers

The parameters of FRNF were tuned, based on the estimation of $Q(x, u)$, as discussed in Section 5.1. The optimization of FRNF output was achieved by updating the parameters of FRNF and maximizing the action-value function $Q(x, u)$ w.r.t. control output u. The parameters of FRNF were tuned using the gradient descent method, described in Section 7.1. The following three different combinations of FRNF algorithms were used in this research work for tuning of parameters.

8.1. FRNF-HBs-LegW Identifier

In this FRNF the identifier was the B-Spline membership function, described in Section 6.1.2, while the consequent part was the Legendre wavelet given in Section 6.2.1. The proposed identifier was a seven-layered scheme, discussed in Section 7. The schematic diagram of the identifier is shown in Figure 7. Both the antecedent and consequent parts were conventional recurrent.

The update equations of all the parameters of the antecedent part and the consequent part, according to (79), used the following chain rules:

$$\frac{\partial E}{\partial \zeta_{ij}} = \frac{\partial E}{\partial \xi_i} \frac{\partial \xi_i}{\partial \varsigma_i} \frac{\partial \varsigma_i}{\partial \zeta_{ij}} \tag{80}$$

where, ζ_{ij} shows the variants of B-Spline membership function defined in (60). The update equations for B-Spline membership function in antecedent part is:

$$\mu_{ij}(p+1) = \mu_{ij}(p) + \gamma \, e_q \, \frac{\xi_j \, (\beta_i - \widehat{u}u)}{\mu_{ij} \sum\limits_{j=1}^{n} \xi_j} \tag{81}$$

where, $\widehat{u}$ is the output of identifier given in Equation (76), $e_q = (e\frac{\partial y}{\partial \zeta})$, ξ_j is output of Equation (70), β is output of Equation (72), update parameters are rules and weights $[\mu]$.

The chain rule for consequent part is given below:

$$\frac{\partial E}{\partial \vartheta_{ij}^{(p)}} = \psi_{ij}^{(p)} \, \frac{\xi_j(x_i)}{\sum\limits_{j=1}^{n} \xi_j(x_i)} \tag{82}$$

where ψ_i represents Leg-W as given in (65) and ϑ shows the Leg-W variants. The update equations are:

$$\vartheta_{ij}^{p}(p+1) \;=\; \vartheta_{ij}^{p}(p) + \gamma \, e_q \, \frac{\xi_j}{\sum\limits_{j=1}^{n} \xi_j} \times \psi_{ij}^{(p)} \tag{83}$$

where, $e_q = (e\frac{\partial y}{\partial \vartheta})$, ξ_j is output of Equation (70), and update parameters are polynomials of Leg-P $[\vartheta_{ij}^{p}]$ given in Equation (67).

8.2. FRNF-MHW Identifier

In this FRNF, the identifier is given in Equation (76), the antecedent part in this control scheme is the Gaussian membership function, described in Section 6.1.1, while the consequent part is the MHW described in Section 6.2.1. The FRNF-MHW identifier has a seven-layered architecture, given in Section 7. The schematic diagram of identifier is shown in Figure 8. The recurrent part in this identifier is adaptive for both antecedent and consequent parts.
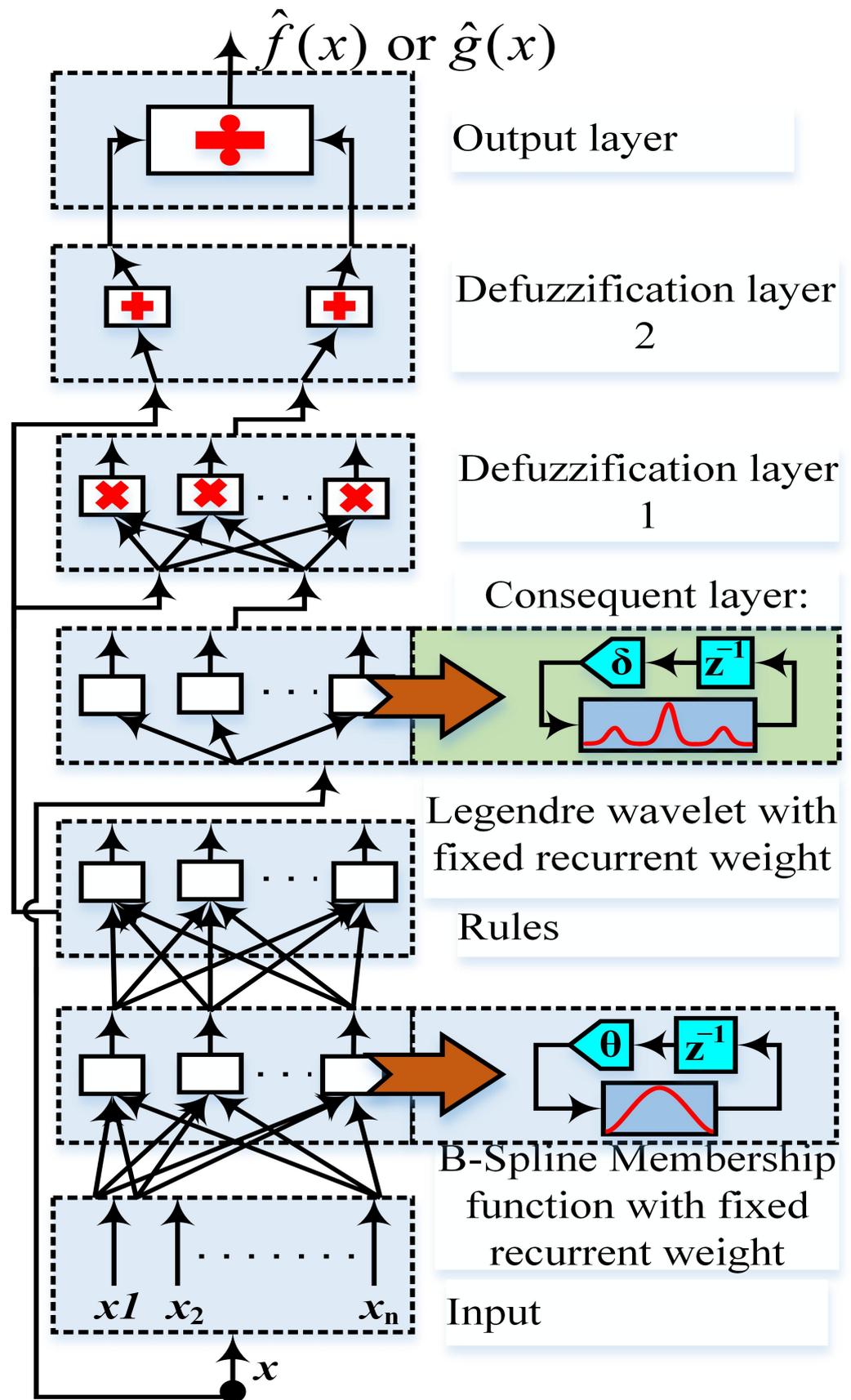
Figure 7. B-Spline membership function and Legendre wavelet.

The update equations of all the parameters of the antecedent part, according to (79), uses the following chain rules:

$$\frac{\partial E}{\partial \zeta_i} = \frac{\partial E}{\partial \xi_i} \frac{\partial \xi_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \zeta_i} \tag{84}$$

where, ζ_i shows the variant like mean, variance and feedback weight of Gaussian membership function.

$$m_i(k+1) = m_i(k) + \gamma e_q \left[\frac{(\beta_i - u)\xi_i}{\sum\limits_{j=1}^{n} \xi_i} \right] \times$$
$$\left[\frac{x_i(k) + \mu_i(k-1)\theta_i - m_i}{(\sigma_i)^2} \right] \tag{85}$$

$$\sigma_i(k+1) = \sigma_i(k) + \gamma e_q \left[\frac{(\beta_i - u)\xi_i}{\sum\limits_{j=1}^{n} \xi_i} \right] \times$$
$$\left[\frac{(x_i(k) + \mu_i(k-1)\theta_i - m_i)}{(\sigma_i)^3} \right] \tag{86}$$

$$\theta_i(k+1) = \theta_i(k) - \gamma e_q \left[\frac{(\beta_i - u)\xi_i}{\sum\limits_{j=1}^{n} \xi_i} \right] \times$$
$$\left[\frac{x_i(k) + \mu_i(k-1)\theta_i - m_i}{(\sigma_i)^2} \right] \tag{87}$$

where, $\widehat{u}$ is the output of identifier, $e_q = (e \frac{\partial y}{\partial \zeta})$, ξ_j is output of Equation (70), β is output of Equation (72), update parameters are mean, variance, and feedback weight $[m \ \sigma \ \theta]$.

The chain rule for the consequent part is given below:

$$\frac{\partial E}{\partial \vartheta_i} = \frac{\partial E}{\partial \xi_i} \frac{\partial \xi_i}{\partial \beta_i} \frac{\partial \beta_i}{\partial \Psi_i} \frac{\partial \Psi_i}{\partial z_i} \frac{\partial z_i}{\partial \vartheta_i} \tag{88}$$

where here Ψ_i represents MHW given in Equation (62), z_i defined below is an intermediate state variable and ϑ shows the variants of MHW like, translation, dilation, feedback weight etc.

$$z_i = \left(\frac{x_i(k) + H_i F_i - t_i}{d_i} \right) \tag{89}$$

$$t_i(k+1) = t_i(k) - \gamma e_q \left(\frac{\xi_i w_i}{\sum\limits_{j=1}^{n} \xi_i} \right) \times$$
$$\left[e^{-0.5(z_i)^2} \frac{\left[\frac{0.5}{z_i} - 3.5 z_i + (z_i)^3 \right]}{|d_i|^{3/2}} \right] \tag{90}$$

$$d_i(k+1) = d_i(k) - \gamma e_q \left(\frac{\xi_i w_i}{\sum\limits_{j=1}^{n} \xi_i} \right) \times$$

$$\left[e^{-0.5(z_i)^2} \frac{\left[0.5 - 3.5z_i^2 + (z_i)^4 \right]}{|d_i|^{3/2}} \right] \tag{91}$$

$$F_i(k+1) = F_i(k) + \gamma e_q \left(\frac{\xi_i w_i}{\sum\limits_{j=1}^{n} \xi_i} \right) \times$$

$$\left[e^{-0.5(z_i)^2} \frac{\left[\frac{0.5}{z_i} - 3.5z_i + (z_i)^3 \right]}{|d_i|^{3/2}} \right] H_i \tag{92}$$

where, $e_q = (e \frac{\partial y}{\partial \vartheta})$, ξ_j is output of Equation (70), and update parameters are translation, dilation, and feedback weight $[t \ \ d \ \ F]$.

The weight of the consequent part is updated according to the following chain rule:

$$\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial \xi_i} \frac{\partial \xi_i}{\partial \beta_i} \frac{\partial \beta_i}{\partial w_i} \tag{93}$$

where w_i represents the weight of the consequent layer.

$$w_i(k+1) = w_i(k) + \gamma e_q \left(\frac{\xi_i}{\sum\limits_{j=1}^{n} \xi_i} \right) \left[\psi_{i+1} + \psi_i \right] \tag{94}$$

where, ψ shows the MHW given in Equation (62).

8.3. FRNF-Mor-W Identifier

In this FRNF, the identifier is given in Equation (76), the antecedent part in this control scheme is the Gaussian membership function, described in Section 6.1.1, while the consequent part is the Morlet wavelet given in Section 6.2.1. The FRNF-Mor-W identifier has a seven-layered architecture, given in Section 7. The schematic diagram of the identifier is shown in Figure 9. The recurrent part in this identifier is adaptive for both antecedent and consequent parts.

The update equations of all the parameters of the antecedent part, according to (79), use the following chain rules:

$$\frac{\partial E}{\partial \zeta_i} = \frac{\partial E}{\partial \xi_i} \frac{\partial \xi_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \zeta_i} \tag{95}$$

where, ζ_i shows the variant like mean, variance and feedback weight of Gaussian membership function.

$$m_i(k+1) = m_i(k) + \gamma e_q \left[\frac{(\beta_i - \hat{u})\xi_i}{\sum\limits_{j=1}^{n} \xi_i} \right] \times$$

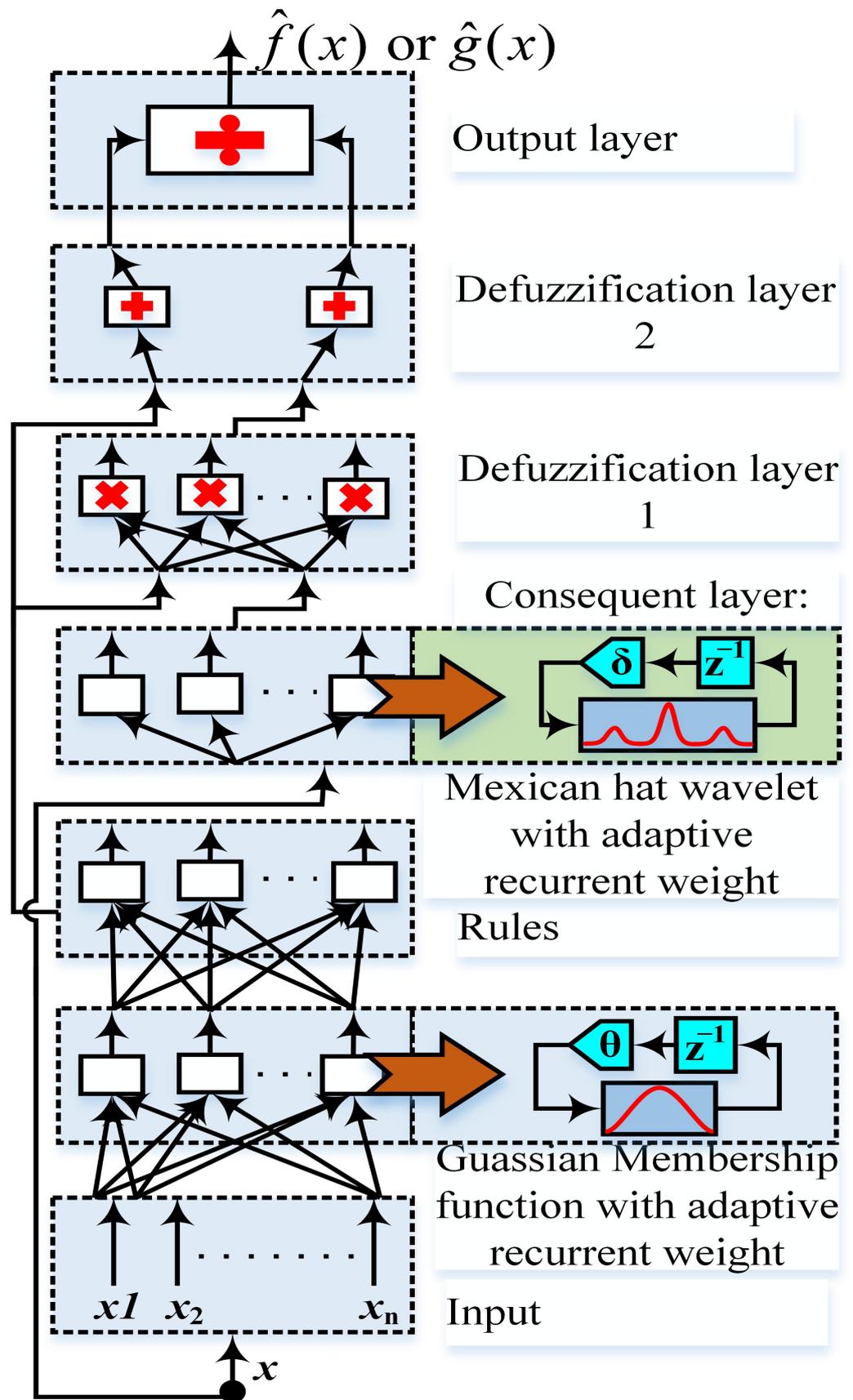$$\left[\frac{x_i(k) + \mu_i(k-1)\theta_i - m_i}{(\sigma_i)^2} \right] \tag{96}$$

Figure 8. Gaussian membership function with Mexican hat wavelet.

$$\sigma_i(k+1) = \sigma_i(k) + \gamma e_q \left[\frac{(\beta_i - u)\xi_i}{\sum\limits_{j=1}^{n} \xi_i} \right] \times$$

$$\left[\frac{(x_i(k) + \mu_i(k-1)\theta_i - m_i)}{(\sigma_i)^3} \right] \qquad (97)$$

$$\theta_i(k+1) = \theta_i(k) - \gamma e_q \left[\frac{(\beta_i - u)\xi_i}{\sum\limits_{j=1}^{n} \xi_i} \right] \times$$

$$\left[\frac{x_i(k) + \mu_i(k-1)\theta_i - m_i}{(\sigma_i)^2} \right] \qquad (98)$$

where, $\widehat{u}$ is the output of identifier given in Equation (76), $e_q = (e\frac{\partial y}{\partial \zeta})$, ξ_j is output of Equation (70), β is output of Equation (72), update parameters are mean, variance, and feedback weight $[m \ \sigma \ \theta]$.

The chain rule for the consequent part is given below:

$$\frac{\partial E}{\partial \vartheta_i} = \frac{\partial E}{\partial \xi_i} \frac{\partial \xi_i}{\partial \varsigma_i} \frac{\partial \varsigma_i}{\partial \Psi_i} \frac{\partial \Psi_i}{\partial \vartheta_i} \qquad (99)$$

where Ψ_i represents Mor-W as given in (63) and ϑ shows the variants of Mor-W-like translation, and dilation as given in (63). The update equations are:

$$t_{ij}(p+1) = t_{ij}(p) + \gamma e_q \mho \qquad (100)$$

$$d_{ij}(p+1) = d_{ij}(p) + \gamma e_q \mho z_{ij} \qquad (101)$$

$$w_{ij}(p+1) = w_{ij}(p) + \gamma \, e_q \, \frac{\xi_j}{\sum\limits_{j=1}^{n} \xi_j} \qquad (102)$$

where, $e_q = (e\frac{\partial y}{\partial \vartheta})$, ξ_j is output of Equation (70), $\mho = \frac{\xi_j w_j^{(5)}}{\sum\limits_{j=1}^{n} \xi_j} \times \frac{\left(cos(5z_{ij})e^{-0.5z_{ij}^2}z_{ij} + 5sin(5z_{ij})e^{-0.5z_{ij}^2} \right)}{d_{ij}^{(5)}}$,

and update parameters are rules, translation, dilation, and weight $[\mu \ t \ d \ w]$.
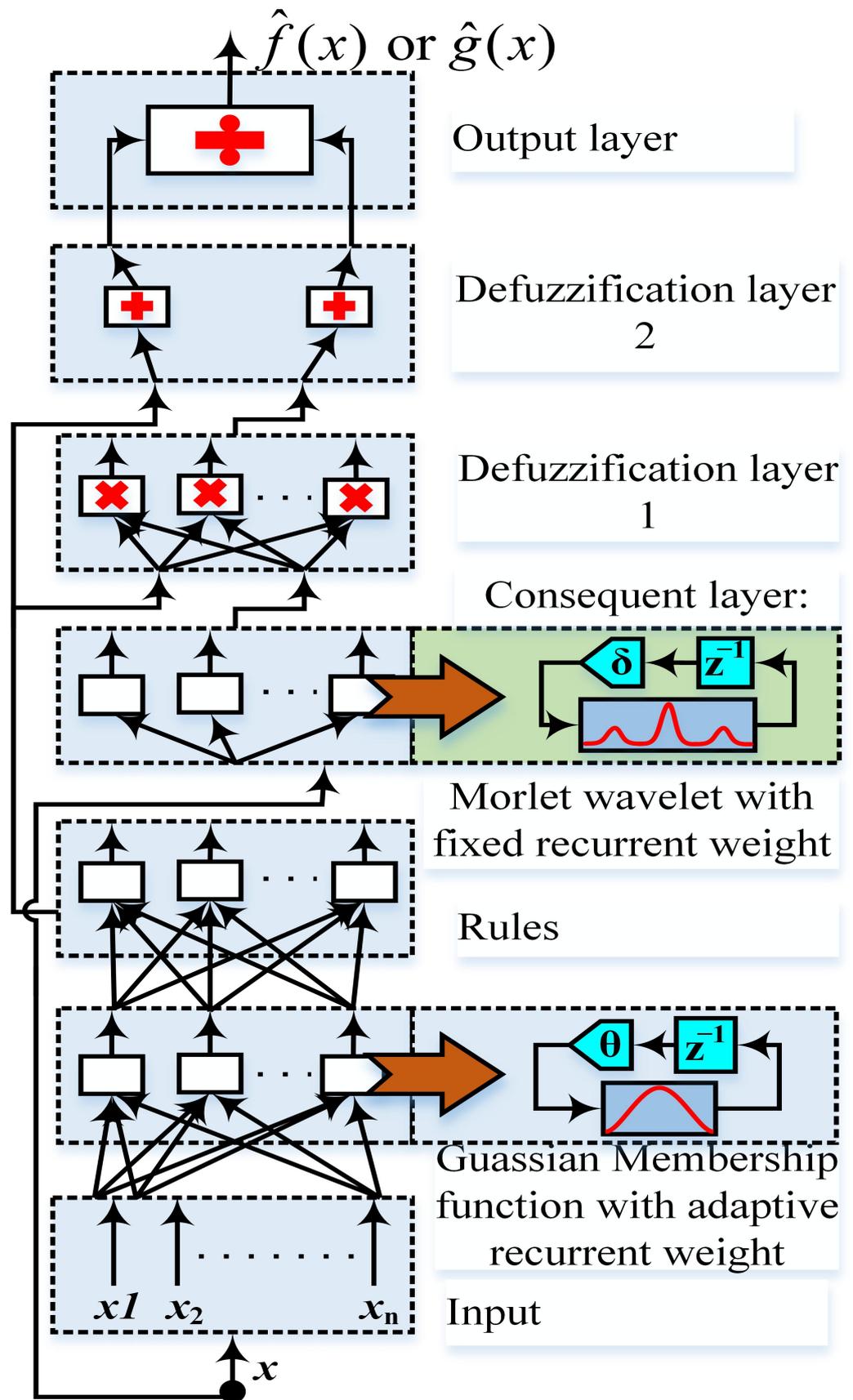
Figure 9. Gaussian membership function with Morlet wavelet.

9. Exploration Policy and Action Modifier

If all the actions from all the states are tried, then $Q(x, u)$ converges to $Q^*(x, u)$, with a probability of 1 [52]. In this research work, an exploration policy was implemented that guaranteed the trail of all possible actions for the control output $\hat{u}$ given by the FRNF. The AEM generated a control output u_c. The u_c was the sum of a disturbed action u_d and $\hat{u}$. The standard deviation of u_d was $\sigma_{Q(t)}$, which was recommended by FRNF and its normal distribution was zero. The problem of exploration was solved by this AEM in reinforcement learning [42].

$$u_c = \hat{u} + u_d \qquad u_d \sim N(0, \sigma_{Q(t)}) \qquad (103)$$

where,

$$\sigma_{Q(t)} = \frac{k}{[1 + 2exp(maxQ(x, a))]} \qquad (104)$$

where, k is a variable that shrinks or expands u_d.

10. Proposed Hybrid Adaptive Neural Network Q-Learning-Based Full Recurrent Adaptive NeuroFuzzy Control Paradigms

Using different combinations of the FRNF identifier algorithms, discussed in Section 8, the following three hybrid adaptive Q-learning control paradigms were proposed:

- Hybrid adaptive full recurrent Legendre wavelet-based Neural Network Q-learning control
  In this control scheme, the NNQLNF was embedded with an FRNF-Leg wavelet identifier, discussed in Sections 5.1 and 8.1.
- Hybrid adaptive full recurrent Mexican hat wavelet-based Neural Network Q-learning control
  In this control scheme, the NNQLNF was embedded with the FRNF-MHW identifier, discussed in Sections 5.1 and 8.2.
- Hybrid adaptive full recurrent molet wavelet-based Neural Network Q-learning-based control
  In this control scheme, the NNQLNF was embedded with the FRNF-Mor-W identifier, discussed in Sections 5.1 and 8.3.

11. Implementation Procedure of Hybrid Adaptive NNQLNF Control Paradigms

The following steps took place for the implementation of the proposed hybrid adaptive NNQLNF control paradigms [42].

- In the first step, the initialization of $Q(x_t, u_t)$, the parameters of FRNF and the weights $w_1 - w_{40}$, $w_{41} - w_{50}$ of QEN took place.
- Control output u_t was obtained from the FRNF identifier.
- u_t was then processed by AEM, according to Equation (103).
- u_c was the actual control output fed to the system.
- The estimated $Q(x_{t+1}, u_{t+1})$ was obtained from QEN, depending on control action, previous and current states.
- The performance of controller r, $Q(x_t, u_t)$ and $Q(x_{t+1}, u_{t+1})$, given in Equation (38), were used to estimate the TD error from Equation (38).
- QEN was updated, based on Equations (43) and (46).
- Parameters of FRNF were updated.
- $Q(x_t, u_t)$ was updated to $Q(x_{t+1}, u_{t+1})$.
- If the parameters of QEN and FRNF were not updated for a specific interval of time, then the learning procedure terminated; otherwise, the algorithm was repeated from step 2 again.

12. PQ Control of Interlinking Inverter Using Hybrid Adaptive NN Based Q-Learning Full Recurrent Adaptive NeuroFuzzy Control Paradigms

This section focuses on the PQ control of the interlinking inverter of SMG-HPS using the adaptive hybrid NNQLNF control paradigms, discussed in Section 5. The main aim of this research was to minimize the THD, frequency, and voltage fluctuations, which arise due to the use of multiple renewable energy sources and non-renewable energy sources along with converters and nonlinear loads. The objective was achieved by PQ control using the three proposed adaptive hybrid NNQLNF-based control schemes on interlinking inverters. The results were compared with an aPID control scheme.

13. Formulation of Control Problem

The active and reactive powers were obtained from the AC-bus. The difference between actual and reference power generated the error signal, *e*, which was used for tuning the parameters of the proposed control schemes.

$$e_p = P_{ref} - P_{AC} \tag{105}$$

$$e_q = Q_{ref} - Q_{AC} \tag{106}$$

where, e_p is the real power error, P_{ref} is the real reference power, P_{AC} is the real AC bus power, e_q is the reactive power error, Q_{ref} is the reactive reference power and Q_{AC} is the reactive AC-bus power.

The error signals were input to the proposed control schemes. The control output obtained from Equation (103) of both real and reactive power was transformed from *dqo* to *abc* and was fed to the PWM generator. The control objective was achieved by the convergence of the following signals:

$$\lim_{t \to +\infty} \begin{cases} P_{AC}(t) \to P_{ref}(t) \\ Q_{AC}(t) \to Q_{ref}(t) \end{cases} \tag{107}$$

The proposed control schemes achieved the objective and, thus, produced minimum THD and negligible fluctuations in frequency and voltage profiles.

14. Results and Discussion

The SMG-HPS testbed was prepared and simulated in Matlab/Simulink R2015a. The technical details of of SMG-HPS are discussed in Section 2. The real-time environmental data (wind speed, ambient temperature, and solar irradiance) were obtained from the Pakistan Meteorological Department (PMD) for a complete solar day at the Islamabad station. The case study taken was the Defense Housing Authority (DHA), Islamabad, Pakistan. The irradiation varied with the appearance of the sun. The average irradiation level during day time was about 1000 W/m^2, while the average temperature level was 20 °C, having a maximum peak of about 42.6 °C during the day time. Figure 10 shows the irradiance profile on the left *y*-axis and the temperature profile on the right *y*-axis, while Figure 11 shows the wind speed profile used for this case study.

In this study, three intelligent hybrid adaptive NNQLNF control schemes were implemented on an interlinking converter in the proposed SMG-HPS. The performances of all the control schemes were compared with an aPID control scheme. Based on the simulation results, the best performing controller was identified.
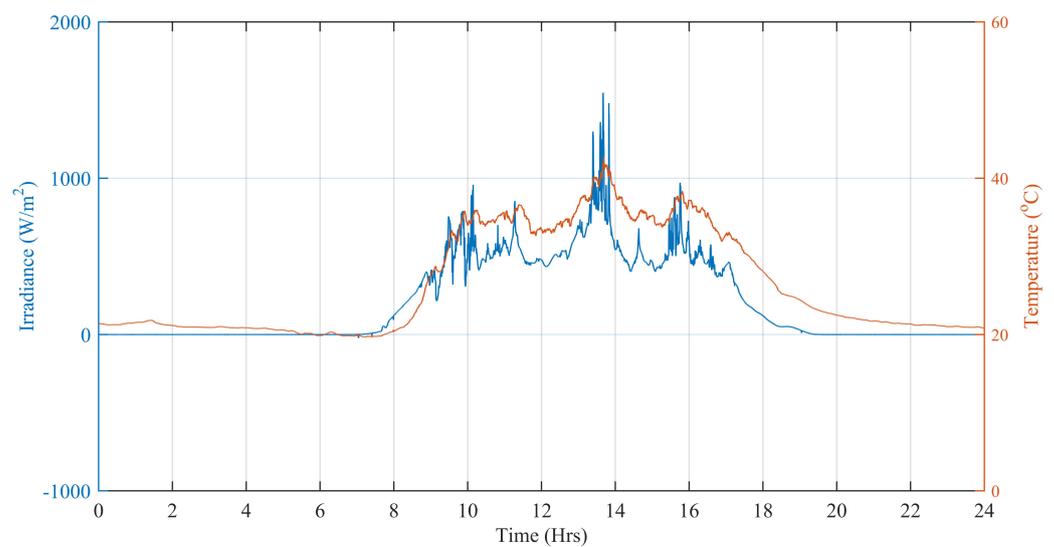
Figure 10. Solar Irradiance Level (W/m^2) (left), Ambient Temperature Level ($^\circ$C) (right).
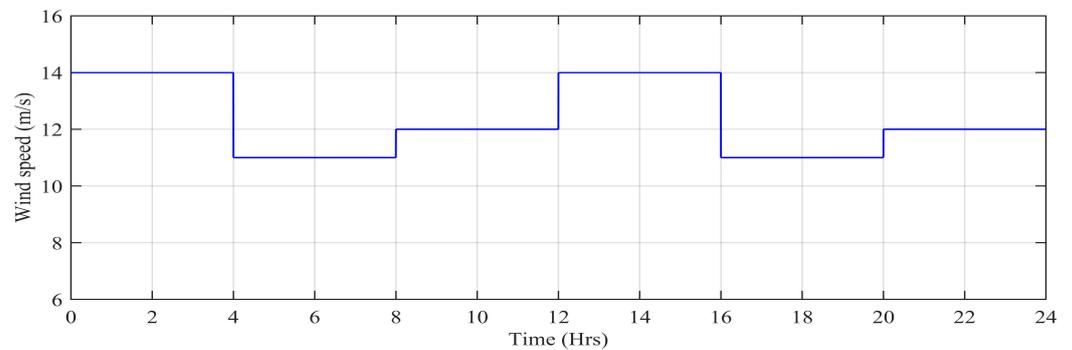
Figure 11. Wind speed (m/s).

For stability of the power system, the net power on the AC-bus should be equal to zero. Figure 12 shows the net active power profiles of the AC-bus for three NNQLNF-based controllers and an aPID control scheme. It was clear from the results that the net power, due to the aPID was of higher magnitude, as compared to the proposed control schemes. The negligible variations in the net real power on the AC-bus was due to the use of power electronic devices, converters, and nonlinear loads. However, the least magnitude was observed for the hybrid adaptive FRNF-MHW-based NNQLNF control scheme.

Figure 13 shows the net reactive power profile of the AC-bus for the proposed control schemes, compared with the aPID. The plots show the variation due to the aPID was of greater magnitude, as compared to the proposed control schemes. The variation in net reactive power of the AC-bus due to the proposed control schemes was of negligible magnitude. Little variation in the results was due to the use of multiple RESs that produced varying outputs, depending on weather conditions. Other factors of the variations in net active and reactive power profiles were due to the use of power electronic converters, CS, and nonlinear loads.
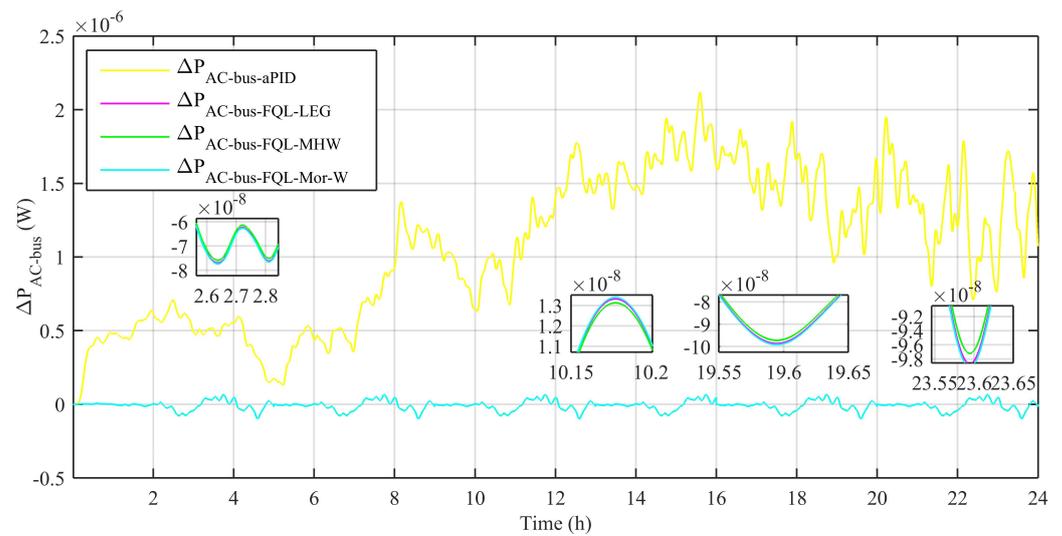
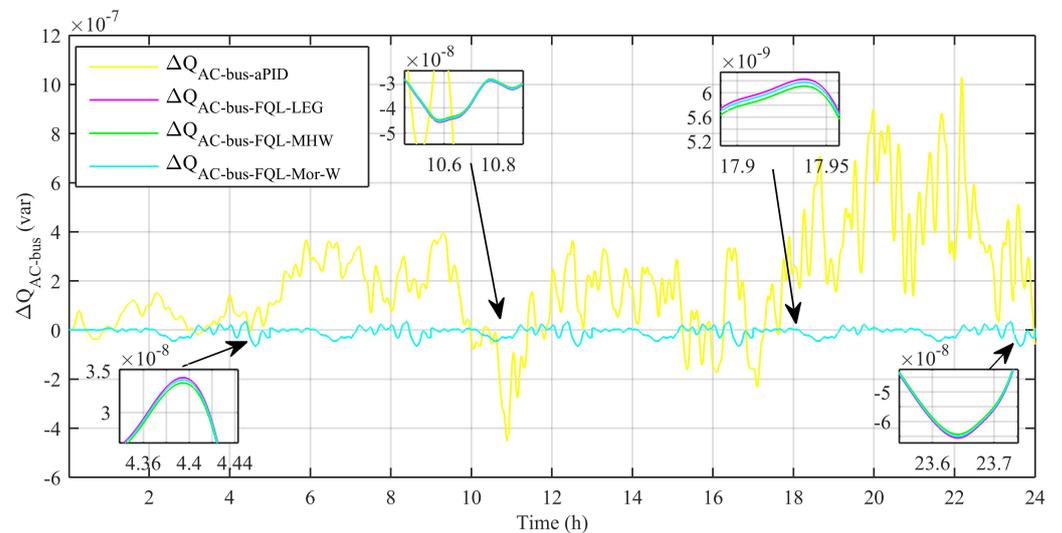**Figure 12.** ΔP_{AC} evolution.

Figure 13. ΔQ_{AC} evolution.

For the stability of the power system and life of various loads, the frequency of the AC-bus must remain close to 50 Hz. Figure 14 shows the frequency of the AC-bus maintained by the interlinking converter during various modes of operation in SMG-HPS. The results revealed that the frequency of the AC-bus for an aPID control scheme had greater variations, which could lead to instability of the power system. The larger magnitude of the frequency at the AC-bus reduced the life of power electronic devices and other sensitive loads. The frequency variations were responsible for imperfect charging/discharging of power storage devices, and, thus reduced their efficiency as well. It must be noted that the frequency maintained by using the proposed intelligent NNQLNF-based control schemes for interlinking converters was nearly 50 Hz all the time. However, negligible variations found in the simulation results were because of continuous switching of power between AC sub-grid and DC sub-grid. Other factors included the use of weather-dependent RES, power electronic converters, and nonlinear loads.
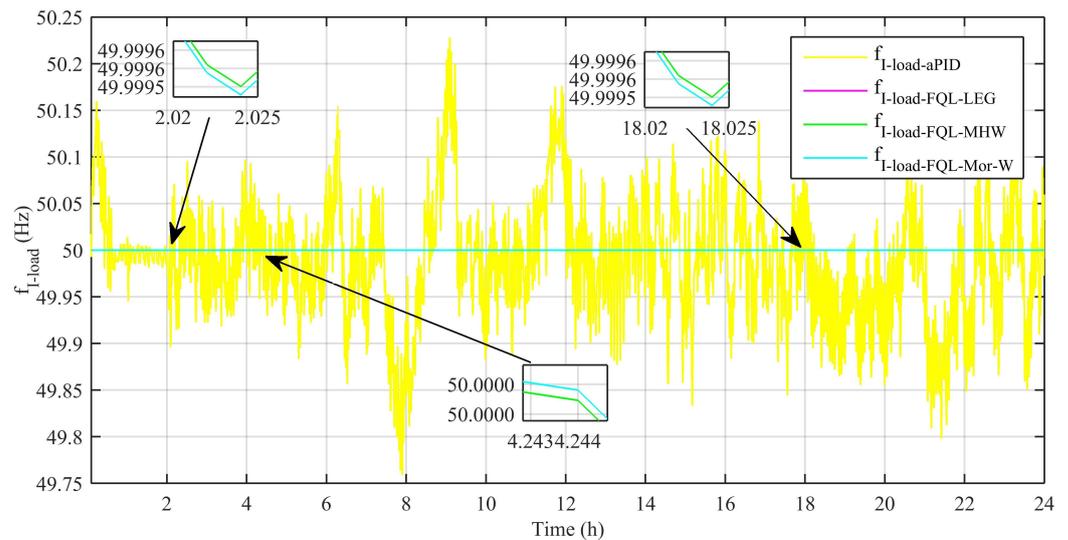
Figure 14. AC bus frequency evolution.

Table 2 shows the average values of $\Delta P_{AC\text{-bus}}$ and $\Delta Q_{AC\text{-bus}}$ variations.

Table 2. Average values of ΔP_{AC-bus} and ΔQ_{AC-bus} variations for all controllers.

Control Scheme	$\mathbf{P}_{AC-bus}$ (μWatts)	$\mathbf{Q}_{AC-bus}$ (μVARs)
aPID	1.0700	0.191100
FRNF-Leg wavelet based NNQLNF control	0.008600	0.008863
FRNF-Mor-W based NNQLNF control	0.008642	0.008794
FRNF-MHW based NNQLNF control	0.008472	0.008698

Figure 15 shows the %age change in load current frequency by various control schemes. The results revealed the performance of the proposed control schemes as being superior over the aPID control scheme. The negligible variation in the results for the proposed control schemes was due to sudden load changes, inverters, regulators, and charging/discharging operation of multiple devices. The variations were also because of power transfer to the utility grid during peak hours and bidirectional power transfer between the charging station and micro-grid. However, the results obtained from the proposed hybrid adaptive FRNF-MHW-based NNQLNF control scheme were more satisfactory and reliable, compared to the other control schemes.

Figure 16 shows the RMS voltage profiles for the various control schemes. The results showed that the RMS voltage for the aPID control scheme was not stable for any specific time interval and was continuously varying. This reduced the life of power electronic devices and other machines connected to the SMG-HPS. However, the RMS voltage for the proposed control schemes showed stable magnitude, compared with an aPID control scheme. The variations were because of switching of power between the AC sub-grid and the DC sub-grid, along with the varying output power of RES due to weather conditions.

Figure 17 shows the comparison of the percentage change in THD for the load current. The %ΔTHD in load current due to the proposed control schemes was of lower magnitude, as compared to an aPID control scheme. The %ΔTHD complied with IEEE standard 1547 [53]. However, the least magnitude was observed for the proposed adaptive FRNF-MHW-based NNQLNF control scheme, which proved its performance over other control schemes.
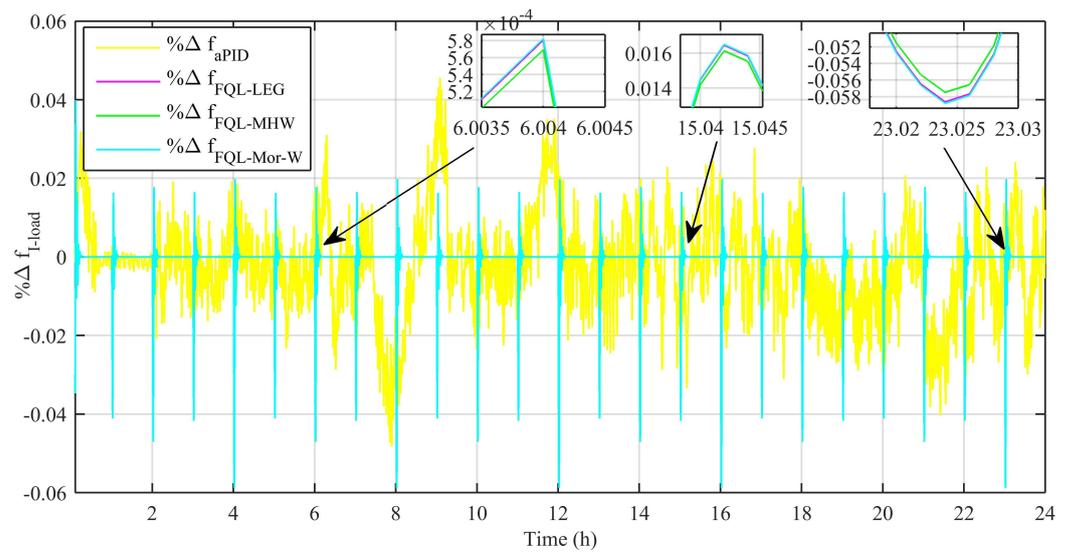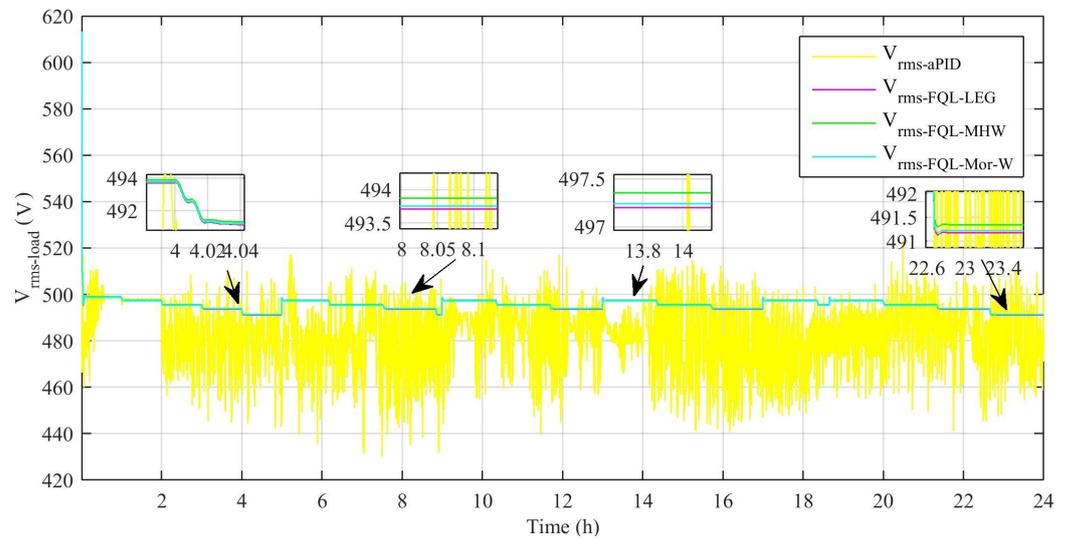
Figure 15. %Δf in load current evolution.

Figure 16. V*rms* load evolution.

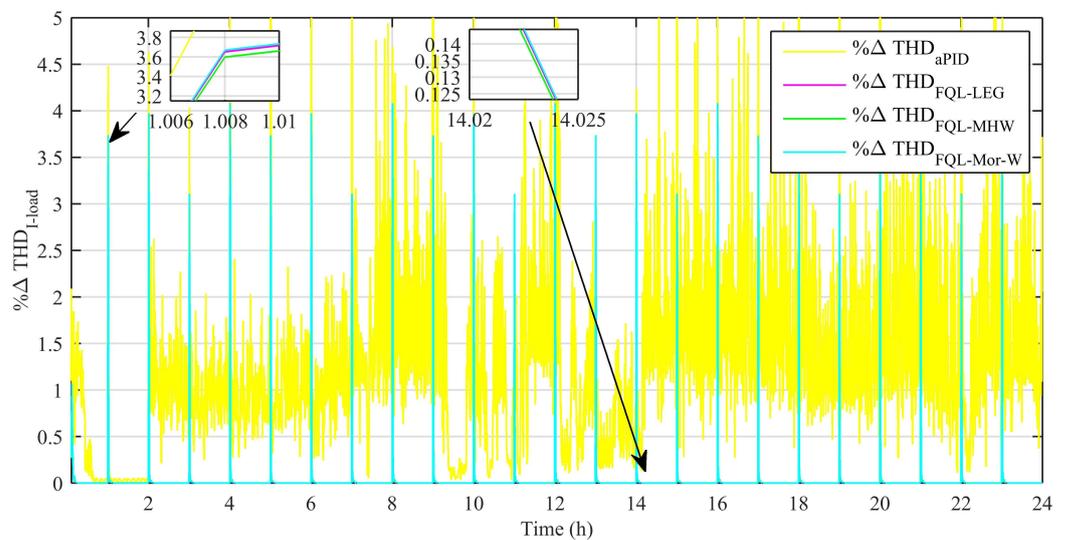**Figure 17.** %ΔTHD in load current evolution.

Table 3 shows the average values of $\%\Delta f_{Load}$ and $\%\Delta THD_{Load}$ in load current.

Table 3. Average values of $\%\Delta f_{Load}$ and $\%\Delta THD_{Load}$ variation in load current for all controllers.

Control Scheme	$\%\Delta f_{Load}$	$\%\Delta THD_{Load}$
aPID	0.0025480	1.29900
FRNF-Leg wavelet based NNQLNF control	0.0006112	0.06519
FRNF-Mor-W based NNQLNF control	0.0006130	0.06551
FRNF-MHW based NNQLNF control	0.0005992	0.06423

Figure 18 shows the maximum output power obtained for the best performing controller given in the previously published article [28], with the same SMG-HPS. The interlinking converter in the research work referred to above was controlled by an aPID controller. However, in this research work, the maximum power output of PV array was obtained with different proposed controllers for the interlinking inverter. It is clear from the figure that the overall maximum power was obtained for the proposed controllers of the interlinking inverter. However, the maximum power was seen for he proposed adaptive FRNF-MHW-based NNQLNF control scheme, which proved its performance over other control schemes.
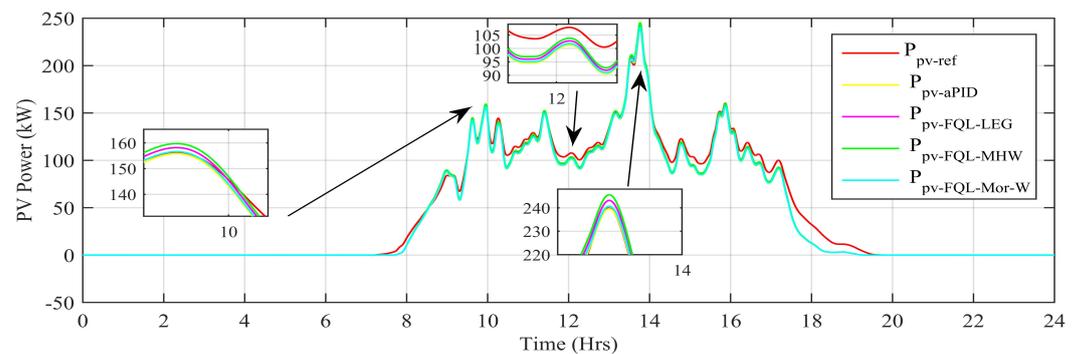

Figure 18. Maximum power point tracking of PV array evolution.

Figure 19 shows the maximum output power obtained for the best performing controller given in a previously published article [29], with the same SMG-HPS. The interlinking converter in the research work referred to above was controlled by an aPID controller. However, in this research work the maximum power output of the wind turbine was obtained with different proposed controllers for the interlinking inverter. It is clear from the figure that the overall maximum power was obtained for the proposed controllers of the interlinking inverter. However, the maximum power was observed for the proposed adaptive FRNF-MHW-based NNQLNF control scheme, which proved its performance over other control schemes.
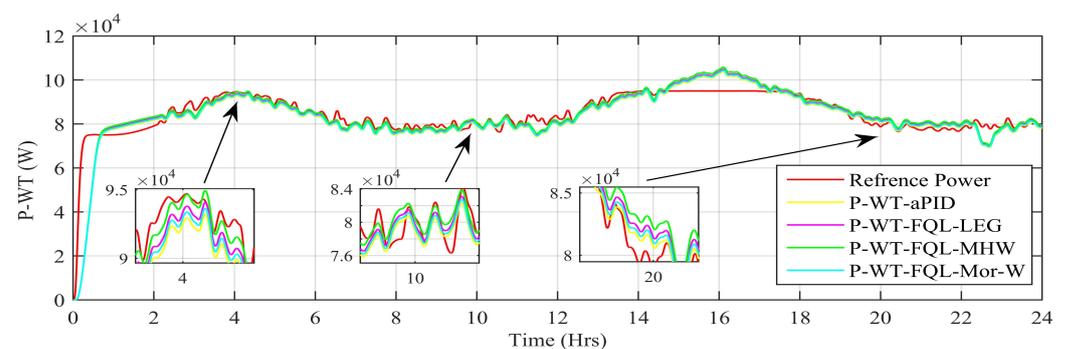
Figure 19. Maximum power point tracking of wind turbine evolution.

Figure 20 shows a spider chart for comparable parameters of interlinking converter controllers. The values were scaled (ΔP_{AC} by 100, ΔQ_{AC} by 20, %Δf by 10, and %ΔTHD by 15) to have a clear picture. The observations and analysis revealed that the hybrid adaptive FRNF-MHW-based NNQLNF control scheme provided superior performance over all other control schemes.
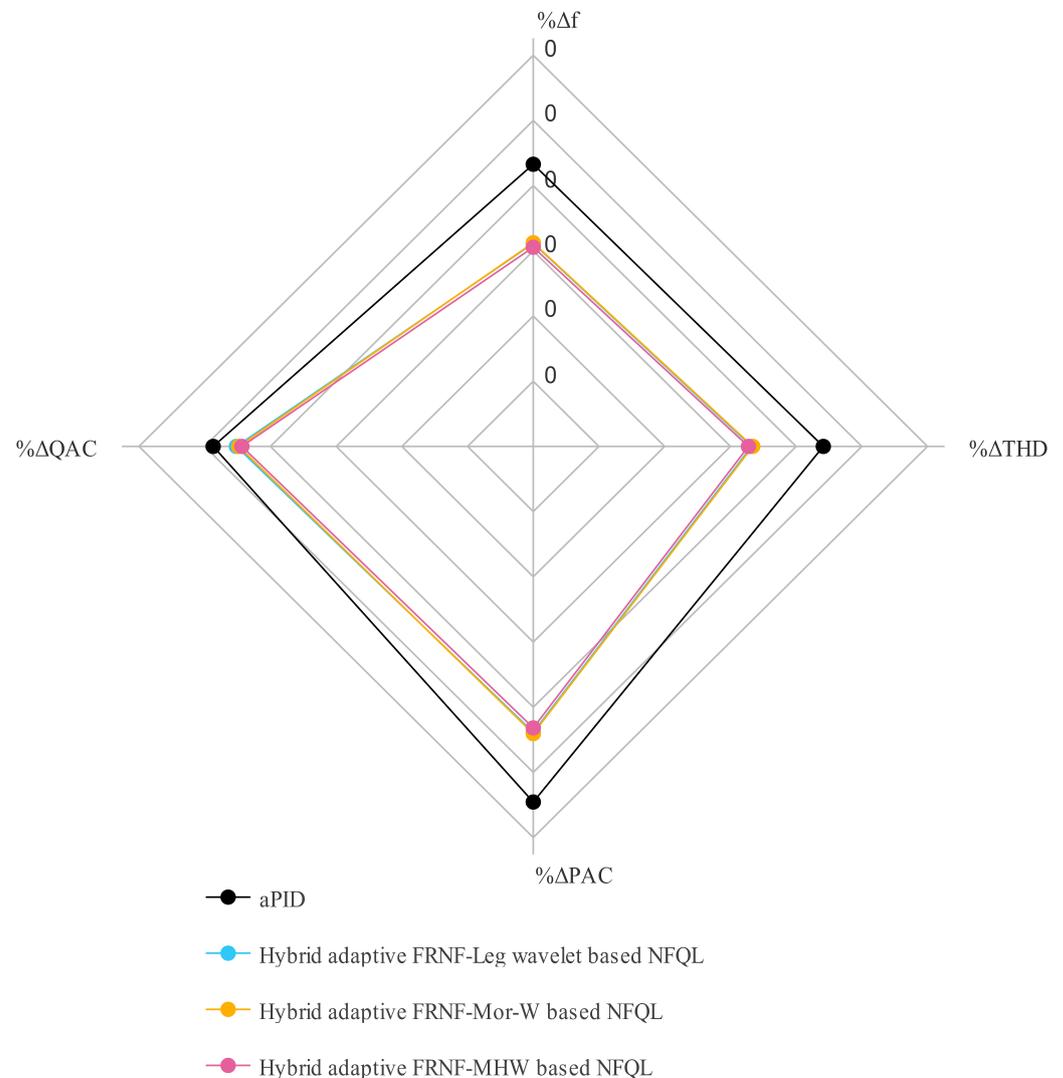
Figure 20. Spider chart of comparable parameters for interlinking converter controllers.

15. Conclusions

In this article, mathematical modeling of the proposed hybrid adaptive FRNF, based on an NN Q-learning control scheme, was discussed in detail. The proposed control schemes used BP NN, FRNF, and AEM for estimation of theoptimal action-value function, parameters update and action exploration modifier. An intelligent supervisory control system was also described in detail with nine modes of operation. The supervisory control was responsible for optimal power flow and ensured the power balance between generation and load.

This research work also focused on three adaptive FRNF-based NN Q-Learning control schemes for the control of a bidirectional interlinking converter between bybrid AC and DC sub-grids in the proposed SMG-HPS. The results were compared with an aPID control scheme. The use of intelligent controllers ensured the following: (i) power system stability, (ii) power quality, and (iii) reliability. These benefits were due to reducing the following: (a) %age ΔTHD and (b) %age Δf in the load current for various modes of the power system

operation. The proposed control schemes improved stability by reducing net real and reactive powers on the AC-buses. However, the overall best performance was observed for the proposed adaptive FRNF-MHW-based NN Q-Learning control scheme.

16. Future Work

Intelligent adaptive supervisory control based on NN/Fuzzy logic, advanced control schemes to schedule BSS/PHEV charging/discharging for profit gain, inclusion of priority-based sensitive loads, like data centers/hospitals, are some interesting and recommended future work.

Abbreviations

The following abbreviations were used in this manuscript:

AEM	Action exploration modifier
aPID	Adaptive PID
BP	Back propagation
BIC	Bidirectional interlinking converter
BSS	Battery storage system
CS	Charging station
DG	Distributed generator
ESS	Energy storage system
FPT	Fuzzy parameter tuning
FQL	Fuzzy Q-learning
FRNF	Full recurrent adaptive Neurofuzzy
Leg-W	Legendre wavelet
MHW	Mexican hat wavelet
Mor-W	Morlet wavelet
MT	Micro-turbine
NFPT	Neurofuzzy parameter tuning
NN	Neural network
NNQLNF	Neural network Q-learning based full recurrent adaptive Neurofuzzy
PHEV	Plug-in-hybrid-vehicle
PQ	P shows real power and Q shows reactive power
PV	Photovoltaic
QEN	estimator network
RES	Renewable energy resource
SC	Supercapacitor
SMG-HPS	Smart microgrid hybrid power system
SOFC	Solid oxide fuel cell
SOC	State of charge
SPF	Single point failure
TD	Temporal difference
THD	Total harmonic distortion
WT	Wind Turbine

Appendix A. Entire System Parameters

Appendix A.1. Parameters of VS-WECS

Table A1 below gives the parameters of VS-WECS (wind turbine) used for this research work:

Table A1. Parameters of VS-WECS.

Type	nED-100
Base wind speed	10 m/s
Rotor speed	3367 rpm
Drive train	2-mass model
Pitch angle	0°
Rated power	100 kW

Appendix A.2. Parameters of SOFC

Table A2 below gives the parameters of SOFC used for this research work.

Table A2. Parameters of SOFC.

Type	Bloom Energy USA ES-5700
Number of cells in seires	768
SOFC stack	4 kW
SOFC array	$5 \times 10 = 50$
SOFC array power rating	$50 \times 4 \text{ kW} = 200 \text{ kW}$

Appendix A.3. Parameters of PV

Table A3 below gives the parameters of PV subsystem used for this research work.

Table A3. Parameters of PV-farm.

Type	SunPower SPR-305-WHT
Module unit	305 W @ 1 kW/m^2, 25 °C
Number of series string/module	13
Number of parallel string/module	66
Power rating	$305 \times 13 \times 66 \approx 262 \text{ kW}$

Appendix A.4. Parameters of Charging Station

Table A4 below gives the parameters of PHEVs and BSS installed at CS used in this research work.

Table A4. Ratings of PHEV's batteries.

Vehicle Company	Battery Type	Battery Capacity (kWh)	Rated Voltage (V)
Nissan	Li-ion	24.0	360
Renault	Li-ion	22.0	300
Mitsubishi	Li-ion	16.0	20
Toyota	Li-ion	6.7	300
Honda	Li-ion	4.4	201

Appendix A.5. Modeling and Parameters of Battery

BSS is an integral part of an SMG-HPS and provides high energy density. It stores electrical power from SMG-HPS during off-peak hours, and, thus, helps in the utilization of RESs at their maximum. The stored electrical energy is returned to the SMG-HPS during

peak hours; thus, making it more flexible and reliable. The important parameters of a BSS are the state of charge (SOC) and the battery voltage. The battery voltage is given as under [54–57].

$$V_{bat} = V_{oc} - r_i \cdot I_{bat} \tag{A1}$$

where V_{bat}, V_{oc}, r_i, and I_{bat} are the terminal voltage, open-circuit voltage, internal resistance and output current of the battery respectively. The output current I_{bat} of the battery can be estimated as:

$$I_{bat} = \frac{V_{oc} - \sqrt{V_{oc}^2 - 4 \cdot r_i \cdot P}}{2r_i} \tag{A2}$$

The following equation shows the Coulomb counting method for the estimation of SOC of a battery:

$$SOC_{bat} = SOC_{bat}^{ini} - \int \frac{\eta \cdot I_{bat}}{q} dt \tag{A3}$$

where SOC_{bat}, SOC_{bat}^{ini}, η, and q represents battery SOC, battery initial SOC, charge/discharge mode and battery capacity (ampere-hour) respectively.

Table A5 below gives the parameters of batteries used in this research work:

Table A5. Parameters of batteries.

Type	CINCO FM/BB12100T
Capacity	200 Ah
Voltage/string	12 V
Number of parallel strings	3
Number of series strings	34
Rated voltage	$12 \times 34 \approx 400$ V

Appendix A.6. Parameters of Electrolyzer

Table A6 below gives the parameters of electrolyzer used in this research work.

Table A6. Parameters of electrolyzer.

Type	QualeanQL-85000
Rated power	150 kW
Rated voltage	380 V
Number of cells in the stack	30
Number of electrolyzers	6

Appendix A.7. Parameters of Microturbine

Table A7 below gives the parameters of the microturbine used in this research work:

Table A7. Parameters of microturbine.

Type	Ingersoll Rand MT250
Rated power	200 kVA, 160 kW
Rated voltage	440 V
Rated frequency	50

Appendix A.8. Parameters of Utility Grid

Table A8 below gives the parameters of UG used in this research work.

Table A8. Parameters of UG.

Parameter	Rating
Phase Voltage	11 kV
Frequency	50 Hz
Rated power	10 MVA

Appendix A.9. Parameters of Interlinking Inverter

Table A9 below gives the parameters of interlinking inverter used in this research work:

Table A9. Parameters of Interlinking Inverter.

Type	Zhejiang, China CHZIRI-2VF
Rated power	400 kW
Rated voltage	200/540 V
Inductance L-filter	2.1 μH

Appendix A.10. Adaptive PID Control System

The aPID control law, used for various comparisons and control in this research work, is given below:

$$u_{aPID}(j) = K_{P-aPID}(j)e(j) + \tag{A4}$$
$$K_{I-aPID}(j) \int (e(j))dt + K_{D-aPID}(j)\frac{d(e(j))}{dt}$$

where, K_{P_aPID}, K_{I-aPID}, and K_{D-aPID} are proportional, integral, and derivative constants. $e(j)$ is the error signal used as cost function. The update equations are:

$$K_{P-aPID}(j+1) = K_{P-aPID}(j) + \alpha e^2(j) \tag{A5}$$

$$K_{I-aPID}(j+1) = K_{I-aPID}(j) + \alpha e(j) \int e(j)dt \tag{A6}$$

$$K_{D-aPID}(j+1) = K_{D-aPID}(j) + \alpha e(j)\frac{d(e(j))}{dt} \tag{A7}$$

References

1. Wang, P.; Jin, C.; Zhu, D.; Tang, Y.; Loh, P.C.; Choo, F.H. Distributed Control for Autonomous Operation of a Three-Port AC/DC/DS Hybrid Microgrid. *IEEE Trans. Ind. Electron.* **2015**, *62*, 1279–1290. [CrossRef]
2. Jin, C.; Dong, C.; Wang, J.; Wang, P. Uniform Control Scheme for the Interlinking Converter Enhancing the Economy and Resilience of Hybrid AC/DC Microgrids. In Proceedings of the 2018 IEEE Innovative Smart Grid Technologies–Asia (ISGT Asia), Singapore, 22–25 May 2018. [CrossRef]
3. Kebede, A.A.; Kalogiannis, T.; Van Mierlo, J.; Berecibar, M. A comprehensive review of stationary energy storage devices for large scale renewable energy sources grid integration. *Renew. Sustain. Energy Rev.* **2022**, *159*, 112213. [CrossRef]
4. Wang, G.; Sadiq, M.; Bashir, T.; Jain, V.; Ali, S.A.; Shabbir, M.S. The dynamic association between different strategies of renewable energy sources and sustainable economic growth under SDGs. *Energy Strategy Rev.* **2022**, *42*, 100886. [CrossRef]
5. Pires, V.F.; Cordeiro, A.; Roncero-Clemente, C.; Rivera, S.; Dragičević, T. DC-DC Converters for Bipolar Microgrid Voltage Balancing: A Comprehensive Review of Architectures and Topologies. *IEEE J. Emerg. Sel. Top. Power Electron.* **2022**, *11*, 981–998. [CrossRef]
6. Ali, M.; Hossain, M.I.; Shafiullah, M. Fuzzy logic for energy management in hybrid energy storage systems integrated DC microgrid. In Proceedings of the 2022 International Conference on Power Energy Systems and Applications (ICoPESA), Singapore, 25-27 February 2022; pp. 424–429.
7. Jovcic, D.; Ahmed, K. *High-Voltage Direct-Current Transmission*; John Wiley & Sons, Ltd.: New York, NY, USA, 2015. [CrossRef]
8. Alzahrani, A.; Ramu, S.K.; Devarajan, G.; Vairavasundaram, I.; Vairavasundaram, S. A review on hydrogen-based hybrid microgrid system: Topologies for hydrogen energy storage, integration, and energy management with solar and wind energy. *Energies* **2022**, *15*, 7979. [CrossRef]
9. Sarwar, S.; Kirli, D.; Merlin, M.; Kiprakis, A.E. Major Challenges towards Energy Management and Power Sharing in a Hybrid AC/DC Microgrid: A Review. *Energies* **2022**, *15*, 8851. [CrossRef]

10. Wang, H.; Li, W.; Yue, Y.; Zhao, H. Distributed economic control for AC/DC hybrid microgrid. *Electronics* **2022**, *11*, 13. [CrossRef]
11. Ullah, Z.; Wang, S.; Lai, J.; Azam, M.; Badshah, F.; Wu, G.; Elkadeem, M.R. Implementation of various control methods for the efficient energy management in hybrid microgrid system. *Ain Shams Eng. J.* **2022**, 101961 . . [CrossRef]
12. Boche, A.; Foucher, C.; Villa, L.F.L. Understanding Microgrid Sustainability: A Systemic and Comprehensive Review. *Energies* **2022**, *15*, 2906. [CrossRef]
13. Fotopoulou, M.; Rakopoulos, D.; Stergiopoulos, F.; Voutetakis, S. A Review on the Driving Forces, Challenges, and Applications of AC/DC Hybrid Smart Microgrids. In *Smart Grids Technology and Applications*; IntechOpen: London, UK, 2022 . . [CrossRef]
14. Prasad, T.N.; Devakirubakaran, S.; Muthubalaji, S.; Srinivasan, S.; Karthikeyan, B.; Palanisamy, R.; Bajaj, M.; Zawbaa, H.M.; Kamel, S. Power management in hybrid ANFIS PID based AC-DC microgrids with EHO based cost optimized droop control strategy. *Energy Rep.* **2022**, *8*, 15081–15094. [CrossRef]
15. Sajid, A.; Sabzehgar, R.; Rasouli, M.; Fajri, P. Control of Interlinking Bidirectional Converter in AC/DC Hybrid Microgrid Operating in Stand-Alone Mode. In Proceedings of the 2019 IEEE Milan PowerTech, Milan, Italy, 23–27 June 2019. [CrossRef]
16. Unamuno, E.; Barrena, J.A. Hybrid AC/DC microgrids—Part II: Review and classification of control strategies. *Renew. Sustain. Energy Rev.* **2015**, *52*, 1123–1134. [CrossRef]
17. Shen, X.; Tan, D.; Shuai, Z.; Luo, A. Control Techniques for Bidirectional Interlinking Converters in Hybrid Microgrids: Leveraging the advantages of both AC and DC. *IEEE Power Electron. Mag.* **2019**, *6*, 39–47. [CrossRef]
18. Ali, S.U.; Aamir, M.; Jafri, A.R.; Subramaniam, U.; Haroon, F.; Waqar, A.; Yaseen, M. Model predictive control—Based distributed control algorithm for bidirectional interlinking converter in hybrid microgrids. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, e12817. [CrossRef]
19. Guo, Y.; Guo, Y.; Sun, H.; Xi, J.; Hao, Y. Overview of Improved Droop Control Methods of Hybrid AC/DC Microgrid Interlinking Converter. In Proceedings of the 2nd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2017), Beijing, China, 8–11 September 2017. [CrossRef]
20. Malik, S.M.; Sun, Y.; Huang, W.; Ai, X.; Shuai, Z. A Generalized Droop Strategy for Interlinking Converter in a Standalone Hybrid Microgrid. *Appl. Energy* **2018**, *226*, 1056–1063. [CrossRef]
21. Wang, C.; Deng, C.; Pan, X. Line impedance compensation control strategy for multiple interlinking converters in hybrid AC/DC microgrid. *IET Gener. Transm. Distrib.* **2022**. [CrossRef]
22. Rault, P.; Guillaud, X.; Colas, F.; Nguefeu, S. Investigation on interactions between AC and DC grids. In Proceedings of the 2013 IEEE Grenoble Conference, Grenoble, France, 16–20 June 2013. [CrossRef]
23. Martinez, S.; Torres, F.; Roa, C.; Lopez, E. Interaction between AC Grids and MTDC Systems Based on Droop Controllers. In Proceedings of the 2018 IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA), Concepcion, Chile, 17–19 October 2018. [CrossRef]
24. Zhang, J.; Guo, D.; Wang, F.; Zuo, Y.; Zhang, H. Control strategy of interlinking converter in hybrid AC/DC microgrid. In Proceedings of the 2013 International Conference on Renewable Energy Research and Applications (ICRERA), Madrid, Spain, 20–23 October 2013. [CrossRef]
25. Adi, F.S.; Song, H.; Kim, J.S. Interlink Converter Controller Design based on System Identification of DC Sub-Grid Model in Hybrid AC/DC Microgrid. *IFAC-PapersOnLine* **2019**, *52*, 45–50. [CrossRef]
26. Watson, J.D.; Lestas, I. Control of Interlinking Converters in Hybrid AC/DC Grids: Network Stability and Scalability. *IEEE Trans. Power Syst.* **2021**, *36*, 769–780. [CrossRef]
27. Zhang, Z.; Fang, J.; Dong, C.; Jin, C.; Tang, Y. Enhanced Grid Frequency and DC-link Voltage Regulation in Hybrid AC/DC Microgrids through Bidirectional Virtual Inertia Support. *IEEE Trans. Ind. Electron.* **2022**, 1–10 . [CrossRef]
28. Awais, M.; Khan, L.; Ahmad, S.; Mumtaz, S.; Badar, R. Nonlinear adaptive NeuroFuzzy feedback linearization based MPPT control schemes for photovoltaic system in microgrid. *PLoS ONE* **2020**, *15*, e0234992. [CrossRef]
29. Awais, M.; Khan, L.; Badar, R.; Ahmad, S.; Mumtaz, S.; Ullah, S. NeuroFuzzy Full-Recurrent Hybrid B-Spline Wavelet Based Feedback Linearization Control for PMSG-WECS in a Grid-connected Hybrid Power System. In Proceedings of the 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan, 12–16 January 2021. [CrossRef]
30. Awais, M.; Khan, L.; Ahmad, S.; Mumtaz, S.; Badar, R.; Ullah, S. Legendre-wavelet embedded NeuroFuzzy feedback linearization based controlscheme for PHEVs charging station in a microgrid. *Turk. J. Electr. Eng. Comput. Sci.* **2021**, *29*, 2046–2066. [CrossRef]
31. Awais, M.; Khan, L.; Ahmad, S.; Jamil, M. Feedback-Linearization-Based Fuel-Cell Adaptive-Control Paradigm in a Microgrid Using a Wavelet-Entrenched NeuroFuzzy Framework. *Energies* **2021**, *14*, 1850. [CrossRef]
32. Awais, M.; Khan, L.; Badar, R.; Mumtaz, S.; Ahmad, S.; Ullah, S. Wavelet-Hybridized NeuroFuzzy Feedback Linearization based Control Strategy for PHEVs Charging Station in a Smart Microgrid. In Proceedings of the 2020 International Symposium on Recent Advances in Electrical Engineering & Computer Sciences (RAEE & CS), Islamabad, Pakistan, 20–22 October 2020. [CrossRef]
33. Dehkordi, N.M.; Sadati, N.; Hamzeh, M. Robust backstepping control of an interlink converter in a hybrid AC/DC microgrid based on feedback linearisation method. *Int. J. Control* **2016**, *90*, 1990–2004. [CrossRef]
34. Zhang, B.; Cao, G.; Ren, C.; Wang, J.; Han, X. Autonomous control strategy of bidirectional AC/DC converter in low voltage hybrid microgrid. In Proceedings of the 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA), Siem Reap, Cambodia, 18–20 June 2017. [CrossRef]

35. Saad, N.H.; El-Sattar, A.A.; Mansour, A.E.A.M. A novel control strategy for grid connected hybrid renewable energy systems using improved particle swarm optimization. *Ain Shams Eng. J.* **2018**, *9*, 2195–2214. [CrossRef]

36. Jiao, J.; Meng, R.; Guan, Z.; Ren, C.; Wang, L.; Zhang, B. Grid-connected Control Strategy for Bidirectional AC-DC Interlinking Converter in AC-DC Hybrid Microgrid. In Proceedings of the 2019 IEEE 10th International Symposium on Power Electronics for Distributed Generation Systems (PEDG), Xi'an, China, 3–6 June 2019. [CrossRef]

37. Yu, H.; Niu, S.; Zhang, Y.; Jian, L. An integrated and reconfigurable hybrid AC/DC microgrid architecture with autonomous power flow control for nearly/net zero energy buildings. *Appl. Energy* **2020**, *263*, 114610. [CrossRef]

38. Sinha, S.; Ghosh, S.; Bajpai, P. Power sharing through interlinking converters in adaptive droop controlled multiple microgrid system. *Int. J. Electr. Power Energy Syst.* **2021**, *128*, 106649. [CrossRef]

39. S., J.N.; Nempu, P.B. Performance Enhancement of a Hybrid AC-DC Microgrid Operating with Alternative Energy Sources Using Supercapacitor. *Int. J. Electr. Comput. Eng. Syst.* **2021**, *12*, 67–76. [CrossRef]

40. Ahmed, M.; Meegahapola, L.; Datta, M.; Vahidnia, A. A Novel Hybrid AC/DC Microgrid Architecture with a Central Energy Storage System. *IEEE Trans. Power Deliv.* **2022**, *37*, 2060–2070. [CrossRef]

41. Antalem, D.T.; Muneer, V.; Bhattacharya, A. Decentralized control of islanding/grid-connected hybrid DC/AC microgrid using interlinking converters. *Sci. Technol. Energy Transit.* **2022**, *77*, 22. [CrossRef]

42. Hu, Y.; Li, W.; Xu, H.; Xu, G. An online learning control strategy for hybrid electric vehicle based on fuzzy Q-learning. *Energies* **2015**, *8*, 11167–11186. [CrossRef]

43. Dai, X.; Li, C.K.; Rad, A.B. An approach to tune fuzzy controllers based on reinforcement learning for autonomous vehicle control. *IEEE Trans. Intell. Transp. Syst.* **2005**, *6*, 285–293. [CrossRef]

44. Khan, L.; Badar, R. Hybrid adaptive neuro-fuzzy B-spline-based SSSC damping control paradigm using online system identification. *Turk. J. Electr. Eng. Comput. Sci.* **2015**, *23*, 395–420. [CrossRef]

45. Badar, R.; Khan, L. Online adaptive Legendre wavelet embedded neurofuzzy damping control algorithm. In Proceedings of the INMIC, Lahore, Pakistan, 19–20 December 2013; pp. 7–12. [CrossRef]

46. Werbos, P.J. Generalization of backpropagation with application to a recurrent gas market model. *Neural Netw.* **1988**, *1*, 339–356. [CrossRef]

47. Tutschku, K. Recurrent Multilayer Perceptrons for Identification and Control: The Road to Applications, 1995. Available online: https://www.researchgate.net/profile/Kurt-Tutschku/publication/229087124_Recurrent_multilayer_perceptrons_for_identification_and_control_The_road_to_applications/links/0912f50bdf57c64002000000/Recurrent-multilayer-perceptrons-for-identification-and-control-The-road-to-applications.pdf (accessed on 30 December 2022).

48. Sun, W.; Wang, Y. A recurrent fuzzy neural network based adaptive control and its application on robotic tracking control. *Neural Inf. Process.-Lett. Rev.* **2004**, *5*, 19–26.

49. Song, J.; Shi, H. Dynamic system modeling based on wavelet recurrent fuzzy neural network. In Proceedings of the 2011 Seventh International Conference on Natural Computation, Shanghai, China, 26–28 July 2011; Volume 2, pp. 766–770. [CrossRef]

50. Liu, Y.; Lin, Y.; Wu, S.; Chuang, C.; Lin, C. Brain Dynamics in Predicting Driving Fatigue Using a Recurrent Self-Evolving Fuzzy Neural Network. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 347–360. [CrossRef]

51. Pratama, M.; Lu, J.; Lughofer, E.; Zhang, G.; Er, M.J. An Incremental Learning of Concept Drifts Using Evolving Type-2 Recurrent Fuzzy Neural Networks. *IEEE Trans. Fuzzy Syst.* **2017**, *25*, 1175–1192. [CrossRef]

52. Watkins, C.J.C.H. Learning from Delayed Rewards. Ph.D. Thesis, University of Cambridge, Cambridge, UK, 1989.

53. *IEEE Std 1547-2018 (Revision of IEEE Std 1547-2003)*; IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces. IEEE: Piscataway, NJ, USA, 2018; pp. 1–138. [CrossRef]

54. Huang, B.; Wang, J. Deep-reinforcement-learning-based capacity scheduling for PV-battery storage system. *IEEE Trans. Smart Grid* **2020**, *12*, 2272–2283. [CrossRef]

55. Mumtaz, S.; Ali, S.; Ahmad, S.; Khan, L.; Hassan, S.Z.; Kamal, T. Energy management and control of plug-in hybrid electric vehicle charging stations in a grid-connected hybrid power system. *Energies* **2017**, *10*, 1923. [CrossRef]

56. Sharma, V.; Haque, M.H.; Aziz, S.M. Energy cost minimization for net zero energy homes through optimal sizing of battery storage system. *Renew. Energy* **2019**, *141*, 278–286. [CrossRef]

57. Braeuer, F.; Rominger, J.; McKenna, R.; Fichtner, W. Battery storage systems: An economic model-based analysis of parallel revenue streams and general implications for industry. *Appl. Energy* **2019**, *239*, 1424–1440. [CrossRef]