Transient Stability Preventive Control via Tuning the Parameters of Virtual Synchronous Generators

Xiaoge Huang, Joon-Young Gwak, Lei Yu, Ziang Zhang
Binghamton University – SUNY
Binghamton, NY, USA
xhuang98@binghamton.edu, jgwak1@binghamton.edu,
lyu@binghamton.edu, zhangzia@binghamton.edu

Hantao Cui
Oklahoma State University
Stillwater, OK, USA
h.cui@okstate.edu

Abstract—**This paper presents an optimal preventive control (OPC) method to improve the power system transient stability via tuning the transient parameter of virtual synchronous generators. The novelty of this work is that we formulate the preventive control as an optimization problem so that inverter parameters can be adjusted at the pre-contingency stage. A reinforcement learning (RL)-driven method is proposed to solve the OPC problem with the fault energy-based reward function. An ANDES-based RL environment is also developed. Versatile functions included in the proposed environment have been presented in this paper. The proposed OPC formulation, the RL-driven method, and the fault energy-based reward function are verified on several standard test systems.**

***Index Terms*—Inverter-based resources, virtual synchronous generators, preventive control, reinforcement learning, training environments**

I. Introduction

The generation of electrical energy systems is currently shifting from centralized thermal units to distributed inverter-based resources (IBRs) [1]. In this context, research on integrating IBRs into the grid has attracted much attention. However, recent studies have pointed out that adding inverters into the system can impact the transient behavior of the system [2]. Improving the transient stability of power grids with mixed virtual synchronous generators (VSGs) and synchronous generators is still an open question.

In general, control methods for improving transient stability can be divided into emergency control and preventive control. Emergency control tracks the system's trajectory after the fault is detected and forces the trajectory to converge to a stable equilibrium point (s.e.p). Preventive control adjusts the operating point in the pre-contingency stage to improve the fault ride-through capability of the system [3]. Unlike emergency control that has occurred during the contingency, preventive control needs to consider all potential contingencies. On one hand, the challenge of plant-level emergency control is to have a control algorithm that could recognize a fault event, make decisions based on limited information, and execute the control action in real-time (less than one second). On the other hand, an ISO/TSO-level preventive control that can perform offline simulations based on estimated fault conditions but has a much larger parameter space needs to be searched and a great uncertainty on the pre-fault power flow conditions.

Preventive control can essentially be modeled as an optimization model that maximizes the overall transient stability under various potential contingencies on the system with a given pre-contingency stage [4]. Several studies have analyzed the transient stability metrics at the pre-contingency stage. For example, basin stability [5] and energy function-based methods such as [6] can provide reliable assessments for transient stability. However, their complex mathematical expression and high computation time make them difficult to apply to tractable optimization paradigms.

With the rise of artificial intelligence methods in recent years, data-driven transient stability metrics such as [7] are also proposed for pre-contingency. Based on these data-driven metrics, the transient stability preventive control is usually formulated as a transient stability constrained-optimal power flow (TSC-OPF) problem [8]. TSC-OPF simultaneously satisfies the systematic economic and stability requirements by formulating the stability requirement as mandatory constraints to the optimal power flow model. The controlled variables of TSC-OPF are limited only within steady-state variables. Many works have verified that tuning the transient parameters of generators and VSGs effectively improves transient stability [2]. However, the transient models of the generators/VSGs are highly non-linear, resulting in poor compatibility with high-speed optimization techniques. Therefore, searching for the optimal transient parameters is still challenging.

This paper proposes a novel optimal preventive control (OPC) model to address the above limitations. At the pre-fault stage, the proposed OPC further enhances the transient stability via adjusting the transient parameters of various machines in the power grid based on the optimal power flow. Our contribution is summarized below:

- A transient-based OPC model is proposed in this paper to maximize the transient stability under various potential contingencies on the system at the pre-contingency stage.
- An RL environment for the OPC problem is proposed. On top of the Andes simulator, this environment

2023 IEEE Power & Energy Society General Meeting (PESGM) | 978-1-6654-6441-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/PESGM52003.2023.10253193

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

provides plentiful user-customized functions for OPC-oriented training, testing, and demonstration.

- A reinforcement learning (RL)-driven solution method is proposed for the OPC model. Also, a novel fault energy-based reward function is proposed, achieving satisfactory training and testing performance.

The rest of the paper is organized as follows: Section II provides a new problem formulation of OPC with respect to tuning control parameters' inner VSG models. The RL-driven OPC is also described in this section. Section III introduces the proposed environment. Section IV presents case studies of the RL-driven OPC on several test systems. Finally, the conclusion is provided in Section V.

II. Problem Formulation of RL-driven OPC

A. Optimal Preventive Control Problem

For a power system with fixed network topology and given pre-fault power flow $\mathbf{p}$, its transient stability can be determined by the contingency Δ and the transient parameters $\boldsymbol{x}$ of machines in the grid. Therefore, we have the following mapping:

$$\begin{cases} s_\Delta = f_\Delta(\boldsymbol{x}, \mathbf{p}) \\ \Delta \in \Pi^\Delta \end{cases}, \tag{1}$$

where s_Δ is a selected metric to measure the power system's stability level, such as transient stability index (TSI), critical clearing time (CCT), etc. We assume a set of pre-defined contingencies Π^Δ is given. Equation (1) indicates that $f_\Delta(\cdot, \mathbf{p})$ measures the stability level with machines' transient parameters as input under various contingencies.

The OPC problem can be described in a concise form:

$$\max_{\boldsymbol{x}} \sum_{\Delta \in \Pi^\Delta} s_\Delta \cdot \gamma_\Delta \tag{2a}$$

$$s.t. \quad s_\Delta = f_\Delta(\boldsymbol{x}, \mathbf{p}) \tag{2b}$$

where γ_Δ is a user-customized weight associated with a pre-defined contingency.

B. Candidate Transient Stability Metrics

The selection of transient stability metric s_Δ is a key to achieving satisfactory OPC performance. In this work, two established and one proposed transient stability metrics forms a candidate set. They are CCT, TSI, and fault energy (FE). These three metrics are introduced as follows:

1) Critical clearing time: The longest contingency duration that a system's trajectory can converge to an s.e.p is called CCT. CCT is deterministic if $\boldsymbol{x}$, $\mathbf{p}$, and Δ are given. To apply CCT to (2), we have the following:

$$f_\Delta(\boldsymbol{x}, \mathbf{p}) = \mathrm{CCT}_\Delta(\boldsymbol{x}, \mathbf{p}) \tag{3}$$

2) Transient stability index: TSI is a well-known transient stability metric which is defined as follows:

$$TSI = \frac{360^\circ - |\delta^{max}|}{360^\circ + |\delta^{max}|} \times 100\%, \tag{4}$$

where δ^{max} is the maximum rotor angle difference between any two generators at any point of the system's trajectory. The system is transient stable when TSI >0. To apply TSI to (2)

$$f_\Delta(\boldsymbol{x}, \mathbf{p}) = TSI_\Delta(\boldsymbol{x}, \mathbf{p}) \tag{5}$$

3) Fault energy: the concept of FE is proposed in this work as a novel transient stability metric. FE measures the scale of oscillating sources in the system. The oscillating source of a single generator or VSG i is the power difference

$$P_{f,i}(t) = P_{m,i}(t) - P_{e,i}(t). \tag{6}$$

P_m and P_e are the mechanical and electrical power of the machine. P_f is also known as the fault power of the machine in this work. The FE for a single machine can thusly be defined as follows:

$$E_{f,i}(t) = \left| \int_{t_f}^{t} P_{f,i}(t)\, dt \right|. \tag{7}$$

t_f is the starting time of a contingency. To apply FE to (2):

$$f_\Delta(\boldsymbol{x}, \mathbf{p}) = -\sum_{i \in G} E_{f,i}(t)\,, G = \{machines\}. \tag{8}$$

C. RL-driven OPC

The OPC model (2) is highly non-linear as it contains the differential equations of transient models. Such constraints are incompatible with most high-speed optimization paradigms, e.g., convex optimization. RL is a machine learning technique that enables agents to learn the optimal control policy offline and execute control action online. To solve model (2), we train an RL agent to obtain the OPC policy and use the well-trained agent as the online optimal preventive controller subject to (2).

The RL agent optimizes its control policy by repeated interference with a simulated power system. To summarize, the RL training has the following major components:

Observation state: The pre-contingency operating point $\mathbf{p}$.
Action: The agent is trained the adjust the [*virtual inertial, virtual damping*, *active power set point*, *reactive power setpoint*] in inner control loops of VSGs. For simplicity, we only adjust the parameter on VSGs. Note that our method is general and can be extended to the control of generators.
Reward: After one-step training, a reward will be returned to evaluate the action for the state. The reward in this work can be the calculation result of (3), (5), and (8).
Reset: A new contingency will be randomly generated and assigned to the environment. This process is also known as resetting the environment to ensure the RL agent can deal with various contingencies.
Training Goal: The training goal of the RL agent is to maximize the reward under different contingencies.

III. The Proposed Environment For RL-driven OPC

A. Modeling and Simulation with ANDES

This work utilizes the ANDES simulator for transient stability simulation [9]. Written in Python, ANDES is ideal for research prototyping of differential-algebraic equations-based models and machine learning algorithms. ANDES is also optimized for performance to meet the need for simulating many scenarios, including the efficient KLU sparse linear solver and the support for compiling compute kernels using Numba. In addition, Andes_gym, an ANDES-based RL environment, is available [10].

B. Implementation of the Environment

Fig. 1 shows the architecture and components of the proposed RL training environment. The architecture consists of an ANDES simulator, a reward calculation module, and an OpenAI Gym-based environment port.

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

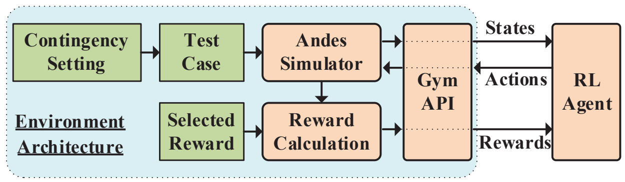

Fig. 1. Architecture of the proposed environment for RL.

The proposed RL environment utilizes ANDES to simulate the system dynamic after the contingency information is given and the test case is prepared. Because all potential contingencies are considered in the OPC, a port for the compactly switch of contingencies is reserved in the environment. Furthermore, considering OPC's objective can be diverse if the environment is open source, we also reserve a port for switching the customized reward function. The reward calculation module calculates the reward based on the selected reward signal and the simulated dynamic. The simulation result and reward are wrapped and organized into OpenAI Gym specifications. OpenAI Gym is a widely used RL toolkit. Gym provides abstractions for environments to adapt to various styles of RL agents.

C. User-Customized Functions

Under the above architecture, the proposed environment provides versatile customized functions for training, testing, and demonstration purposes. Major functions include:

1) RL basic components: initialize(), step(), and reset() are three basic functions for general-purpose RL environments, which are used to set up the environment, conduct one-step training and testing, and reset the environment.

2) CCT calculation: We provide a mature function to search the closest unstable equilibrium point (u.e.p) to the pre-contingency s.e.p in time via the bi-section algorithm.

3) Rotor angle difference plotting: rotor angle difference stability is an important stability criterion for power system transient stability. We provide a function to plot rotor angle differences with the customized reference machine number.

4) A curriculum learning-based training framework: We notice that the designed rewards have limited capability to reflect an action's contribution to the transient stability if the contingency duration can be picked in a wide range. For example, a system with a 0.2 s contingency duration will have a greater chance to return a higher reward than a system with a 0.5 s contingency duration. That is, an agent gets a higher reward in the 0.2 s contingency system doesn't mean its control actions are better than the agent in the 0.5 s contingency system, simply because the tasks for the two agents are different. To this end, we also include a function to invoke a proposed curriculum learning-based training framework. Curriculum learning is a concept in training that trains a model from easier tasks to harder tasks [11]. The proposed framework separates the training into several difficulty levels d_l, of which each d_l corresponds to a narrow interval for sampling random contingency durations. The d_l will gradually increase if the user-customized increase requirement is satisfied. The flowchart of the proposed framework is shown in Fig. 2.

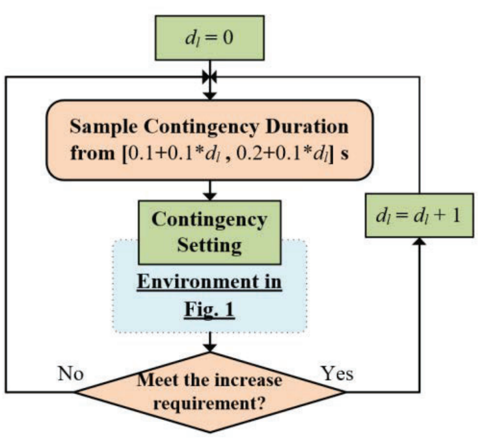

Fig. 2. A curriculum learning-based training framework.

IV. Case Studies

This section verifies the effectiveness of the proposed OPC on two designed test systems. We will first introduce the general experiments setting and then analyze the case studies on three test systems.

The proposed OPC and RL-driven solution is trained and tested on the proposed environment. The transient model and parameters of generators and VSGs used in the simulation can be found in ANDES [9]. The Advantage Actor Critic (A2C)-based RL agents are applied via the RL implementation tool Stable-Baselines. The maximum training episodes is set to 1000. The proposed curriculum learning-based training framework is utilized in all cases, and the difficulty level will increase if the stable ratio is larger than 70% in the last 10 training episodes.

A. Case I: A Single Generator Plus Single VSG system

We start with a simple system in our first case study, the design of which is shown in Fig. 3. The system consists of a classical generator model and a voltage-controlled VSG model. Two transmission lines connect two machines. The generation and load on the two buses are symmetrical. Assume the three-phase-to-ground fault can occur on location 1 or 2. The 5% and 90% in Fig. 3 are ratios of electrical distance. An RL-based controller is applied to preventively control the transient parameters of VSG2.

In RL training, it is natural to track the learning progress of the agent by monitoring the reward varies with the training episode. Fig. 4 shows the reward vs. episodes in training using the TSI-based reward function on the G+VSG system. The moving average reward for the latest 10 episodes is also included to remove inevitable perturbations on reward caused by exploration. As shown, the TSI-based RL agent has converged to an optimum after 600 episodes of training. Besides, the difficulty level of the training will be gradually increased as the curriculum learning framework is utilized. We represent increasing difficulty levels in the figure by using different-colored backgrounds. Since different difficulty levels

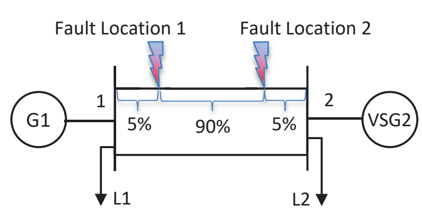

Fig. 3. The designed single generator plus single VSG (G+VSG) system.

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

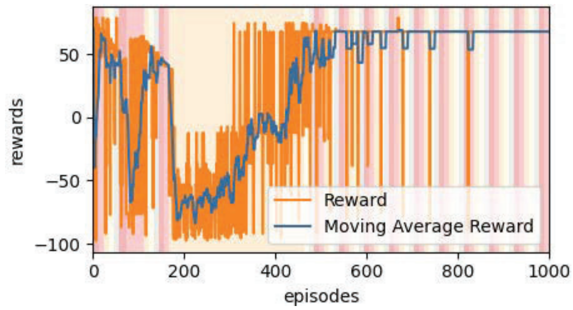

Fig. 4. The reward vs. episodes in training using TSI-based reward on the G+VSG system. Different-colored backgrounds denote the increasing difficulty levels in the proposed curriculum learning framework.

TABLE I. CCT RESULTS ON THE G+VSG SYSTEM

Fault Location	Reward Function	Training Time	Original CCT (s)	CCT after OPC(s)
1	FE	49 minutes	0.71	2.71
1	TSI	45 minutes	0.71	3.00
2	FE	37 minutes	0.55	2.92
2	TSI	40 minutes	0.55	3.00

directly correspond to different CCTs, Fig. 4 also enables users to predict the current CCT of the system in the training process.

We train RL-based optimal preventive controllers with different reward functions and test the well-trained controllers under various contingencies. CCT is one of the most reliable and straightforward transient stability metrics. However, the CCT is computationally complex, and training time can be multiplied relative to using TSI or FE as the reward function. Therefore, we train controllers with TSI and FE as the reward and evaluate the trained agents using CCT as the transient stability metric. The evaluation results are summarized in Table I. As shown, the G+VSG system with a trained controller can improve at least 2 seconds in CCT.

To better understand the impact of the proposed OPC on transient dynamics, we also compare the system's time domain simulations before and after the application of the OPC.

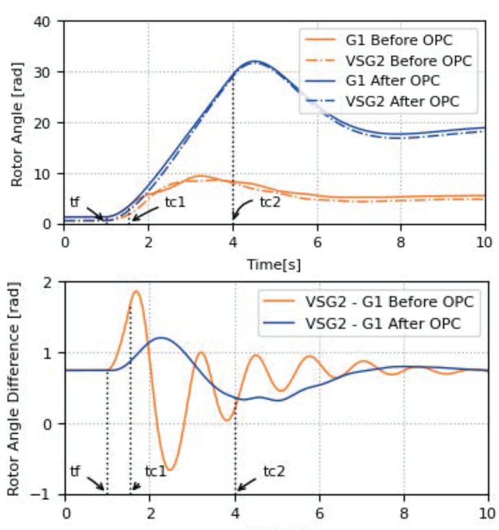

Fig. 5. The impact of OPC on the when fault durations on the G+VSG system equal their own CCTs: Comparison of rotor angle dynamic (Upper). Comparison of rotor angle difference dynamic (Lower).

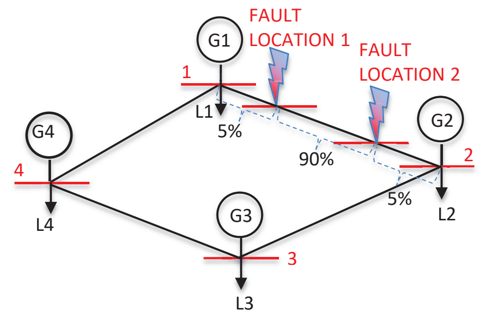

Fig. 6. The designed 4-bus 4-machine (4B4M) system.

Limited by space, the last tested case in Table I (Fault location = 1, Reward = TSI) will be used as an example to show time domain results. We compare the dynamics with and without OPC when fault durations equal their own CCTs in Fig. 5. The t_f denotes the starting time of the fault. t_{c1} is the fault clearing time without OPC, and t_{c2} is the clearing time with OPC. Since t_{c2} is much larger than t_{c1}, the oscillation on angles inevitably increases on the system supported by OPC. However, OPC can still mitigate the oscillation on angle difference.

B. Case II: A 4-Bus 4-Machine System

The proposed OPC is also tested on a 4-bus 4-machine (4B4M) system, as shown in Fig. 6. The system consists of four classical generator models. Four transmission lines connect four machines. The generation and load on four buses are symmetrical. Assume the three-phase-to-ground fault can occur on location 1 or 2. An RL-based controller is applied to control the transient parameters of one or more than one VSG.

We assume every generator in the system can be replaced by a VSG model supported by OPC. For example, controlled VSG on bus 2 denotes that an OPC VSG replaces the generator on bus 2. If the system contains more than one VSG, the RL-based controller will be trained and tested to control multiple VSGs simultaneously. As shown in Table II, OPC can effectively improve CCTs on the 4B4M system.

We also compare the time domain simulations before and after the OPC was applied to the 4B4M system. The last tested case in Table II (Controlled VSGs on Bus = [2, 3, 4], Fault location = 1, Reward = TSI) is used to demonstrate time domain results. Fig. 7 compares the dynamics with and without

TABLE II. CCT RESULTS ON THE 4B4M SYSTEM

Controlled VSGs on Bus	Fault on Bus	Reward Function	Training Time (minute)	Origin CCT (s)	CCT after OPC(s)
2	1	FE	29	0.74	0.80
2	1	TSI	33	0.74	1.85
2	2	FE	52	1.31	1.58
2	2	TSI	49	1.31	2.22
2, 4	1	FE	32	0.54	1.14
2, 4	1	TSI	39	0.54	0.97
2, 4	2	FE	46	1.59	2.99
2, 4	2	TSI	57	1.59	2.99
2, 3, 4	1	FE	55	0.32	0.71
2, 3, 4	1	TSI	64	0.32	1.56

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

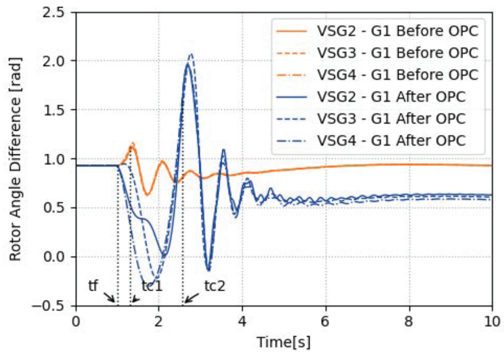

Fig. 7. Comparison of angle difference dynamics in the 4B4M system.

OPC when fault durations on the 4B4M system equal their own CCTs. As shown, the two tested systems can be stabilized in almost the same amount of time, while the fault duration of the system with OPC is much longer than that of the system without OPC. It is also observed that the oscillation increase in angle differences is inevitable when the system gets larger compared with the result in the G+VSG system.

C. Case III: The 140-bus NPCC System

The proposed OPC is also verified with two scenarios on the NPCC system [9], as shown in Fig. 8. In scenario 1, the generator on bus 79 is replaced as a VSG, and the fault occurs near bus 79. In scenario 2, the generator on bus 36 is replaced as a VSG, and the fault occurs near bus 36. Fig. 9 shows a simple case of the NPCC system with a 0.3s fault in scenario 2. Initially, the system lost synchronization. But it can be stabilized after applying the proposed OPC. We further tested the CCTs of the two different scenarios. Table III shows that the OPC can effectively improve CCTs in the tested scenarios.

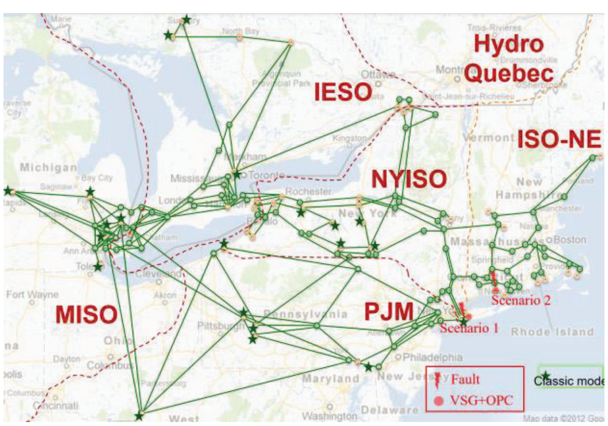

Fig. 8. The 140-bus NPCC test system.

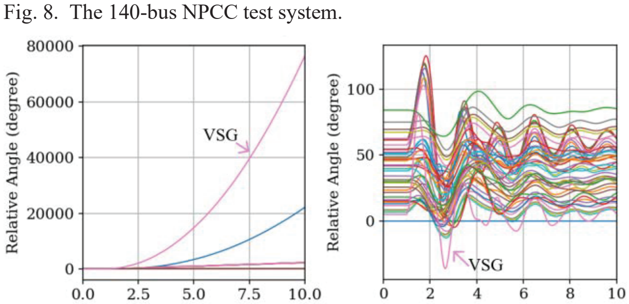

Fig. 9. 50-machine angle difference dynamics of the NPCC system: Dynamics without OPC (left). Dynamic with OPC (right).

TABLE III. CCT RESULTS ON THE NPCC SYSTEM

Scenario#	Original CCT (s)	CCT after OPC(s)	Improvement
1	0.145	0.233	60.69 %
2	0.238	0.378	58.82 %

V. CONCLUSION

This paper proposed an OPC problem formulation, which provides a new method of power system preventive control. In addition to adjusting the steady-state operating point (P, Q set points), the transient model parameters (virtual inertia and virtual damping of the VSG) are also allowed to be tuned at the pre-contingency stage via our model. An RL-driven method is proposed to solve the OPC problem using a novel reward function called the fault energy function. We have verified the proposed method on various systems. An RL training environment for OPC is developed. User-customized functions embedded in the proposed environment for training, testing, and demonstration are presented in this paper.

ACKNOWLEDGMENT

This work was supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office Award Number DE-EE0009341.

REFERENCES

[1] M. Pertl, T. Weckesser, M. Rezkalla, and M. Marinelli, "Transient stability improvement: a review and comparison of conventional and renewable-based techniques for preventive and emergency control," *Electr. Eng.*, vol. 100, no. 3, pp. 1701–1718, Oct. 2017.

[2] E. Hammad, A. Farraj and D. Kundur, "On Effective Virtual Inertia of Storage-Based Distributed Control for Transient Stability," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 327-336, Jan. 2019.

[3] S. Yang, Z. Hao, B. Zhang and M. Hojo, "An Accurate and Fast Start-Up Scheme for Power System Real-Time Emergency Control," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3562-3572, Sept. 2019.

[4] D. Gan, R. J. Thomas, and R. D. Zimmerman, "Stability-constrained optimal power flow," *IEEE Trans. Power Syst.*, vol. 15, no. 2, pp. 535-540, May 2000.

[5] Z. Liu, X. He, Z. Ding and Z. Zhang, "A Basin Stability Based Metric for Ranking the Transient Stability of Generators," *IEEE Trans. Ind. Inf.*, vol. 15, no. 3, pp. 1450-1459, Mar. 2019.

[6] P. J. Hart, M. Gong, H. Liu, Z. Chen, Y. Zhang and Y. Wang, "Provably-Stable Overload Ride-Through Control for Grid-Forming Inverters using System-Wide Lyapunov Function Analysis," *IEEE Trans. Energy Convers.*, 2022.

[7] L. Zhu and D. J. Hill, "Networked Time Series Shapelet Learning for Power System Transient Stability Assessment," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 416-428, Jan. 2022.

[8] Y. Fu, X. Zhang, L. Chen, Z. Tian, K. Hou and H. Wang, "Analytical Representation of Data-driven Transient Stability Constraint and Its Application in Preventive Control," *J. Mod. Power Syst.*, vol. 10, no. 5, pp. 1085-1097, Sept. 2022.

[9] H. Cui, F. Li, and K. Tomsovic, "Hybrid Symbolic-Numeric Framework for Power System Modeling and Analysis," *IEEE Trans. Power Syst.*, vol. 36, no. 2, Mar. 2021.

[10] H. Cui and Y. Zhang, "Andes_gym: A Versatile Environment for Deep Reinforcement Learning in Power Systems," in *Proc. PES General Meeting*, Denver, USA, Jul. 2022, pp. 1–5.

[11] X. Huang, Z. Ding and Z. Zhang. "A Guided Deep Reinforcement Learning Method For Distribution Voltage Regulation via Battery Systems." in *Proc. 2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2021, pp. 1-5.