Reinforcement Learning for Volt-Var Control: A Novel Two-stage Progressive Training Strategy

Si Zhang, Mingzhi Zhang, Rongxing Hu, David Lubkeman, Yunan Liu, and Ning Lu
Department of Electrical and Computer Engineering
North Carolina State University, Raleigh, NC
{szhang56, mzhang33, rhu5, dllubkem, yliu48, nlu2}@ncsu.edu

***Abstract*—This paper develops a *reinforcement learning* (RL) approach to solve a cooperative, multi-agent Volt-Var Control (VVC) problem for high solar penetration distribution systems. The ingenuity of our RL method lies in a novel two-stage progressive training strategy that can effectively improve training speed and convergence of the machine learning algorithm. In Stage 1 (individual training), while holding all the other agents inactive, we separately train each agent to obtain its own optimal VVC actions in the action space: {consume, generate, do-nothing}. In Stage 2 (cooperative training), all agents are trained again coordinatively to share VVC responsibility. Rewards and costs in our RL scheme include (i) a system-level reward (for taking an action), (ii) an agent-level reward (for doing-nothing), and (iii) an agent-level action cost function. This new framework allows rewards to be dynamically allocated to each agent based on their contribution while accounting for the trade-off between control effectiveness and action cost. The proposed methodology is tested and validated in a modified IEEE 123-bus system using realistic PV and load profiles. Simulation results confirm that the proposed approach is robust and computationally efficient; and it achieves desirable volt-var control performance under a wide range of operation conditions.**

***Index Terms*—Distribution systems, inverter-based resources, machine learning, multi-agent, progressive training, reinforcement learning, smart inverter, volt-var control.**

I. Introduction

Solar photovoltaic (PV) systems equipped with smart inverters have superior continuous reactive power (Q) regulation capabilities compared with capacitor banks and voltage regulators. Therefore, developing control strategies for distributed PV systems to provide Volt-Var control (VVC) is gaining increasing attention. In general, there are three popular VVC approaches: rule-based, optimization-based, and more recently, machine learning-based. Although rule-based approaches are widely used in the field due to the ease of implementation, they lack the ability to adapt to fast-changing operational conditions. The major drawbacks of optimization-based approaches are their strict requirement of accurate network models and complex computational platforms for implementation. Futuremore, the computational complexity increases exponentially as the system scale (e.g. number of controllable devices) increases.

Machine learning, especially *reinforcement learning* (RL), has been proven effective to generate optimal voltage control policies via offline and online training [1-3]. Comparing to conventional rule-based VVC controls, main advantages of the RL-based VVC are its ease of implementation and high adaptability in a fast-changing operational environment. Zhang *et al.* [1] and Sun and Qiu [2] proposed multi-agent reinforcement learning (MARL) solutions for training VVC agents in both centralized and decentralized environments. However, under this setting, the decentralized agent does not have learning capability - it only executes. Wang *et al.* [3] formulated the VVC problem as Markov game for solving the Voltage Violation problem using (one-shot) static environmental data as an episode. In order for agents to evolve their policies in response to a nonstationary environment, Lowe *et al.*. [4] developed a multi-agent deep deterministic policy gradient (MADDPG) method. However, inferring other agents' actions requires training additional neural networks, causing the design of VVC increasingly complex when the number of VVC agents increases.

Centralized RL-based control design approaches often suffer from the so-called *curse of dimension*. Convergence and stability in training are usually difficult to achieve when many agents need to coordinate their operations in a fast changing environment. For example, an common scenario that often occurs in VVC training is passing clouds accompanied by rapid load changes in a distribution circuit with many PV systems.

To address the aforementioned issues, in this paper, we develop a two-stage RL approach to train multiple VVC agents progressively on a distribution feeder. Our contributions are two-fold. *First*, we propose a novel reward design and allocation mechanism to account for the contributions of all agents; we aim to trade-off between control effectiveness and cost. In particular, aach VVC agent can take one of three basic actions: "generate-Q", "consume-Q" and "do-nothing". The system's performance score is calculated by the degree of system-wide voltage violations for assessing VVC performance achieved by all VCC agents. Immediate reward is defined as the score of take-an-action (i.e., generate or consume) minus the score of do-nothing. At the agent-level, the action cost is calculated according to the efforts committed by an agent. The "do-nothing" reward allows us to include "do-nothing" as a "wise" action when the value of take-an-action diminishes. Note that the agent-level reward plays an important role in a decentralized, co-operative training environment. Rewards are not shared uniformly among all agents but rather dynamically

This project is supported by the DOE ARPA-E EDEGPRO project.

assigned according to their efforts, which takes into account an agent's contribution while considering the cost for taking an action.

Second, we propose a novel two-stage, progressive training strategy. In Stage 1 (individual training), each agent is learn to take three basic control actions: "generate-Q", "consume-Q", and "do-nothing", assuming all other agents are inert. Because the training of the agents can be conducted in parallel, the training time is unaffected by the number of agents. In absence of interventions from the other agents' actions, the agent currently being trained can focus on selecting one of the three actions with a fixed Q. This guarantees that our algorithm converges fast and is robust. In Stage 2 (cooperative training), as all agents have gained understanding of when to "generate-Q", "consume-Q", and "do-nothing", the training can now focus on learning the "optimal" magnitude of Q an agent needs to provide in the presence of the other agents, i.e. learning coordination only. Thus, the training complexity is significantly reduced. Our results show that this 2-stage, progressive training approach is computationally much more efficient than the state-of-the-art methods, leading to faster convergence and more robust performance.

II. Problem Formulation

A. Assumptions

First, all actions are taken in fast control intervals (i.e., at 1- or 5- minute), so they are immediately observable to all VVC agents at time t. We can use the persistence model instead of policy inference to predict the other agents' actions by assuming that observations of the environment at $t-1$ are sufficient to predict the states of the environment at t. *Second*, the communication among agents is via the system operator, who is responsible for letting a VVC agent "know" of required information at time t (e.g., actions taken by all other agents at $t-1$). However, the other agents' actions at t are hidden (and only revealed at the next step) so each agent will make its own decision independently. The parameters of the policy network of a VVC agent are also unknown to other agents. *Third*, there is no other VVC devices on the feeder so that the PVs are the only resources for reactive regulation. *Fourth*, the only objective of an VVC agent is to control the nodal voltages to be within the defined operational range.

B. Problem Formulation

To formulate the VVC problem as a Markov decision process (MDP), in that we define the global state, $\mathcal{S}^t$, partial observation $\mathcal{O}_i^t$, and action set, $\mathcal{A}$, of the i^{th} VVC agent at time t as:

$$\begin{aligned}
\mathcal{S}^t &:= \begin{bmatrix} \mathcal{V}^t & \mathcal{P}_{\text{pv}}^t & P_{\text{feeder}}^t & Q_{\text{feeder}}^t \end{bmatrix} && (1)\\
\mathcal{O}_i^t &:= \begin{bmatrix} \mathcal{V}^t & P_{\text{pv},i}^t & P_{\text{feeder}}^t & Q_{\text{feeder}}^t & \mathcal{A}^{t-1} \end{bmatrix} && (2)\\
a_i^t &:= Q_{\text{pv},i}^t && (3)\\
\mathcal{A}^{t-1} &= \left[a_1^{t-1} \dots a_i^{t-1} \dots a_N^{t-1}\right], \quad \forall i \in [1, N] && (4)\\
\mathcal{A}^t &= \left[a_1^t \dots a_i^t \dots a_N^t\right], \quad \forall i \in [1, N] && (5)\\
\mathcal{V}^t &= \left[V_1^t \dots V_k^t \dots V_M^t\right], \quad \forall k \in [1, M] && (6)
\end{aligned}$$

where $\mathcal{V}^t$ is the nodal voltage set; $\mathcal{P}_{\text{pv}}^t$ is the active power output Set of PV farms at step t; $P_{\text{pv},i}^t$ is the i^{th} PV real power output at t, P_{feeder}^t and Q_{feeder}^t are the active and reactive power output at the feeder head t, respectively; a_i^t is the action taken by the i^{th} agent for generating (positive) or consuming (negative) reactive power of $Q_{\text{pv},i}^t$; M is the number of nodes being monitored; N is the number of VVC agents. Note that $\mathcal{S}^t$ represents the global view of the environment and $\mathcal{O}_i^t$ describes the agent's local view of the environment.

Note that $\mathcal{A}^{t-1} = 0$ when $t = 1$. The action space of a centralized VVC controller is an action set because all agents' actions need to be considered. However, the action space of the i^{th} distributed VVC agent is a scalar, which is the reactive power output of the i^{th} PV farm, $Q_{\text{pv},i}^t$ (see (3)).

C. VVC Performance Score Calculation

The ANSI standard requires the distribution system voltage to be maintained within the interval $[V^-, V^+]$, with $V^- = 0.95$ and $V^+ = 1.05$ p.u. However, a utility may choose to hold system voltage to be within another designated interval $[V^{\text{Hlim}}, V^{\text{Llim}}]$. Inspired by [3], we revised the control target from a single voltage reference to a set of piece-wise linear score functions, as shown in Fig. 1. Note that s_k^t is the voltage score calculated for node k at time t.

$V_k^t(p.u)$	s_k^t
$[V^{\text{Llim}}, V^{\text{Hlim}}]$	1.0
$[V^{\text{Hlim}}, V^+]$	$\frac{1.05 - V_k^t}{1.05 - V^{\text{Hlim}}}$
$[V^-, V^{\text{Llim}}]$	$\frac{V_k^t - 0.95}{V^{\text{Llim}} - 0.95}$
$[1.05, 1.10]$	$\frac{5 \times (V_k^t - V^{\text{Hlim}})}{V^{\text{Hlim}} - 1.10}$
$[0.90, 0.95]$	$\frac{5 \times (V^{\text{Llim}} - V_k^t)}{0.9 - V^{\text{Llim}}}$
$[1.10, +\infty]$	-1×5
$[-\infty, 0.90]$	-1×5

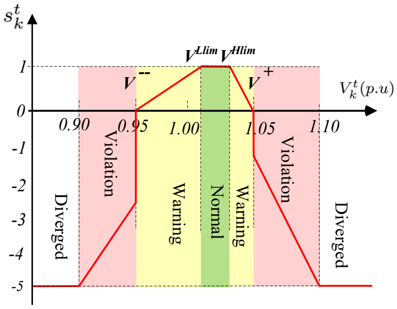


Fig. 1. The setup of the nodal voltage score curve.

The system score, $Score_A^{AS}$, is defined as the average nodal voltage score, that is,

$$Score_A^{AS} = \frac{1}{M} \sum_{k=1}^{M} s_k^t, \tag{7}$$

where AS is the joint action space and A is the joint action set. Note that the score curve is capped at 5.0 after the system voltage drops below 0.9 p.u. or surpasses 1.10 p.u.

D. Design and Allocation of Reward

The system-level reward r_s is defined as

$$r_s := Score_A^{AS} - Score_{DN}^{AS}, \tag{8}$$

where DN indicates the action of "do-nothing".

Our reward definition follows the idea of the *Advantage Actor Critic* (A2C) method [5]. We deduct the actual reward from a baseline to reduce the variance of policy gradient so that policy network can be trained easily. The baseline score can be computed from (7). This advantage reward can be explicitly formulated and meaningful to show the effectiveness of taking actions.

In the single-agent setting, the agent can seek effective actions by bench-marking against a predetermined baseline action. However, when there are many VVC agents in the system, the dimensionality of the (joint) action space increases drastically. In addition, intervention between actions taken by different agents makes the baseline action selection much more complicated. For simplicity, we use the action of "do-nothing" as a unified baseline action in multi-agent setting. In power distribution systems, "do-nothing", in absence of voltage violations, is indeed often preferred as a baseline action for controlling a VVC device.

III. TWO-STAGE PROGRESSIVE TRAINING

A. Stage 1: Individual Training

The goal of the first stage is to train a VVC agent by considering two simplified basic control strategies: i) when the voltage violation cannot be alleviated by it action, do-nothing, and ii) when taking an action, what is the polarity of the action, i.e., generate (+) or consume (-).

The agent-level reward for the i^{th} agent in stage 1, $r_{1,i}$, when taking an action, a_i^t, is expressed as

$$\begin{aligned} Cost_i &= w_{\text{cost}} \times |Q_i^t| && (9)\\ r_{DN} &= \begin{cases} 1 \times 10^{-3} & |a_i^t| \leq a^{\text{th}} \\ 0 & \text{otherwise} \end{cases} && (10)\\ r_{1,i} &= r_s - Cost_i + r_{\text{DN}} && (11) \end{aligned}$$

where $Cost_i$ is the "cost" for taking an action for the i^{th} agent, w_{cost} is the weighting factor for computing $Cost$ from $|Q_i^t|$, a^{th} is the action threshold, and r_{DN} is the do-nothing reward.

Note that (11) aims to trade-off between reward and cost. The r_{DN} term can not only encourage the agent to take no action while the effectiveness of its effort diminishes, it also forms a non-action zone to avoid unnecessary oscillatory actions. The above setting of the reward structure constitutes one of the major novelties of our approach.

When regulating nodal voltages on a distribution feeder, each VVC agent has an effective control range determined by the network topology and the location of the PV farm. Thus, when an agent's action is ineffective or only marginally effective for mitigating voltage violations, the optimal strategy is do-nothing. However, when multiple agents are being trained jointly in a nonstationary environment, two main challenges arise: i) the lack of appropriate baseline actions for assessing performance improvements and ii) the lack of a fair performance-driven reward allocation mechanism for each agent. Consequently, the training process becomes lengthy and unstable. Convergence to the optimal VVC control and coordination strategy for all agents is therefore difficult to achieve.

Main advantages. If there exists only one VVC, the training converges quickly. This is because the radial distribution network topology ensures a relatively linear V-Q relationship. Furthermore, as the agent receives full credits/penalties for its action as specified in (11), learning the polarity of an action is straightforward. As the first stage training can be conducted in parallel, having multiple agents will not slow down the training process. After the first-stage training, all agents should "understand" when their actions are effective. This is a very important feature for reward allocation in the second-stage training, because all action-taking agents are considered effective contributors. Through an appropriate credit-sharing mechanism, agents can learn to contribute the right amount of Q in presence of other agents.

B. Stage 2: Cooperative Training

The goal of the second stage is to train all VVC agents jointly in the same environment so that each VVC agent can learn to *generate/consume its own share* of reactive power when coordinating with the other agents for reducing nodal voltage violations. Our assumption is that, after stage 1, each agent has learned to take only effective actions, i.e., an agent will be inert when its action will not help to alleviate voltage violations and will know when to generate/consume Q. As defined in (2), in the second stage, the actions taken by all agents, $\mathcal{A}^{t-1}$, at $t-1$ will serve as an input for the observation space of all VVC agents at t.

In the second stage, the agent-level action reward, $r_{2,i}$, is calculated as

$$r_{2,i} := CF_i \times r_s - Cost_i \tag{12}$$

where the cost is calculated using (9) and the contribution factor CF_i can be calculated as

$$CF_i := \frac{|Q_i^t|}{\sum_{j=1}^{N} |Q_j^t|} \tag{13}$$

Note that here we use the absolute value of Q because there may be cases in which one agent is generating Q for boosting its local voltage while another agent is consuming Q for suppressing its local voltage. In this case the two agents are collaborating with each other to remove the voltage violations. Let $a_{\text{S}_1,i}$ and $a_{\text{S}_2,i}$ be the action outputs by the stage-1 policy network, S_1, and stage-2 policy network, S_2 by agent i, respectively. The agent's final action, $a_{2,i}$, is determined by

$$a_{2,i} := sign(a_{\text{S}_1,i}) \times \mathbb{1}(|a_{\text{S}_1,i}| > a_{th}) \times a_{\text{S}_2,i} \tag{14}$$

Note that (14) generates desired control strategies at two levels: i) Stage-1 policy network determines a "raw" action: whether an action is needed and if so, its polarity; ii) Stage-2 policy network provides a "complete" action by prescribing the magnitude of the action when such an action is needed.

C. Algorithm Implementation

We solve the distributed RL problem using DDPG proposed in [6] following the workflow shown in Fig. 2. Note that $\mathcal{O}_{\text{S}1,i}$ and $\mathcal{O}_{\text{S}2,i}$ are the partial observations to agent i in stages 1 and 2, respectively. In (2), for $\mathcal{O}_{\text{S}1,i}$, $\mathcal{A}^{t-1} = [a_i^{t-1}]$ since there is no other agent in the system; for $\mathcal{O}_{\text{S}2,i}$, $\mathcal{A}^{t-1} = [a_1^{t-1}...a_i^{t-1}...a_N^{t-1}]$, given the second assumption in Section II-A.

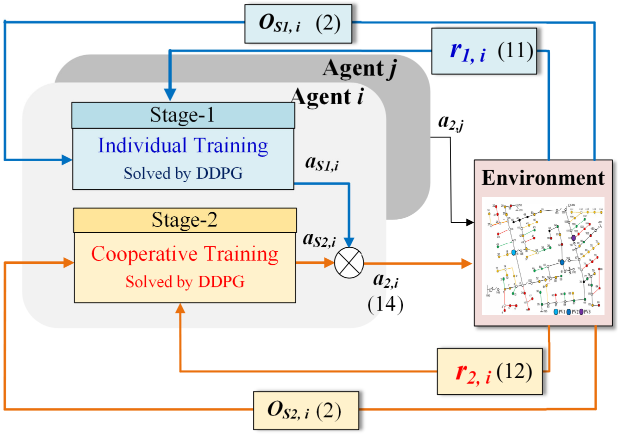


Fig. 2. An illustration of the two-stage training process using DDPG.

IV. Numerical Studies

The training is conducted on a testbed developed using the topology of the IEEE 123-bus system, as shown in Fig. 3. The back-end of the environment is OpenDSSDirect running on Python. The RL agents are trained using Pytorch. The annual load and PV data are generated from the PECAN street data set [7]. We consider a 5-minute control interval and a 30-minute learning episode. Table I lists the locations and capacitys of all PV farms. All PV inverters are oversized so $S_{pv} = 1.08 P_{pv}$. According to [8], the inverter regulates Q within $[-44\%, 44\%]$ of the PV rated capacity, S_{pv}. Voltage regulators set at 1:1 ratio mode and is inert during the training. The utility preferred voltage operation range is determined by $V^{\text{Hlim}} = 1.03$ p.u. and $V^{\text{Llim}} = 1.01$ p.u., the values of which are within the ANSI limit.

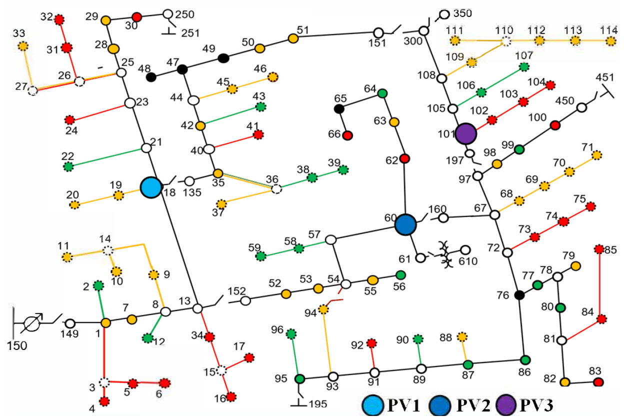


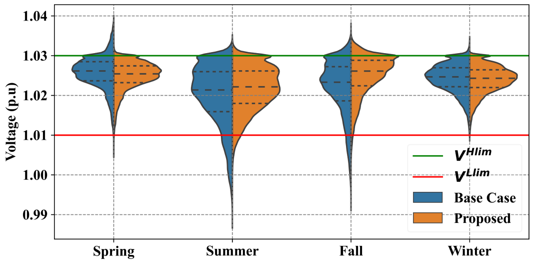
Fig. 3. Configuration of the training environment. (Test feeder topology: modified IEEE 123-bus syste. Green, red, blue, and black lines: a, b, c, and 3-phase circuits, respectively. Empty circles: buses without loads.)

TABLE I
Location and Capacity of the PV Farms

Inverter	Connected Bus(Phase)	Installment Capacity
PV1	18(a,b,c)	800 kW
PV2	60(a,b,c)	600 kW
PV3	101(a,b,c)	300 kW
Total PV capacity		1700 kW

A. RL-based VVC Performance in four Seasons

In each season, 20 days are selected for training and 3 days for testing. The base case is the "do-nothing" case. The nodal voltage distributions and the three-day-average voltage scores of the base case and the proposed VVC cases for the four seasons are summarized in Fig. 4. The left side of the violin plot represents the base case and the right side is the VVC case. The base case results show that in summer, the nodal voltages often drop below V^{Llim} (i.e. 1.01 p.u.) and in spring, the second worst, the nodal voltages often go above V^{Hlim} (i.e. 1.03 p.u.). As shown by the inner quartiles of the plots, the nodal voltage for the VVC case are significantly improved and most of the time the nodal voltages are within the preferred operation zone. As shown in Fig. 4, in winter, the voltage violations are rare so for the remaining studies, we only show the results obtained in the summer and spring.



Voltage Score	Spring	**Summer**	Fall	Winter
Basecase	0.987563	**0.967055**	0.978325	0.999242
Stage1	0.999850	**0.992984**	0.993532	0.999951
Stage2	0.999513	**0.995537**	0.991858	0.999952

Fig. 4. Distributions of nodal voltages and summary of voltage scores.

B. Performance Comparison with the Decentralized VVC

The parameters of a set of conventional inverter-based decentralized VVC control curves [8] are shown in Fig. 5. As shown in Table II, the conventional decentralized VVC takes the least number of actions, which is measured by the cumulative Q consumption, $\sum Q$. However, it receives the lowest Voltage score, showing an inferior voltage regulation performance. Stage-1 policy does not consider coordination. Thus, PV1 always generates Q, causing more V^{Hlim} violations. Stage-2 policy has the highest voltage score, showing superior VVC control performance. By coordinating with other agents, $\sum Q$ is significantly reduced in stage-2.

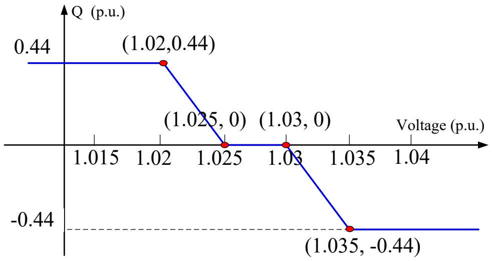


Fig. 5. Conventional decentralized Volt-Var control curve.

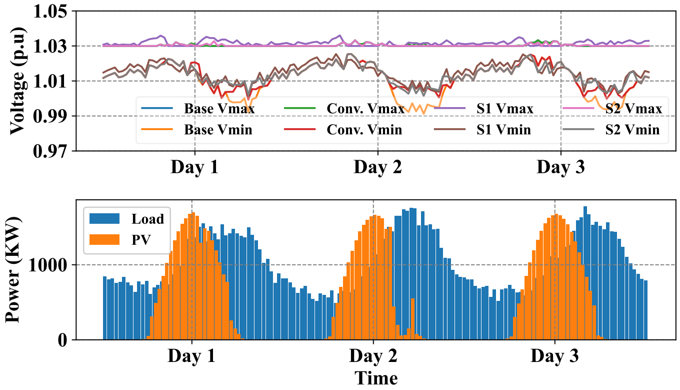


Fig. 6. Voltage, PV, and load profiles in the three testing days.

TABLE II

VVC PERFORMANCE COMPARISON (THE SUMMER CASE) EVALUATION

Algorithm	Voltage Score	Q_{total} (kVAR)	Q_{pv1} $(p.u)$	Q_{pv2} $(p.u)$	Q_{pv3} $(p.u)$
Base Case	0.98756	-	-	-	-
Conventional	0.98995	93.859	0.01688	0.07611	0.09242
Stage-1	0.99286	452.43	0.41250	0.12527	0.04583
Stage-2	**0.99556**	**144.03**	**0.01660**	**0.14361**	**0.11305**

If some nodal voltages fall outside of the designated interval $[V^{\mathrm{Llim}} V^{\mathrm{Hlim}}]$ in a control interval, we consider this interval to be a voltage violation event. Then, we compare the duration of such voltage events in four use cases: base case, conventional, stage-1 policy, and stage-2 policy in the summer season. Table III and Fig. 7 summarize the statistics of the durations of all voltage violation events in the three summer testing days. Conventional VVC is effective in reducing longer voltage violations while leading to many shorter voltage violations. This results in a large number of cumulative violations. Overall, the stage-2 policy exhibits optimal performance in terms of reducing the total voltage violation duration. Nevertheless, from time to time all PVs have inevitably reached their maximum regulating capability, as shown in Fig. 6.

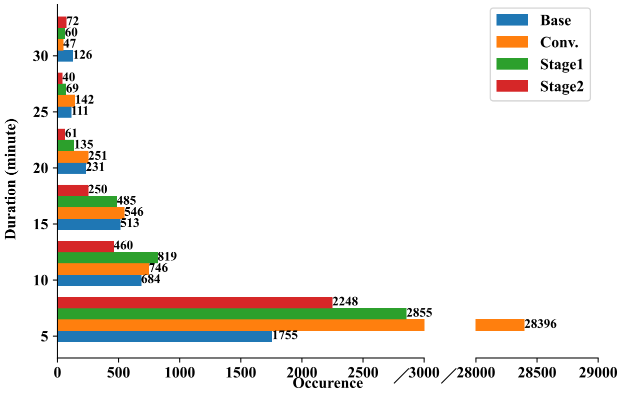


Fig. 7. Distribution of voltage event duration

C. Impact of Action Cost

Because action-taking incurs a cost (specified by (10)), we next conduct a sensitivity analysis in C_w. To observe the impact of action cost (i.e., the value of w_{cost}) on VVC control performance, we present the performance comparison using different w_{cost} values in three spring days. As shown in Table IV, there is a noticeable decrease of Q_{total} when $w_{\mathrm{cost}} = 0.005$, but the degradation in voltage score seems less evident. However, if w_{cost} increases to 0.01, the voltage score declines significantly due to the lack of action from the agents.

TABLE III

STATISTICS OF THE VOLTAGE EVENT

Statistics	Base	Conv.	Stage 1	Stage 2
Count	4031	30219	4831	**3314**
Mean	6.74	1.16	2.64	**2.08**
Std	15.83	1.075	4.16	**2.86**
25 percentile	1	1	1	**1**
50 percentile	2	1	1	**1**
75 percentile	4	1	2	**2**
MaxDuration	95	43	47	**44**
Nodes of MaxDuration	2	5	1	**2**
Integration Sum	27176	34940	12759	**6914**

TABLE IV

IMPACT OF ACTION COST ON VVC PERFORMANCE

C_w	0.001	0.002	0.003	**0.005**	**0.01**
Voltage Score	0.999935	0.999935	0.999777	**0.999789**	**0.997897**
$Q_{total}(kVar)$	225.106	224.310	216.976	**206.443**	**72.033**

V. CONCLUSION

In this paper, we develop a two-stage progressive training strategy for improving the training speed and convergence when training multiple RL-based VVC agents in high PV-penetration distribution systems. Simulation results substantiate that stage-1 training can make agents effectively learn when their actions are effective, while stage-2 training can further strengthen the agents's understanding on how to coordinate with others to achieve satisfactory VVC performance. Most importantly, the policy obtained by the stage-1 can also serve as a backup strategy in case communication may be disconnected. Our follow-up journal paper will present the algorithm in detail with extensive testing results on actual feeder models.

REFERENCES

[1] Y. Zhang, X. Wang, J. Wang, and Y. Zhang, "Deep Reinforcement Learning Based Volt-VAR Optimization in Smart Distribution Systems," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 361–371, Jan. 2021.

[2] X. Sun and J. Qiu, "Two-Stage Volt/Var Control in Active Distribution Networks With Multi-Agent Deep Reinforcement Learning Method," *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 2903–2912, Jul. 2021.

[3] S. Wang, J. Duan, D. Shi, C. Xu, H. Li, R. Diao, and Z. Wang, "A Data-Driven Multi-Agent Autonomous Voltage Control Framework Using Deep Reinforcement Learning," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4644–4654, Nov. 2020.

[4] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, "Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments," *arXiv:1706.02275 [cs]*, Mar. 2020, arXiv: 1706.02275.

[5] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, "Asynchronous Methods for Deep Reinforcement Learning," *arXiv:1602.01783 [cs]*, Jun. 2016, arXiv: 1602.01783.

[6] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv:1509.02971 [cs, stat]*, Jul. 2019, arXiv: 1509.02971.

[7] "PECAN STREET." [Online]. Available: https://www.pecanstreet.org/

[8] "IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces," IEEE, Tech. Rep., iSBN: 9781504446396.