A review of self-learning and adaptive techniques for grid balancing

Antonella TANNOUS, Ferréol BINOT, Bruno FRANCOIS
Univ. Lille, Arts et Metiers Institute of Technology, Centrale Lille, Junia, ULR 2697 - L2EP, F-59000 Lille, France
antonella.tannous@centralelille.fr, ferreol.binot@centralelille.fr, bruno.francois@centralelille.fr

ABSTRACT – The cost of energy and climate change due to the emission of greenhouse gases have yielded a large penetration of renewable energy sources in the grid and new types of loads (electric vehicles, ...). The increasing variability and uncertainty in sources and demand lead to significant impacts on small electrical distribution networks. Thus, a stable, and reliable operation in local areas becomes more difficult to achieve. Adapted grid control according to the observed or expected variability in power flows and in model parameters of the electrical system, is required while maintaining the costs and CO_2 emissions to low levels. This paper presents a state of the art on adaptive techniques and new self-learning methods for achieving grid balancing.

Keywords – Power system operation, artificial intelligence, power reserves, grid control, machine learning, distribution grids, review.

1. INTRODUCTION

In conventional power systems with synchronous generators, balance between power generation and demand is indicated by the grid frequency. The frequency value is the same everywhere on the grid, and is imposed by the angular speed of connected synchronous alternators. To achieve a power balance and maintain the frequency, the frequency deviation is sensed and mitigated by a hierarchical power system control composed of primary, secondary, and tertiary control.

To minimize operating costs, controllable generators' set points are planned one day ahead with renewable generation and load demand predictions, in addition to the required operating power reserves (PR) for compensating unexpected events and forecast errors. In conventional power systems, PR is provided by large synchronous generators. PR are usually sized to provide the capacity of the largest connected generator and/or a fragment of the load. For the management and use of the available PR, required contingency reserves are classed into three types according to prescribed time responses for the power system operation (see Fig.1) : Primary or spinning reserves are provided by fast power sources (within 30 seconds) to stop the frequency drop following an event (connection of a large load, fault, etc.). Secondary or supplementary reserves' full power are provided within 15 minutes to bring the frequency into margins, hence satisfying the balance between generation and load demand. Tertiary or replacement reserves are provided within 30 minutes to share the required additional power among slow synchronous generators (cheaper to operate) by modifying their setpoints.

Nonetheless, the planned retirement of thermal power plants and the large penetration of renewable energy sources (RES) require a larger capacity of PR due to the intermittent nature of the newly introduced sources in local areas. A part of PR must compensate the unexpected variations of the RES-based production affected by weather conditions and solve frequency stability issues [1].

In section 2, challenges facing power systems are described. Section 3, presents a classification of advanced techniques for grid balancing. The use of these techniques for direct frequency regulation, PID controller tuning and distributed control architectures are discussed in section 4, 5 and 6, respectively. To conclude, a table of the limitations, communication requirements, and computation time of the proposed methods is derived in section 7.

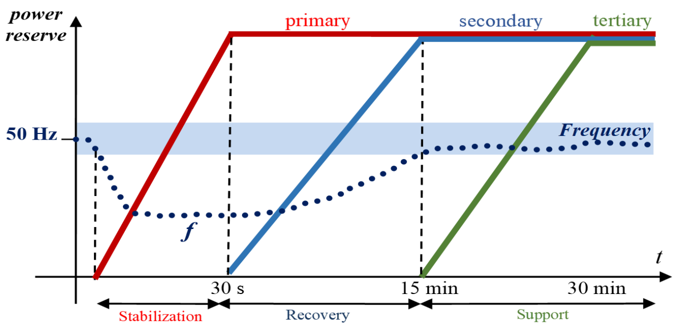

FIG. 1. Power reserve operation for active power balance.

2. CHALLENGES

2.1. *Stochastic operational planning of power sources*

PR must be properly sized and allocated to minimize economic operating costs while maintaining a satisfactory level of security. Traditional deterministic operational planning approaches are being replaced by stochastic ones due to the variability of RES-based generation. These sources are highly variable but predictable with residual uncertainties [2]. However, even with the best forecasting tools, unexpected events might occur such as a passing cloud. Hence, the availability of PR during the next hour of the day is harder to predict. As a result, calculating the exact setpoints for each generator to meet the demand becomes challenging.

2.2. *Adapted output power of power electronics grid connected converters in AC grids*

Power electronics (PE) converters between RES with or without battery energy storage system (BESS) and the grid have very fast response times, because of their high frequency switching operation. In AC grids, the increased integration of RES leads to a reduced inertia and large frequency deviations. Thus, conventional PR with a least share of synchronous generators will not be efficient in future power systems with a large amount of grid connected PE converters. Additionally, a growing amount of grid codes require that RES with or without storage devices must participate in inertia response and frequency control to adapt their output power to achieve the grid balance [3].

2.3. *Balancing DC grids*

DC grids present a hot topic due to the better efficiency in power flows between sources of the same DC nature such as photovoltaic (PV) panels, light emitting diodes (LED), electrical vehicles (EVs), etc. without reactive power. However, DC grids with zero frequency cannot be balanced with the frequency signal, but with the voltage as a power balance indicator. Energy management and balancing of DC grids clusters rely extensively on communication networks, which presents information security and latency challenges among others [4]. Since communication failure problems are inevitable, alternative techniques with local implementation and without communication are mandatory.

2.4. Requirements

With all that in mind, the resilience and robustness of modern power systems are being questioned. The use of a frequency droop controller seems to be limited to some operating range and dynamic scenarios in the context of synchronous generators or emulated ones by grid connected PE converters. New external inputs such as solar irradiation, wind speed, forecasting errors, and others have large effects on the system's operation. These inputs can be used in new control techniques to increase the controllability and maintain a good performance. The linearity and stationarity of these inputs with the electrical quantities are still to be studied. Moreover, to ensure the energy demand at the lowest cost with less complexities, one solution is to organize a local regulating area (local energy communities) that achieves a local balancing. To implement this method, available flexibilities such as electric vehicles, distributed storage systems and others must be managed in real-time by advanced control techniques.

Due to nonlinearities, increasing variabilities and uncertainties, the mathematical model of such electrical systems is extremely complex and includes non-stationary parameters. Based on this complex model (supposedly known), the design of the control system must include nonlinear dynamic responses. It should also be able to change according to external and internal uncertain excitatory influences. This adaptability is classically implemented with an adjustment of various parameters (in the control law) such that a deviation error from an objective function is minimized. This is the principle of a learning algorithm, which is fed in real-time with updated data in this changing environment.

The increased amount of data in power systems due to advanced measurement and communication systems (smart meters, etc.) has led to a greater focus on data driven methods (especially artificial intelligence) in power system operation. Moreover, the enhanced computational capacity allowed for powerful and real-time applications. Artificial intelligence (AI) is defined as an information-processing system that learns from the environment and adapts. It is a field comprising various techniques and methods that allow a machine to perform intelligent tasks like those performed by human intelligence.

3. Classification of advanced techniques for grid balancing

The mostly applied AI methods to power systems can be divided into three main categories : rule-based, metaheuristic, and machine learning methods [5]. In this section, a classification of AI methods adapted to power systems with examples is proposed.

3.1. Rule-based methods

Rule-based methods or expert systems require a set of if-then rules predefined by human knowledge, and thus the system can make decisions.

To implement the interface with the analog world, fuzzy logic is often used. It involves three steps : fuzzification to convert a crisp input value to a fuzzy value representing a membership value to an input set, inference to apply rules and generate the decision and/or the belonging index of an output set, and defuzzification to convert the obtained fuzzy quantity to a precious value. In most grid balancing applications, fuzzy logic is used to solve multi-input problems such as frequency deviation and its derivative. It is then combined with neural networks (NN) to create an adaptive neuro-fuzzy controller whose parameters are trained via an adaptive neuro-fuzzy inference system [6].

3.2. Metaheuristic methods

Metaheuristic methods are defined as algorithms able to solve complex optimization problems. They rely on a general set of rules rather than mathematical approaches, and can be either population-based or trajectory-based. Population-based methods maintain a population of possible solutions and try to improve them iteratively. Trajectory-based methods try to improve a single possible solution by small adjustments and evaluation of resulting solution. In power systems, they are mostly used to tune parameters of models or controllers.

3.3. Machine learning algorithms

Machine learning techniques include supervised, unsupervised, and reinforcement learning.

Supervised learning (SL) methods require input and output datasets (labeled data). They calculate parameters of the input-output model (training phase) and validate it (testing phase). In the training phase, an optimizer is used to calculate parameters that minimize the error between the actual and target output values.

Unsupervised learning (UL) does not require labeled data. Therefore, the algorithm needs to identify structure and patterns in provided data to make decisions.

Reinforcement learning (RL) is an incorporation of human-level control into systems. It is an agent-based method, meaning that an agent (computational entity) learns to make decisions by interacting with the environment and trying to maximize its reward. The main difference between RL and the previous machine learning approaches, is that the agent does not rely on labeled data but on the consequences of the actions it took (see Fig.2). RL requires a representation of the environment via a Markov Decision Process (MDP) consisting of a state space, an action space, the probability of transition, a discounting factor, the initial distribution over states (probability of being in each state at the beginning) and the reward. The agent observes the environment's state, chooses the action to receive the best reward and transition to a new state. The choice of the action is defined by a policy . The total expected reward for a given policy can be estimated by a value function, which is a state-value function or an action-value function, also known as Q-function [7]. Various RL methods exist, but they can be grouped into value-based, policy-based, actor-critic or other methods :

1. Value-based methods use value functions to learn an optimal policy. They estimate the anticipated value of each state and/or action through a value function and make decisions that maximize the expected reward. Q-learning, Deep Q-network, and State-Action-Reward-State-Action (SARSA) are the most common methods used in power systems.

2. Policy-based methods learn a policy directly without the need for the value function. They parameterize the policy as a function of the state and find the policy parameters via optimization. These methods are less popular than value-based methods in energy systems.

3. Actor-critic methods are a combination of both policy-based methods (actors) and value-based methods (critics). The critics estimate the value function, and the actors learn the policy and take actions. The critic network assesses the quality of each state-action combination and

gives input to the actor network to adjust the policy accordingly. Deep deterministic policy gradient (DDPG) is a widely used algorithm in energy systems.

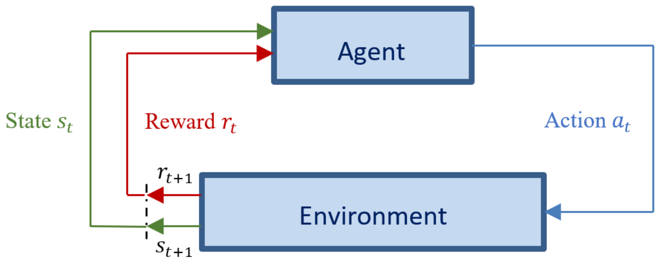

FIG. 2. Reinforcement learning.

4. ADAPTIVE CONTROLLERS OF THE FREQUENCY REGULATOR

Controlling energy systems has become more difficult with the integration of RES and the limitations of physics-based models. Therefore, an evolution of traditional energy system control is necessary to face different arising challenges such as uncertainties, security, complexity, etc. Due to the non-linearity of AC microgrids, the knowledge of accurate model parameters is essential for their control design. To reduce the complexity, usually a linearization is applied. The obtained model is a simplified representation with modelling errors. Data-driven based RL approaches are currently being investigated because of their ability to learn the variable non-linear models or control laws, and ensure better dynamic performances and robustness of frequency control. These approaches employ an agent that takes input from the environment and makes decisions accordingly. Data-driven models are tuned by taking part in the decision-making procedure and maximizing its overall reward. They have proven to be successful methods in handling the complexities and uncertainties of models that cannot be easily managed with explicit detailed mathematical models.

In [8], a data-driven frequency regulation technique is proposed for an islanded MG without the need for accurate parameters. An improved scheme of adaptive dynamic programming called distributed dual heuristic dynamic programming algorithm (DDHP) is used. A DDHP controller generates frequency controller inputs for the generator units to define their participation in secondary frequency control. Hence, the frequency on each MG bus is regulated and the load is optimally shared among the generators. This is done by sensing frequency deviation and active power difference between the MG's nodes. A local error for each bus is then formulated and sampled. The error along with the control action of the generator on the bus are used to train a model NN (MNN) to approximate the dynamics of the system. A critic neural network is trained with the output of the MNN and approximates the value function while an actor neural network outputs the optimal control input. This input is integrated and added to the droop control output. Compared to the distributed averaging PI controller, the DDHP controller converges faster.

In the same context, a load frequency controller based on a goal representation dynamic programming algorithm (GRADP) in [9] is introduced to control a micro-turbine and EVs of an islanded smart grid. In fact, the GRADP takes the frequency deviation and its two previous values as inputs and provides a supplementary control to the PID controller of each unit for better achieving a frequency stability. The GRADP method consists of :

1. A critic network that estimates the cost-to-go function (1),

$$J[x(i), i] = \sum_{t=i}^{\infty} \gamma^{t-i} U[x(t), u(t), i] \tag{1}$$

   Where γ represents the discount factor, U the utility function, $x(t)$ the state vector and $u(t)$ the control action.
2. An action network that tries to minimize the function by generating the appropriate control policy,
3. A goal network that tries to improve the learning capability of the critic.

The GRADP based controller is compared to a particle swarm optimization (PSO) based fuzzy controller under power disturbances and signal transmission delays. The former has shown better learning abilities and more robust control. The main difference between DDHP and GRADP is the representation of their control objective and the learning of control policy. DDHP leverages heuristic functions to solve large-scale MDPs in a distributed manner, while GRADP focuses on learning a goal representation to facilitate exploration in high-dimensional state spaces.

The third method proposed in literature is deep RL (DRL), which is a subfield of RL that uses deep NN to approximate the optimal policy or the value function. In other words, it combines the principles of RL with the power of deep NN to allow for an efficient learning of complex behaviors and decision-making processes that cannot be handled by the traditional RL schemes. Authors in [10] proposed a data driven load frequency controller based on DRL to minimize frequency deviations in stochastic power systems. The DRL based controller adjusts the power references of a power plant to regulate the frequency deviations due to load and wind power variations. The DRL based controller here is model-free and policy-based. It is offline trained where the explorations are made continuously. The performance of the controller is then evaluated by an action-value function and based on DDPG, the policy agent is updated. Since the input features to the DRL based controller are crucial, a stacked denoising autoencoder method (SDAE) is implemented as a feature analyzer. SDAE extracts features from data rather than feeding frequency deviation as an input to the controller like most common methods.

5. ARTIFICIAL INTELLIGENCE TUNING BASED PID CONTROLLER

5.1. *Metaheuristic population based approach*

Optimizing the power flow and variation of distributed generators as well as minimizing frequency deviation in an islanded MG, are examined in [11]. Authors applied a PSO algorithm with a cost function to optimize the PI parameters of the voltage control loop of a grid-support/grid-forming voltage source inverter. This type of inverter is implemented on the PV and the BESS due to its flexibility in different MG operation modes. It regulates the voltage and frequency references based on active and reactive power levels in the MG. Three control loops are implemented (from slowest to fastest) : the droop, voltage, and current control loops. A virtual impedance control is added to the voltage control loop to increase the accuracy of the power sharing among generators. PI controller parameters of the control loops are determined via the Zeigler Nichols method. Hence, the overall stability of the control loops is established and analyzed using the Root Locus method.

Additionally, a power management control strategy is developed to allow PV units to supply active and reactive power with regards to the MG's voltage and frequency. Also, the BESS is allowed to charge/discharge based on its safe operating range, PV production and MG's demand. Consequently, a balance between production and demand is achieved despite load and solar irradiance changes.

The cost function (2) is formulated using an integral of the squared voltage error e and a penalty function p. The former optimizes the voltage error signals of the voltage control loop to regulate the reference currents and hence ensure optimal active and reactive power flow. The latter solves frequency variation constraints, so it does not exceed ± 0.15 Hz.

$$cost\,function = \int_0^T e^2(t) \cdot \mathrm{d}t + p \tag{2}$$

The goal of optimizing control parameters and the power flow, as well as reducing frequency deviation and power variations of the units is achieved. Compared to a PI based control of the unit's inverter, PSO proved better performance.

5.2. Supervised learning approach

Traditionally, the tuning of PI controllers is based on the primary conditions of a system. However, system conditions can highly deviate from their nominal values in microgrids where the inertia is low, and the production is highly variable. Authors in [12] compared the traditional PI controller to a self-tuning controller based on Artificial NN (ANN). The intelligent unit collects response data from a plant, adjusts the ANN weights via an error back-propagation and proposes a control signal to tune the PI parameters online. Nonlinear blocks, delays due to communication and limiters in the model are considered. When subject to load disturbance, wind and solar power variations, the proposed controller showed a better frequency control than the conventional one.

Other authors [13] proposed PID controllers whose parameters are obtained by ANN. Each controller charges/discharges a BESS or a plug-in hybrid electric vehicle to perform frequency control due to their fast response compared to the utility grid. In this work, frequency deviation of a linearized MG is fed as an input to the PID controllers where a multi-layer feedforward NN fine-tunes the PID parameters. The range of parameter values is set to reduce the computation time while offering acceptable performance. In the training phase, random values are chosen for the inputs (PID parameters) and output of the NN (3).

$$IAE = \int_0^\infty |P_g - P_l| \cdot t \cdot \mathrm{d}t \tag{3}$$

Where IAE represents the integral absolute error, P_g the production fluctuations, and P_l the load deviations.

The goal is to minimize the error, so the corresponding inputs are the optimal parameters. When compared to a PSO based PID controller, the proposed controller reduces frequency deviations more effectively. However, it does not consider load deviation and power variations into dynamic parameter tuning.

5.3. Rule based with Reinforcement learning approach

Frequency regulation of a MG consisting of a tidal power unit, EV, PV, diesel generators and load is investigated in [14]. Since tidal energy is intermittent, even if it can be predicted in the long term, spinning reserves are required from energy storage devices. In this study, a fuzzy logic controller (FLC) is proposed as the main load frequency controller and a supplementary control is provided by an RL based controller. While the former ensures basic load frequency controller performance (stabilize the frequency fluctuations), the latter reinforces the control by adaptations according to the uncertainties in the MG. The FLC is type 2 with a single input to provide a better control because of its fuzzy membership functions. Typically, FLC is optimized either by heuristic methods or NN that fail to learn online. Hence, RL can be ideal where an agent learns a policy and maximizes the reward by interacting with the environment.

Regarding the main controller, a baseline PID controller whose coefficients are adjusted heuristically, is cascaded with the FLC. The footprint of uncertainty (FOU) is an important concept in type 2 fuzzy logic because it indicates how much uncertain the membership of a value is in a fuzzy set. The greater its area, the greater the uncertainty is. Hence, fractional gradient descent is applied to regulate the FOU coefficient to its lowest value.

As for the supplementary controller, a deep deterministic policy gradient (DDPG) algorithm is applied for its ability to work in a continuous action space. The method consists of a critic network, and an actor network. The critic network analyzes the quality of the control signal by considering a state and an action as inputs, and then generates the Q-value. The frequency deviation and its derivative are fed as inputs to the actor network which then generates the control action. To assess the performance of the controller, a reward is defined as such (4) :

$$r(s,a) = -\left(\sum_{i=t-1}^{t} |\Delta f_i|\right) \tag{4}$$

Where $r(s,a)$ represents the cumulative reward of state s and action a, and Δf_i the frequency deviation.

When the frequency deviation increases, the reward is reduced and vice versa. Hence, weights are updated, and the supplementary controller reduces the effect of MG uncertainties by generating adaptive control actions.

6. Distributed control for the balancing

6.1. Centralized learning

Multi-agent RL (MARL) can be used to implement decentralized control. Authors in [15] implemented a load frequency control as a MARL problem. The controllers of the synchronous generators are denoted as agents. These agents interact with each other and the environment to learn optimal policies. The MARL approach needs a mathematical formalization of the load frequency control problem. Thus, MDP is used. It consists of a state space, an action space, a transition function, and a reward function. The state space includes two states that are the current control action and the deviation from the synchronous speed of the generator. The action space is defined by two actions, one that increases and the other decreases the control action depending on the states. The transition function determines the environment's dynamics when transitioning between states, by using a set of equations. The reward function represents the effect of being present in a state and executing an action. Consequently, the agent tries to minimize the deviation as much as possible by maximizing its reward. Therefore, finding the optimal policy to maximize the reward is defined as the MARL's goal. Since the global performance of the agent cannot be assessed via the reward, an action-value function Q_π is used to determine the expected reward in the long run. This function is approximated by Deep Q-learning due to the high number of states and actions. Solving the MARL problem is done using the multi-agent deep deterministic policy gradient algorithm. The critics and actors are modelled by long short-term memory networks :

1. The role of the critics is to teach the actors the behavior of the other agents and the environment's dynamics, using central information. The speed deviation, current and predicted control action of all agents are fed as input to the critic network. The network then computes Q_π.
2. The role of the actors is to generate the action based on local information only as they have already been trained and know the behavior of other actors. The speed deviation and the current control action of the agent are fed as inputs to the network. The output is the predicted control of the agent itself.

The proposed scheme learns in a central way and implements the control actions in a distributed manner. Hence, communication infrastructure is no longer needed.

6.2. Distributed learning

Distributed learning [16] is a machine learning approach where a model is trained using data that are distributed across several machines, instead of just one machine. This subfield of machine learning can solve complex problems and handle large data sets. Data security and privacy is ensured, while single point failure is avoided. Hence, distributed learning is interesting for distributed RES applications because it reduces the data stored and exchanged between sources. In traditional learning, data are collected and treated in a central way to derive a model. However, in distributed learning, data provided from a data center is partitioned with an independent and identical unknown distribution over several edge devices. Edge devices represent communication links between nodes. Local training is achieved in nodes and the gradients are sent to the central server whose role is to update them. Communication between agents is allowed (see Fig.3).

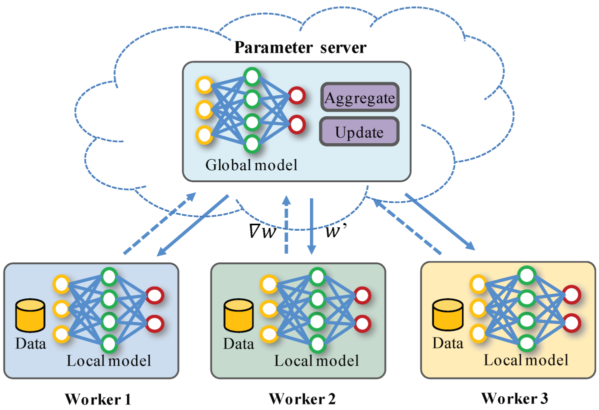

FIG. 3. Distributed learning [17].

Distributed learning can be applied to deep learning, reinforcement learning, etc. It has been implemented in the energy management of MGs. In a specific application, each MG element's controller was supervised by a central server. The dynamic behavior of the MG elements was learned by utilizing localized Hamiltonians that incorporated operational variables, such as cost and voltage-current relationship. A virtual control element was responsible for interacting with other elements through an optimal control process that was subject to constraints. The interaction was repeated multiple times by using a 2-person Pareto game until an equilibrium in the MG was reached.

Assisted learning is a variant where agents assist each other through the exchange of non-sensitive data. Here, communication with the central server is not established. Federated learning (FL) is another variant of distributed learning where communication between agents is not allowed. It is the role of the central server to coordinate the agents in parallel to achieve a specific outcome. In [18], a self-adaptive PID controller based on FL fractional order recurrent NN is applied for frequency regulation. To optimize the PID parameters, the neurons send their updated weights and receive the weights from their neighboring neurons. The neurons keep updating their weights until they formulate the aggregated model.

In [19], a distributed RL actor-critic NN was developed to regulate the frequency of several interconnected MGs while searching for the optimal control scheme. In addition to local power plants, distributed energy resources are controlled to provide frequency regulation. Moreover, tie-line bias control is implemented to coordinate the active power flow between MGs.Typically, actor and critic networks are separated. However, to enhance the control performance and establish the relationship between the two outputs, both networks are integrated in this study (see Fig.4). Optimizing the short-term performance and achieving stability requires an approximation of the desired control output $\hat{u}_{d,i}$ by a deterministic learning algorithm. The desired control output should be able to damp frequency deviation exponentially. On the other hand, the long-term performance can be enhanced by estimating a strategic utility function $\hat{Q}_i$ that imposes penalties for large control outputs and system states. The goal is to avoid deteriorating power devices due to large frequency deviations. Hence, the critic network responsible for the prediction of $\hat{Q}_i$ takes as inputs the control output, the MG's state vector x_i and its neighboring's state vector x_j.

The state vector includes the frequency deviation, the plant's power deviation, the main frequency controller output, and the regional demand of the MG. $N'(i)$ corresponds to the set of MGs cyber connected to the MG i. The proposed control scheme was tested on a power grid of 7 and 12 interconnected MGs, and proved better performance than conventional methods.

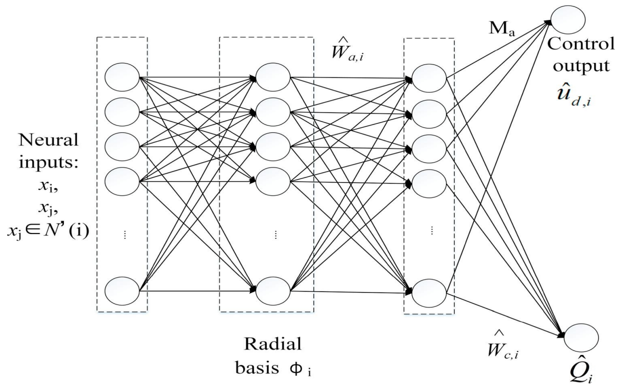

FIG. 4. Integrated actor-critic network [19].

7. CONCLUSION

Modern power systems are facing challenges due to the integration of RES into the power grid and the planned retirement of thermal power plants. The variability and uncertainty in power production and consumption have made stable and reliable operation of small electrical distribution networks more difficult to achieve. Stochastic operational planning is required to account for the variability of renewable generation, while the implementation of power electronics converters, and the management of DC grids introduce new complexities.

The resilience and robustness of modern power systems are being questioned due to external inputs such as solar irradiation, wind speed, and forecasting errors, which can significantly impact system operation. To address these challenges, new control techniques incorporating these inputs are being explored to achieve grid balancing while minimizing costs and CO_2 emissions. Additionally, the concept of local energy communities can help achieve local balancing by managing flexibilities such as EVs and distributed storage systems in real-time.

The complexity of electrical systems, non-stationary parameters, and increasing uncertainties call for the application of adaptive techniques and machine learning methods. Artificial intelligence, with its ability to learn from the environment, offers promising solutions for power system operation. Data-driven approaches, enabled by advanced measurement and communication systems, have gained prominence, leveraging the computational power for real-time applications. Continued research and development in these areas are essential to address the evolving needs of modern power systems.

This review of some self-learning and adaptive techniques for grid balancing groups them into three major methods : adaptive controllers of frequency regulators, AI tuning based PID controllers, and distributed control. Limitations, communication requirements, and computation time of the methods discussed in this article are presented in table 1.

TABLE 1 – Limitations, communication requirements, and computation time of presented articles.

Article	Limitations	Communication requirements	Computation time
[8]	Tradeoff between computational efficiency and algorithm performance. Small number of neurons to reduce the convergence time.	Frequency, droop coefficient, and injected power of each node is communicated with its neighbors.	Time complexity $\mathcal{O}$(nodes $\times$ sample size $\times$ iterations).
[9]	Need for coordination strategy between PID controllers. Need for effective methods to represent signal transmission delays.	Sensors for each generation unit and load is required to communicate data with the local smart grid dispatch system.	When the PID controllers are coordinated, the computation time increases.
[10]	Lack of adaptability and inability to capture real-time dynamics. Limited application for a single plant.	No communication required.	SDAE uses SL and UL which is time-consuming. Trial and error method is computationally expensive.
[11]	Actual values of PI controller parameters are needed to extend the level of search boundaries. The selection of swarm size and number of iterations is not justified.	No communication required.	Method requires time to converge.
[12]	Need for a large dataset. The selection of number of neurons and hidden layers is not justified.	No communication required.	Computation time depends on the size of the dataset, network architecture, etc.
[13]	The parameters of the PID controller are not adaptively fine-tuned. No continuous learning or updates to the model.	No communication required.	If the range of PID parameters is limited, the computation time is reduced.
[14]	The self-tuned FLC requires more time to stabilize the frequency. The supplementary controller does not provide a great enhancement.	Centralized control requires communication.	The adaptation of RL is time-consuming.
[15]	Limited application to synchronous generators. Scalability challenges when the number of agents increases.	No communication between agents required.	Curse of dimensionality due to the number of agents.
[19]	Coordination and synchronization challenges between agents. Scalability challenges when the number of agents increases.	Communication protocols and delays between MGs are not detailed or considered.	Curse of dimensionality due to the number of agents.

8. REFERENCES

[1] E. Ela and al., « Evolution of operating reserve determination in wind power integration studies », *IEEE PES General Meeting*, Minneapolis, MN, USA, 2010, pp. 1-8, doi: 10.1109/PES.2010.5589272.

[2] X. Yan, B. Francois, and D. Abbes, « Operating power reserve quantification through PV generation uncertainty analysis of a microgrid », *IEEE PowerTech*, Eindhoven, Netherlands, 2015, pp. 1-6, doi: 10.1109/PTC.2015.7232577.

[3] M. Dreidy, H. Mokhlis, and S. Mekhilef, « Inertia response and frequency control techniques for renewable energy sources : A review », *Renewable Sustainable Energy Review*, Elsevier, vol. 69, pp. 144- 155, 2017, doi: 10.1016/j.rser.2016.11.170.

[4] L. Meng and al., « Review on Control of DC Microgrids and Multiple Microgrid Clusters », *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 5, no. 3, pp. 928-948, Sept. 2017, doi: 10.1109/JESTPE.2017.2690219.

[5] S. Stock, D. Babazadeh, and C. Becker, « Applications of Artificial Intelligence in Distribution Power System Operation », *IEEE Access 9*, 2021, doi: 10.1109/ACCESS.2021.3125102.

[6] R. Dhanalakshmi, and S. Palaniswami, « ANFIS Based Neuro-Fuzzy Controller in LFC of Wind-Micro Hydro-Diesel Hybrid Power System », *International Journal of Computer Applications*, vol. 42, no. 6, pp. 28-35, Mar. 2012, doi: 10.5120/5698-7745.

[7] A. T. D. Perera, and P. Kamalaruban, « Applications of reinforcement learning in energy systems », *Renewable Sustainable Energy Review*, vol. 137, Mar. 2021, doi: 10.1016/j.rser.2020.110618.

[8] N.-L. Mo and al., « Data-driven based optimal distributed frequency control for islanded AC microgrids », *Intern. Journal of Electr. Power & Energy Syst.*, vol. 119, 2020, doi: 10.1016/j.ijepes.2020.105904.

[9] Y. Tang, J. Yang, J. Yan, and H. He, « Intelligent Load Frequency Controller Using GrADP for Island Smart Grid with Electric Vehicles and Renewable Resources », *Neurocomputing 170*, Elsevier, pp. 406- 416, 2015, doi: 10.1016/j.neucom.2015.04.092.

[10] Z. Yan, and Y. Xu, « Data-Driven Load Frequency Control for Stochastic Power Systems : A Deep Reinforcement Learning Method with Continuous Action Search », *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1653–1656, Mar. 2019, doi: 10.1109/TPWRS.2018.2881359.

[11] A. Vinayagam, and al., « Intelligent control strategy in the islanded network of a solar PV microgrid », *Electric power systems research*, Elsevier, vol. 155, pp. 93–103, 2018, doi: 10.1016/j.epsr.2017.10.006.

[12] F. Habibi, S. Shokoohi, and H. Bevrani, « Designing a self-tuning frequency controller using ann for an isolated microgrid », *International Symposium On Smart Grid Operation and Control (ISSGOC)*, 2012.

[13] S. K. Akula, and H. Salehfar, « Frequency control in microgrid communities using neural networks », *North American Power Symposium (NAPS)*, Wichita, KS, USA, pp. 1-6, 2019, doi: 10.1109/NAPS46351.2019.9000219.

[14] M. H. Khooban and M. Gheisarnejad, « A Novel Deep Reinforcement Learning Controller Based Type-II Fuzzy System : Frequency Regulation in Microgrids », *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 5, no. 4, pp. 689–699, Aug. 2021, doi: 10.1109/TETCI.2020.2964886.

[15] S. Rozada, D. Apostolopoulou, and E. Alonso, « Deep Multi-agent Reinforcement Learning for Cost-efficient Distributed Load Frequency Control », *IET Energy Systems Integration*, pp. 327-343, Sept. 2021, doi: 10.1049/esi2.12030.

[16] N. Gholizadeh, and P. Musilek, « Distributed Learning Applications in Power Systems : A Review of Methods, Gaps, and Challenges », *Energies*, MDPI, vol. 14, no. 12, Jun. 2021, doi: 10.3390/en14123654.

[17] Y. Ko, and S-W Kim, « SHAT : A Novel Asynchronous Training Algorithm That Provides Fast Model Convergence in Distributed Deep Learning », *Applied Sciences*, MDPI, Dec. 2021, doi: 10.3390/app12010292.

[18] V. Veerasamy and al., « Blockchain-Based Decentralized Frequency Control of Microgrids Using Federated Learning Fractional-Order Recurrent Neural Network », *IEEE Transactions on Smart Grid*, 2023, doi: 10.1109/TSG.2023.3267503.

[19] J. Sun and al., « An integrated critic-actor neural network for reinforcement learning with application of DERs control in grid frequency regulation, » *International Journal of Electrical Power & Energy Systems*, vol. 111, pp. 286–299, 2019, doi: 10.1016/j.ijepes.2019.04.011.