UNIVERSIDADE DE SÃO PAULO – USP
ESCOLA DE ENGENHARIA DE SÃO CARLOS – EESC
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Pedro Inácio de Nascimento e Barbalho

Controle Supervisório de Microrredes Utilizando *Deep Reinforcement Learning*

São Carlos – SP
2021

Pedro Inácio de Nascimento e Barbalho

Controle Supervisório de Microrredes Utilizando *Deep Reinforcement Learning*

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia de São Carlos, como parte dos requisitos para obtenção do título de Mestre em Ciências.

Área de Concentração: Sistemas Elétricos de Potência

Orientador: Prof. Tit. Denis Vinicius Coury

São Carlos – SP

2021

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da
EESC/USP com os dados inseridos pelo(a) autor(a).

B228c Barbalho, Pedro Inacio de Nascimento e
 Controle Supervisorio de Microrredes Utilizando
 Deep Reinforcement Learning / Pedro Inacio de
 Nascimento e Barbalho; orientador Denis Vinicius Coury.
 São Carlos, 2021.

 Dissertação (Mestrado) - Programa de
 Pós-Graduação em Engenharia Elétrica e Área de
 Concentração em Sistemas Elétricos de Potência --
 Escola de Engenharia de São Carlos da Universidade de
 São Paulo, 2021.

 1. Aprendizado por reforço. 2. Controle
 Inteligente. 3. Controle Supervisorio. 4. Microrredes.
 I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **PEDRO INÁCIO DE NASCIMENTO E BARBALHO**.

Título da dissertação: "Controle supervisorío de microrredes utilizando Deep Reinforcement Learning".

Data da defesa: 17/02/2021.

Comissão Julgadora

Resultado

Prof. Titular **Denis Vinicius Coury**
(Orientador)
(Escola de Engenharia de São Carlos – EESC/USP)

APROVADO

Prof. Dr. **Renato Machado Monaro**
(Escola Politécnica/EP-USP)

APROVADO

Prof. Dr. **Fernando Pinhabel Marafão**
(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP-Sorocaba)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Ass. **João Bosco London Junior**

Presidente da Comissão de Pós-Graduação:
Prof. Titular **Murilo Araujo Romero**

AGRADECIMENTOS

Agradeço a todos que contribuíram diretamente e indiretamente no desenvolvimento deste trabalho. Todo o suporte e orientação do Professor Denis V. Coury foram essenciais para a conclusão desta Pesquisa. Agradeço também ao Vinícius Lacerda por todas as sugestões e pelo auxílio fornecido para esses dois anos. Agradeço, também, ao Professor Ricardo Augusto Souza Fernandes por todas as contribuições no grupo de pesquisa do Laboratório.

Ainda, incluindo o Professor Denis Coury, agradeço ao Professor José Carlos de Melo Vieira Junior e ao Professor Mario Oleskovicz, integrantes e coordenadores do Laboratório de Sistemas de Energia Elétrica (LSEE) do Departamento de Engenharia Elétrica e de Computação da Escola de Engenharia de São Carlos. Graças a eles, o nosso laboratório apresenta um excelente ambiente para o avanço da Pesquisa dos vários alunos orientados. Vale ser ressaltado que, além da excelente coordenação dos Professores, sou grato a todos os amigos e colegas do Laboratório que sempre se prontificam a contribuir de alguma forma.

Por fim, agradeço à Universidade de São Paulo, à Escola de Engenharia de São Carlos e ao Departamento de Engenharia Elétrica pela estrutura fornecida. Além disso, o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES) - Código de Financiamento 001.

RESUMO

BARBALHO, P. I. N.. **Controle Supervisório de Microrredes Utilizando *Deep Reinforcement Learning***. 71 p. Dissertação de Mestrado – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2021.

As microrredes surgiram devido à necessidade de se coordenar um novo modelo de geração em ascensão, a geração distribuída, com sistemas de armazenamento de energia e cargas locais. Esse tipo de rede é um sistema complexo e há diversas proposições de como controlá-lo de forma a permitir a sua operação conectada ao sistema de distribuição ou ilhada. Além disso, para um melhor desempenho do controle da microrrede, é preciso que este se adapte aos seus diferentes pontos de operação. O controle de microrredes pode ser dividido em três níveis hierárquicos que diferem em tempo de resposta, objetivos de controle e necessidade de comunicação. Por tratar-se de um tema recente, ainda há espaço para novas análises e contribuições. Neste sentido, o objetivo deste estudo foi modelar o controlador supervisor de uma microrrede ilhada, utilizando um algoritmo de *deep reinforcement learning* denominado de *deep deterministic policy gradient*. A microrrede modelada para o estudo, possui um gerador síncrono, dois sistemas de armazenamento de energia por baterias e um gerador fotovoltaico. O controlador proposto faz a leitura das tensões terminais dos recursos distribuídos de energia e da frequência da rede e atua nas potências ativas e reativas com o intuito de manter as variáveis de controle o mais próximo possível da referência. O desempenho do controlador proposto foi comparado a um esquema utilizando o método *droop* convencional, por meio da simulação de três cenários: i) desconexão e conexão programada de alimentado; ii) Aplicação de curto-circuito com posterior abertura de disjuntor para isolar a falha; e iii) Desconexão da maior carga do sistema. *A priori*, almejava-se uma melhora expressiva do controlador inteligente em comparação ao *droop*. Pelos resultados obtidos, não houve uma superioridade do controlador inteligente com relação ao *droop*, mas pode-se dizer que seu desempenho foi equivalente. Em alguns casos, o *droop* teve um desempenho mais adequado na minimização das oscilações de frequência e tensão, enquanto que, em outros, o comportamento da metodologia abordada foi mais conveniente.

Palavras-chave: Aprendizado por reforço. Controle inteligente. Controle supervisor. Microrredes.

ABSTRACT

BARBALHO, P. I. N.. **Supervisory Control of Microgrids Using Deep Reinforcement Learning**. 71 p. Master Dissertation – São Carlos School of Engineering, University of São Paulo, São Carlos, 2021.

The microgrids emerged due to the need to coordinate an ascending new generation model, the distributed generation, with energy storage systems and local loads. This type of grid is a complex system and several proposals have been made to control it and ensure that the microgrid operates both connected to the main grid and islanded. Besides that, the microgrid control needs to adapt to its different operating points to maintain acceptable performance. In order to obtain promising results, the microgrid control is usually divided into three hierarchical levels, that differ in terms of speed response, control objectives, and communication requirements. However, this subject is recent and there is still room for new contributions and analysis about microgrid control. Taking that into account, this research project aims to model an intelligent supervisory controller of an islanded microgrid using the Deep Deterministic Policy Gradient, which is a deep reinforcement learning algorithm. The modelled microgrid comprises a synchronous generator, two battery energy storage systems, and a photovoltaic generator. The proposed controller reads the energy resources terminal voltages and the grid frequency and varies the storage elements' power to keep these variables near their nominal values. The proposed controller performance was compared to the droop control, using the simulation of three scenarios: i) Disconnection and reconnection of a feeder; ii) A short-circuit event with subsequent breaker tripping to isolate it; and iii) Disconnection of the microgrid's biggest load. Initially, a noticeable improvement in the intelligent controller was expected compared to the droop. However, the results showed that the intelligent control performed similarly to the droop control. While the droop controller has minimised the voltage and frequency oscillations more conveniently in some scenarios, the opposite occurred considering the proposed methodology.

Keywords: Intelligent techniques. Microgrids. Reinforcement learning. Supervisory control.

LISTA DE ILUSTRAÇÕES

Figura 1 – Representação de uma microrrede genérica.	6
Figura 2 – Modelo de interação ambiente-agente no aprendizado por reforço.	19
Figura 3 – Microrrede modelada.	30
Figura 4 – Topologia do BESS.	32
Figura 5 – Topologia do VSC.	32
Figura 6 – Modelo do BESS simplificado.	34
Figura 7 – Malha de controle do modelo completo do BESS.	35
Figura 8 – Diagrama do gerador síncrono.	37
Figura 9 – Malha de controle do regulador de velocidade.	37
Figura 10 – Malha de controle do regulador de tensão.	38
Figura 11 – Representação do agente no Simulink.	41
Figura 12 – Microrrede modelada no Simulink/Matlab.	43
Figura 13 – Treinamento do agente com duas camadas escondidas na rede Ator e Crítico.	44
Figura 14 – Treinamento do agente com apenas uma camada escondida na rede Ator.	44
Figura 15 – Potência do PV.	45
Figura 16 – Frequência da rede para o caso 1.	46
Figura 17 – Tensão do GS para o caso 1.	47
Figura 18 – Tensão de B1 para o caso 1.	47
Figura 19 – Tensão de B2 para o caso 1.	47
Figura 20 – Potências ativas das baterias para o caso 1.	48
Figura 21 – Potências reativas das baterias para o caso 1.	49
Figura 22 – Representação do local do curto-circuito aplicado no caso 2.	50
Figura 23 – Frequência da rede para o caso 2.	50
Figura 24 – Tensão do GS para o caso 2.	51
Figura 25 – Tensão de B1 para o caso 2.	51
Figura 26 – Tensão de B2 para o caso 2.	51
Figura 27 – Potências ativas das baterias para o caso 2.	52
Figura 28 – Potências reativas das baterias para o caso 2.	53
Figura 29 – Localização da carga desconectada.	54
Figura 30 – Frequência da rede para o caso 3.	54
Figura 31 – Tensão do GS para o caso 3.	55
Figura 32 – Tensão de B1 para o caso 3.	55
Figura 33 – Tensão de B2 para o caso 3.	55

Figura 34 – Potências ativas das baterias para o caso 3. 56

Figura 35 – Potências reativas das baterias para o caso 3. 56

Figura 36 – EQM normalizado da tensão do GS para o caso 1. 57

Figura 37 – EQM normalizado da frequência da rede para o caso 1. 58

LISTA DE TABELAS

Tabela 1 – Trabalhos que implementaram sistemas de controle associados ao aprendizado por reforço.	27
Tabela 2 – Especificações técnicas dos elementos armazenadores.	31
Tabela 3 – Dados nominais do BESS.	36
Tabela 4 – Parâmetros do gerador síncrono.	36
Tabela 5 – Dados do transformador do GS	37
Tabela 6 – Parâmetros do gerador solar fotovoltaico.	39
Tabela 7 – Parâmetros do DDPG no Matlab.	40
Tabela 8 – Erro quadrático médio de cada variável controlada para cada caso estudado.	58

LISTA DE SIGLAS

AD	<i>Armazenamento Distribuído</i>
ACO	<i>Ant Colony Optimization</i>
BESS	<i>Battery Energy Storage system</i>
CIGRÉ	<i>Conseil International des Grands Réseaux Électriques</i>
CA	<i>Corrente Alternada</i>
CC	<i>Corrente Contínua</i>
DDPG	<i>Deep Deterministic Policy Gradient</i>
DQN	<i>Deep Q-Learning</i>
DRL	<i>Deep Reinforcement Learning</i>
DER	<i>Distributed Energy Resources</i>
GD	<i>Geração Distribuída</i>
GS	<i>Gerador Síncrono</i>
IGBT	<i>Insulated Gate Bipolar Transistor</i>
MPPT	<i>Maximum power point tracking</i>
MR	<i>Microrrede</i>
MDP	<i>Markov Decision Process</i>
PAC	<i>Ponto de Acoplamento Comum</i>
PO	<i>Perturb and Observe</i>
PMU	<i>Phasor Measurement Unit</i>
PSCAD	<i>Power System Computer Aided Design</i>
PSO	<i>Particle Swarm Optimization</i>
PV	<i>Painel Fotovoltaico</i>
PI	<i>Proporcional-Integrativo</i>
SMA	<i>Sistema Multiagente</i>
SWAPSC	<i>System Wide Adaptive Predictive Supervisory Control</i>
VERITAS	<i>Veneziana Energia Risorse Idriche Territorio Ambiente Servizi</i>

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Motivação e Objetivo da Pesquisa	1
1.2	Estrutura da Dissertação	2
1.3	Lista de Publicações	3
2	CONTROLE HIERÁRQUICO DE MICRORREDES	5
2.1	Desenvolvimento das Microrredes	5
2.2	Estrutura de Controle das Microrredes	8
2.3	Controle Supervisório de Microrredes	9
2.3.1	Controle Clássico	10
2.3.2	Controle Inteligente	13
2.4	Considerações Finais	16
3	DEEP REINFORCEMENT LEARNING	17
3.1	Aprendizado por Reforço	17
3.1.1	Principais componentes do aprendizado por reforço	19
3.1.2	Algoritmos de aprendizado por reforço	20
3.1.2.1	Métodos Tabulares	20
3.1.2.2	Métodos baseados em redes neurais artificiais	22
3.2	Estado da Arte da Ferramenta Aplicada a Controle de Microrredes	24
3.3	Considerações Finais	28
4	MODELAGEM DA MICRORREDE	29
4.1	Microrrede para Estudo	29
4.2	Modelagem dos Recursos Distribuídos de Energia	30
4.2.1	Sistema de Armazenamento de Energia à Baterias	31
4.2.2	Gerador Síncrono	36
4.2.3	Gerador Fotovoltaico	38
4.3	Modelagem do Controle Supervisório	39
4.4	Considerações Finais	42
5	OPERAÇÃO DA MICRORREDE COM O CONTROLE SUPERVISÓRIO INTELIGENTE	43
5.1	Caso 1 - Desconexão programada de alimentador	46

5.2	Caso 2 - Aplicação de um curto-circuito	49
5.3	Caso 3 - Desconexão de carga	53
5.4	Análise quantitativa de desempenho	57
5.5	Considerações finais	59
6	CONCLUSÃO	61
6.1	Trabalhos Futuros	62
	REFERÊNCIAS	63
	APÊNDICE A CÓDIGO DO DDPG NO MATLAB	69

INTRODUÇÃO

A Geração Distribuída (GD) é caracterizada pela instalação de fontes de energia próximas às cargas, podendo estar conectadas à rede de distribuição ou não. Esses pequenos centros geradores autônomos existem há décadas e surgiram com a necessidade de atender consumidores que se encontram distantes do sistema de fornecimento principal (OLIVARES *et al.*, 2014). Embora, os principais insumos utilizados por esse tipo de geração tenham sido, inicialmente, os combustíveis fósseis, nos últimos anos, nota-se a predominância do uso das fontes renováveis e alternativas de energia para tal finalidade (OLIVARES *et al.*, 2014).

Atualmente, a GD com o Armazenamento Distribuído (AD) compõem os recursos distribuídos de energia (do inglês, *Distributed Energy Resources* (DER)). A conexão desses recursos tem sido incentivada pela busca da redução da emissão de gases de efeito estufa, aumento da confiabilidade e melhoria da qualidade da energia entregue aos consumidores (HATZIARGYRIOU *et al.*, 2007). Contudo, o sistema de distribuição com elevada penetração de DERs torna-se mais complexo e sua coordenação a partir de um controle centralizado é dificultada. Por isso, a coordenação desses sistemas vem sendo amplamente estudada nos últimos anos (SRIVASTAVA; KUMAR; SCHULZ, 2012; OLIVARES *et al.*, 2014).

1.1 Motivação e Objetivo da Pesquisa

Nesse contexto, surge o conceito de Microrrede (MR), que representa um conjunto de cargas e DERs conectados a um Ponto de Acoplamento Comum (PAC) e que são coordenados com o objetivo de aumentar a confiabilidade do fornecimento de energia. Além disso, as MRs podem operar conectadas à rede de distribuição ou até mesmo ilhadas. Dessa forma, é possível reduzir a complexidade da rede principal já que a MR pode ser caracterizada como uma única carga controlável (HATZIARGYRIOU *et al.*, 2007; OLIVARES *et al.*, 2014). Para operar uma MR conectada à rede de distribuição ou ilhada, é comum dividir seu sistema de controle hierarquicamente em três níveis denominados de primário, secundário e terciário. Esses níveis

diferem em tempo de resposta, objetivos de controle e necessidade de comunicação (OLIVARES *et al.*, 2014).

Essa estrutura de controle faz com que a MR seja uma solução adequada para a integração de DERs. Dessa forma, o controle de MRs é tema de diversos estudos, visando que sua operação seja segura, confiável e economicamente viável. Esses requisitos devem ser atendidos tanto para a MR ilhada quanto conectada na rede de distribuição (OLIVARES *et al.*, 2014). Deve ser ressaltado que, na bibliografia revisada, há grande emprego de técnicas inteligentes, bem como sistemas multi-agentes na coordenação de DERs em MRs.

Um exemplo disso, foi o método de sintonia do controle de frequência de uma MR ilhada proposto por Bevrani *et al.* (2012). A MR em questão era composta por um gerador a diesel, uma célula combustível, sistema de armazenamento de energia baseado em baterias, painéis fotovoltaicos, aerogeradores e sistemas de armazenamento de energia baseados em volantes de inércia (*flywheel*). A partir da função de transferência da rede, um controlador Proporcional-Integrativo (PI) foi sintonizado. Em seguida, foram realizados testes com um sistema *fuzzy* que modificava os ganhos do controlador em tempo real para melhorar a resposta do sistema. Posteriormente, foi proposta a implementação da otimização por enxame de partículas para a especificação das funções de pertinência do sistema *fuzzy* que sintonizava os ganhos do controlador. Por fim, comparou-se o desempenho das metodologias abordadas. Neste trabalho, foi possível notar uma melhora no desempenho do sistema com a sintonia proposta, mesmo considerando variação de topologia.

Os controladores de uma MR devem ser robustos para que haja uma melhor integração dos DERs, ou seja, que se adaptem à mudança de topologia ou aos pontos de operação do sistema (BEVRANI *et al.*, 2012). Portanto, se faz necessário desenvolver controladores que sejam adaptativos. Dessa forma, há trabalhos que utilizam técnicas inteligentes para aumentar a robustez do controlador e incrementar uma característica adaptativa, como foi feito por Bisht *et al.* (2014).

Dessa maneira, o objetivo geral desta Dissertação de Mestrado é a implementação do controle supervisor de uma MR ilhada utilizando técnicas inteligentes. Embora já existam trabalhos relacionados, o controle supervisor de MRs ainda é um tema que carece de contribuições e novas análises. Além disso, é de suma importância a realização de pesquisas que avaliem o desempenho desses sistemas com a utilização de técnicas inteligentes, para que seja possível comparar o emprego dessas técnicas com as convencionalmente adotadas, oferecendo uma alternativa a elas e buscando superar suas limitações.

1.2 Estrutura da Dissertação

No Capítulo 2 foram apresentados o contexto no qual o desenvolvimento das MRs está inserido; exemplos de MRs reais que já foram desenvolvidas; a estrutura de controle de uma

MR e o conceito de controle supervisório; e uma breve revisão bibliográfica sobre técnicas tradicionais e inteligentes aplicadas ao controle supervisório das MRs.

No Capítulo 3 estão descritos os conceitos de aprendizado por reforço e do algoritmo utilizado nesta Pesquisa de Mestrado. Além disso, encontra-se uma revisão bibliográfica relacionada à aplicações de aprendizado por reforço no contexto de MRs e sistemas inteligentes.

No Capítulo 4 foram apresentados os passos tomados para a modelagem da MR e do controlador implementados neste trabalho.

No Capítulo 5 encontra-se uma discussão relacionada aos resultados obtidos na aplicação do algoritmo inteligente no controle supervisório de uma MR com diferentes tipos de fontes de energia.

Por fim, as conclusões da pesquisa são apresentadas no Capítulo 6.

1.3 Lista de Publicações

No Mestrado, foram produzidos artigos em congressos nacionais e internacionais, bem como um artigo de periódico.

Publicação em periódico:

- LACERDA, V.A.; BARBALHO, P.I.N.; MONARO, R.M.; COURY D.V. Signal processing techniques for synchrophasors considering short-circuit signals: a comparative study. In: **2020 IET Generation, Transmission Distribution 14 (19)**.

Publicações em Congressos Internacionais:

- BARBALHO, P. I.; MENEZES, T. S.; COURY, D. V.; VIEIRA JUNIOR, J. C. M.; OLESKOVICZ, M. (2019). Analysis of Distributed Energy Resources Impact on Islanding Detection Techniques. **2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)**.
- MENEZES, T.S.; BARBALHO, P.I.N.; COURY D.V.; FERNANDES, R.A.S. Islanding Detection for Distributed Generators Based on Artificial Neural Network and S-transform. In: **2019 IEEE PES Innovative Smart Grid Technologies Conference-Latin America (ISGT Latin America)**. IEEE, 2019. p. 1-5.
- MOMESSO, A.E.C.; BARRA, P.H.A.; BARBALHO, P.I.N.; ASADA, E.N.; COURY D.V.; VIEIRA JUNIOR, J.C.M.; OLESKOVICZ, M.; BICZKOWSKI, M.; Impact Analysis of a Transportable BESS on the Short-Circuit Behavior of a Distribution System. **2020 IEEE PES General Meeting**. 2020. Montreal - CA.

Publicações em Congressos Nacionais:

- MOMESSO, ANTONIO EDUARDO CEOLIN et al. Análise transitória da conexão de um sistema de armazenamento de energia transportável em um sistema de distribuição. **XXV SNPTEE Seminário Nacional de Produção E Transmissão de Energia Elétrica**. 2019. Belo Horizonte-MG.
- SANTOS, G. P.; SANTOS, G. G.; MENEZES, T. S.; VIEIRA JUNIOR, J. C. M.; BARBALHO, P. I. (2019). A New Approach for Fault Classification Based on S-Transform and Artificial Neural Network. **XIII CBQEE - Conferência Brasileira sobre Qualidade de Energia Elétrica**.

CONTROLE HIERÁRQUICO DE MICRORREDES

Conforme foi discutido, o desenvolvimento das MRs tem sido muito importante para a integração dos recursos distribuídos de energia. Sendo assim, este Capítulo apresenta as principais características desses sistemas e as implementações reais que já foram feitas. Ainda, como o escopo desta pesquisa está relacionado ao controle das MRs, foi apresentada a principal estrutura de controle utilizada e como ela é dividida.

2.1 Desenvolvimento das Microrredes

Uma MR é uma região composta por DERs e cargas locais. Nessa região, a coordenação desses elementos é realizada por um controle próprio, de forma que seja possível fornecer energia às cargas conectadas a ela por tempo considerável mesmo após seu ilhamento, ou seja, desconexão da rede de distribuição (LU; WANG; GUO, 2016). Além disso, a fim de melhorar a qualidade e confiabilidade da energia produzida, as MRs possuem sistema de armazenamento que, juntamente ao controle, podem tornar o PAC, do ponto de vista da rede principal, uma carga constante (HATZIARGYRIOU *et al.*, 2007). Ademais, devido à sua estrutura, a MR pode ser conectada a outras MRs vizinhas e assim realizar intercâmbio de energia entre as áreas. A Figura 1 apresenta o esquema de uma MR genérica.

Figura 1 – Representação de uma microrrede genérica.

Fonte – Adaptado de Katiraei *et al.* (2008).

Tendo em vista a composição de uma MR e sua estrutura de controle, o emprego desse sistema se justifica pelo cenário mundial. Nesse cenário, a penetração da geração distribuída tem aumentado nos últimos anos, seja em países desenvolvidos ou emergentes (ACKERMANN; ANDERSSON; SÖDER, 2001). As principais causas desse crescimento e do desenvolvimento de outros DERs tem sido a busca por maior confiabilidade e eficiência do fornecimento de eletricidade, a busca pela redução da emissão de gases responsáveis pelo efeito estufa e o desenvolvimento de formas eficientes de geração de calor combinadas à de energia elétrica (HATZIARGYRIOU *et al.*, 2007; ACKERMANN; ANDERSSON; SÖDER, 2001).

Entretanto, a conexão indiscriminada desses centros geradores pode causar instabilidade na rede principal de distribuição. Isso se deve ao número de fontes renováveis, cujo uso para este tipo de aplicação vem apresentando uma tendência de crescimento (HATZIARGYRIOU *et al.*, 2007). Esses geradores são, em maioria, dependentes das condições climáticas instantâneas, podendo gerar mais ou menos energia em função da disponibilidade da sua fonte. Por exemplo, as turbinas eólicas e painéis fotovoltaicos nem sempre conseguem fornecer sua potência máxima instalada ao consumidor conectado a eles. Neste caso, a geração dependerá da velocidade do vento e do nível de irradiação solar na área (KANEKO *et al.*, 2011), (TIWARI; MISHRA;

SOLANKI, 2011). Da mesma maneira, sem o controle devido, esses geradores podem acabar gerando mais ou menos potência do que está sendo consumido. Assim, ou o excedente de energia é enviado à rede comum ou, em situações críticas como a de ilhamento, a MR pode perder a estabilidade de tensão devido ao desbalanço de carga. Dessa forma, o controle de fluxo de potência do sistema de distribuição numa situação de elevada penetração de geração distribuída torna-se mais complexo caso as instalações desse tipo de geração não possuam sistemas de controle próprios (HATZIARGYRIOU *et al.*, 2007).

Há três tipos principais de MRs, que são a Corrente Alternada (CA), a Corrente Contínua (CC) e as Híbridas. As redes CA são as mais utilizadas e há um vasto número de estudos acerca dela (ALFERGANI *et al.*, 2018). As redes CC possuem características que compensam limitações da CA, a saber: a falta da necessidade de sincronização com a rede; não há controle de frequência e potência reativa; e há menores perdas associadas à transformação de sinais alternados para contínuos e vice-versa (SAHOO; SINHA; KISHORE, 2018). Contudo, esse tipo de rede necessita de construção de cabos CC privados, sua proteção é mais desafiadora e a tensão é controlada apenas por potência ativa, enquanto em redes CA, pode ser controlada pela potência reativa (SAHOO; SINHA; KISHORE, 2018).

Atualmente, é possível apontar diversas implementações reais de MRs. Em Veneza, como consequência da parceria da *Università degli Studi di Padova* e a empresa *Veneziana Energia Risorse Idriche Territorio Ambiente Servizi* (VERITAS), desenvolveu-se um projeto de uma MR de baixa tensão em que seriam instalados dois geradores fotovoltaicos e um gerador biodiesel (GUARNIERI *et al.*, 2018). Além disso, era prevista a conexão de 5 sistemas armazenadores de energia, baseados em cinco diferentes tecnologias, a saber: Li-ion, Pb-A, Na-Ni-Cl, célula combustível e uma bateria redox de vanádio. O projeto previa um controle hierárquico, sendo que o nível secundário é descentralizado e o nível primário composto por controladores com técnica *droop*.

Na União Europeia, um dos principais projetos de desenvolvimento dedicados unicamente a MRs foi o intitulado “Microrredes: Integração em Larga Escala de Micro-Geração em Redes de Baixa Tensão”. Nele, houve investimento de cerca de € 4,5 milhões e o envolvimento de instituições de ensino, governamentais e empresariais de 7 países europeus (HATZIARGYRIOU *et al.*, 2007). Esse programa tinha como objetivo desenvolver métodos de controle, a fim de otimizar a operação de MRs, estudar o comportamento da MR durante ilhamento e faltas, desenvolver metodologias de proteção que garantissem ao sistema maior proteção e melhor resposta a curtos-circuitos e perdas de conexão com a rede principal. Ainda, foi estudada a viabilidade econômica da implementação de MRs ao sistema de distribuição, a fim de aumentar a penetração de sistemas renováveis de GD. Para isso, o projeto de MR foi testado e demonstrado em um laboratório de pesquisa (HATZIARGYRIOU *et al.*, 2007). No final, o programa obteve resultados positivos, em que surgiram novas estratégias de operação de MR tanto conectadas ao fornecedor comum quanto em situações de ilhamento. Também foram desenvolvidos métodos de

proteção e resposta a situações de falta e métodos capazes de estimar a confiabilidade de uma MR (HATZIARGYRIOU *et al.*, 2007).

Outros exemplos de implementação de MRs podem ser encontrados no Japão. Um deles é o Projeto Aomori, em Hachinohe, cujo objetivo foi avaliar a qualidade da energia, o custo de uma MR e a redução de emissão de gases responsáveis pelo efeito estufa (HATZIARGYRIOU *et al.*, 2007). Essa MR em questão foi composta por fontes renováveis de energia e responsável por alimentar sete cargas na cidade de Hachinohe. Implementou-se um sistema de controle capaz de coordenar a energia gerada pelas fontes renováveis com a energia absorvida ou liberada por um centro de armazenamento. Dessa maneira, as cargas utilizaram menos energia proveniente da rede principal e a emissão de gás carbônico foi mitigada (HATZIARGYRIOU *et al.*, 2007).

2.2 Estrutura de Controle das Microrredes

Para que uma MR cumpra sua função, atualmente, a estrutura de controle hierárquica é a mais aceita e utilizada. Essa estrutura é dividida em níveis primário, secundário e terciário. O nível primário é o que possui a resposta mais rápida às mudanças e se baseia em leituras locais e mínima comunicação. Dessa forma, pode ser necessário implementar nesse nível uma detecção de ilhamento e controles de potência ativa e reativa, a fim de regular a tensão e frequência do sistema dependendo do comando enviado pelos níveis superiores de gerenciamento da MR. Um exemplo de técnica utilizada nesse nível é o controle *droop* que auxilia na manutenção dos níveis de frequência e tensão da rede. Esse tipo de controle pode provocar desvios de frequência e tensão durante o modo ilhado e devem ser compensados pelo nível secundário (OLIVARES *et al.*, 2014; BIDRAM; DAVOUDI, 2012).

O controle secundário é o responsável pela operação segura, confiável e por garantir que a MR opere segundo limites econômicos estabelecidos, até mesmo quando essa rede estiver ilhada. Para isso, o nível secundário altera os valores de referência rastreados pelo primário, devendo possuir então uma resposta mais lenta. Além disso, deve ser garantida a transição adequada da MR do modo conectado para o ilhado e vice-versa. Essencialmente, as estratégias adotadas no nível secundário podem ser divididas em duas: controle centralizado e descentralizado (OLIVARES *et al.*, 2014).

O controle centralizado exige muitas informações da MR, tais como: parâmetros das linhas; valores estimados de irradiação solar e velocidade dos ventos; cargas locais; custos; e modos de operação. Todas essas informações devem ser enviadas a um ponto central onde será realizada a tomada de decisão (KATIRAEI *et al.*, 2008). Nessa estratégia, os cálculos podem ser realizados de forma *online* ou *offline*. O emprego desse tipo de controle é interessante quando trata-se de pequenas MRs, permitindo: um dispêndio menor com os equipamentos de processamento de dados e com a estrutura de comunicação; que todos os componentes da MR possuam os mesmos objetivos, ou seja, o proprietário das cargas e dos geradores deseja otimizar

a produção e consumo daquela região; que a MR seja mais segura a ataques cibernéticos; que a configuração da MR seja praticamente fixa, fazendo com que não haja nenhuma alteração do seu perfil de tipo de carga ou de geradores (MENG *et al.*, 2016). Contudo, quando trata-se de sistemas inteligentes, é possível que algumas desvantagens do controle centralizado sejam mitigadas e sua adoção seja interessante para redes onde há mudança topológica frequente. Essa situação será discutida mais adiante.

Adicionalmente, o controlador centralizado necessita de dados provenientes das unidades geradoras e consumidoras da MR e da rede de distribuição (MENG *et al.*, 2016). O atraso e qualidade dos sinais enviados/recebidos pelo controlador centralizado influenciam na estabilidade da MR. Assim, é importante o emprego de equipamentos que sejam capazes de realizar a aquisição adequada desses dados e que possam se comunicar de forma sincronizada. Nesse sentido, pode-se destacar, para aplicações que exigem comunicação sincronizada e precisa, a Unidade de Medição Fasorial (do inglês, *Phasor Measurement Unit* (PMU)). Atualmente, diversos trabalhos propõem soluções que levam em consideração a utilização desse equipamento e que demonstram os benefícios de uso (ZANJANI; MAZLUMI; KAMWA, 2018; ELBANA *et al.*, 2019).

Considerando o controle descentralizado, esse almeja garantir a maior autonomia possível dos DERs. Dessa forma, as decisões são feitas localmente. Adicionalmente, é comum o emprego de sistemas multi-agentes para a coordenação e troca de informações entre as fontes distribuídas, mas ainda mantendo a autonomia de cada controle local (OLIVARES *et al.*, 2014). O controle descentralizado pode ser empregado quando: (i) a MR for de maior porte ou no caso de existirem cargas e geradores muito dispersos, o que dificultaria um sistema de aquisição de dados centralizado; (ii) os componentes da MR forem de consumidores com necessidades operacionais diferentes, o que exigiria uma tomada de decisão local; (iii) houver a possibilidade de mudanças da configuração da rede.

O nível terciário de controle da MR é o que possui a resposta mais lenta do sistema e especifica valores de referência para os outros níveis, dependendo da necessidade da rede de distribuição. O controle terciário coordena múltiplas MRs e pode ser considerado como parte da rede principal. Dessa forma, quando a MR opera ilhada, não há ação do nível terciário (OLIVARES *et al.*, 2014).

2.3 Controle Supervisório de Microrredes

Ao estabelecer os três níveis de controle de uma MR, é possível caracterizar o nível mais alto de comando que é usualmente denominado de controle supervisório. O supervisório é o agente responsável pelo gerenciamento do sistema, visto que recebe as funções do nível terciário relacionadas à operação da MR e também as funções do nível secundário.

As funções do controle supervisório são fornecer regulação de potência, serviços anci-

lares, participação no mercado de energia e otimização da operação do sistema, dando-lhe um caráter inteligente (MENG *et al.*, 2016). Além disso, o supervisor deve coordenar as fontes geradoras e cargas, atender à demanda de energia, adaptar-se ao crescente número de conexões de recursos distribuídos de energia e garantir os níveis de qualidade de energia da MR (MENG *et al.*, 2016). Portanto, o gerenciamento de uma MR trata-se de um problema multi-objetivo e, por isso, a modelagem de um sistema supervisor robusto e confiável é desafiador.

Tendo em vista que os problemas de qualidade de energia são uma grande preocupação em sistemas elétricos, sendo um agravante a alta penetração de recursos distribuídos interfaceados por eletrônica de potência, uma das principais funções do supervisor é a manutenção da tensão e frequência da MR dentro de uma faixa aceitável. Sendo assim, este trabalho se concentrou em revisar os estudos que propuseram controladores capazes de regular tensão e frequência da MR, utilizando técnicas de controle convencionais e inteligentes.

Cabe mencionar que, na revisão de estudos focados em técnicas inteligentes, notou-se que as soluções de controladores foram capazes de resolver problemas multi-objetivos e não se limitaram a regular tensão e frequência, mas também fizeram a previsão de geração de fontes intermitentes de energia e gerenciaram a carga e descarga dos sistemas de armazenamento de energia.

2.3.1 Controle Clássico

Rey *et al.* (2017) propuseram um controle chaveado para o nível secundário da MR ilhada, sem a necessidade de comunicação. A proposição se baseia na técnica *droop* empregada nos controles locais. O controle *droop* pode ser descrito pelas Equações (2.1) e (2.2) (REY *et al.*, 2017):

$$\omega = \omega_0 - mP, \quad (2.1)$$

$$V = V_0 - nQ, \quad (2.2)$$

sendo que, P e Q são as potências ativas e reativas, enquanto ω e V são a frequência angular e tensão de referência do sistema a serem alteradas pelo controle *droop*; e ω_0 e V_0 são a frequência angular e tensão nominais do sistema. As constantes m e n são definidas pelas especificações do sistema e caracterizam as funções de reta que relacionam as potências e as variáveis de controle. Assim, o controle secundário atua de forma a corrigir os erros de frequência angular e tensão que surgem devido ao *droop*. Define-se portanto uma variável $\delta(t)$ que atua para corrigir esse erro, como pode ser visto na Equação (2.3):

$$\omega = \omega_0 - mP + \delta. \quad (2.3)$$

O sistema de controle secundário pode assumir duas dinâmicas após a detecção de um evento que provoque o desvio de frequência angular. A alternância entre essas duas dinâmicas

depende da variável $k(t)$ que varia com o tempo, segundo uma função pré-estabelecida. A princípio, após a detecção, a variável $k(t)$ é maior do que zero e $\delta(t)$ pode ser descrita no domínio da frequência ($\Delta(s)$) pela Equação (2.4):

$$\Delta(s) = \frac{k_i}{s + k \cdot k_i}(\omega_0 - \omega). \quad (2.4)$$

Após um certo instante de tempo definido na modelagem, δ assume um valor constante e $k(t) = 0$. Observa-se que δ , na Equação (2.4), possui o comportamento de um filtro passa-baixa. A variação de k no tempo se dá pelo tipo de resposta da função de transferência para diferentes valores da variável. Para valores altos de $k(t)$, o controle *droop* apresentará alta largura de banda e uma resposta inexata de regime permanente. Quando o valor de $k(t)$ é baixo, a resposta da malha torna-se mais lenta, porém mais precisa. Dessa forma, para efeitos de sintonia, após a detecção do evento, $k(t)$ assume um valor alto e começa a decair linearmente até atingir zero (REY *et al.*, 2017). Ademais, k_i é um parâmetro constante calculado de forma a alocar os polos do sistema. O desempenho desse controle secundário proposto por Rey *et al.* (2017) foi testado experimentalmente e apresentou boa resposta transitória, robustez e permitiu a restauração da frequência para seu valor nominal.

O controle a nível secundário da MR utilizando abordagens tradicionais, tais como a obtenção do modelo matemático do sistema e o posicionamento de polos e zeros existe em abundância na literatura. Simpson-Porco *et al.* (2015) propuseram um método semelhante ao de Rey *et al.* (2017). Entretanto, o termo que é incrementado na Equação (2.1) possui contribuições de todas as leituras locais das outras GDs da MR. Os autores argumentam que a comunicação entre as GDs pode trazer um melhor desempenho para o sistema de controle secundário (SIMPSON-PORCO *et al.*, 2015).

Uma estrutura de controle hierárquica aprimorada foi proposta por Han *et al.* (2017) para a coordenação de conversores do tipo *Voltage Source Converter* (VSC) de uma MR ilhada. Essa estrutura hierárquica foi dividida em níveis primário e secundário, sendo que cada nível cumpre uma função e o primeiro está subordinado ao segundo. No nível primário, implementou-se o controle *droop* com o intuito de variar os valores de referência das potências ativa e reativa dos conversores dependendo dos desvios de frequência e tensão que surgem na MR, sem a necessidade de comunicação entre as fontes. O diferencial apresentado nesse nível hierárquico está na utilização de malhas de controle de impedância virtual com o objetivo de aumentar o amortecimento do sistema ao tornar as malhas de controle PQ dos conversores mais desacopladas sem acréscimo de perdas elétricas. Por fim, embora o controle *droop* seja capaz de mitigar os desvios de frequência e tensão presentes no sistema, é preciso que seja implementado um nível de controle capaz de garantir erro de regime igual a zero. Dessa forma, modelou-se o nível secundário de controle utilizando controladores PI, com a intenção de compensar os sinais de frequência e tensão gerados pelo controle *droop*. Essa estrutura hierárquica proposta foi capaz de garantir que a frequência e as tensões dos terminais dos conversores se mantivessem em seus

níveis nominais para regime permanente e, também, garantir uma resposta transitória amortecida para variações de cargas e desconexão de um dos conversores do sistema.

Mohammadi, Vanashi e Feliachi (2018) modelou em espaço de estados de uma MR ilhada composta por conversores CC-CA dos tipos fonte de tensão e fonte de corrente. Para garantir que a frequência do sistema se mantivesse no valor nominal, implementou-se um controle secundário que foi responsável por alterar a frequência de referência dos conversores, dos tipos fonte de tensão e fonte de corrente, que não dependiam de sistemas de armazenamento e que possuíam controle local do tipo *droop*. Dessa forma, a estabilidade do sistema não depende do estado de carga desses elementos armazenadores. Para verificar a efetividade do modelo da MR e do controle sintonizado, utilizou-se um sistema de 11 barras implementado no MATLAB. Os controladores locais e o do nível secundário foram capazes de manter os erros de regime da frequência do sistema iguais a zero.

Liu *et al.* (2017) propuseram um controle secundário distribuído de uma MR utilizando um Sistema Multiagente (SMA) e um conjunto de inversores em modo de controle PQ e *droop*. Devido à capacidade de comunicação do SMA, os geradores distribuídos foram aglomerados em *clusters*. Dessa forma, foi possível reduzir os erros de regime de frequência e tensão da MR provocados pelo controle *droop*. Ademais, o método de controle facilitou a integração de outras MRs.

Guo *et al.* (2015) propuseram o controle de frequência e tensão de uma MR de maneira descentralizada. Na estrutura proposta, é essencial a comunicação entre os controladores locais para que o objetivo global seja atingido por meio de um consenso. Foi apresentado um estudo de caso de uma MR composta por fontes de energia baseadas em inversores, em que foi possível validar o desempenho do método de controle proposto. Chen e Feng (2015) também propuseram uma estrutura de controle de uma MR de forma descentralizada. Foram projetados o controle primário, o secundário e um controlador de despacho ótimo de potência. O controle secundário realiza uma ação no nível primário do sistema, que é baseado no método *droop*, e é capaz de restaurar os níveis de tensão e frequência. Não se limitando a esses trabalhos, outros, como Pashajavid, Ghosh e Zare (2018), Li *et al.* (2015) e Rezaei e Kalantar (2015) apresentaram soluções de controle em nível secundário para a regulação de tensão e/ou frequência de uma MR, mantendo uma estrutura hierárquica de gerenciamento.

Nota-se que há uma tendência de implementar estratégias convencionais de controle em nível secundário para MRs, os quais se baseiam na aplicação de uma perturbação no controle *droop* dos conversores para garantir erro de frequência e de tensão iguais a zero em regime permanente. Contudo, é possível encontrar ainda trabalhos como Nasirian *et al.* (2016) e Mohiuddin e Qi (2019) que propuseram um controlador capaz de realizar as funções dos níveis primário e secundário sem a utilização do método *droop*.

Além disso, nas estruturas de controle clássico, são empregados controladores PI sintonizados a partir de um modelo linear da planta que pretende-se controlar. Destarte, o principal

inconveniente dessa estratégia é a dependência do desempenho do sistema do ponto de operação para o qual os controladores foram modelados. Para superar isso, há extensa utilização de técnicas de controle não lineares que se mostraram eficazes para realizar a tarefa de gerenciamento da MR a nível secundário (DING; HAN; ZHANG, 2019), (GU *et al.*, 2017). Adicionalmente, podem ser implementados sistemas capazes de adicionar uma característica adaptativa ao nível secundário da MR. Nesse sentido, na sequência, serão descritos alguns trabalhos que buscaram adicionar essa característica utilizando técnicas inteligentes e que reforçam a importância da presente pesquisa.

2.3.2 Controle Inteligente

A modelagem de inteligência biológica a partir de algoritmos computacionais, os chamados sistemas inteligentes, tem obtido sucesso para soluções de problemas cada vez mais complexos (ENGELBRECHT, 2007). Esses sistemas podem ser baseados em redes neurais artificiais, computação evolutiva, inteligência de enxames, sistemas *fuzzy*, entre outros (ENGELBRECHT, 2007). No contexto de controle de MR, é possível encontrar trabalhos que implementaram esses sistemas inteligentes. Alguns exemplos são descritos a seguir.

O controle inteligente de um sistema de geração composto por painéis fotovoltaicos, uma célula combustível e um sistema de armazenamento de energia baseado em baterias (do inglês, *Battery Energy Storage system* (BESS)) foi proposto por Chettibi e Mellit (2018). Os painéis atuaram como a fonte principal de energia e a célula combustível e o BESS atuaram como auxiliares. Neste trabalho, foram apresentados três controladores: (i) o controlador responsável por rastrear o ponto de máxima geração do painel fotovoltaico; (ii) o controlador das células combustíveis e do BESS; e (iii) o controlador de fluxo de potência entre o sistema de geração, as cargas e a rede de distribuição, denominado sistema de gerenciamento de energia. Dentre esses controladores, vale destacar o segundo e o terceiro. Nesses dois controladores, que eram do tipo PI, seus ganhos foram alterados dinamicamente por um sistema *neuro-fuzzy*. Com isso, mesmo em casos de afundamentos de tensão, de variação de carga, irradiação solar e temperatura, o controlador foi capaz de manter a MR em operação (CHETTIBI; MELLIT, 2018).

A modelagem do controle de uma MR com múltiplas fontes renováveis de energia, utilizando métodos estocásticos dinâmicos é proposto por Han *et al.* (2015). Trata-se de um Controle Supervisório Preditivo Adaptativo do Sistema (do inglês, *System Wide Adaptive Predictive Supervisory Control* (SWAPSC)) e sua função foi suavizar a potência de saída dos painéis fotovoltaicos, aerogeradores e banco de baterias, além de garantir a tensão do barramento em seu valor nominal e reduzir as perdas da MR. O SWAPSC é composto por três camadas. A primeira utiliza um modelo preditivo de controle e é responsável por garantir a estabilidade de cada módulo de geração para os diferentes pontos de operação. A segunda é o controle local que é dividido em modos de operação que dependem dos estados de geração dos aerogeradores e painéis fotovoltaicos e do estado de carga do banco de baterias. A terceira camada utiliza

métodos heurísticos cujo objetivo é otimizar o desempenho do controle de tensão do sistema e suavizar as potências dos painéis e aerogeradores. Nessa última camada, usou-se redes neurais artificiais para que fosse obtida uma solução ótima. Esse esquema de controle proposto por Han *et al.* (2015) demonstrou melhor desempenho se comparado àqueles que possuem apenas um objetivo a ser otimizado.

Paridari *et al.* (2013) propuseram o controle de tensão de uma MR sob condições de cargas desequilibradas e não lineares. Para tal, foi utilizado um controlador ressonante cujos parâmetros foram determinados pelo método de otimização por enxame de partículas (do inglês, *Particle Swarm Optimization* (PSO)). Foi utilizado um modelo simples de MR que consistia numa fonte baseada em inversor com um filtro LC nos seus terminais. Além disso, havia uma carga não linear desequilibrada conectada nos terminais de saída do filtro. Nesse caso, de posse do modelo matemático do sistema, o método de otimização tinha como objetivo encontrar os melhores valores de ganhos para reduzir o erro de regime da tensão no barramento, segundo as restrições impostas. Essas restrições foram margem de ganho, margem de fase e largura de banda. Como resultado, as cargas não lineares foram alimentadas com tensões balanceadas e senoidais sem que houvesse perda de estabilidade e considerável distorção harmônica dos sinais de tensão. Nota-se, portanto, que os ganhos encontrados pelo PSO atenderam aos requisitos impostos, garantindo a estabilidade da MR para a condição projetada. Contudo, deve ser destacado que, caso a topologia da MR se altere ou sejam adicionados outros tipos de cargas, não há garantia que o controlador apresentará o mesmo desempenho.

Um sistema de gerenciamento de energia de uma MR foi proposto por Colson, Nehrir e Pourmousavi (2010), utilizando otimização por colônia de formigas (do inglês, *Ant Colony Optimization* (ACO)). Além disso, esse trabalho faz uma descrição breve com relação às teorias de métodos de otimização como o ACO e o PSO. Também, foram destacadas suas vantagens, como a fácil implementação e o bom desempenho em obter uma solução ótima em situações onde se deseja atingir múltiplos objetivos. O sistema proposto almejava reduzir a emissão de gás carbônico e minimizar o custo de operação de uma MR, utilizando as características das fontes de geração, disponibilidade de recursos e demanda das cargas. As informações eram atualizadas a cada 30 segundos e o sistema de gerenciamento precisou operar dentro desse período. O diferencial deste trabalho é o emprego da ACO em tempo real. Com os testes realizados, observou-se que há um limite onde a precisão do método aumenta pouco com o aumento do custo computacional. Dessa forma, utilizando um menor número de colônias, mantendo então a precisão e reduzindo o custo computacional, foi possível encontrar soluções otimizadas dentro da janela de atualização de dados (30 segundos). Entretanto, mesmo que tenha sido possível atingir os objetivos da otimização dentro dessa janela temporal, deixou-se em evidência que, para operações em tempo real de gerenciamento de energia, faz-se necessário obter soluções que apresentem respostas mais rápidas e robustas (COLSON; NEHRIR; POURMOUSAVI, 2010).

O uso de sistemas inteligentes para a sintonia de ganhos dos controladores tradicionais

pode ser encontrado também em outros trabalhos. Habibi, Shafiee e Bevrani (2019) propuseram o emprego de redes neurais artificiais para o ajuste dos coeficientes do controle *droop* de uma MR, a fim de controlar a frequência do sistema durante um ilhamento. Safari, Babaei e Farrokhifar (2018) propuseram o controle de frequência de uma MR composta por veículos elétricos, um aerogerador e um gerador a diesel. Na solução, uma rede neural treinada por PSO é capaz de ajustar os ganhos de um controlador PID para que o balanço de geração e consumo de energia da MR seja mantido.

Dentre os tipos de sistemas inteligentes, as redes neurais artificiais, muito empregadas para classificação de padrões, tem sido cada vez mais utilizadas para auxiliar controladores ou até mesmo atuarem como um. Isso foi feito não apenas por Safari, Babaei e Farrokhifar (2018), mas também por Duan *et al.* (2020), onde desenvolveu-se um controlador baseado em *Deep Reinforcement Learning* (DRL) para a regulação da tensão de sistemas elétricos. Para isso, foram projetados dois tipos de redes neurais, a saber: *Deep Q-Learning* (DQN) e *Deep Deterministic Policy Gradient* (DDPG). O controlador resultante varia a tensão no terminal dos geradores a fim de manter as tensões das barras dentro da faixa de 0.95 p.u. a 1.05 p.u. Tanto o controlador com a DQN quanto o controlador com a DDPG conseguiram garantir os níveis de tensão do sistema. Ainda, foi possível notar que o último tipo de rede possui maior velocidade de convergência. Embora Duan *et al.* (2020) não tenham trabalhado especificamente com MRs, é possível avaliar a potencialidade da ferramenta de controle empregada para a regulação de frequência de uma rede elétrica.

O aprendizado de máquina baseado em redes neurais e empregado por Duan *et al.* (2020) tem apresentado bom desempenho para o controle de sistemas. O *Deep Learning* e o *Reinforcement Learning* têm sido utilizados em diversos estudos. A exemplo disso, Kuznetsova *et al.* (2013) propuseram o gerenciamento de energia da MR via algoritmo baseado em aprendizado por reforço que estima a geração eólica e, assim, decide o horário ótimo para carregar um banco de baterias. Adicionalmente, Jafari *et al.* (2019) propuseram o emprego de um controlador secundário inteligente, treinado a partir de uma metodologia de *reinforcement learning*. Tal controlador é capaz de manter o sistema operando de forma adequada mesmo na presença de ruído e atraso nas leituras dos sinais e durante distúrbios eventuais no sistema.

No cenário atual, a penetração da geração distribuída cresce exponencialmente e a implementação de MRs é primordial para a coordenação dessas novas formas de geração com os sistemas de armazenamento distribuídos e as cargas locais. Dessa forma, é de grande importância o desenvolvimento de estudos que propõem maneiras de se controlar esses sistemas. Analisando a literatura, nota-se que as soluções dependentes da linearização do sistema em um ponto de operação perdem desempenho quando a rede é sujeita a condições adversas. Essas soluções podem ser, por exemplo, controladores PI com ganhos fixos e que foram calculados a partir da função de transferência da MR com uma configuração fixa. Conforme comprovado em outros estudos, as técnicas inteligentes podem agregar uma característica adaptativa aos controladores,

permitindo que não haja perda de desempenho quando a MR é submetida a condições de operação adversas ou à mudança de topologia.

2.4 Considerações Finais

Tendo em vista a potencialidade do emprego de MRs no sistema de distribuição para a integração de DERs, desenvolver um adequado gerenciamento das MRs é crucial para sua operacionalização e consolidação. Assim, sendo uma MR um sistema multi-objetivo, não linear e de controle complexo, o incremento de inteligência biológica a partir de algoritmos computacionais no controlador supervisor apresenta-se como uma solução promissora, como os estudos têm demonstrado. Dessa forma, a relevância do escopo desta pesquisa é reforçado. Nos Capítulos seguintes, estão descritos os modelos de recursos distribuídos e rede elétrica, as ferramentas computacionais e a parametrização dos controladores utilizados. Ressalta-se que, embora a comunicação e envio de dados sejam importantes para o controle supervisor com o auxílio de PMUs, tais conceitos não serão contemplados neste trabalho, uma vez que o foco é a análise da contribuição do sistema inteligente.

No Capítulo 3 foi realizado o desenvolvimento teórico da técnica inteligente adotada neste trabalho, que é baseada no *Deep Reinforcement Learning*. Ademais, foi realizada uma revisão bibliográfica com o intuito de encontrar trabalhos que já abordaram técnicas com o mesmo princípio e também de encontrar outras aplicações da técnica em outras áreas.

DEEP REINFORCEMENT LEARNING

Como discutido anteriormente, as MRs são um importante agente na transformação do sistema elétrico. Sendo assim, o estudo de sistemas de controle que viabilizem a implementação desses modelos de geração são importantes. Nesse sentido, este trabalho apresentou alguns dos diversos estudos que procuraram propor estratégias de controle. A partir do que foi exposto, discutiu-se que as estratégias tradicionais de controle podem não ser suficientes para atender a demanda das MRs. Por esse motivo, existem diversos trabalhos que vem tentando utilizar técnicas inteligentes. Dentro dessas técnicas, apresenta-se o aprendizado por reforço e o crescente uso do *Deep Reinforcement Learning*.

Neste Capítulo, as teorias do aprendizado por reforço e do *Deep Reinforcement Learning* são apresentadas. Ademais, são apontadas as vantagens e desvantagens de técnicas baseadas nesses modelos e, posteriormente, há uma revisão bibliográfica mais focada nos estudos relacionados a essa aplicação. Ao fim deste Capítulo, ficará claro como o *Deep Reinforcement Learning* é útil para atingir objetivos de controle complexos e, conseqüentemente, atender aos objetivos desta dissertação de mestrado.

3.1 Aprendizado por Reforço

O aprendizado por reforço é um elemento de *Machine Learning* que tem ganhado destaque nos estudos relacionados a controle inteligente. Diferentemente do aprendizado supervisionado e o não-supervisionado, a aprendizagem por reforço se baseia no treinamento por experiência. No caso de sistemas de controle, a planta controlada é submetida a diversos cenários e os parâmetros do controlador recebem ajustes baseado no quão bem a tarefa foi realizada.

O que viria a ser chamado como aprendizado por reforço se baseou no desenvolvimento de duas áreas de pesquisa que são estudadas desde meados do século XX. Uma área está relacionada com o aprendizado por tentativa e erro e que se baseia no desenvolvimento animal.

A outra área está relacionada ao desenvolvimento da teoria de controle ótimo e a busca por um controlador capaz de manipular um determinado estado de um sistema dinâmico.

No que se trata do controle ótimo, Richard Bellman foi um dos primeiros que relacionou os estados de um sistema dinâmico a uma função de valor. Posteriormente, essa função seria denominada de Bellman e sua solução é importante para avaliar a otimalidade de um problema de controle ótimo. No caso de problemas da mesma natureza discretos, Richard Bellman definiu o processo de decisão de Markov (do inglês, *Markov Decision Process* (MDP)).

Com o tempo, percebeu-se que era possível aliar os conceitos de controle ótimo e de aprendizado. Como é apontado em Sutton e Barto (2018), desde meados do século XX até o fim da década de 80, houve muito avanço relacionado à união dessas áreas. Em sequência, em 1996, o termo programação neurodinâmica foi utilizado pela primeira vez por Dimitri Bertsekas e John Tsitsiklis.

Em período similar que a teoria de controle ótimo e de programação dinâmica foi se aprimorando, houve intensa pesquisa relacionada ao aprendizado de máquinas utilizando conceitos de treinamento por tentativa e erro. Há relatos de implementações de redes neurais que aprenderam por tentativa e erro já na década de 1950, como é apontado em Sutton e Barto (2018). Nessa época, existia uma discussão relacionada à diferença entre aprendizado supervisionado e o aprendizado por reforço. Em alguns trabalhos, acreditava-se estar empregando um treinamento supervisionado, embora estivessem sendo usados conceitos claros de aprendizado por reforço. Sutton e Barto (2018) ainda apontam como alguns livros de redes neurais demoraram a caracterizar certos tipos de algoritmos de treinamento baseados em tentativa e erro como aprendizado por reforço.

Hoje o conceito está bem consolidado e, embora exista, há uma menor confusão na distinção entre os tipos de treinamento aplicados ao aprendizado de máquinas. No contexto de redes neurais, esses modelos podem ser caracterizados como Sutton e Barto (2018):

- Aprendizado supervisionado – baseado no ajuste dos pesos da rede neural a partir do erro entre a saída obtida pela rede e o valor que se deseja obter. Nesse caso, o desenvolvedor da rede tem conhecimento de qual é o conjunto de saídas que a rede neural deve alcançar de forma que seus pesos são alterados até que o erro seja minimizado;
- Aprendizado não supervisionado – baseado no ajuste dos pesos da rede neural de forma que ela separe um banco de dados conhecido em padrões, características, grupos ou estruturas que são específicas daquele conjunto; e
- Aprendizado por reforço – baseado no ajuste dos pesos da rede neural de forma que uma função, denominada de recompensa, seja maximizada. O desenvolvedor não tem conhecimento de qual é a saída ideal da rede neural, mas ele sabe qual seria a resposta

adequada do sistema. Esse tipo de aprendizado se assemelha ao supervisionado pelo fato de o ajuste de pesos ainda ser baseado numa função erro. Contudo, o procedimento adotado e seu princípio são substancialmente diferentes.

3.1.1 Principais componentes do aprendizado por reforço

Até aqui, já foi realizada uma breve contextualização acerca do desenvolvimento do aprendizado por reforço e quais as diferenças principais entre ele e os aprendizados supervisionado e não supervisionado. Agora, já é possível detalhar melhor os elementos que compõem esse aprendizado e quais são os conceitos fundamentais para a implementação dessa técnica.

O aprendizado por reforço é composto por política, sinal de recompensa, função de valor e modelo do ambiente. Com esses 4 elementos, é possível caracterizar o agente do sistema e como ele se comunicará com o ambiente, tanto recebendo as informações, quanto enviando as ações de controle. Na Figura 2, pode ser visualizada a representação gráfica de como o sistema é estruturado.

Figura 2 – Modelo de interação ambiente-agente no aprendizado por reforço.

Fonte – Elaborada pelo autor.

Nota-se, por meio da Figura 2, que o agente recebe um conjunto de variáveis provenientes do ambiente e essas são os estados (S). Ademais, o agente envia para o ambiente um conjunto de ações (A). Por fim, dependendo dos estados atingidos pelo ambiente devido às ações empregadas nele, o agente recebe uma recompensa. A forma como o agente atua no ambiente segundo os estados recebidos da iteração anterior é denominado de política. A política pode ser tanto estocástica, onde o agente escolhe o conjunto de ações que possuem a maior probabilidade de produzirem um resultado positivo no ambiente, quanto determinística.

Ainda na Figura 2, observa-se que o sinal de recompensa é enviado para o agente apenas durante o treinamento. Além disso, ele será uma função cuja saída deve refletir se o conjunto de ações tomado pelo agente é ruim ou não. A partir disso, é possível construir a função de valor. Essa função traz a informação de qual será a provável recompensa acumulada pelo agente no longo prazo.

3.1.2 Algoritmos de aprendizado por reforço

Existem, essencialmente, duas categorias de algoritmos baseados em aprendizado por reforço. São aqueles algoritmos que estimam a função de valor por combinações de vetores ou tabelas e aqueles que estimam de forma aproximada. Para que o agente opere de forma adequada, é importante entender quais os princípios básicos desses algoritmos. Por isso, esta subseção está destinada a definir brevemente as duas categorias citadas e onde as redes neurais artificiais se encaixam.

3.1.2.1 Métodos Tabulares

Para entender como os métodos tabulares funcionam, deve-se ter conhecimento das principais funções e equações que dão base à teoria de aprendizado por reforço. Inicialmente, considere um processo de decisão de Markov finito, cujos estados, recompensas e ações são todos definidos. Nesse sistema, as variáveis S_t e R_t , apresentadas na Figura 2, possuem uma distribuição de probabilidade finita bem definida, e dependem apenas dos estados e ações anteriores. Essa característica pode ser definida pela Equação (3.1) (SUTTON; BARTO, 2018).

$$p(s', r|s, a) = \Pr\{S_t = s' | S_{t-1} = s, A_{t-1} = a\} = \sum_{r \in \mathcal{R}} p(s', r|s, a). \quad (3.1)$$

Essa é a equação de transição de estados. Ela retorna a probabilidade do agente receber determinada recompensa r , para um determinado estado s' atingido pelo sistema após transicionar de um estado anterior s devido a uma ação a .

Ainda, o conjunto de ações e transição de estados pode ser dado, por exemplo, por intervalos de tempos igualmente espaçados. Como o processo é finito, esses passos são limitados a um valor e ao fim de cada conjunto de ações tomadas pelo agente, define-se o que é chamado por episódio. A recompensa final esperada, denominada de Retorno, para uma dada política tomada, partindo de um estado inicial, é dada pela Equação (3.2):

$$G_t = R_{t+1} + R_{t+2} + R_{t+3} + R_{t+4} \dots + R_T, \quad (3.2)$$

sendo que T é o ponto final da operação do agente. Note que, para que a recompensa total obtida seja finita, é preciso que o processo tenha um fim. No caso de sistemas onde os episódios não possuem um final definido, é preciso definir um ajuste na Equação (3.2) para que a recompensa final não tenda a infinito. Esse ajuste, denotado por γ , pode ser dado ao descontar a recompensa de cada transição de estados, conforme mostrado pela Equação (3.3):

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \gamma^3 R_{t+4} \dots = \sum_{k=0} \gamma^k R_{t+k+1}. \quad (3.3)$$

Após a formalização de alguns conceitos, é possível agora definir duas funções fundamentais no aprendizado por reforço. Pelas Eqs. (3.4) e (3.5), tem-se as funções de valor do estado e de valor da ação, respectivamente:

$$v_{\pi}(s) = E_{\pi} \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s \right], \quad (3.4)$$

$$q_{\pi}(s, a) = E_{\pi} \left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s, A_t = a \right], \quad (3.5)$$

em que a função E_{π} retorna a esperança de uma determinada variável, para uma determinada política π e um instante específico. No caso da Equação (3.4), será retornado o valor que se espera obter para um determinado estado. E, no caso da Equação (3.5), retorna-se o valor que se espera receber após tomar certa ação a em um determinado estado s . Desenvolvendo a Equação (3.4), é possível obter a Equação (3.6) que é também conhecida como a Equação de Bellman para v_{π} :

$$v_{\pi}(s) = \sum_a \pi(a|s) \sum_{s'} \sum_r p(s', r|s, a) [r + \gamma E_{\pi}[G_{t+1} | S_{t+1} = s']]. \quad (3.6)$$

Para MDP's finitos, é possível afirmar que a política ótima será aquela que retornar o máximo valor encontrado de v_{π} , a qual será a função de valor de estado ótima, dada pela Equação (3.7).

$$v_*(s) = \max_{\pi} v_{\pi}(s). \quad (3.7)$$

Dessas políticas ótimas, também podem ser obtidas funções de valor de ação ótimas, dadas pela Equação (3.8).

$$q_*(s, a) = \max_{\pi} q_{\pi}(s, a). \quad (3.8)$$

Considerando o desenvolvimento previamente realizado, tem-se que a busca pela política ótima para um MDP se dá por meio da solução das equações de Bellman. Os métodos que realizam a estimação da solução desse conjunto de equações por meio de tabelas ou matrizes, são denominados de métodos tabulares. Atualmente, um dos métodos tabulares mais relevantes é o *Q-Learning*, o qual se baseia na Equação (3.9):

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha [R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]. \quad (3.9)$$

Neste algoritmo, estima-se a função valor de ação de forma que não é preciso ter conhecimento da política que está sendo adotada e, ainda assim, alcança-se uma aproximação de q_* . O *Q-Learning* é estruturado conforme mostrado no Algoritmo 1.

Algoritmo 1: Algoritmo de treinamento *Q-Learning*.**Entrada:** Estados (S_t) e recompensa (R).**Saída:** A_t **início**Inicializar os valores de $Q(s, a)$ em uma matriz de forma aleatória;Definir tamanho do episódio N ; $t = 0$;**enquanto** $t \leq N$ **faça**Escolher uma ação A para o estado S a partir da tabela Q , utilizando determinado critério. Por exemplo, adotar a ação com o maior valor de Q para aquele estado;Tomar a ação A , observar a recompensa R e o estado atingido S' ; $Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)]$;Atualizar: $S \leftarrow S'$; $t = t + 1$;**fim****fim****retorna** R .

Nota-se que, apesar de um algoritmo simples, seu custo computacional aumenta em função das dimensões do ambiente. Conforme aumentam os números de estados e ações, a tabela Q também aumenta, de forma que em algum momento é inviável alocar a memória necessária para desenvolvê-la. Nesse sentido, existem soluções que procuram estimar as funções de valor por meio de aproximações não lineares. Sabe-se que as redes neurais artificiais são capazes de aproximar funções e generalizar seu treinamento para diferentes condições de operação. Assim, ao considerar essas afirmações, desenvolveu-se o DRL, que se baseia no emprego de redes neurais artificiais em problemas de aprendizado por reforço.

3.1.2.2 Métodos baseados em redes neurais artificiais

Os algoritmos de aprendizado por reforço buscam treinar um controlador/agente de forma que, com a experiência e iterações sucessivas, uma política ótima seja alcançada. Uma das estratégias dos algoritmos para atingir este objetivo se dá pela aproximação da função de valor. Contudo, os métodos tabulares não são suficientes para atender as necessidades de sistemas com elevado número de ações e estados. Ainda, em aplicações de sistemas contínuos, é preciso realizar uma série de adequações dos algoritmos para que seja possível aplicar esses métodos.

Devido a essas limitações, o *DRL* tem sido uma importante ferramenta no avanço dos estudos com o aprendizado por reforço. Em vez de matrizes ou aproximações por funções lineares, é possível realizar uma estimação com as redes neurais artificiais de forma que as dimensões do sistema não se tornem um problema. Atualmente, existem diversas estratégias e tipos de DRL que podem ser aplicadas a controle de sistemas dinâmicos. Nesta dissertação de Mestrado, optou-se por implementar uma técnica que tem demonstrado bons resultados em outras áreas e seu desenvolvimento foi voltado a controle de sistemas contínuos. Essa técnica é a DDPG.

A DDPG foi proposta pela primeira vez por Lillicrap *et al.* (2016), de forma a unir os princípios do *Q-Learning* a uma estratégia ator-crítico. Neste caso, o agente é composto por dois elementos, o Ator ($\mu(s|\theta^\mu)$) e o Crítico ($Q(s,a|\theta^Q)$), onde os pesos dessas duas redes são θ^μ e θ^Q , respectivamente. Para melhorar a convergência do algoritmo, Lillicrap *et al.* (2016) propuseram o acréscimo de outras duas redes que são uma cópia das anteriormente citadas, representadas por ($\mu'(s|\theta^\mu)$) e ($Q'(s,a|\theta^Q)$). Essas cópias são usadas para calcular os valores almejados e seus pesos são atualizados de forma mais suave que os das redes originais.

O Ator original toma as ações (a_t) que serão enviadas ao ambiente e o Crítico avalia se aquela decisão foi positiva ou não, estimando a função valor de ação, a partir dos estados recebidos (s_t) e a recompensa (r). Essas transições, (s_t, a_t, r, s_{t+1}) , são armazenadas em um lote (R_b). Assim, os dados para o treinamento são retirados de um conjunto de (n) transições do lote R_b .

Durante o treinamento, define-se a variável da Equação (3.10):

$$y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'}), \quad (3.10)$$

sendo que, i está associado ao índice do conjunto n .

Na Equação (3.10), a variável y_i é composta pela recompensa de índice i somada ao que está sendo estimado pela cópia do Crítico descontado por um fator γ . Dessa forma, os pesos θ^Q são atualizados pela minimização de uma função de perda dada pela Equação (3.11):

$$L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i|\theta^Q))^2. \quad (3.11)$$

Enquanto isso, o Ator é treinado através da maximização do gradiente de política. Esse gradiente, é aproximado pela Equação (3.12):

$$\nabla_{\theta^\mu} J \approx \frac{1}{N} \sum_i \nabla_a Q(s_i, \mu(s_i)|\theta^Q) \nabla_{\theta^\mu} \mu(s_i|\theta^\mu), \quad (3.12)$$

sendo que, N é o número de transições contidas em n ; e J é a função de política.

Na Equação (3.12), é possível observar que foi feita uma somatória dos gradientes calculados dentro do conjunto i com a respectiva normalização pelo total de transições N . Isso é utilizado para treinar o Ator a partir da média do gradiente da política.

Por fim, como foi comentado, são atualizados os pesos das cópias do Crítico do Ator. As atualizações dos pesos são dadas pelas Equações (3.13) e (3.14):

$$\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'}, \quad (3.13)$$

$$\theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'}. \quad (3.14)$$

Todo o processo de treinamento do DDPG está representado no Algoritmo 2 (LILLICRAP *et al.*, 2016).

Algoritmo 2: Algoritmo de treinamento do DDPG.

Entrada: Estados (S_t) e recompensa (R).

Saída: A_t

início

 Inicializar a variável N_t que será responsável pela exploração do algoritmo;

 Receber o estado inicial s_0 ;

 Definir o tamanho do episódio N ;

$t = 0$;

 Calcular a ação por meio de: $a_t = \mu(s_t | \theta^\mu) + N_t$;

enquanto $t \leq N$ **faça**

 Receber a recompensa e o novo estado atingido;

 Armazenar a transição em R_b ;

 Amostrar um pequeno conjunto aleatório de n transições no bloco R_b ;

 Definir $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1} | \theta^{\mu'}) | \theta^{Q'})$;

 Atualizar os pesos do Crítico por meio da minimização do erro quadrático médio entre y_i e a saída calculada pelo crítico;

 Atualizar os pesos do Ator utilizando o gradiente de política;

$\theta^{Q'} \leftarrow \tau \theta^{Q'} + (1 - \tau) \theta^{Q'}$;

$\theta^{\mu'} \leftarrow \tau \theta^{\mu'} + (1 - \tau) \theta^{\mu'}$.

$t = t + 1$;

 Calcular a ação por meio de: $a_t = \mu(s_t | \theta^\mu) + N_t$;

fim

fim

retorna R .

Nota-se que neste algoritmo surge uma variável aleatória (N_t) que deve causar um deslocamento nas ações calculadas pelo Ator. Essa variável é o ruído e ele delimita o quanto o agente vasculhará todo o espaço de operação do ambiente. Um dos maiores desafios no projeto desses algoritmos é definir um balanço entre manter a exploração no ambiente e aproveitar a região onde foi encontrada uma recompensa localmente ótima. Conforme o treinamento se prolonga, esse ruído decairá, de forma que, após um tempo, não existirá mais processos de exploração do ambiente e o agente poderá ficar parado num ótimo local.

3.2 Estado da Arte da Ferramenta Aplicada a Controle de Microrredes

Há diversos estudos relacionados a controle de MRs que propuseram a implementação de controladores com sistemas inteligentes. Dentre a vasta gama de ferramentas disponíveis, foi apontada uma que tem surgido em estudos recentes de forma mais evidente e que esta dissertação de Mestrado optou por utilizar. Nesta seção, serão apresentados alguns dos estudos encontrados

durante a revisão bibliográfica e que demonstraram maior detalhamento tanto do sistema de controle quanto da MR utilizada.

Khorramabadi e Bakhshai (2015) propuseram um sistema de controle em uma MR, utilizando RNAs associadas à lógica *fuzzy*. Esse sistema consistia num agente *fuzzy* responsável por ajustar os pesos do controlador neutro-*fuzzy* a partir de um algoritmo de aprendizado por reforço. Esse controlador, por sua vez, varia os valores de potências ativa e reativa de referência das GDs da MR. Neste trabalho, os modelos de GDs implementados foram baseados em inversores. O controlador, então, realiza o ajuste de potência de forma indireta pelo controle de corrente, operando na malha interna da estrutura de controle dos conversores. Para avaliar o desempenho do controlador, foram testados 5 cenários diferentes de operação de uma MR e também foi implementado um controlador PI para comparar com o proposto. Pelos resultados, o tempo de resposta teve uma redução expressiva com o controlador neutro-*fuzzy* quando comparado ao PI. Além disso, foi calculado o erro médio absoluto das potências de saída das GDs em comparação aos valores de referência estabelecidos. Nesse caso, o controlador neuro-*fuzzy* mostrou-se melhor, embora o índice calculado não tenha sido expressivamente maior que o obtido para o controlador PI. Destaca-se que, como este trabalho foi publicado em 2015, o aprendizado por reforço foi utilizado como coadjuvante ao sistema de controle e sua contribuição foi na forma de ajustar parâmetros do controlador neuro-*fuzzy*. Como será mostrado pelas próximas publicações, o aprendizado por reforço assumiu um maior protagonismo nas proposições recentes.

Qiu, Nguyen e Crow (2016) propuseram um algoritmo de gerenciamento dos estados de carga de armazenadores de energia de uma MR. Esse algoritmo é baseado no *Q-learning* e suas ações são ligar ou desligar os armazenadores, com base nos respectivos estados de carga e a potência de saída total. Para validação do algoritmo, foi idealizada uma MR que possui uma bateria redox de vanádio e uma de chumbo ácido. As fontes de energia são um conjunto de painéis fotovoltaicos conectados ao mesmo elo CC dos armazenadores. Para o treinamento e validação, foram considerados estágios de operação da MR, o que não exigiu a implementação de modelos complexos das fontes de energia, armazenadores e cargas. Nesses estágios, foram avaliadas apenas a potência da carga, a potência de saída do conjunto de Painel Fotovoltaico (PV) e qual a potência de saída total dos armazenadores. Pelos resultados, o algoritmo foi capaz de manter equilíbrio de geração e consumo, além de minimizar as perdas nos bancos de baterias, visto que foi capaz de evitar descargas muito profundas desses elementos. Neste trabalho, um maior número de cenários de operação poderiam ser simulados, onde ainda poderia ser realizada a comparação desse algoritmo com alguma técnica tradicional. Ainda é possível observar que o principal componente para o gerenciamento de energia é uma técnica baseada no aprendizado por reforço.

Kofinas, Dounis e Vouros (2018) propuseram um controlador *fuzzy*, associado ao *Q-Learning*, para realizar o gerenciamento de energia de uma MR ilhada. Utilizando uma estrutura

descentralizada, os agentes de cada componente da rede compartilham suas variáveis de controle, mas atuam de forma independente. Esses agentes foram implementados tanto nas fontes geradoras quanto nos armazenadores e cargas variáveis. Como o objetivo é realizar o gerenciamento de energia, foi apresentada a geração e consumo da MR ao longo de semanas de simulação. Com o tempo, o sistema conseguiu otimizar a geração das fontes intermitentes e também o consumo das cargas. Os autores apontaram também que futuros trabalhos de comparação com outras técnicas de controle seriam relevantes. Vale ser ressaltado que eles propuseram uma estratégia diferente de empregar o sistema *fuzzy* apenas pelo ajuste das funções de pertinência a partir do conhecimento do especialista. Esse processo foi realizado utilizando conceitos do *Q-learning*, em que o sistema *fuzzy* toma a melhor decisão para determinado estado, procurando sempre maximizar a recompensa.

Xia *et al.* (2020) propuseram o controle de tensão do barramento e do intercâmbio de corrente de uma MR em CC. A DDPG foi empregada para atuação a nível secundário nas fontes de energia. A proposta apresentou uma estrutura de controle distribuída, onde cada agente tem conhecimento apenas dos dados das fontes de energia vizinhas. A fim de constatar a eficácia do controlador, foi realizada uma simulação de uma MR em CC com quatro GDs. O software utilizado foi o MATLAB e a DDPG foi estruturada pela ferramenta já implementada nele. O controlador foi capaz de regular a tensão do barramento e garantiu a equalização das correntes de cada geração. Embora tenha bons resultados, o trabalho possui limitações como a falta de comparação de desempenho para diferentes cenários de operação e também pela ausência de implementação de uma estratégia de controle tradicional, a qual poderia ser empregada para fins de comparação.

Khooban e Gheisarnejad (2020) propuseram o controle carga-frequência de uma MR. Com ação primária, foi implementado um controlador PID-*fuzzy* cuja atuação se dá no gerador a diesel do sistema. Ainda, foi associada uma rede DDPG para melhorar o amortecimento da frequência pelo controlador proposto. A MR em análise possui, além do gerador a diesel, um veículo elétrico, um painel fotovoltaico e um gerador síncrono alimentado por energia maremotriz. Essa variação no tipo de fonte de energia dos geradores síncronos tem impacto na constante de tempo das turbinas. Para validar o desempenho do controlador, realizou-se a variação paramétrica do sistema *fuzzy* e também foi implementado um controlador PID. Ademais, o sistema foi simulado num esquema *Model in the Loop*, onde atrasos de comunicação e ruídos puderam ser adicionados e analisados. Pelos dados obtidos, o controlador proposto teve desempenho similar ao do PID, uma vez que esse controlador garantiu um erro de regime pequeno e uma resposta transitória amortecida.

Younesi, Shayeghi e Siano (2020) propuseram um algoritmo baseado em *Q-learning* para realizar o controle supervisorio de uma MR em CA. Foram desenvolvidos dois agentes, dos quais o primeiro seria responsável pelo controle de tensão e o segundo pelo controle de frequência. A saída do algoritmo inteligente é utilizada para aprimorar a ação de um controlador

PID, o qual atua diretamente nos geradores a gás do sistema. Para validar o desempenho da proposta, foi implementada uma MR baseada num sistema de distribuição dinamarquês e outros métodos de controle foram implementados, tais como um PID clássico e um PID *fuzzy*. Foram testados dois cenários diferentes e o controlador proposto proporcionou um maior amortecimento ao sistema. Para aumentar a contribuição do trabalho, um maior número de cenários poderia ter sido analisado.

Ressalta-se que, como é proposto por Xia *et al.* (2020) e Khooban e Gheisarnejad (2020), os estudos mais recentes têm avançado em relação ao emprego de *deep learning* associado ao aprendizado por reforço, como é o caso de algoritmos como o DDPG. É notável ainda que o aprendizado por reforço torna-se protagonista no controle inteligente. A Tabela 1 sumariza os trabalhos discutidos, destacando o objetivo do controle, a técnica inteligente utilizada, as técnicas utilizadas para fins de comparação e a estrutura/elementos da MR.

Tabela 1 – Trabalhos que implementaram sistemas de controle associados ao aprendizado por reforço.

Referência	Objetivo de Controle	Técnica Utilizada	GD	AD	Tipo de MR	Análise de casos	Técnicas comparativas
Khorramabadi e Bakhshai (2015)	Potências ativa e reativa	Neuro-fuzzy	3 Fontes chaveadas	-	CA	5 Cenários	PI
Qiu, Nguyen e Crow (2016)	Gerenciamento de energia	Q-Learning	PV	VRB, Chumbo-Ácido	CC	1 Cenário	Variação paramétrica
Kofinas, Dounis e Vouros (2018)	Gerenciamento de energia	Fuzzy + Q-Learning	FV, FC, Diesel	Banco de baterias	CA	1 Cenário	-
Younesi, Shayeghi e Siano (2020)	Frequência e Tensão	PID + Q-Learning	CHP + WTG	-	CA	2 Cenários	PID e Fuzzy-PID
Khooban e Gheisarnejad (2020)	Frequência	PID-Fuzzy + DDPG	VE, FV, 2 GS	-	CA	2 Cenários	Variação paramétrica e PID
Xia <i>et al.</i> (2020)	Tensão e Equalização de Corrente	DDPG	4 Fontes chaveadas	-	CC	1 Cenário	-

Fonte – Elaborada pelo autor.

Nota-se, pela Tabela 1, que os estudos trabalharam com MRs mais complexas, com mais de um recurso distribuído e, em alguns casos, com cargas especiais. Essa característica não é muito comum em grande parte dos trabalhos, pois normalmente os modelos são simplificados a fim de analisar o controlador que está sendo proposto. Contudo, considerando que MRs são sistemas complexos, quanto melhor for caracterizada a MR, mais relevante será o estudo para o avanço das técnicas de controle na área. Além disso, também pode ser observado que não são todos os trabalhos da Tabela 1 que compararam o controlador proposto com outras técnicas. Além disso, poucos cenários de simulação são realizados para validarem o desempenho. Por esse motivo, a área ainda carece de mais estudos relacionados ao assunto e que façam um detalhamento maior do sistema usado para validar proposições de controle inteligente.

3.3 Considerações Finais

Neste Capítulo foi realizado um breve desenvolvimento teórico acerca do aprendizado por reforço. Essa área de pesquisa é extensa e para maior aprofundamento do conteúdo, sugere-se verificar Sutton e Barto (2018).

Nesta dissertação de Mestrado, houve uma maior preocupação em apresentar um sistema de controle utilizando uma técnica inteligente baseada em DRL e que fosse capaz de manter a estabilidade de uma MR com razoável penetração de GD. Sendo assim, como será detalhado no próximo Capítulo, foi modelada uma MR de média tensão com 4 fontes de energia. Além disso, o software escolhido para as simulações foi o Matlab, devido à sua biblioteca de aprendizado por reforço. Dessa forma, as especificações e construção do controlador no software também estão presentes no próximo Capítulo.

MODELAGEM DA MICRORREDE

Neste Capítulo, será descrita a metodologia adotada nesta pesquisa de Mestrado. Dessa forma, serão considerados os modelos de linhas e cargas da rede, o equacionamento e parametrização necessários para a implementação dos DERs, a parametrização do controle secundário e as ferramentas computacionais utilizadas.

Inicialmente, toda a MR foi modelada no PSCAD. Com o desenvolvimento do trabalho, foi necessário migrar de *software* de simulação, uma vez que não seria possível implementar em tempo hábil o controlador inteligente em linguagem C no PSCAD. Sendo assim, a MR foi modelada novamente no Simulink/Matlab, onde há uma biblioteca completa de DRL. Além disso, no Simulink, foi possível reduzir a complexidade do modelo da MR ao mínimo necessário e com máxima velocidade de simulação, melhorando a eficiência do treinamento do controlador.

4.1 Microrrede para Estudo

Para a concepção da MR, uma rede de distribuição foi implementada no *software Power System Computer Aided Design* (PSCAD). Os dados dessa rede estão presentes em um *benchmark* realizado pelo *Conseil International des Grands Réseaux Électriques* (CIGRÉ) e equivale a um sistema de distribuição alemão de média tensão. Essa rede foi modelada de forma a contribuir em estudos que empregam fontes alternativas de energia. A representação da rede é apresentada na Figura 3.

Figura 3 – Microrrede modelada.

Fonte – Adaptado de CIGRE Task Force C6.04.02 (2014).

Por tratar-se de uma rede europeia, sua frequência de *benchmark* é de 50 Hz. Contudo, esse valor foi alterado para 60 Hz, que é o padrão adotado no Brasil. Com a mudança, os valores de impedâncias das linhas se mantiveram e as indutâncias e capacitâncias sofreram alteração. Ademais, não foi necessário trabalhar no Matlab com os *taps* dos transformadores T_1 e T_2 , pois apenas a região compreendida pela MR foi modelada.

Para compor a MR, foram consideradas as instalações de quatro DERs, sendo: um PV (nó 8); dois BESS (nós 7 e 11); e um gerador síncrono (PAC). Uma vez que o estudo envolve a operação da MR ilhada, definiu-se o gerador síncrono como o de referência e os bancos de baterias foram implementados com controle *droop* para dar suporte à manutenção dos valores nominais de frequências e tensões do sistema. Enquanto isso, o gerador fotovoltaico está configurado com controle de potência e exporta sempre a máxima energia disponível de suas respectivas fontes para a rede.

4.2 Modelagem dos Recursos Distribuídos de Energia

Os DERs são os principais agentes de uma MR que garantem a confiabilidade e robustez do sistema. Assim, é de suma importância garantir a coordenação dessas fontes de energia a partir de controladores robustos e estratégias adequadas. Há basicamente duas estratégias de controle em que um recurso pode trabalhar na MR: a *grid-forming* e a *grid-following* (KATIRAEI *et al.*, 2008). Em *grid-forming*, o objetivo do DER é garantir os níveis de tensão e frequência do PAC.

Por outro lado, em *grid-following*, o recurso distribuído apenas absorve ou injeta potência da rede, a partir de uma estratégia local, não havendo monitoramento da tensão e frequência do PAC. Levando isso em consideração, é possível desenvolver a estratégia de coordenação dos recursos distribuídos em uma MR. Nas seções seguintes, os DERs empregados serão apresentados, bem como suas características dinâmicas e especificações técnicas.

4.2.1 Sistema de Armazenamento de Energia à Baterias

Os sistemas de armazenamento distribuídos desempenham importante papel na MR. Suas aplicações podem ser para garantir níveis de qualidade de energia, gerenciamento de energia ou até mesmo alimentar as cargas de forma independente (LUO *et al.*, 2015). Para cada uma dessas, existem requisitos para a eficiência, capacidade de armazenamento, velocidade de resposta, vida útil e impacto ambiental (JARNUT; WERMIŃSKI; WAŚKOWICZ, 2017).

Assim como em Divya e Østergaard (2009), Luo *et al.* (2015), Jarnut, Werminiński e Waśkowicz (2017), há diversos estudos que apresentam as diferentes tecnologias de armazenamento existentes e suas respectivas características. Conforme mostra a Tabela 2, o BESS de íons de Lítio possui uma das maiores eficiências energéticas, velocidade de resposta e durabilidade. Sendo assim, atendendo aos requisitos citados por Jarnut, Werminiński e Waśkowicz (2017), o BESS é uma excelente solução para ser integrada como armazenamento distribuído em MRs.

Tabela 2 – Especificações técnicas dos elementos armazenadores.

Elemento Armazenador	Tempo de Resposta	Durabilidade (anos)	Capacidade de Armazenamento (MWh)	Capacidade de Potência	Eficiência do Ciclo (%)
SMES	ms, <1/4 ciclo	>20	0,0008-0,001	100 kW - 10 MW	95-98
BESS (Íon de Lítio)	ms, <1/4 ciclo	5-15	0,004 - 10	0,005 - 100 MW	75-97
Supercapacitor	ms, <1/4 ciclo	10-30	0,0005	0,001-0,3 MW	84-97
FES	<1 ciclo, s	15-20	0,0052; >5	>5 MW	90-95
TES	-	10-30	-	0,1-300 MW	30-60
BESS(VRB)	<1/4 ciclo	5-20	<60; 2; 3,6	0,03-3 MW	65-85
CAES	minutos	>20	<1000	>300 MW	42; 54
PHS	minutos	>40	500-8000	100 -5000 MW	70-85
FC-HES	s, <1/4 ciclo	5-15, >20	0,312	<50 MW	20-66

Fonte – Adaptado de Luo *et al.* (2015)

Um BESS é composto por um conversor do tipo fonte de tensão (VSC, do inglês, *Voltage-source Converter*). Além disso, pode ser instalado, no elo CC, um conversor *buck-boost* ou então o banco de baterias é conectado diretamente ao elo. O sistema de armazenamento é conectado à rede junto a um filtro LCL na saída do conversor e um transformador. Na MR, foram empregados dois BESS (cada um possui 1 MVA). Dessa forma, como será discutido posteriormente, foram inseridos saturadores nas saídas do controlador inteligente, com o objetivo de manter sempre o limite de potência aparente. A topologia de BESS implementada é apresentada na Figura 4.

Figura 4 – Topologia do BESS.

Fonte – Elaborada pelo autor.

O VSC é capaz de realizar a interface entre a tensão CA da rede e a CC da bateria. O conversor empregado é trifásico de dois níveis e suas chaves são transistores do tipo *Insulated Gate Bipolar Transistor* (IGBT). O VSC, além de ser capaz de injetar ou absorver potência ativa da rede, pode ser empregado para dar suporte de reativo à MR. Devido ao uso de um gerador síncrono na MR, responsável pelo controle de tensão e frequência do PAC, os elementos de armazenamento modelados cumprirão um papel de *grid-forming*. Assim, o VSC do BESS controla as potências ativas e reativas, sendo que os valores de referência das potências serão determinados pelo método *droop*. Por sua vez, o *droop* foi utilizado para que o BESS auxilie o gerador síncrono, garantindo maior robustez ao sistema, mesmo quando utilizadas fontes intermitentes de energia como os painéis fotovoltaicos. O modelo de VSC adotado pode ser visualizado por meio da Figura 5.

Figura 5 – Topologia do VSC.

Fonte – Elaborada pelo autor.

O controle de potências reativas e ativas está relacionado à corrente do indutor do VSC que varia conforme a modulação enviada aos transistores. As equações dinâmicas que

representam um conversor fonte de tensão com uma indutância no seu terminal são dadas por:

$$\begin{aligned} L \frac{di_a(t)}{dt} &= -(R + r_{on})i_a(t) + v_{TA}(t) - v_{sA}(t) \\ L \frac{di_b(t)}{dt} &= -(R + r_{on})i_b(t) + v_{TB}(t) - v_{sB}(t), \\ L \frac{di_c(t)}{dt} &= -(R + r_{on})i_c(t) + v_{TC}(t) - v_{sC}(t) \end{aligned} \quad (4.1)$$

sendo que, $(R + r_{on})$ é a resistência equivalente no terminal do VSC, L é a indutância no terminal do conversor e v_{sA} , v_{sB} e v_{sC} são as tensões do lado da rede elétrica. Ademais, v_{TA} , v_{TB} e v_{TC} são as tensões nos terminais do conversor e podem ser definidas como:

$$v_{TA}(t) = m_a(t) \frac{V_{cc}}{2}, \quad (4.2)$$

$$v_{TB}(t) = m_b(t) \frac{V_{cc}}{2}, \quad (4.3)$$

$$v_{TC}(t) = m_c(t) \frac{V_{cc}}{2}, \quad (4.4)$$

onde, m_a , m_b e m_c representam o ciclo de trabalho dos transistores do conversor. No caso de um VSC de dois níveis trifásico que utiliza modulação senoidal, essas variáveis seguem funções senoidais. Aplicando a Transformada de Park nas equações dinâmicas do conversor, tem-se as Equações (4.5) e (4.6):

$$L \frac{di_d(t)}{dt} = -(R + r_{on})i_d(t) + u_d(t), \quad (4.5)$$

$$L \frac{di_q(t)}{dt} = -(R + r_{on})i_q(t) + u_q(t), \quad (4.6)$$

sendo que, u_d e u_q são as variáveis de controle de uma malha interna, onde o objetivo é regular as correntes trifásicas nos terminais do VSC.

Aplicando a transformada de Park e adotando uma ação de controle apropriada, é possível controlar a potência ativa do conversor a partir da corrente de eixo direto (i_d), enquanto a potência reativa é controlada pela corrente de eixo em quadratura (i_q). Assim, cria-se uma nova malha de controle que é a externa, para regular as potências ativa e reativa do conversor. Como a topologia de BESS implementada possui um conversor *buck-boost*, o VSC é encarregado de controlar a tensão do elo. Uma vez que o conversor CC-CC gerencia a energia do banco de baterias, o VSC irá demandar ou injetar a potência ativa necessária para a rede.

Nesse conversor CC-CC, a potência ativa pode ser controlada ao variar o ciclo de trabalho dos transistores, de forma a regular a corrente do indutor conectado no terminal do banco de baterias. Para qualquer nível de tensão de saída da bateria e diferentes níveis de estado de carga, o controlador irá alterar a corrente do indutor para manter a potência constante, conforme valor de referência. Dessa forma, quando deseja-se analisar o regime transitório com esse tipo de conversor, nesse modo de controle, é possível representá-lo como uma fonte de corrente conectada ao elo CC do VSC. O modelo do BESS simplificado, considerando essa simplificação utilizando a fonte de corrente, está representado na Figura 6.

Figura 6 – Modelo do BESS simplificado.

Fonte – Elaborada pelo autor.

Para maiores detalhes acerca do controle de potência de um VSC, é possível encontrar o desenvolvimento teórico em Yazdani e Iravani (2010). Adicionalmente, a metodologia adotada para a modelagem do BESS foi similar à abordada em Barbalho (2018). Neste trabalho, foi projetado um filtro LCL, utilizando o procedimento apresentado em Zeng *et al.* (2010), no terminal do VSC para mitigar os harmônicos. A malha de controle equivalente dos conversores está representada na Figura 7. Destaca-se que a sintonia dos ganhos dos controladores foi realizada de forma empírica.

O VSC, no modelo completo, controla a tensão do elo CC e a potência reativa de saída do BESS. Com relação às malhas de controle desse conversor, é possível dividir em duas. A malha interna é responsável por controlar a corrente terminal e a malha externa é responsável por gerar os sinais de referência da malha interna para atingir os níveis de tensão do elo CC e potência reativa exigidos. O controle do CC-CC é responsável por extrair ou injetar a potência ativa de referência do banco de baterias. Por fim, o nível mais externo de controle é o *droop*. O *droop*, a partir dos desvios de tensão e frequência dos valores de referência, envia os sinais para as respectivas malhas do VSC e CC-CC.

Para determinar o ganho do *droop* de frequência e tensão, foram consideradas as equações (4.7) e (4.8):

$$K_d = \frac{P_0}{\Delta f}, \quad (4.7)$$

$$K_q = \frac{Q_0}{\Delta V}, \quad (4.8)$$

sendo que P_0 e Q_0 são os limites das potências ativa e reativa, respectivamente; Δf e ΔV são os desvios de frequência e tensão que poderão ocorrer durante a operação da MR e foram estabelecidos como 0,5 Hz e .5 p.u., respectivamente.

Como o conversor CC-CC é o responsável por controlar a potência ativa do BESS, o *droop* de frequência gera a potência reativa de referência do controlador desse conversor, enquanto o *droop* de tensão especifica a potência de referência do controle da malha externa do VSC.

Vale ser ressaltado que o sistema de armazenamento foi implementado com os ganhos do *droop* calculados conforme previamente descrito, porém, houve um ajuste empírico que garantiu

Figura 7 – Malha de controle do modelo completo do BESS.

Fonte – Elaborada pelo autor.

uma resposta mais rápida do BESS a variações de frequência e tensão. Menciona-se ainda que os ganhos $K_d = K_q = 10$. Os dados elétricos dos BESS projetados estão presentes na Tabela 3.

Conforme este trabalho se desenvolveu, foi possível concluir que a velocidade de simulação da rede no Simulink/Matlab não seria rápida o suficiente mesmo com o modelo médio. Dessa forma, a partir do comportamento já conhecido desses conversores utilizando o modelo detalhado, foi feita uma adequação da rede para possibilitar a simulação de forma fasorial no Simulink. Com isso, as equações dinâmicas são reduzidas ao máximo e a velocidade de simulação aumenta consideravelmente. Na simulação fasorial, as fontes de energia renováveis (PV e baterias) são equivalentes a uma fonte de potência controlada. Ademais, os modelos de fonte de potência não são conectados na MR através de um transformador de interface. Por isso, a tensão CA do BESS apontada na Tabela 3, vale apenas para o modelo completo. Para o modelo fasorial, deve-se considerar a tensão nominal do BESS igual aos 20 kV da MR.

Tabela 3 – Dados nominais do BESS.

BESS	
Tensão CA nominal	4,16 kV
Tensão nominal do banco de baterias	2,313 kV
Capacidade de armazenamento	1530 Ah
Tensão do elo CC	9,0 kV
Potência aparente nominal	1 MVA
Limite de potência ativa	1 MW
Limite de potência reativa	1 MVA _r
C_{eq}	1235 μ F
L_f	13,5 mH
L_g	12,1 mH
C_f	0,42 μ F
Frequência de chaveamento do VSC	6 kHz

Fonte – Elaborada pelo autor.

4.2.2 Gerador Síncrono

Conforme previamente mencionado, o Gerador Síncrono (GS) na MR é responsável por garantir a referência de tensão e frequência no PAC em um caso de ilhamento. Além disso, uma das vantagens de utilizar máquinas rotativas é que, com elas, adiciona-se à MR uma tendência intrínseca de manter a frequência nominal devido à inércia Katiraei *et al.* (2008). O modelo de gerador e dados elétricos utilizados podem ser encontrados em Moura *et al.* (2010), conforme mostrado na Tabela 4.

Tabela 4 – Parâmetros do gerador síncrono.

Gerador Síncrono	
Potência nominal	5 MVA
Tensão nominal	6,6 kV
Constante de inércia	1 s
Número de polos	2

Fonte – Adaptado de Moura *et al.* (2010).

Vale destacar que, em Moura *et al.* (2010), o gerador possui 4 polos e assume-se que é um hidrogerador. Contudo, o modelo de gerador empregado no PSCAD possui apenas 2 polos. Por isso, no Simulink, esse último valor foi mantido no modelo. Ademais, a dinâmica da turbina foi inserida como uma função de transferência de primeira ordem com uma constante de tempo de 1 s. O modelo de excitatriz foi baseado no trabalho de Moura *et al.* (2010). Na Figura 8, encontra-se o diagrama do gerador no Simulink.

Figura 8 – Diagrama do gerador síncrono.

Fonte – Elaborada pelo autor.

O gerador síncrono foi conectado à MR por meio de um transformador de interface de 6,6 kV/20 kV cuja configuração é Y-Y. Os dados do transformador no Simulink estão na Tabela 5.

Tabela 5 – Dados do transformador do GS

Dados Gerais			
Potência nominal	5 MVA	Frequência	60 Hz
Dados do Primário		Dados do Secundário	
R_1 (p.u.)	0,00125	R_2 (p.u.)	0,00125
L_1 (p.u.)	0,035	L_2 (p.u.)	0,035

A malha do regulador de velocidade está representada na Figura 9, a qual foi adaptada de Report (2013) e os ganhos do PI (K_{p1} e K_{i1}) foram definidos de forma empírica e valem 0,143 e 0,04, respectivamente.

Figura 9 – Malha de controle do regulador de velocidade.

Fonte – Elaborada pelo autor.

A malha do regulador de tensão está representada na Figura 10. Essa malha de controle

foi adaptada de IEEE (2006). Ressalta-se que os ganhos de todas as funções de transferência foram definidos de forma empírica.

Figura 10 – Malha de controle do regulador de tensão.

Fonte – Elaborada pelo autor.

O controle secundário da MR é, então, de responsabilidade do gerador síncrono. Devido a isso, a máquina está conectada ao PAC e a tensão desse nó, bem como a frequência do sistema são controlados. Assim, o controle inteligente atuará nos BESSs para melhorar a estabilidade do sistema e auxiliar o controle do gerador síncrono.

4.2.3 Gerador Fotovoltaico

Os PVs têm ganhado destaque no contexto de geração distribuída mundial. Apenas o Brasil, que possui elevada irradiação solar, nota-se um crescimento considerável da potência instalada. Sendo assim, nesta pesquisa de Mestrado, optou-se por implementar também um gerador deste tipo, a fim de diversificar as DERs da MR.

O gerador fotovoltaico, tipicamente, possui topologia similar à do BESS. A diferença é que o conversor CC-CC dessa tecnologia é do tipo *boost* e, por isso, a transferência de energia é unidirecional. O PV é um recurso distribuído não despachável, uma vez que são implementados algoritmos que rastreiam a máxima geração de potência possível, dada uma irradiação específica. Há diversos algoritmos desse tipo, denominados de *Maximum power point tracking* (MPPT), sendo que o empregado foi o *Perturb and Observe* (PO). Esse algoritmo se baseia na variação da tensão de referência do controlador do conversor *boost* até que a potência injetada no elo CC deixe de aumentar e comece a diminuir.

O modelo de PV, detalhado em PSCAD (2018), está disponível na base de projetos do PSCAD. Não foram feitas grandes alterações, a não ser pela mudança do transformador CA de interface do DER. Sendo assim, os dados elétricos do gerador fotovoltaico podem ser encontrados na Tabela 6.

Tabela 6 – Parâmetros do gerador solar fotovoltaico.

Gerador Fotovoltaico	
Potência nominal	1 MW
Tensão nominal	480 V
Frequência de chaveamento do VSC	8000 Hz
Irradiação solar para potência nominal	1000 W/m ²

Fonte – Elaborado pelo autor.

Apesar de ter sido realizada a modelagem do PV no PSCAD, houve uma migração para o MATLAB por motivos previamente descritos. Além disso, o modelo do PV foi reduzido para uma fonte de potência controlada (modelo fasorial). Devido à essa mudança, a tensão nominal do PV, como fonte de potência, é de 20 kV. A tensão nominal da Tabela 6 foi considerada apenas na modelagem completa desse gerador distribuído.

4.3 Modelagem do Controle Supervisório

Como foi relatado, essa pesquisa de Mestrado implementou a DDPG para cumprir uma das principais funções do controle supervisório de uma MR ilhada, que é controlar a frequência e tensão do sistema. Para isso, deve-se considerar a estrutura do modelo da DDPG proposta em Lillicrap *et al.* (2016). As modificações feitas neste trabalho foram todas relacionadas a este modelo base. Sendo assim, em Lillicrap *et al.* (2016), o Ator possui duas camadas escondidas com 400 e 300 neurônios, respectivamente. O Crítico, devido a receber tanto os estados quanto as ações como entradas, possui dois caminhos. No primeiro caminho, há uma camada escondida e é por onde as ações são inseridas. No segundo caminho, são inseridos os estados e há duas camadas escondidas. A saída desses dois caminhos servem como entrada para mais uma camada escondida de união, mas nessa camada há apenas as funções de ativação. No caminho dos estados, há 400 e 300 neurônios em cada camada, respectivamente, enquanto no caminho das ações há 300 neurônios. Considerando agora esses números de neurônios como o valor base, os resultados que serão apresentados foram obtidos utilizando uma rede mais simplificada, em que o maior número de neurônios utilizado foi 25 vezes menor que o número existente no caso base.

A partir dessas definições, é preciso determinar quais são as ações do agente. No caso, como deseja-se controlar a tensão e frequência a nível secundário, o agente controlará as potências ativas e reativas das baterias B1 e B2. As saídas do agente foram saturadas de forma que a potência aparente de cada bateria não ultrapasse 1 p.u. Ademais, os estados observados pelo agente são os erros das tensões e os próprios valores das tensões de cada barra que possui uma recurso de energia distribuído, o erro de frequência do sistema e a frequência do sistema e, por fim, também são inseridos os valores das tensões e frequência do passo anterior. Esses valores foram inseridos com o intuito de permitir que a rede conseguisse realizar operações de

derivadas, o que poderia melhorar o desempenho do controlador.

A DDPG foi implementada no Matlab, através de uma biblioteca de aprendizado por reforço do software. Com essa biblioteca, é preciso apenas inserir os parâmetros de treinamento e quantos neurônios e camadas o Ator e Crítico possuem. Ademais, o algoritmo de treinamento já está incluso nessa biblioteca. Os parâmetros base utilizados nos estudos deste trabalho estão apresentados na Tabela 7.

Tabela 7 – Parâmetros do DDPG no Matlab.

Número de ações	4	Variância do ruído	0,1
Número de estados	12	Taxa de decaimento do Ruído	1e-5
γ	0,99	Passo de atuação do controlador	0,2 s
Taxa de aprendizado (Ator)	1e-4	Passos por episódio	500
Taxa de aprendizado (Crítico)	1e-3	Tamanho da janela móvel	20 (Base)
τ	1e-3	Tamanho do <i>minibatch</i>	250

Fonte – Elaborada pelo autor.

Por último, a função de recompensa é expressa pela Equação (4.9):

$$R = -16(V_{GS} - V_{GS,ref})^2 - 12(V_1 - V_{1,ref})^2 \dots - 12(V_2 - V_{2,ref})^2 - 18(freq - freq_{Ref})^2 - 20 \left| \frac{d(freq - freq_{Ref})}{dt} \right| - 500P + 30 \frac{T_s}{T_f}, \quad (4.9)$$

sendo que, V_{GS} e $V_{GS,ref}$ são a tensão lida no terminal do GS e a tensão de referência, respectivamente; V_1 e $V_{1,ref}$ são a tensão lida no terminal de B1 e a tensão de referência, respectivamente; V_2 e $V_{2,ref}$ são a tensão lida no terminal de B2 e a tensão de referência, respectivamente; T_s e T_f são o passo de atuação do controlador e T_f é o tempo total de simulação; e P é uma variável binária que indica que ocorreu alguma variação muito grande de tensão ou frequência na rede. Assim, P assume valor igual a 1, caso algumas das tensões controladas assumam valor inferior a 0,2 p.u. ou superior a 2 p.u. ou caso a frequência da rede exceda 1,5 p.u. ou fique inferior a 0,5 p.u. Além disso, se alguma dessas condições é atingida, o episódio é interrompido e inicia-se outro. Caso nenhuma dessas condições seja atendida, P assume valor igual a zero.

A função de recompensa foi construída a partir da lógica que prioriza manter certas variáveis mais próximas ao valor de referência do que outras. Por exemplo, pela Equação (4.9), nota-se que o erro da tensão do gerador síncrono recebeu um peso maior do que as outras tensões controladas. Neste caso, entende-se que é mais importante regular a tensão dessa unidade geradora do que a das outras. Ainda, ao realizar algumas simulações com o sistema, notou-se a necessidade de dar um peso maior ao erro de frequência e, não apenas isso, foi preciso punir mais o agente para ações que causavam respostas de grande oscilação. No caso, para punir comportamentos muito oscilatórios, foi adicionado o termo de módulo da derivada. Por fim, o único termo positivo da função de recompensa é responsável por bonificar o agente por manter

o sistema estável durante todo o episódio. Assim, se todos os valores de tensão e frequência fossem mantidos nas respectivas referências, sem comportamentos oscilatórios, o agente recebe no máximo uma recompensa igual a 30.

Considerando as entradas e saídas do agente, é possível observar na Figura 11 como o mesmo foi modelado no Simulink.

Figura 11 – Representação do agente no Simulink.

Fonte – Elaborada pelo autor.

Na Figura 11, os blocos em azul representam as entradas do agente. Dentre as variáveis lidas, vale ressaltar que wm é a frequência da rede ($freq$) lida a partir da máquina síncrona. O índice k está associado ao passo de atuação do controlador, o que indica que os valores das variáveis do passo imediatamente anterior estão representadas com o índice $k - 1$. Os blocos em vermelho são as saídas e servem para indicar o nível de potência que deve ser injetada ou consumida pelas baterias. Na saída do agente, foi inserida uma transformação cartesiano-polar para limitar o nível de potência aparente em 1 p.u. Ainda, resalta-se que as variáveis com subscrito “e” são os erros. Por exemplo, $V1_e$ é o erro de tensão do terminal de B1 ($V_{1,ref} - V_1$).

Como está representado na Figura 11, o agente possui uma comunicação com o código do MATLAB que, por sua vez, inicializa os ganhos das malhas de controle do GS e os tempos de atuação dos disjuntores, quando for o caso. Além disso, no código, o agente é estruturado, bem como são configurados os parâmetros de treinamento. O código utilizado para um dos agentes treinados foi adicionado no Apêndice A.

4.4 Considerações Finais

Neste Capítulo, foram apresentados os modelos completos dos recursos distribuídos e da MR implementados. Posteriormente, os modelos foram reduzidos de forma que fosse possível simular a MR no Simulink/Matlab. Assim, as fontes chaveadas foram modeladas como uma fonte de potência controlada e o modelo de gerador síncrono foi mantido. A interface entre o GS e a MR foi realizada através de um transformador de 20 kV/6,6 kV configurado em Y-Y. Como o modelo das outras fontes foi reduzido para um modelo de fonte de potência controlada, não foram considerados transformadores dessas fontes. Assim, a tensão nominal dos equivalentes das fontes distribuídas de energia no modelo fasorial é de 20 kV. O passo de simulação foi de 2,5 ms e o tipo de *solver* foi o *backward-euler*.

OPERAÇÃO DA MICRORREDE COM O CONTROLE SUPERVISÓRIO INTELIGENTE

O principal objetivo desta dissertação de Mestrado foi o estudo e a implementação de um controlador inteligente que fosse capaz de regular tensão e frequência de uma MR. Com o desenvolvimento da pesquisa, decidiu-se adotar a DDPG para o controle dos armazenadores da MR, de forma a auxiliar a máquina síncrona. Conforme foi apontado, a MR estudada foi toda modelada no Simulink, utilizando a simulação fasorial. Na Figura 12, encontra-se a região efetivamente modelada no *software*.

Figura 12 – Microrrede modelada no Simulink/Matlab.

Fonte – Elaborada pelo autor.

Inicialmente, implementou-se no Matlab, uma DDPG com os mesmos parâmetros do que foi proposto em Lillicrap *et al.* (2016). Em seguida, foram alterados os números de neurônios nas camadas escondidas do Ator e do Crítico; o número de camadas escondidas; a variância do ruído; e o tamanho da janela para cálculo da média móvel. Com o avanço, foi possível notar

que o número de neurônios das redes era muito grande para a dimensão do problema. Nas Figuras 13 e 14, encontra-se o processo de treinamento de duas redes parametrizadas de forma similar, diferindo apenas no número de neurônios em suas camadas escondidas.

Figura 13 – Treinamento do agente com duas camadas escondidas na rede Ator e Crítico.

Fonte – Elaborada pelo autor.

Figura 14 – Treinamento do agente com apenas uma camada escondida na rede Ator.

Fonte – Elaborada pelo autor.

Na Figura 13, foi mantida a quantidade de camadas escondidas. No entanto, os neurônios foram reduzidos em 25 vezes, se comparado ao proposto em Lillicrap *et al.* (2016). Nota-se que o agente alcançou um nível de recompensa e permaneceu estável ao longo do treinamento. Ao reduzir ainda mais o número de neurônios e retirando a camada escondida no caminho dos estados no Crítico e uma camada escondida no Ator, foi obtido o resultado da Figura 14. É possível observar que um maior número de camadas e neurônios não melhora consideravelmente

o desempenho do controlador, visto que com o tempo ambos atingem recompensas mínimas semelhantes. Dessa maneira, por não haver necessidade de uma estrutura mais complexa do agente, o modelo mais simples foi utilizado, cujo treinamento está representado na Figura 14.

Adicionalmente, o treinamento do agente foi baseado na abertura e fechamento de D1 de forma intercalada no intervalo de 100 s. Dentro desse período, o disjuntor começa fechado, em seguida abre e, por fim, fecha novamente. Os tempos de abertura e fechamento de D1 foram definidos de forma aleatória para que o agente não memorizasse o evento. A única regra definida para o cálculo dos intervalos é que entre o tempo de abertura e o fechamento, fosse de no mínimo 30 s. Ainda, a potência do PV não ficou constante durante o treinamento, mas oscila de forma periódica, como mostra a Figura 15.

Figura 15 – Potência do PV.

Fonte – Elaborada pelo autor.

Por fim, vale ressaltar que a recompensa média do agente neste caso está variando mais, pois sua janela móvel possui tamanho de 20 episódios, enquanto que, para o caso da Figura 14, o tamanho é de 50 episódios. Sendo assim, optou-se por utilizar a composição mais simples, visando reduzir a complexidade do controlador, uma vez que ele traz desempenho similar.

Após os estudos de variação dos parâmetros de treinamento, foram realizadas simulações para validar o desempenho do controlador. Os estudos foram divididos em 3 casos principais, a saber:

- Caso 1 – desconexão programada do disjuntor D1 com a intenção de simular uma manutenção programada da rede;
- Caso 2 – abertura de D1 após um evento de curto-circuito;
- Caso 3 – desconexão de uma carga de grande porte do sistema.

5.1 Caso 1 - Desconexão programada de alimentador

Para avaliar o desempenho do controlador inteligente treinado, foi realizada a abertura programada de D1 aos 30 s e posterior fechamento aos 60 s. No total, foram 100 s de simulação e os estados de D1 foram: fechado, aberto e fechado. Esse é um caso bem grave na MR estudada, uma vez que a abertura de D1 ocasiona uma perda de metade da carga. Com isso, já era esperada uma grande perturbação na rede. Ademais, foram comparados os desempenhos do controlador inteligente com um controlador *droop*, apresentando ainda o caso onde os armazenadores de energia não operam, o que equivale à não existência do controle secundário.

A princípio, pode-se realizar uma análise qualitativa por meio de avaliação gráfica, em que o comportamento da frequência da rede pode ser observado na Figura 16.

Figura 16 – Frequência da rede para o caso 1.

Fonte – Elaborada pelo autor.

Nota-se, pela Figura 16, que tanto o *droop* quanto a DDPG conseguiram diminuir o pico da frequência do sistema. Contudo, é possível observar que o *droop* demonstrou um desempenho superior ao DDPG quando o disjuntor D1 foi reconectado. Seguindo com a análise qualitativa, faz-se importante observar as tensões medidas na MR, que correspondem aos nós 3 (GS), 7 (B2) e 11 (B1) e que podem ser visualizadas respectivamente pelas Figuras 17, 18 e 19.

Ao verificar a suavização da tensão dos nós apresentadas nas Figuras 17, 18 e 19, nota-se que o DDPG conseguiu minimizar o pico de tensão atingido devido à oscilação, embora pareça ter desempenho inferior ao *droop*. Ressalta-se que a diferença de desempenho entre o *droop* e o DDPG é mais evidente na tensão do GS, onde há maior discrepância. Nas tensões de B1 e B2, os dois métodos de controle desempenham de forma semelhante quando observado visualmente.

Para compreender como o DDPG atua nas baterias para regular a tensão e frequência da rede, é válido observar as potências ativas de saída dos armazenadores B1 e B2, conforme mostrado na Figura 20.

Na Figura 20, nota-se que o treinamento do DDPG fez com que ele designasse às baterias

Figura 17 – Tensão do GS para o caso 1.

Fonte – Elaborada pelo autor.

Figura 18 – Tensão de B1 para o caso 1.

Fonte – Elaborada pelo autor.

Figura 19 – Tensão de B2 para o caso 1.

Fonte – Elaborada pelo autor.

tarefas diferentes. No caso de B1, o armazenador mantém, por padrão, sempre uma geração de energia. Durante uma variação positiva de frequência, o DDPG faz com que esse armazenador

Figura 20 – Potências ativas das baterias para o caso 1.

Fonte – Elaborada pelo autor.

varie sua potência de maneira mais brusca. No caso de B2, o DDPG deixou a variação de potência desse armazenador mais sensível às oscilações provocadas pelo PV e, ainda, garantiu uma variação de potência mais expressiva no fechamento de D1, aos 60 s.

A interpretação que se tem sobre esse comportamento está relacionada a dois fatos. O primeiro é a proximidade do armazenador a uma perturbação que ocorre periodicamente na MR. Isso significa que B2 pode ter sua potência mais sensível às oscilações de frequência e tensão, por estar mais próxima ao PV. Outro fato é o estado de geração imediatamente anterior ao evento. Antes da desconexão do alimentador, B1 está próximo à sua máxima potência de geração e B2 está no máximo do seu consumo. Ao abrir D1, grande parte da carga da MR é desconectada, o que demanda uma ação que compense essa perda. Por isso, o DDPG alterou mais a potência de B1, que é o armazenador com maior capacidade de consumir naquele instante. No caso do fechamento de D1, o princípio é o mesmo, fazendo com que B2 diminua seu consumo momentaneamente a fim de suavizar as oscilações.

O agente atua de forma similar ao variar as potências reativas dos armazenadores, conforme apresentado na Figura 21.

Figura 21 – Potências reativas das baterias para o caso 1.

Fonte – Elaborada pelo autor.

A bateria B2 atua muito mais do que B1 e a atuação dessa última é mais perceptível durante a abertura de D1. Vale ser ressaltado que as potências reativas não se aproximam de 1 p.u., uma vez que existem saturadores nas saídas do DDPG, garantindo sempre o limite de potência aparente de 1 p.u.

5.2 Caso 2 - Aplicação de um curto-circuito

No segundo caso estudado, foi aplicado um curto-circuito trifásico na barra 4 do sistema. Na Figura 22, encontra-se uma representação do local.

Figura 22 – Representação do local do curto-circuito aplicado no caso 2.

O curto-circuito, com impedância de 20Ω , foi aplicado aos 30 s de simulação. Em seguida, considerou-se que o disjuntor D1 estava associado a um sistema de proteção com capacidade de detectar este tipo de evento em 6 ciclos de onda. Assim, aos 30,1 s, D1 abre e os controladores do nível secundário devem restaurar a tensão e frequência da MR. Vale ressaltar que esse valor de impedância foi escolhido na tentativa de não permitir que houvesse um afundamento de tensão grande durante o curto-circuito. Conforme os resultados forem apresentados, essa restrição será discutida.

Inicialmente, na Figura 23, é mostrado o comportamento da frequência da rede após atuação dos controladores.

Figura 23 – Frequência da rede para o caso 2.

Observando a Figura 23, é possível notar que o pico da frequência é similar ao do caso 1. Contudo, na transição de estado de operação, há uma oscilação maior do sistema que pode ter sido provocado pela existência de um curto-circuito no instante anterior à abertura de D1. Neste

caso, o *droop* e o DDPG atuaram de forma que o pico da frequência fosse o mesmo. Entretanto, o *droop* eliminou a oscilação inicial observada no cenário sem armazenadores.

Agora, observe as tensões do GS, de B1 e B2 nas Figuras 24, 25 e 26.

Figura 24 – Tensão do GS para o caso 2.

Fonte – Elaborada pelo autor.

Figura 25 – Tensão de B1 para o caso 2.

Fonte – Elaborada pelo autor.

Figura 26 – Tensão de B2 para o caso 2.

Fonte – Elaborada pelo autor.

Conforme mostrado para o caso 1, também é importante verificar o comportamento das tensões dos nós 3, 7 e 11. Nota-se que, na transição dos estados de operação (mostrados nas Figuras 24, 25 e 26), há um pico que extrapola os limites do eixo das ordenadas. Após comparar os resultados das simulações de todos os controladores e até mesmo no caso que não há bateria, considerou-se que foi um erro numérico do Simulink devido ao chaveamento. Portanto, essa transição foi desconsiderada nas análises. Ainda em relação às Figuras 24, 25 e 26, verifica-se desempenho similar dos controladores DDPG e do *droop*. Ambos conseguiram manter os níveis de tensão o mais próximo possível de 1 p.u.

Em termos de potências ativas e reativas, seus comportamentos podem ser analisados por meio das Figuras 27 e 28.

Figura 27 – Potências ativas das baterias para o caso 2.

Fonte – Elaborada pelo autor.

Figura 28 – Potências reativas das baterias para o caso 2.

Fonte – Elaborada pelo autor.

Novamente, é notável a similaridade com o ocorrido no caso 1. Contudo, destaca-se um aumento da variação de potência ativa, uma vez que o curto-circuito provocou um desvio maior dos valores de tensão e frequência controlados.

5.3 Caso 3 - Desconexão de carga

Na última análise, foi realizada a desconexão da maior carga do sistema aos 30 s de simulação. Na Figura 29, é possível encontrar a localização da carga que foi desconectada por meio de D2.

Figura 29 – Localização da carga desconectada.

Conforme mostrado na Figura 30, o DDPG não conseguiu reduzir de forma acentuada o pico da frequência, sendo que o *droop* mostrou-se mais adequado.

Figura 30 – Frequência da rede para o caso 3.

Quanto às tensões dos nós 3, 7 e 11 (mostradas nas Figuras 31, 32 e 33), nota-se que, embora o DDPG e o *droop* tenham reduzido o pico de variação de forma muito similar para as tensões nos terminais de B1 e B2, o DDPG apresentou desempenho superior para regular a tensão do gerador síncrono.

Conforme esperado, a variação de potência para esse caso foi menor se comparado aos casos anteriores, uma vez que a variação de carga também foi menor. Esse comportamento pode ser notado por meio das Figuras 34 e 35, em que a variação da potência reativa de B1 não foi semelhante ao que ocorreu nos demais casos. Em vez de diminuir a geração, ou até mesmo consumir potência reativa para diminuir o pico das tensões, a geração se manteve. Esse fato tem as seguintes possíveis explicações, a saber: (i) não houve aumento suficiente de tensão para que isso refletisse na saída do DDPG; (ii) devido ao controlador inteligente possuir 12 entradas, a interação entre elas pode ter provocado essa saída para o caso analisado; ou (iii) o controlador

Figura 31 – Tensão do GS para o caso 3.

Fonte – Elaborada pelo autor.

Figura 32 – Tensão de B1 para o caso 3.

Fonte – Elaborada pelo autor.

Figura 33 – Tensão de B2 para o caso 3.

Fonte – Elaborada pelo autor.

demanda mais episódios de treinamento para aprender que esse tipo de resposta não trará uma recompensa tão boa.

Figura 34 – Potências ativas das baterias para o caso 3.

Fonte – Elaborada pelo autor.

Figura 35 – Potências reativas das baterias para o caso 3.

Fonte – Elaborada pelo autor.

5.4 Análise quantitativa de desempenho

Embora pela análise gráfica seja possível notar que o desempenho do DDPG foi similar ao *droop* em alguns casos, foi realizada uma comparação quantitativa de desempenho pelo cálculo do erro quadrático médio (EQM). O EQM foi calculado em janelas de 10 s, tanto para o caso com o algoritmo inteligente, quanto para o caso com o *droop*. Além disso, os valores foram normalizados ao EQM calculado para o caso onde não há sistema de armazenamento na MR, ou seja, nem o DDPG opera nem o *droop*. Na Figura 36, encontra-se a representação do EQM normalizado da tensão do GS.

Figura 36 – EQM normalizado da tensão do GS para o caso 1.

Fonte – Elaborada pelo autor.

Por tratar-se do EQM normalizado, quando o seu valor é maior do que 1, significa que aquela variável desviou mais do valor de referência quando comparado ao caso sem armazenadores de energia. Sendo assim, nota-se, pela Figura 36, que o DDPG e o *droop* provocaram um aumento do EQM normalizado, como entre (20 s – 30 s). Neste caso, o aumento do EQM não é tão relevante, pois o valor de tensão não se desviou tanto de 1 p.u. Considerando os instantes onde há maior desvio, podem ser citados os períodos de transição do evento do caso 1, que são 30 s – 40 s e 60 s – 70 s. No primeiro período, tanto o *droop* quanto o DDPG reduziram o valor de EQM. Contudo, no segundo, apenas o DDPG reduziu, enquanto o *droop* provocou um aumento.

Apesar do DDPG ter sido superior nas transições mais relevantes, é importante analisar o valor do EQM para a frequência da rede, representado na Figura 37.

Pela Figura 37, é possível notar que o *droop* e o DDPG foram capazes de reduzir os erros de frequência nos períodos de maior desvio (30 s – 40 s e 60 s – 70 s). Entretanto, vale ser ressaltado que a redução proporcionada pela ação do DDPG foi menor comparada ao *droop*.

Pelas Figuras 36 e 37, pode-se avaliar a variação do EQM normalizado a cada 10 s durante o caso 1. As diferenças entre os desempenhos do DDPG e do *droop* observadas nesse

Figura 37 – EQM normalizado da frequência da rede para o caso 1.

Fonte – Elaborada pelo autor.

gráfico surgem também em uma análise mais geral, a partir do cálculo do EQM em toda a janela de simulação. Dessa forma, pelo cálculo do EQM de toda a janela de simulação de cada caso, a Tabela 8 foi construída.

Tabela 8 – Erro quadrático médio de cada variável controlada para cada caso estudado.

Caso 1	freq.	V_{GS}	V_{B1}	V_{B2}
S/ BESS	1,54e-03	3,16e-03	2,72e-03	2,58e-03
DDPG	1,29e-03	8,63e-04	2,14e-03	2,10e-03
Droop	7,03e-04	1,25e-03	1,55e-03	1,42e-03
Caso 2	freq.	V_{GS}	V_{B1}	V_{B2}
S/ BESS	1,07e-03	5,77e-03	2,26e-03	2,16e-03
DDPG	7,45e-04	2,83e-03	1,73e-03	1,71e-03
Droop	4,86e-04	2,83e-03	1,35e-03	1,26e-03
Caso 3	freq.	V_{GS}	V_{B1}	V_{B2}
S/ BESS	4,38e-04	1,97e-03	1,22e-03	1,09e-03
DDPG	3,27e-04	7,91e-04	1,08e-03	1,04e-03
Droop	1,43e-04	8,16e-04	8,99e-04	7,80e-04

Fonte – Elaborada pelo autor.

Para definir cada variável da Tabela 8, tem-se que $freq$ é a frequência da rede; V_{GS} é a tensão do gerador síncrono; V_{B1} é a tensão de B1; e V_{B2} é a tensão de B2.

Nota-se que os valores de EQM são baixos, até mesmo para o caso sem os armazenadores. Contudo, deve-se destacar que o cálculo desse índice foi realizado por toda a janela de simulação. Ou seja, a percepção do impacto da oscilação causada pelos eventos é amortizado. Ainda, os erros de tensão e frequência foram todos inferiores a 1 p.u., o que justifica os valores da Tabela 8.

Comparando os desempenhos dos controladores por meio do EQM, nota-se que o DDPG garantiu a suavização da frequência e da tensão de forma muito semelhante ao *droop*. Por

exemplo, como pode ser visto no Caso 1, o *droop* foi muito melhor no ajuste de frequência. Porém, o DDPG obteve um menor valor de EQM para a tensão do gerador síncrono. Ademais, os dois controladores conseguem reduzir o EQM do sistema quando comparado ao caso sem os armazenadores.

5.5 Considerações finais

Pelos resultados obtidos, observa-se que foi possível desenvolver um controlador inteligente com desempenho muito similar a um *droop* convencional. Em alguns cenários, embora o *droop* tenha desempenhado melhor, o DDPG conseguiu manter outras variáveis de controle mais próximas ao seu valor de referência. Ainda, pode ser destacada uma característica interessante do controlador inteligente, que foi o de acionar de forma diferente cada um dos bancos de baterias. Adicionalmente, apesar de ser uma técnica de difícil implementação, o DDPG exigiu baixo custo computacional durante sua operação.

Ademais, vale ser destacado que, em todos os casos, o DDPG garantiu um EQM menor ou igual ao *droop* com relação à tensão do GS. Esse é um ponto interessante, pois os dois métodos de controle atuam diretamente nas baterias. Contudo, pelo DDPG é possível inserir o erro da tensão do GS, uma vez que é uma proposta de controle centralizado.

Apesar de, ao final do estudo, não ter sido possível encontrar um ajuste ótimo dos parâmetros do DDPG de forma que seu desempenho superasse o controle *droop*, destaca-se a importância de um estudo maior acerca da estrutura do DDPG, bem como em torno da sintonia de seus parâmetros. Não limitado a isso, existem outros algoritmos de DRL que se encaixam bem a aplicações de controle e que poderiam ser implementados.

CONCLUSÃO

O objetivo principal deste trabalho foi a implementação de um controle supervisão de MRs utilizando uma técnica inteligente. O controlador central deve ser capaz de prever a geração de fontes intermitentes, variação de carregamento e outros eventos, de forma a garantir a estabilidade e confiabilidade do sistema. Para atestar o desempenho do supervisão, foram observadas a frequência e a tensão do PAC de forma que seus valores devem ser mantidos dentro de uma faixa aceitável, definida por normas técnicas.

Encontrou-se na literatura diversas técnicas inteligentes que já foram usadas no controle hierárquico das MRs, como redes neurais, sistemas *fuzzy*, algoritmos baseados em comportamento biológico (por exemplo, PSO e ACO). Além disso, foram encontrados sistemas que usam a combinação dessas técnicas para operarem como um controlador inteligente ou para sintonizar ganhos de controladores PI durante o funcionamento da planta. Nessa revisão, foi encontrada uma técnica que tem obtido bons resultados em outras áreas como a de jogos, carros e robôs autônomos. Essa técnica é a DRL, que une as redes neurais artificiais com a metodologia de aprendizado por reforço. Assim, por existir trabalhos que demonstraram, com êxito, a potencialidade da DRL para cumprir alguns objetivos de controle de MRs, pretende-se aplicar, neste trabalho, essa técnica para operar como o controlador supervisão da rede. Dentro da DRL, foi identificado o DDPG, um algoritmo que tem se mostrado útil para o controle de sistemas dinâmicos contínuos. Por esse motivo, este trabalho se voltou ao uso desse algoritmo em específico.

Neste trabalho de Mestrado, foi realizada toda a modelagem de uma MR, com suas respectivas fontes de geração, no Matlab. A MR consistia em um gerador síncrono, um gerador fotovoltaico e dois bancos de baterias. Enquanto a máquina síncrona é a maior responsável por garantir os níveis de tensão e frequência do PAC, um controlador inteligente, baseado em DDPG, foi implementado de forma a auxiliar essa máquina. A partir das leituras das tensões dos barramentos onde havia geração distribuída conectada e da frequência da rede, o controlador

inteligente envia às baterias o valor de potência ativa e reativa que deve ser gerado.

Os resultados obtidos não demonstraram uma superioridade do método inteligente com relação ao método *droop* que é comumente utilizado. Contudo, foi possível observar que o controlador delegou diferentes níveis de geração para as baterias. Por exemplo, enquanto uma bateria atua mais para uma condição de perda de alimentador e carga, a outra atua para auxiliar no momento de reconexão deste alimentador. Sendo assim, apesar de não ter sido obtido um grande avanço na suavização das oscilações de tensão e frequência, ainda é possível considerar os estudos que trabalham com a DRL de forma a investigar o que ainda pode ser melhorado. Além disso, este estudo de Mestrado abre uma frente de investigação para a melhoria da utilização de algoritmos de aprendizado por reforço, quando comparados aos convencionais, para aplicações na área de controle de microrredes.

6.1 Trabalhos Futuros

Ainda há muito o que ser investigado nessa área de pesquisa. Por exemplo, técnicas como DDPG são muito propensas a estagnarem em pontos de ótimo local. Essa característica foi observada nos resultados obtidos no trabalho, ao notar que após os primeiros 1000 episódios, o agente não conseguiu obter recompensas muito maiores conforme o treinamento prosseguia. Por esse motivo, é ainda preciso investigar se esses resultados surgiram devido às limitações conhecidas do DDPG ou se ainda é possível obter recompensas melhores. Adicionalmente, já existem algoritmos que prometem compensar algumas limitações do DDPG, o que dá abertura para novos estudos a fim de avaliar a aplicabilidade deles para o controle de MRs.

Outras pesquisas futuras podem se voltar ao aprimoramento da validação do desempenho do controlador. Encontrar um método de validação confiável e robusto é muito importante não apenas para o desenvolvimento dos estudos relacionados ao aprendizado por reforço, mas também para garantir a possibilidade de aplicar esses algoritmos em MRs ou sistemas elétricos reais. Complementarmente, é importante que os modelos de MRs sejam cada vez mais diversificados no que se trata aos tipos de geração. Dessa forma, os resultados obtidos retratarão melhor a efetividade desses algoritmos na coordenação de fontes de energia e de garantia de estabilidade das MRs.

REFERÊNCIAS

- ACKERMANN, T.; ANDERSSON, G.; SÖDER, L. Distributed generation: a definition. **Electric Power Systems Research**, Elsevier, v. 57, n. 3, p. 195–204, 2001. Citado na página 6.
- ALFERGANI, A.; ALFAITORI, K. A.; KHALIL, A.; BUAOSSA, N. Control Strategies in AC Microgrid : A Brief Review. 2018. Citado na página 7.
- BARBALHO, P. I. **Análise da Contribuição de um Banco de Baterias para a Regulação de Tensão em um Sistema de Distribuição**. 2018. 114 p. Trabalho de Conclusão de Curso (Bacharel em Engenharia Elétrica), EESC-USP, São Carlos, SP. Citado na página 34.
- BEVRANI, H.; HABIBI, F.; BABAHAYANI, P.; WATANABE, M.; MITANI, Y. Intelligent frequency control in an ac microgrid: Online pso-based fuzzy tuning approach. **IEEE Transactions on Smart Grid**, IEEE, v. 3, n. 4, p. 1935–1944, 2012. Citado na página 2.
- BIDRAM, A.; DAVOUDI, A. Hierarchical structure of microgrids control system. **IEEE Transactions on Smart Grid**, v. 3, n. 4, p. 1963–1976, 2012. ISSN 19493053. Citado na página 8.
- BISHT, M. S. *et al.* Fuzzy based intelligent frequency control strategy in standalone hybrid ac microgrid. In: IEEE. **2014 IEEE Conference on Control Applications (CCA)**. [S.l.], 2014. p. 873–878. Citado na página 2.
- CHEN, G.; FENG, E. Distributed secondary control and optimal power sharing in microgrids. **IEEE/CAA Journal of Automatica Sinica**, IEEE, v. 2, n. 3, p. 304–312, 2015. Citado na página 12.
- CHETTIBI, N.; MELLIT, A. Intelligent control strategy for a grid connected PV/SOFC/BESS energy generation system. **Energy**, v. 147, p. 239–262, 2018. ISSN 03605442. Citado na página 13.
- CIGRE Task Force C6.04.02. **Technical Brochure 575: Benchmark systems for network integration of renewable and distributed energy resources**. [S.l.]: CIGRE, 2014. Citado na página 30.
- COLSON, C. M.; NEHRIR, M. H.; POURMOUSAVI, S. A. Towards real-time microgrid power management using computational intelligence methods. **IEEE PES General Meeting, PES 2010**, p. 1–8, 2010. ISSN 1944-9925. Citado na página 14.
- DING, L.; HAN, Q.-L.; ZHANG, X.-M. Distributed secondary control for active power sharing and frequency regulation in islanded microgrids using an event-triggered communication mechanism. **IEEE Transactions on Industrial Informatics**, Institute of Electrical and Electronics Engineers (IEEE), v. 15, n. 7, p. 3910–3922, jul 2019. Citado na página 13.
- DIVYA, K.; ØSTERGAARD, J. Battery energy storage technology for power systems—an overview. **Electric Power Systems Research**, Elsevier, v. 79, n. 4, p. 511–520, 2009. Citado na página 31.

DUAN, J.; SHI, D.; DIAO, R.; LI, H.; WANG, Z.; ZHANG, B.; BIAN, D.; YI, Z. Deep-reinforcement-learning-based autonomous voltage control for power grid operations. **IEEE Transactions on Power Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 35, n. 1, p. 814–817, jan 2020. Citado na página 15.

ELBANA, M. S.; ABBASY, N.; MEGHED, A.; SHAKER, N. μ pmu-based smart adaptive protection scheme for microgrids. **Journal of Modern Power Systems and Clean Energy**, Springer, p. 1–12, 2019. Citado na página 9.

ENGELBRECHT, A. P. **Computational Intelligence: An Introduction**. [S.l.]: Wiley, 2007. ISBN 978-0-470-03561-0. Citado na página 13.

GU, W.; LOU, G.; TAN, W.; YUAN, X. A nonlinear state estimator-based decentralized secondary voltage control scheme for autonomous microgrids. **IEEE Transactions on Power Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 32, n. 6, p. 4794–4804, nov 2017. Citado na página 13.

GUARNIERI, M.; BOVO, A.; GIOVANNELLI, A.; MATTAVELLI, P. A real multitechnology microgrid in venice: A design review. **IEEE Industrial Electronics Magazine**, Institute of Electrical and Electronics Engineers (IEEE), v. 12, n. 3, p. 19–31, sep 2018. Citado na página 7.

GUO, F.; WEN, C.; MAO, J.; SONG, Y.-D. Distributed secondary voltage and frequency restoration control of droop-controlled inverter-based microgrids. **IEEE Transactions on Industrial Electronics**, Institute of Electrical and Electronics Engineers (IEEE), v. 62, n. 7, p. 4355–4364, jul 2015. Citado na página 12.

HABIBI, F.; SHAFIEE, Q.; BEVRANI, H. Online generalized droop-based demand response for frequency control in islanded microgrids. **Electrical Engineering**, Springer Science and Business Media LLC, v. 101, n. 2, p. 409–420, may 2019. Citado na página 15.

HAN, J.; KHUSHALANI-SOLANKI, S.; SOLANKI, J.; LIANG, J. Adaptive Critic Design-Based Dynamic Stochastic Optimal Control Design for a Microgrid with Multiple Renewable Resources. **IEEE Transactions on Smart Grid**, v. 6, n. 6, p. 2694–2703, 2015. ISSN 19493053. Citado 2 vezes nas páginas 13 e 14.

HAN, Y.; SHEN, P.; ZHAO, X.; GUERRERO, J. M. Control strategies for islanded microgrid using enhanced hierarchical control structure with multiple current-loop damping schemes. **IEEE Transactions on Smart Grid**, IEEE, v. 8, n. 3, p. 1139–1153, 2017. Citado na página 11.

HATZIARGYRIOU, N.; ASANO, H.; IRAVANI, R.; MARNAY, C. Microgrids. **IEEE Power and Energy Magazine**, IEEE, v. 5, n. 4, p. 78–94, 2007. Citado 5 vezes nas páginas 1, 5, 6, 7 e 8.

IEEE. Ieee recommended practice for excitation system models for power system stability studies. **IEEE Std 421.5-2005 (Revision of IEEE Std 421.5-1992)**, p. 1–93, April 2006. Citado na página 38.

JAFARI, M.; GHASEMKHANI, A.; SARFI, V.; LIVANI, H.; YANG, L.; XU, H. Biologically inspired adaptive intelligent secondary control for MGs under cyber imperfections. **IET Cyber-Physical Systems: Theory & Applications**, Institution of Engineering and Technology (IET), v. 4, n. 4, p. 341–352, dec 2019. Citado na página 15.

- JARNUT, M.; WERMIŃSKI, S.; WAŚKOWICZ, B. Comparative analysis of selected energy storage technologies for prosumer-owned microgrids. **Renewable and Sustainable Energy Reviews**, Elsevier, v. 74, p. 925–937, 2017. Citado na página 31.
- KANEKO, T.; UEHARA, A.; SENJYU, T.; YONA, A.; URASAKI, N. An integrated control method for a wind farm to reduce frequency deviations in a small power system. **Applied Energy**, Elsevier, v. 88, n. 4, p. 1049–1058, 2011. Citado na página 6.
- KATIRAEI, F.; IRAVANI, R.; HATZIARGYRIOU, N.; DIMEAS, A. Microgrids management. **IEEE Power and Energy Magazine**, IEEE, v. 6, n. 3, 2008. Citado 4 vezes nas páginas 6, 8, 30 e 36.
- KHOOBAN, M. H.; GHEISARNEJAD, M. A Novel Deep Reinforcement Learning Controller Based Type-II Fuzzy System: Frequency Regulation in Microgrids. **IEEE Transactions on Emerging Topics in Computational Intelligence**, p. 1–11, 2020. ISSN 2471-285X. Disponível em: <<https://ieeexplore.ieee.org/document/8970451/>>. Citado 2 vezes nas páginas 26 e 27.
- KHORRAMABADI, S. S.; BAKHSHAI, A. Intelligent Control of Grid-Connected Microgrids: An Adaptive Critic-Based Approach. **IEEE Journal of Emerging and Selected Topics in Power Electronics**, IEEE, v. 3, n. 2, p. 493–504, jun 2015. ISSN 2168-6777. Disponível em: <<http://ieeexplore.ieee.org/document/6840285/>>. Citado 2 vezes nas páginas 25 e 27.
- KOFINAS, P.; DOUNIS, A. I.; VOUIROS, G. A. Fuzzy Q-Learning for multi-agent decentralized energy management in microgrids. **Applied Energy**, v. 219, n. March, p. 53–67, 2018. ISSN 03062619. Citado 2 vezes nas páginas 25 e 27.
- KUZNETSOVA, E.; LI, Y.-F.; RUIZ, C.; ZIO, E.; AULT, G.; BELL, K. Reinforcement learning for microgrid energy management. **Energy**, Elsevier BV, v. 59, p. 133–146, sep 2013. Citado na página 15.
- LI, Q.; CHEN, F.; CHEN, M.; GUERRERO, J. M.; ABBOTT, D. Agent-based decentralized control method for islanded microgrids. **IEEE Transactions on Smart Grid**, Institute of Electrical and Electronics Engineers (IEEE), p. 1–1, 2015. Citado na página 12.
- LILLICRAP, T. P.; HUNT, J. J.; PRITZEL, A.; HEESS, N.; EREZ, T.; TASSA, Y.; SILVER, D.; WIERSTRA, D. Continuous control with deep reinforcement learning. In: BENGIO, Y.; LECUN, Y. (Ed.). **4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings**. [s.n.], 2016. Disponível em: <<http://arxiv.org/abs/1509.02971>>. Citado 5 vezes nas páginas 23, 24, 39, 43 e 44.
- LIU, W.; GU, W.; XU, Y.; WANG, Y.; ZHANG, K. General distributed secondary control for multi-microgrids with both PQ-controlled and droop-controlled distributed generators. **IET Generation, Transmission & Distribution**, Institution of Engineering and Technology (IET), v. 11, n. 3, p. 707–718, feb 2017. Citado na página 12.
- LU, X.; WANG, J.; GUO, L. Using microgrids to enhance energy security and resilience. **The Electricity Journal**, Elsevier, v. 29, n. 10, p. 8–15, 2016. Citado na página 5.
- LUO, X.; WANG, J.; DOONER, M.; CLARKE, J. Overview of current development in electrical energy storage technologies and the application potential in power system operation. **Applied Energy**, Elsevier, v. 137, p. 511–536, 2015. Citado na página 31.

MENG, L.; SANSEVERINO, E. R.; LUNA, A.; DRAGICEVIC, T.; VASQUEZ, J. C.; GUERRERO, J. M. Microgrid supervisory controllers and energy management systems: A literature review. **Renewable and Sustainable Energy Reviews**, Elsevier BV, v. 60, p. 1263–1273, jul 2016. Citado 2 vezes nas páginas 9 e 10.

MOHAMMADI, F. D.; VANASHI, H. K.; FELIACHI, A. State-space modeling, analysis, and distributed secondary frequency control of isolated microgrids. **IEEE Transactions on Energy Conversion**, IEEE, v. 33, n. 1, p. 155–165, 2018. Citado na página 12.

MOHIUDDIN, S. M.; QI, J. Droop-free distributed control for AC microgrids with precisely regulated voltage variance and admissible voltage profile guarantees. **IEEE Transactions on Smart Grid**, Institute of Electrical and Electronics Engineers (IEEE), p. 1–1, 2019. Citado na página 12.

MOURA, F. A.; CAMACHO, J. R.; CHAVES, M. L.; GUIMARÃES, G. C. Independent power producer parallel operation modeling in transient network simulations for interconnected distributed generation studies. **Electric Power Systems Research**, Elsevier BV, v. 80, n. 2, p. 161–167, feb 2010. Citado na página 36.

NASIRIAN, V.; SHAFIEE, Q.; GUERRERO, J. M.; LEWIS, F. L.; DAVOUDI, A. Droop-free distributed control for AC microgrids. **IEEE Transactions on Power Electronics**, Institute of Electrical and Electronics Engineers (IEEE), v. 31, n. 2, p. 1600–1617, feb 2016. Citado na página 12.

OLIVARES, D. E.; MEHRIZI-SANI, A.; ETEMADI, A. H.; CAÑIZARES, C. A.; IRAVANI, R.; KAZERANI, M.; HAJIMIRAGHA, A. H.; GOMIS-BELLMUNT, O.; SAEEDIFARD, M.; PALMA-BEHNKE, R.; JIMÉNEZ-ESTÉVEZ, G. A.; HATZIARGYRIOU, N. D. Trends in microgrid control. **IEEE Transactions on Smart Grid**, v. 5, n. 4, p. 1905–1919, 2014. ISSN 19493053. Citado 4 vezes nas páginas 1, 2, 8 e 9.

PARIDARI, K.; HAMZEH, M.; EMAMIAN, S.; KARIMI, H.; BAKHSHAI, A. 07-A new decentralized voltage control scheme of an autonomous microgrid under unbalanced and nonlinear load conditions. **Proceedings of the IEEE International Conference on Industrial Technology**, p. 1812–1817, 2013. Citado na página 14.

PASHAJAVID, E.; GHOSH, A.; ZARE, F. A multimode supervisory control scheme for coupling remote droop-regulated microgrids. **IEEE Transactions on Smart Grid**, Institute of Electrical and Electronics Engineers (IEEE), v. 9, n. 5, p. 5381–5392, sep 2018. Citado na página 12.

PSCAD. **Photovoltaic Example**. [S.l.], 2018. Citado na página 38.

QIU, X.; NGUYEN, T. A.; CROW, M. L. Heterogeneous Energy Storage Optimization for Microgrids. **IEEE Transactions on Smart Grid**, IEEE, v. 7, n. 3, p. 1453–1461, 2016. ISSN 19493053. Citado 2 vezes nas páginas 25 e 27.

REPORT, T. Jan 2013 Dynamic Models for Turbine-Governors in Power System Studies. p. 26–27, 2013. Citado na página 37.

REY, J. M.; MARTI, P.; VELASCO, M.; MIRET, J.; CASTILLA, M. Secondary Switched Control with no Communications for Islanded Microgrids. **IEEE Transactions on Industrial Electronics**, v. 64, n. 11, p. 8534–8545, 2017. ISSN 02780046. Citado 2 vezes nas páginas 10 e 11.

- REZAEI, N.; KALANTAR, M. Smart microgrid hierarchical frequency control ancillary service provision based on virtual inertia concept: An integrated demand response and droop controlled distributed generation framework. **Energy Conversion and Management**, Elsevier BV, v. 92, p. 287–301, mar 2015. Citado na página 12.
- SAFARI, A.; BABAEI, F.; FARROKHIFAR, M. A load frequency control using a PSO-based ANN for micro-grids in the presence of electric vehicles. **International Journal of Ambient Energy**, Informa UK Limited, p. 1–13, dec 2018. Citado na página 15.
- SAHOO, S. K.; SINHA, A. K.; KISHORE, N. K. Control techniques in AC, DC, and hybrid AC–DC microgrid: A review. **IEEE Journal of Emerging and Selected Topics in Power Electronics**, Institute of Electrical and Electronics Engineers (IEEE), v. 6, n. 2, p. 738–759, jun 2018. Citado na página 7.
- SIMPSON-PORCO, J. W.; SHAFIEE, Q.; DORFLER, F.; VASQUEZ, J. C.; GUERRERO, J. M.; BULLO, F. Secondary Frequency and Voltage Control of Islanded Microgrids via Distributed Averaging. **IEEE Transactions on Industrial Electronics**, v. 62, n. 11, p. 7025–7038, 2015. ISSN 02780046. Citado na página 11.
- SRIVASTAVA, A. K.; KUMAR, A. A.; SCHULZ, N. N. Impact of distributed generations with energy storage devices on the electric grid. **IEEE Systems Journal**, v. 6, n. 1, p. 110–117, 2012. ISSN 19328184. Citado na página 1.
- SUTTON, R. S.; BARTO, A. G. **Reinforcement learning: An introduction**. [S.l.]: MIT press, 2018. Citado 3 vezes nas páginas 18, 20 e 28.
- TIWARI, G.; MISHRA, R.; SOLANKI, S. Photovoltaic modules and their applications: a review on thermal modelling. **Applied energy**, Elsevier, v. 88, n. 7, p. 2287–2304, 2011. Citado na página 7.
- XIA, Y.; XU, Y.; WANG, Y.; DASGUPTA, S. A Distributed Control in Islanded DC Microgrid based on Multi-Agent Deep Reinforcement Learning. p. 2359–2363, 2020. Citado 2 vezes nas páginas 26 e 27.
- YAZDANI, A.; IRAVANI, R. **Voltage-sourced converters in power systems: modeling, control, and applications**. [S.l.]: John Wiley & Sons, 2010. Citado na página 34.
- YOUNESI, A.; SHAYEGHI, H.; SIANO, P. Assessing the use of reinforcement learning for integrated voltage/Frequency control in AC microgrids. **Energies**, v. 13, n. 5, 2020. ISSN 19961073. Citado 2 vezes nas páginas 26 e 27.
- ZANJANI, M. G. M.; MAZLUMI, K.; KAMWA, I. Application of μ pmus for adaptive protection of overcurrent relays in microgrids. **IET Generation, Transmission & Distribution**, IET, v. 12, n. 18, p. 4061–4068, 2018. Citado na página 9.
- ZENG, G.; RASMUSSEN, T. W.; MA, L.; TEODORESCU, R. Design and control of LCL-filter with active damping for active power filter. In: **2010 IEEE International Symposium on Industrial Electronics**. [S.l.]: IEEE, 2010. Citado na página 34.

CÓDIGO DO DDPG NO MATLAB

Neste Apêndice, encontra-se o código do MATLAB que configura os parâmetros do agente DDPG e os parâmetros de treinamento. Neste algoritmo, estão os parâmetros do agente final utilizado, ou seja, o que possui estrutura mais simples.

```

1 %% Parâmetros do treinamento
2 meanavg = 50; %Tamanho da janela da média móvel
3 maxepisodes = 10000; % Número máximo de episódios simulados
4
5 %% Neurônios por camadas
6 lyr1 = round(400/25);
7 lyr2 = round(300/25);
8
9 %% Criar interface com o ambiente
10 obsInfo = rlNumericSpec([12 1],...
11     LowerLimit ,[-inf] ,...
12     UpperLimit ,[ inf]);
13 obsInfo.Name = observations ;
14 numObs = obsInfo.Dimension(1);
15
16 %Limitar as referências de potências reativas e as ativas em 1;
17 actInfo = rlNumericSpec([4 1], LowerLimit ,[-1 -1 -1 -1] ,...
18     UpperLimit ,[1 1 1 1] );
19 actInfo.Name = powers ;
20 numActions = actInfo.Dimension(1);
21
22 %Construindo o ambiente com o simulink pronto
23 env = rlSimulinkEnv( MRagent_2015a_phasor2 , MRagent_2015a_phasor2/RL...
24     Agent ,obsInfo,actInfo);
25 rng(0)

```

```

26
27 env.ResetFcn = @(in) MResetFcn(in); %Cria um sistema de inicialização...
    aleatória de algumas variáveis do ambiente
28
29 %% Criando o agente
30 % Configuração do Critic
31
32 statePath = [
33     featureInputLayer(numObs, Normalization , none , Name , State )
34     fullyConnectedLayer(lyr1, Name , CriticStateFC2 )];
35 actionPath = [
36     featureInputLayer(numActions, Normalization , none , Name , ...
        Action )
37     fullyConnectedLayer(lyr2, Name , CriticActionFC1 )];
38 commonPath = [
39     additionLayer(2, Name , add )
40     reluLayer( Name , CriticCommonRelu )
41     fullyConnectedLayer(1, Name , CriticOutput )];
42
43 criticNetwork = layerGraph();
44 criticNetwork = addLayers(criticNetwork,statePath);
45 criticNetwork = addLayers(criticNetwork,actionPath);
46 criticNetwork = addLayers(criticNetwork,commonPath);
47 criticNetwork = connectLayers(criticNetwork, CriticStateFC2 , add/in1...
    );
48 criticNetwork = connectLayers(criticNetwork, CriticActionFC1 , add/...
    in2 );
49
50 criticOpts = rlRepresentationOptions( LearnRate ,1e-03, ...
    GradientThreshold ,1);
51
52 critic = rlQValueRepresentation(criticNetwork,obsInfo,actInfo, ...
    Observation ,{ State }, Action ,{ Action },criticOpts);
53
54 % Actor
55
56 actorNetwork = [
57     featureInputLayer(numObs, Normalization , none , Name , ...
        observation )
58     fullyConnectedLayer(4, Name , ActorFC2 )
59     reluLayer( Name , ActorRelu2 )
60     fullyConnectedLayer(numActions, Name , ActorFC3 )
61     tanhLayer( Name , ActorTanh )
62     scalingLayer( Name , ActorScaling , Scale ,max(actInfo.UpperLimit...
        ))];
63
64 actorOpts = rlRepresentationOptions( LearnRate ,1e-04, ...

```

```
        GradientThreshold ,1);
65
66 actor = rlDeterministicActorRepresentation(actorNetwork,obsInfo,...
        actInfo, Observation ,{ observation }, Action ,{ ActorScaling },...
        actorOpts);
67
68 % Configurações do agente
69 agentOpts = rlDDPGAgentOptions(...
70     SampleTime ,Ts,...
71     TargetSmoothFactor ,1e-3,...
72     DiscountFactor ,.99, ...
73     MiniBatchSize ,250, ...
74     ExperienceBufferLength ,1e6);
75 agentOpts.NoiseOptions.Variance = 0.1;
76 agentOpts.NoiseOptions.VarianceDecayRate = 1e-5;
77
78 agent = rlDDPGAgent(actor,critic,agentOpts);
79
80 %% Configurações de treinamento
81
82 maxsteps = ceil(Tf/Ts); % Tf = 100 s -> Tempo de simulação
83 trainOpts = rlTrainingOptions(...
84     MaxEpisodes ,maxepisodes, ...
85     MaxStepsPerEpisode ,maxsteps, ...
86     ScoreAveragingWindowLength ,meanavg, ...
87     Verbose ,false, ...
88     Plots , training-progress ,...
89     StopTrainingCriteria , AverageReward ,...
90     StopTrainingValue ,10, ...
91     SaveAgentCriteria , EpisodeReward , ...
92     SaveAgentValue ,-100);
93
94 %% iniciar treinamento
95 trainingStats = train(agent,env,trainOpts);
```