Data-driven Coordinated Voltage Control Method of Distribution Networks with High DG Penetration

Yanda Huo, Peng Li, *Senior Member*, *IEEE*, Haoran Ji*, *Member*, *IEEE*, Hao Yu, *Member*, *IEEE*, Jinyue Yan, Jianzhong Wu, *Member*, *IEEE,* Chengshan Wang, *Senior Member*, *IEEE*

***Abstract*—The highly penetrated distributed generators (DGs) aggravate the voltage violations in active distribution networks (ADNs). The coordination of various regulation devices such as on-load tap changers (OLTCs) and DG inverters can effectively address the voltage issues. Considering the problems of inaccurate network parameters and rapid DG fluctuation in practical operation, multi-source data can be utilized to establish the data-driven control model. In this paper, a data-driven voltage control method with the coordination of OLTC and DG inverters on multiple time-scales is proposed without relying on the accurate physical model. First, based on the multi-source data, a data-driven voltage control model is established. Multiple regulation devices such as OLTC and DG are coordinated on multiple time-scales to maintain voltages within the desired range. Then, a critical measurement selection method is proposed to guarantee the voltage control performance under the partial measurements in practical ADNs. Finally, the proposed method is validated on the modified IEEE 33-node and IEEE 123-node test cases. Case studies illustrate the effectiveness of the proposed method, as well as the adaptability to DG uncertainties.**

***Index Terms*—active distribution network (ADN), distributed generator (DG), data-driven, coordinated voltage control, critical measurements.**

NOMENCLATURE

Sets

$\mathcal{N}$, $\mathcal{N}_m$	Set of all nodes of ADN/area m
Λ	Set of typical scenarios
C	Set of categories of similar nodes
c_i	Set of nodes similar to node i

Indices

t	Indices of instants
i, j, g	Indices of nodes
m	Indices of areas
n	Indices of DG clusters
ζ	Indices of typical scenarios
a_i, h_i	Row and column indices of the similarity matrix related to node i

Variables

$\boldsymbol{U}[t]$	Measurement of nodal voltage amplitude at instant t
$\widetilde{\boldsymbol{U}}[t]$	Estimation of nodal voltage amplitude at instant t
$\boldsymbol{O}[t]$, $\boldsymbol{O}[t-\Delta T]$	Tap step of OLTC at instant t and $t-\Delta T$
$\boldsymbol{E}[t+\Delta T]$	Prediction information during instant t to $t+\Delta T$
$\boldsymbol{E}'[t]$	Information of DGs and loads at instant t
$\widehat{\boldsymbol{\Phi}}_{\mathrm{O}}[t]$	Estimation of Pseudo-Jacobi matrix of OLTC at instant t
$\widehat{\boldsymbol{\Phi}}_{\mathrm{E}}[t]$	Estimation of Pseudo-Jacobi matrix of the prediction information at instant t
$\boldsymbol{P}_{m,n}^{\mathrm{DG}}[t], \boldsymbol{S}_{m,n}^{\mathrm{DG}}[t]$	Power production/capacity of DG cluster n in area m at instant t
$\widetilde{\boldsymbol{U}}_m[t], \boldsymbol{U}_m[t]$	Estimation/measurement of nodal voltage amplitude at instant t in area m
$\boldsymbol{X}_{m,n}[t]$	Vector of the reactive power output of DG cluster n at instant t in area m
$\widehat{\boldsymbol{\Phi}}_{m,n}[t]$	Estimation of Pseudo-Jacobi matrix of DG cluster n at instant t in area m
$\Delta P_g^*(\zeta)$, $\Delta Q_g^*(\zeta)$	Variation of active/reactive power at node g in scenario ζ
$S_{i,g}(\zeta)$	Voltage-power sensitivity between node i and node g in scenario ζ
$\Delta U_i^*(\zeta)$	Variation of voltage amplitude measurement at node i in scenario ζ
$\boldsymbol{F}_i$, $\boldsymbol{F}_g$	Eigenvectors of nodes i and g
$\overline{\boldsymbol{F}}_i$, $\overline{\boldsymbol{F}}_g$	Average values of eigenvectors of nodes i and g
$\boldsymbol{H}_i$	Sensitivity matrix of node i
$\boldsymbol{\Gamma}_{c,i}$	Similarity matrix of category c_i
b_i	Similarity index of node i
γ_{d}, γ_{c}, γ_{o}	Voltage deviation indices of the data-driven approach/ centralized method/ original state of ADN

Parameters

N, N_m	Number of nodes of ADN/area m
N_m^{DG}	Number of DG clusters in area m
N_{s}	Number of typical scenarios
$N_{c,i}$	Number of nodes in the category c_i
ε	Threshold of similarity
ΔT	Slow time-scale

This work was supported by the National Natural Science Foundation of China (U1866207, 52007131) and Swedish Research Council (2018-06007).

Y. Huo, P. Li, H. Ji, H. Yu, and C. Wang are with the Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin 300072, China (email: jihaoran@tju.edu.cn). (*Corresponding author: Haoran Ji*)

J. Yan is with the School of Business, Society and Engineering, Mälardalen University, Västerås 72123, Sweden (email: jinyue.yan@mdh.se).

J. Wu is with the Institute of Energy, School of Engineering, Cardiff University, Cardiff CF24 3AA, U.K. (email: wuj5@cardiff.ac.uk).

Δt	Fast time-scale
T	Control time horizon
$\mu_{\mathrm{O}}, \eta_{\mathrm{O}}, \lambda_{\mathrm{O}}, \rho_{\mathrm{O}}, \mu_{\mathrm{E}}, \eta_{\mathrm{E}}, \mu_{\mathrm{X}}, \eta_{\mathrm{X}}, \lambda_{\mathrm{X}}, \rho_{\mathrm{X}}, w$	Weight factors
$\boldsymbol{O}^{\max}, \boldsymbol{O}^{\min}$	Maximum/minimum tap steps of OLTC
k	Maximum change of tap steps of OLTC
ϵ	Small positive constant
$\boldsymbol{U}^{\mathrm{ref}}$	Voltage reference of ADN

I. Introduction

With the highly penetrated distributed generators (DGs) [1], issues of voltage violations in active distribution networks (ADNs) are exacerbated by the rapid fluctuation of DGs [2]. Various types of voltage control devices have been utilized to optimize the operation performance of ADNs [3]. However, because of the discrete and slow regulation speed, conventional regulators such as the on-load tap changers (OLTCs) are incapable of suppressing frequent voltage fluctuations caused by DGs [4].

The regulation of reactive power generated from DG inverters is also a promising approach to realizing real-time voltage control [5]. Since DG inverters provide the rated power only a fraction of the time, the surplus capacity is able to continuously regulate voltages [6]. Limited by the capacity of the DG inverter, it is necessary to involve both continuous and discrete adjustments, such as DG inverters and OLTC to efficiently regulate voltages. Therefore, a coordinated control approach involving multiple time-scales is required.

Generally, the model-based voltage control methods with accurate network parameters are adopted to regulate voltage profiles in ADNs [7]. Ref. [8] proposed a novel analytical expression to curtail the renewable energy for voltage control, which effectively maximized the technical benefits of DG units in distribution systems. In Ref. [9], the voltage control was formulated as a convex quadratic problem with linear constraints for efficient solving.

To coordinate various devices with different response times and control characteristics, hierarchical voltage control methods are widely utilized. In Ref. [10], a two-stage control framework was proposed to address the problem of optimally coordinated control. Ref. [11] proposed a two-stage voltage-load sensitivity matrix-based demand response algorithm for regulating the voltages. In Ref. [12], a multi-stage voltage support optimization method was proposed for safeguarding the ADN operation and providing flexible power delivery capacity. A three-layer hierarchical voltage control strategy was proposed in Ref. [13], considering the customized charging navigation of electric vehicles. Ref. [14] proposed a coordination algorithm for OLTC-based voltage control and reactive power compensation to mitigate the overvoltage. In Ref. [15], a two-stage approach was introduced to coordinate the available reactive power of DG and OLTC considering the multiple optimal objectives. Ref. [16] proposed a coordination control technique to utilize DG and minimize the interaction of OLTC based on the concept of control zone. These hierarchical control methods can effectively solve the voltage deviation when the accurate model parameters of ADNs can be obtained.

However, the main obstacle to the application of model-based optimization approaches is lacking accurate network parameters due to the complex operational environments of ADNs. In addition, integrated with various DGs and flexible loads, the operation state and topology information of ADNs change frequently [17]. Thus, the traditional model-based optimization approaches are challenged by inaccurate parameters and frequently changed states of ADNs.

With the digitalization of distribution networks, massive real-time operational data is available to the distribution network operators [18]. Important information can be revealed, including the feature of user behaviors and the variation of system states [19]. Information utilization is of vital importance to enhance operational performances. Hence, data-driven optimization methods obtain widespread attention to cope with the problem of lacking accurate parameters [20]. It only utilizes the operational or historical data instead of the detailed physical models with accurate parameters. In addition, auxiliary assistance provided by data-driven methods can improve the effectiveness of model-based operation strategies.

Preliminary studies of data-driven-based optimization have been conducted in ADNs. Data-based iteratively adaptive control and deep reinforcement learning-based approaches are the two main forms of data-driven methods. To simulate the complex nonlinear characteristics of ADNs, a massive amount of offline data is utilized by deep reinforcement learning-based methods. Then neural networks are trained and implemented to achieve optimal operation. Ref. [21] established an information-theoretic reinforcement learning framework for phase identification in power distribution systems. A deep reinforcement learning method was proposed in Ref. [22] to regulate nodal voltage without physical models. However, the training process of reinforcement learning is time-consuming and may be ineffective in adapting to state changes of ADNs.

Iteratively adaptive control is an alternative data-driven method. It models the features of ADN on the basis of operational data and can realize real-time control through the iterative interaction with ADN [23]. Model-free adaptive control (MFAC) represents a classic iteratively adaptive control algorithm [24]. Ref. [25] proposed a basic framework of MFAC, of which convergence and stability could be guaranteed under some practical assumptions. Ref. [26] focused on the data-driven operation strategy of soft open point (SOP) based on MFAC, which could improve the operational performance of ADN without accurate parameters. However, data-driven voltage control with various devices on multiple time-scales can be further considered.

In addition, from the perspective of economy and engineering feasibility, full coverage installation of real-time measurement devices is still difficult for large-scale ADNs [27]. How to use critical measurements to regulate voltage has become a challenge for data-driven control of ADNs. Thus, it is essential to take incomplete measurements into account in practical voltage control.

To effectively reduce the voltage violation with various voltage control devices, a data-driven coordinated voltage control method on multiple time-scales is proposed in this paper. The main contributions are outlined as follows:

1) Based on the multi-source data, a data-driven voltage control model is established for the coordination of OLTC and DG inverters on multiple time-scales to maintain voltages within the desired range. Considering the change rate of OLTC, the data-driven control on the slow time-scale is conducted mainly based on the prediction information. By utilizing the real-time measurements, the reactive power of DG inverters is adjusted rapidly on the fast time-scale, to suppress voltage fluctuation.

2) A critical measurement selection method is proposed to guarantee the voltage control performance under the incomplete measurements in practical ADNs. Based on the critical measurements, the data-driven control model is simplified while obtaining an approximate optimal solution.

The organization of this paper is stated as follows. In Section II, the data-driven coordinated voltage control method on multiple time-scales is elaborated. Section III describes the selection method of critical measurements in ADNs. In Section IV, the case studies are conducted to verify the effectiveness of the proposed data-driven control method. Finally, conclusions are drawn in Section V.

II. Data-driven Coordinated Voltage Control

The incorporation of multiple voltage control devices may create operational conflicts, which brings challenges to voltage control [28]. Taking the coordination of discrete and continuous regulation devices into account, a framework of data-driven coordinated voltage control is formulated on multiple time-scales.

A. Framework of Data-driven Coordinated Voltage Control

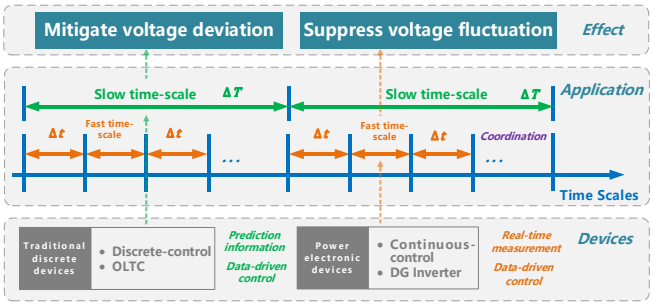


Fig. 1. Framework of data-driven coordinated voltage control.

Fig. 1 shows the framework of the proposed coordinated control method. It includes multiple time-scales: slow time-scale ΔT and fast time-scale Δt. The role of OLTC is to adjust the tap step to prevent severe voltage deviation on the slow time-scale ΔT. Considering the change frequency of OLTC, the data-driven control on the slow time-scale is mainly conducted based on the prediction information. Whereas the reactive power of DG inverters is adjusted rapidly on the fast time-scale Δt. By utilizing the real-time measurements, the data-driven control on the fast time-scale can adaptively suppress rapid voltage fluctuation. Since the slow time-scale may decrease the control accuracy, it is necessary to coordinate those regulation devices with fast control devices such as DG inverters. Thus, discrete and continuous regulation devices are coordinated on multiple time-scales to maintain voltages within the desired range.

1) Modelling of data-driven voltage control problem

On the basis of measurement information, the relationship between the nodal voltage and strategies of voltage control devices is established in Eq. (1).

$$\boldsymbol{y} = f(\boldsymbol{x}) \tag{1.a}$$

where $f(\cdot)$ denotes the unknown function. $\boldsymbol{x}$ denotes operation strategies of voltage control devices including discrete and continuous regulation devices. $\boldsymbol{y}$ denotes measurements of ADN such as the voltage amplitude of each node.

Assumption (1): The controlled system satisfies the Lipschitz condition. For $\forall\, k_1, k_2$, the Lipschitz condition of Eq. (1.a) is shown as follows. L is a constant.

$$\begin{aligned} &\left|\left|\boldsymbol{y}[t+k_1\Delta t]-\boldsymbol{y}[t+k_2\Delta t]\right|\right|_2 \\ &\le L\left|\left|\boldsymbol{x}[t+(k_1-1)\Delta t]-\boldsymbol{x}[t+(k_2-1)\Delta t]\right|\right|_2 \end{aligned} \tag{1.b}$$

Assumption (2): The partial derivative of $f(\cdot)$ is continuous for the input $\boldsymbol{x}[t]$.

In Eq. (1.a), ADN is equivalent to a black-box system, of which the inputs are the strategies of voltage control devices, and outputs are the measurements of nodal voltage. Thus, the relationship between the nodal voltage and strategies of voltage control devices is established based on a non-linear unknown function $f(\cdot)$.

In the voltage control problem of ADN, the physical essence of $\frac{\left|\left|\boldsymbol{y}[t+k_1\Delta t]-\boldsymbol{y}[t+k_2\Delta t]\right|\right|_2}{\left|\left|\boldsymbol{x}[t+(k_1-1)\Delta t]-\boldsymbol{x}[t+(k_2-1)\Delta t]\right|\right|_2}$ is the voltage-to-power sensitivity during $[t+k_1\Delta t, t+k_2\Delta t]$. According to the definition of voltage-to-power sensitivity, it is easy to obtain the following equation.

$$\frac{\left|\left|\boldsymbol{y}[t+k_1\Delta t]-\boldsymbol{y}[t+k_2\Delta t]\right|\right|_2}{\left|\left|\boldsymbol{x}[t+(k_1-1)\Delta t]-\boldsymbol{x}[t+(k_2-1)\Delta t]\right|\right|_2} \le L \tag{1.c}$$

Eq. (1.c) can be satisfied by the practical ADNs according to the operational constraints of ADNs. Thus, Assumption (1) can be satisfied. In addition, Assumption (2) is generally considered valid in practical ADNs.

Theorem A1: If Assumptions (1) and (2) are satisfied in the non-linear control systems, $\boldsymbol{\Phi}[t] \in \mathbf{R}^{m\times m}$ exists. Then Eq. (1.a) can be linearized to the form of Eq. (2) on different time scales. The proof of Theorem A1 is shown in Appendix.

$$\widetilde{\boldsymbol{U}}[t+\Delta T] = \boldsymbol{U}[t] + \boldsymbol{\Phi}_{\mathrm{O}}[t](\boldsymbol{O}[t]-\boldsymbol{O}[t-\Delta T]) \tag{2.a}$$

$$\widetilde{\boldsymbol{U}}[t+\Delta t] = \boldsymbol{U}[t] + \boldsymbol{\Phi}_{\mathrm{X}}[t](\boldsymbol{X}[t]-\boldsymbol{X}[t-\Delta t]) \tag{2.b}$$

where $\widetilde{\boldsymbol{U}}[t+\Delta T]$ and $\widetilde{\boldsymbol{U}}[t+\Delta t]$ denote the nodal voltage estimation at instant $t+\Delta T$ and $t+\Delta t$. $\boldsymbol{U}[t]$ is the nodal voltage measurement at instant t. $\boldsymbol{O}[t]$ and $\boldsymbol{O}[t-\Delta T]$ denote the strategies of discrete voltage regulation devices at instant t and $t-\Delta T$. $\boldsymbol{X}[t]$ and $\boldsymbol{X}[t-\Delta t]$ are the strategies of continuous regulation devices at instant t and $t-\Delta t$. $\boldsymbol{\Phi}_{\mathrm{O}}[t]$ and $\boldsymbol{\Phi}_{\mathrm{X}}[t]$ are the Pseudo-Jacobi Matrix (PJM) at instant t, which represents the relationship between voltage measurements and

operation strategies of multiple voltage control devices.

Eq. (2.a) represents the data-driven model of discrete devices, which is established on a slow time-scale ΔT. The strategies of discrete devices remain unchanged during ΔT to avoid frequent operations. Eq. (2.b) denotes the data-driven model of continuous regulation devices, which is established to respond to the voltage fluctuation on a fast time-scale Δt.

2) Coordination on multiple time-scales

The discrete and continuous regulation devices can be coordinated on multiple time-scales. Taking OLTC and DG inverters as an example, OLTC adjusts the tap step to prevent severe voltage deviation on a slow time-scale ΔT. Whereas the reactive power of DG inverters is adjusted rapidly on a fast time-scale Δt, which can adaptively suppress rapid voltage fluctuation caused by DGs. Thus, discrete and continuous regulation devices are coordinated on multiple time-scales to maintain voltages within the desired range.

For decentralized and efficient voltage control, the network of ADN is partitioned into multiple sub-areas based on the electrical distance. Since the partition of ADN can be described as a clustering problem, the clustering algorithm can be utilized to determine the range and number of sub-areas. The proposed data-driven method models the features of sub-areas in ADN based on real-time operational data and can realize the real-time control through the iterative interaction with ADN without training process. Based on the dynamic linearization, the proposed method creates a data model of a nonlinear complex system without accurate parameters.

The coordination among regulation devices in the same area is also taken into consideration. Taking DG inverters in area m as an example, adjacent DGs can be divided into several clusters. As the voltage control effect of DG clusters in the same area is coupling, the relationship of DG clusters in the same area is mainly considered in this paper. Since the data-driven control model of continuous regulation devices is linearized dynamically at each Δt [23], the voltage control effect of each DG cluster satisfies linear superposition principles at each Δt in the same area.

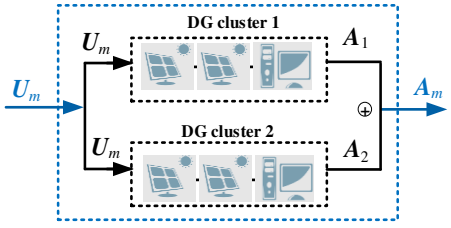


Fig. 2. The parallel relationship between DG clusters in the same area.

Therefore, it is reasonable to assume that the DG clusters in each area are paralleled. Taking an area with two DG clusters as an example, the parallel relationship between two DG clusters is shown in Fig. 2. $\boldsymbol{U}_m$ denotes the measurements of nodal voltage amplitude in area m. $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are the voltage control effect of DG clusters 1 and 2. $\boldsymbol{A}_m$ represents the voltage control effect of DG clusters in area m.

B. Data-driven Coordinated Voltage Control Model

1) Modelling of data-driven control on the slow time scale

Taking OLTC as an example, the data-driven model on the slow time-scale ΔT is established. To determine the strategy of OLTC during each optimization horizon ΔT, the influence of prediction information of DGs and loads is considered. Then Eq. (2.a) can be transformed to a prediction model at each ΔT:

$$\begin{aligned}\widetilde{\boldsymbol{U}}[t+\Delta T] = \boldsymbol{U}[t] + \boldsymbol{\Phi}_{\mathrm{O}}[t](\boldsymbol{O}[t] - \boldsymbol{O}[t-\Delta T]) \\ + w\boldsymbol{\Phi}_{\mathrm{E}}^{\mathrm{T}}[t](\boldsymbol{E}[t+\Delta T] - \boldsymbol{E}'[t])\end{aligned} \tag{3}$$

where $\widetilde{\boldsymbol{U}}[t+\Delta T]$ denotes the estimation of nodal voltage amplitude at instant $t+\Delta T$. $\boldsymbol{\Phi}_{\mathrm{O}}[t]$ denotes the PJM which represents the relationship between the tap steps of OLTC and nodal voltages. $\boldsymbol{\Phi}_{\mathrm{E}}[t]$ is the PJM at instant t, which represents the relationship between prediction information and nodal voltage. $\boldsymbol{O}[t]$ and $\boldsymbol{O}[t-\Delta T]$ are tap steps of OLTC at instant t and $t-\Delta T$. $\boldsymbol{E}[t+\Delta T]$ denotes prediction information during instants t to $t+\Delta T$, including prediction information of DGs and loads. $\boldsymbol{E}'[t]$ denotes information of DGs and loads at instant t. w represents the influence of the voltage measurement from the prediction information of DGs and loads. w can be adjusted according to the practical experiences of the distribution system operator.

There exists two unknown PJMs in Eq. (3), which need to be estimated, namely $\boldsymbol{\Phi}_{\mathrm{O}}[t]$ and $\boldsymbol{\Phi}_{\mathrm{E}}[t]$. Based on measurement information, the estimation functions of $\boldsymbol{\Phi}_{\mathrm{O}}[t]$ and $\boldsymbol{\Phi}_{\mathrm{E}}[t]$ can be represented as follows.

$$\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t] = \operatorname{argmin}\left(\left\|\boldsymbol{U}[t] - \widetilde{\boldsymbol{U}}[t]\right\|_2^2 + \mu_{\mathrm{O}}\left\|\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t] - \hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t-\Delta T]\right\|_2^2\right) \tag{4}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t] = \operatorname{argmin}\left(\left\|\boldsymbol{U}[t] - \widetilde{\boldsymbol{U}}[t]\right\|_2^2 + \mu_{\mathrm{E}}\left\|\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t] - \hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t-\Delta t]\right\|_2^2\right) \tag{5}$$

Then based on the gradient descent algorithm, the iterative expressions for solving $\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t]$ and $\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t]$ are formulated.

$$\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t] = \hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t-\Delta T] + \frac{\eta_{\mathrm{O}}(\Delta\boldsymbol{U}[t] - \hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t]\Delta\boldsymbol{O}[t-\Delta T] - \hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t]\Delta\boldsymbol{E}'[t-\Delta t])\Delta\boldsymbol{O}[t-\Delta T]}{\mu_{\mathrm{O}} + \|\Delta\boldsymbol{O}[t-\Delta T]\|_2^2} \tag{6}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t] = \hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t_0], \text{ if } \hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t] \le \epsilon \text{ or } \Delta\boldsymbol{O}[t-\Delta T] \le \epsilon \text{ or } \operatorname{sign}(\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t]) \ne \operatorname{sign}(\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t_0]) \tag{7}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t] = \hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t-\Delta t] + \frac{\eta_{\mathrm{E}}(\Delta\boldsymbol{U}[t] - \hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t]\Delta\boldsymbol{E}'[t-\Delta t])\Delta\boldsymbol{E}'^{\mathrm{T}}[t-\Delta t]}{\mu_{\mathrm{E}} + \|\Delta\boldsymbol{E}'[t-\Delta t]\|_2^2} \tag{8}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t] = \hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t_0], \text{ if } \hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t] \le \epsilon \text{ or } \Delta\boldsymbol{E}'[t-\Delta t] \le \epsilon \text{ or } \operatorname{sign}(\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t]) \ne \operatorname{sign}(\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t_0]) \tag{9}$$

where $\Delta\boldsymbol{U}[t]$ denotes $\boldsymbol{U}[t] - \boldsymbol{U}[t-\Delta t]$. $\Delta\boldsymbol{O}[t-\Delta T]$ denotes $\boldsymbol{O}[t-\Delta T] - \boldsymbol{O}[t-2\Delta T]$. $\Delta\boldsymbol{E}'[t-\Delta t]$ denotes $\boldsymbol{E}'[t-\Delta t] - \boldsymbol{E}'[t-2\Delta t]$. $\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t_0]$ and $\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t_0]$ are the initial values of $\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t]$ and $\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t]$. $\hat{\boldsymbol{\Phi}}_{\mathrm{O}}[t_0]$ is assigned based on the sensitivities of nodal voltage and tap steps of OLTC. $\hat{\boldsymbol{\Phi}}_{\mathrm{E}}[t_0]$ is assigned on the basis of sensitivities of nodal voltage and information of DGs and loads at instant t_0. Eqs. (7) and (9) are introduced to enable Eqs. (6) and (8) be with a stronger parameter tracking ability [24].

Note that both the gradient descent method and Newton method are typical iterative methods for finding the minimum of an objective function [29]. The Newton method is an efficient algorithm for solving the day-ahead optimization problems of ADNs [30]. From the perspective of applicability, the gradient descent method is suitable for the iterative calculation

in the proposed data-driven method. The reasons are elaborated as follows. a) If the second derivative of the objective function is undefined in the objective function's root, gradient descent methods can still be applied. b) The time complexity of the gradient descent method is O(n). The Hessian matrix and its inverse matrix are unnecessary to be calculated in the gradient descent method, which is time-saving. c) In addition, the gradient descent method is not sensitive to its initial values.

In the proposed data-driven method, it puts forward higher requirements for the calculation speed in each iteration for the real-time control, which requires a sample and time-saving method. Thus, the gradient descent method is utilized to iteratively solve equations such as Eqs. (4) and (5).

Considering the minimum voltage deviation and the change rate of the tap steps of OLTC, the objective function of OLTC is expressed as:

$$J(\boldsymbol{O}[t]) = \min \left(\left\| \boldsymbol{U}^{\mathrm{ref}} - \widetilde{\boldsymbol{U}}[t+\Delta T]\right\|_2^2 + \lambda_{\mathrm{O}} \|\boldsymbol{O}[t] - \boldsymbol{O}[t-\Delta t]\|_2^2\right) \tag{10}$$

The iterative formulation for calculating $\boldsymbol{O}[t]$ based on the gradient descent algorithm is expressed as Eq. (11).

$$\boldsymbol{O}[t] = \boldsymbol{O}[t-\Delta T] + \frac{\rho_{\mathrm{O}} \widehat{\boldsymbol{\Phi}}_{\mathrm{O}}[t](\boldsymbol{U}^{\mathrm{ref}} - \widetilde{\boldsymbol{U}}[t] - \widehat{\boldsymbol{\Phi}}_{\mathrm{E}}[t]\Delta \boldsymbol{E}[t+\Delta T])}{\lambda_{\mathrm{O}} + \left\|\widehat{\boldsymbol{\Phi}}_{\mathrm{O}}[t]\right\|_2^2} \tag{11}$$

where $\widehat{\boldsymbol{\Phi}}_{\mathrm{O}}[t]$ and $\widehat{\boldsymbol{\Phi}}_{\mathrm{E}}[t]$ can be solved by Eq. (4) and Eq. (5).

Considering the variation range limits of the tap steps, $\boldsymbol{O}[t]$ are constrained by the following equations.

$$\boldsymbol{O}[t] = \boldsymbol{O}[t-\Delta T] + k, \text{ if } \boldsymbol{O}[t] - \boldsymbol{O}[t-\Delta T] > k \tag{12.a}$$

$$\boldsymbol{O}[t] = \boldsymbol{O}[t-\Delta T] - k, \text{ if } \boldsymbol{O}[t] - \boldsymbol{O}[t-\Delta T] < -k \tag{12.b}$$

$$\boldsymbol{O}[t] = \boldsymbol{O}^{\max}, \text{ if } \boldsymbol{O}[t] > \boldsymbol{O}^{\max} \tag{13.a}$$

$$\boldsymbol{O}[t] = \boldsymbol{O}^{\min}, \text{ if } \boldsymbol{O}[t] < \boldsymbol{O}^{\min} \tag{13.b}$$

Eqs. (12.a) and (12.b) limit the maximum variation of the tap steps during the considered time horizon. Eqs. (13.a) and (13.b) represent the variation range of the discrete tap steps.

2) Modelling of data-driven control on the fast time scale

Taking paralleled DG clusters in area m as an example, the data-based model of DG cluster n is established during a fast time-scale Δt. Influence from other DG clusters in the same area is taken into consideration, as shown in Eq. (14).

$$\widetilde{\boldsymbol{U}}_m[t+\Delta t] = \boldsymbol{U}_m[t] + \sum_{n=1}^{N_m^{\mathrm{DG}}} \boldsymbol{\Phi}_{m,n}^{\mathrm{T}}[t](\boldsymbol{X}_{m,n}[t] - \boldsymbol{X}_{m,n}[t-\Delta t]) \tag{14}$$

where $\widetilde{\boldsymbol{U}}_m[t+\Delta t]$ represents the estimation of nodal voltage amplitude in area m at instant $t+\Delta t$. $\boldsymbol{U}_m[t]$ denotes nodal voltage measurement in area m at instant t. N_m^{DG} represents the number of DG clusters in area m. $\boldsymbol{\Phi}_{m,n}[t]$ is the PJM of DG cluster n in area m at instant t, which denotes the relationship between nodal voltage measurement and the reactive power output of DG cluster n in area m. $\boldsymbol{X}_{m,n}[t]$ is the reactive power output vector of DG cluster n at instant t in area m.

In Eq. (14), the PJM $\boldsymbol{\Phi}_{m,n}[t]$ can be estimated as Eq. (15). Different from the Jacobi Matrix of ADN, the PJM utilized in this paper is calculated based on the multi-source data. It represents the relationship between voltage measurements and operation strategies of multiple voltage control devices.

$$\widehat{\boldsymbol{\Phi}}_{m,n}[t] = \operatorname{argmin} \left(\left\|\boldsymbol{U}_m[t] - \widetilde{\boldsymbol{U}}_m[t]\right\|_2^2 + \mu_{\mathrm{X}} \left\|\widehat{\boldsymbol{\Phi}}_{m,n}[t] - \widehat{\boldsymbol{\Phi}}_{m,n}[t-\Delta t]\right\|_2^2\right) \tag{15}$$

Considering the influence of other DG clusters in area m, $\widehat{\boldsymbol{\Phi}}_{m,n}[t]$ can be calculated at each instant. The iterative expression for estimating $\widehat{\boldsymbol{\Phi}}_{m,n}[t]$ is shown as follows.

$$\widehat{\boldsymbol{\Phi}}_{m,n}[t] = \widehat{\boldsymbol{\Phi}}_{m,n}[t-\Delta t] + (\Delta \boldsymbol{U}_m[t] - \sum_{i=1}^{N_m^{\mathrm{DG}}} \boldsymbol{\Phi}_{m,i}[t]\Delta \boldsymbol{X}_{m,i}[t-\Delta t]) \frac{\eta \Delta \boldsymbol{X}_{m,i}^{\mathrm{T}}[t-\Delta t]}{\mu + \left\|\Delta \boldsymbol{X}_{m,n}[t-\Delta t]\right\|_2^2} \tag{16}$$

$$\widehat{\boldsymbol{\Phi}}_{m,n}[t] = \widehat{\boldsymbol{\Phi}}_{m,n}[t_0], \text{ if } \widehat{\boldsymbol{\Phi}}_{m,n}[t] \le \epsilon \text{ or } \Delta \boldsymbol{X}_{m,n}[t-\Delta t] \le \epsilon \text{ or } \operatorname{sign}(\widehat{\boldsymbol{\Phi}}_{m,n}[t]) \neq \operatorname{sign}(\widehat{\boldsymbol{\Phi}}_{m,n}[t_0]) \tag{17}$$

where $\Delta \boldsymbol{U}_m[t]$ represents $\boldsymbol{U}_m[t] - \boldsymbol{U}_m[t-\Delta t]$. $\Delta \boldsymbol{X}_{m,n}[t-\Delta t]$ represents $\boldsymbol{X}_{m,n}[t-\Delta t] - \boldsymbol{X}_{m,n}[t-2\Delta t]$. $\widehat{\boldsymbol{\Phi}}_{m,n}[t_0]$ denotes the initial value of $\widehat{\boldsymbol{\Phi}}_{m,n}[t]$, which can be initialized by sensitivities between nodal voltages and reactive power of DG clusters.

Considering the minimum voltage deviation and the change rate of DG outputs, the objective function of DG cluster n in area m is expressed as:

$$J(\boldsymbol{X}_{m,n}[t]) = \min \left(\left\|\boldsymbol{U}^{\mathrm{ref}} - \widetilde{\boldsymbol{U}}_m[t+\Delta t]\right\|_2^2 + \lambda_{\mathrm{X}} \left\|\boldsymbol{X}_{m,n}[t] - \boldsymbol{X}_{m,n}[t-\Delta t]\right\|_2^2\right) \tag{18}$$

In the data-driven voltage control of DG, the control strategies of DG's reactive power are obtained with an iterative process. If the reactive power output changes too fast at the adjacent control instants, it may lead to rapid voltage fluctuation to impact on the secure operation of ADN. Thus, the change rate of DG outputs is included in the objective function.

Considering the control strategies of other DG clusters in area m, $\boldsymbol{X}_{m,n}[t]$, the control strategy of DG cluster n at instant t in area m, is iteratively calculated as follows.

$$\boldsymbol{X}_{m,n}[t] = \boldsymbol{X}_{m,n}[t-\Delta t] + \left(\boldsymbol{U}^{\mathrm{ref}} - \boldsymbol{U}_m[t] - \sum_{l=1,l\neq n}^{N_m^{\mathrm{DG}}} \widehat{\boldsymbol{\Phi}}_{m,l}[t]\left(\boldsymbol{X}_{m,l}[t] - \boldsymbol{X}_{m,l}[t-\Delta t]\right)\right) \frac{\rho_{\mathrm{X}} \widehat{\boldsymbol{\Phi}}_{m,n}[t]}{\lambda_{\mathrm{X}} + \left\|\widehat{\boldsymbol{\Phi}}_{m,n}[t]\right\|_2^2} \tag{19}$$

where $\widehat{\boldsymbol{\Phi}}_{m,n}[t]$ can be solved by Eq. (16).

To further consider the limitation of DG inverters' capacities, operational constraints of DGs are shown as follows.

$$\begin{aligned} \boldsymbol{X}_{m,n}[t] &= \sqrt{\left(\boldsymbol{S}_{m,n}^{\mathrm{DG}}\right)^2 - \left(\boldsymbol{P}_{m,n}^{\mathrm{DG}}[t]\right)^2}, \\ &\text{if } \boldsymbol{X}_{m,n}[t] > \sqrt{\left(\boldsymbol{S}_{m,n}^{\mathrm{DG}}\right)^2 - \left(\boldsymbol{P}_{m,n}^{\mathrm{DG}}[t]\right)^2} \\ \boldsymbol{X}_{m,n}[t] &= -\sqrt{\left(\boldsymbol{S}_{m,n}^{\mathrm{DG}}\right)^2 - \left(\boldsymbol{P}_{m,n}^{\mathrm{DG}}[t]\right)^2}, \\ &\text{if } \boldsymbol{X}_{m,n}[t] < -\sqrt{\left(\boldsymbol{S}_{m,n}^{\mathrm{DG}}\right)^2 - \left(\boldsymbol{P}_{m,n}^{\mathrm{DG}}[t]\right)^2} \end{aligned} \tag{20}$$

Constraint (20) represents the variation range of reactive power outputs of DG clusters at each instant. Limited by the capacity of DG inverter, it is necessary to involve both continuous and discrete adjustments, such as DG inverters and OLTC to efficiently regulate voltages. Thus, when DG inverters achieve their maximum allowable reactive power, the proposed coordinated control approach method will dispatch volt-

age regulation devices involving multiple time-scales to mitigate the voltage violation.

Remark 1: In practical ADNs, there are mainly two forms of the system structure updating. a) The integration of new devices, such as DGs, switches on the original structure of ADNs. b) The expanded planning of ADNs.

The integration of new devices does not change the observability of ADN. However, as the controlled resources have been integrated, the communication links with these new devices need to be supplemented. With the application of 5G technology, local communication will be convenient to realize. Then, the proposed method can be compatible with new device integration by properly extending the data-driven model. As for the expanded planning of ADNs, the observability of ADNs will be influenced. Thus, new measurements should be added to incorporate the observability of expanded ADNs. The proposed method is still scalable after the supplement of measurements devices.

Remark 2: The proposed data-driven method mainly includes two kinds of parameters: time-varying parameters ($\boldsymbol{\Phi}_{\mathrm{O}}[t]$ and $\boldsymbol{\Phi}_{\mathrm{X}}[t]$) and fixed parameters ($\mu_{\mathrm{O}}, \eta_{\mathrm{O}}, \lambda_{\mathrm{O}}, \rho_{\mathrm{O}}, \mu_{\mathrm{E}}, \eta_{\mathrm{E}}, \mu_{\mathrm{X}}, \eta_{\mathrm{X}}, \lambda_{\mathrm{X}}, \rho_{\mathrm{X}}, w$). As for the influence of parameters, the time-varying parameters in the proposed method represent the internal structure and parameters of ADNs.

Parameter tuning is important for the control performance of the data-driven method. To adapt to the variation of system states, the time-varying parameters ($\boldsymbol{\Phi}_{\mathrm{O}}[t]$ and $\boldsymbol{\Phi}_{\mathrm{X}}[t]$) are estimated by Eqs. (4)-(8) and Eq. (15), according to the multiple-source data. As for the fixed parameters, it is proved in [23] that stability can be guaranteed when the fixed parameters are within the predefined range, namely $\eta_{\mathrm{O}}, \eta_{\mathrm{X}}, \eta_{\mathrm{E}} \in (0,1]$, $\mu_{\mathrm{O}}, \mu_{\mathrm{X}}, \mu_{\mathrm{E}} > 0$, $\lambda_{\mathrm{O}}, \lambda_{\mathrm{X}} > 0$ and $\rho_{\mathrm{O}}, \rho_{\mathrm{X}} \in (0,1]$.

III. Selection of Critical Measurements

The coordinated voltage control approach described in Section II mainly relies on the system's whole voltage measurements. Considering the incomplete measurements in practical ADNs, it is indispensable to reduce the requirement of total measurement amounts while obtaining an approximate optimal solution. Thus, critical measurements should be selected beforehand.

A. Critical Measurement Selection

To reduce the requirements of measurement amount, nodes with similar voltage variation to the fluctuation of DGs and loads are clustered into a category. Then critical measurements of nodes can be selected from each category.

1) Clustering of measurement nodes

To quantify the similarities between nodes, the modified cosine similarity is utilized, which is widely used to measure the difference between two vectors [31]. Taking node i in area m as an example, by using the modified cosine similarity, the similarity between node i with other nodes in area m can be quantified as follows.

$$\cos(\boldsymbol{F}_i, \boldsymbol{F}_g) = \frac{\sum_{k=1}^{\sigma}(\boldsymbol{F}_i(k)-\overline{\boldsymbol{F}}_i)(\boldsymbol{F}_g(k)-\overline{\boldsymbol{F}}_g)}{\sqrt{\sum_{k=1}^{\sigma}(\boldsymbol{F}_i(k)-\overline{\boldsymbol{F}}_i)^2}\sqrt{\sum_{k=1}^{\sigma}(\boldsymbol{F}_g(k)-\overline{\boldsymbol{F}}_g)^2}}\, i, g \in \mathcal{N}_m \quad (21)$$

where $\cos(\boldsymbol{F}_i, \boldsymbol{F}_g)$ indicates the cosine similarity of node i and node g. $\boldsymbol{F}_i$ and $\boldsymbol{F}_g$ are eigenvectors of nodes i and g, which denote the voltage variation in typical scenarios. $\overline{\boldsymbol{F}}_i$ and $\overline{\boldsymbol{F}}_g$ represent the average values of the eigenvectors. σ is the dimension of the eigenvector $\boldsymbol{F}_i$, and $\sigma = N_m \times N_{\mathrm{s}}$. N_{s} is the number of typical scenarios, which are clustered based on the historical data [32].

The eigenvector $\boldsymbol{F}_i$ is transformed by the vectorization from a sensitivity matrix $\boldsymbol{H}_i$ of node i, which contains sensitivities of node i in typical scenarios.

$$\boldsymbol{H}_i = \begin{bmatrix} S_{i,1}(1) & \cdots & S_{i,1}(N_{\mathrm{s}}) \\ \vdots & S_{i,g}(\zeta) & \vdots \\ S_{i,N_m}(1) & \cdots & S_{i,N_m}(N_{\mathrm{s}}) \end{bmatrix}, i, g \in \mathcal{N}_m, \zeta \in \Lambda \quad (22)$$

where $S_{i,g}(\zeta)$ indicates the voltage-to-power sensitivity [33] between nodes i and g in scenario ζ, which can be obtained based on the historical data.

$$S_{i,g}(\zeta) = \alpha\partial\Delta U_i^*(\zeta)/\partial\,\Delta P_g^*(\zeta) + \beta\partial\Delta U_i^*(\zeta)/\partial\,\Delta Q_g^*(\zeta) \quad i, g \in \mathcal{N}_m, \zeta \in \Lambda \quad (23)$$

Note that the number of clusters is not a predefined parameter. In the clustering step, categories are formed in the clustering process based on node similarities. As the similarities between nodes are quantified, nodes with larger cosine similarity value than threshold ε can be classified into a category denoted by $c_i | c_i \in C$, as shown in Eq. (24). Then the number of clusters can be determined.

$$i, g \in c_i \quad \text{if } \cos(\boldsymbol{F}_i, \boldsymbol{F}_g) > \varepsilon \quad (24)$$

2) Selection of critical measurements

Then a similarity matrix $\boldsymbol{\Gamma}_{c,i}$ can be formed based on category $c_i | c_i \in C$, which contains a group of similar nodes including node i.

$$\boldsymbol{\Gamma}_{c,i} = \begin{bmatrix} \cos(\boldsymbol{F}_i, \boldsymbol{F}_i) & \cdots & \cos(\boldsymbol{F}_i, \boldsymbol{F}_j) \\ \vdots & \ddots & \vdots \\ \cos(\boldsymbol{F}_j, \boldsymbol{F}_i) & \cdots & \cos(\boldsymbol{F}_j, \boldsymbol{F}_j) \end{bmatrix} i, j \in c_i \quad (25)$$

where element $\cos(\boldsymbol{F}_i, \boldsymbol{F}_j)$ in the matrix $\boldsymbol{\Gamma}_{c,i}$ denotes the cosine similarity between node i and node j.

To select representative nodes as critical measurements, the similarity index is defined as follows.

$$b_i = \sum_{h_i=1}^{N_{c,i}} \boldsymbol{\Gamma}_{c,i}(a_i, h_i)\; i \in c_i \quad (26)$$

where b_i denotes the similarity index of node i. N_c is the number of nodes in a category c_i. $\boldsymbol{\Gamma}_{c,i}(a_i, h_i)$ is the element in row a_i and column h_i.

The nodes in a category are sorted according to the similarity index values. Then the most representative nodes with the highest similarity index in a category are recommended as the critical measurement nodes. Based on the critical measurements, the data-driven coordinated voltage control model described in Section II can be simplified, and the requirement of total measurement amounts is reduced.

There are mainly two application conditions for the selection of the critical measurements, namely the existing ADN, and the expanded planning of ADN. a) In the existing ADN, if the historical operation data can be obtained, the critical measurement selection can be carried out with the proposed method. If the historical data is insufficient, the critical measurement selection can also be expanded based on the state estimation and physical model of ADN. The voltage-to-power sensitivities can be calculated by the differential of nodal voltage variation and injection power variation based on multiple state estimation results. With the voltage-to-power sensitivities of all nodes, critical measurements can be obtained. Note that, the model-based state estimation is only utilized in the critical measurement selection, instead of the data-driven control process. b) As for the extended planning of ADN, operation data can be simulated through the physical model of ADN. The network structure can be provided from the planning information of ADN. The photovoltaic (PV), wind turbine (WT) and load curves can be obtained by forecasting information. Based on the physical model, the voltage-to-power sensitivities of the extended ADN are calculated. Then the critical measurement selection can be carried out.

Through critical measurement selection based on historical data, a data-driven control model can be established without the requirement of the measurements in the whole network. It helps to reduce the measurement demand as well as the dimension of the data-driven control model, which is beneficial to the convergence speed and the control performance.

B. Implementation of the Coordinated Voltage Control

Fig. 3 shows the flowchart of data-driven coordinated voltage control on multiple time-scales.

1) The similarities of nodes are quantified based on the modified cosine similarity. Nodes with high similarity values are classified as one category. In each category, the most representative nodes are recommended as critical measurements.

2) Concerning the critical measurements at instant t and the prediction information during ΔT, a data-driven model of discrete regulation devices on the slow time scale is established. The operation strategies of discrete regulation devices are solved and implemented to mitigate the voltage deviation on the slow time-scale ΔT.

3) Based on the critical measurements in each area, the data-driven models of continuous regulation devices are respectively established on each fast control period Δt. The operation strategies of continuous regulation devices are calculated and implemented on each Δt, which can adaptively suppress frequent voltage fluctuations.

4) The operation strategies of continuous regulation devices are adjusted continuously until ΔT is reached. Then Steps 2) and 3) are repeated, until the total control horizon T is reached.

Remark 3: The proposed data-driven voltage control method is mainly oriented for medium-voltage distribution networks, which are three-phase balanced and can be generally equivalent to a single phase. To address the voltage control problem in unbalanced three-phase ADN, a three-phase data-driven control method can be further extended. Correspondingly, the data-driven model and variables will be expanded from the single-phase to three-phase. Considering the asymmetric integration of DGs in unbalanced ADNs [34], it is assumed that three-phase data-driven voltage control is conducted separately. First, nodal voltage measurements are obtained to establish the data-driven voltage control model in each phase. Then, operational strategies of the regulation devices can be calculated and implemented in each phase. Thus, voltage control problem of unbalanced ADNs can be solved by conducting data-driven control in each phase.

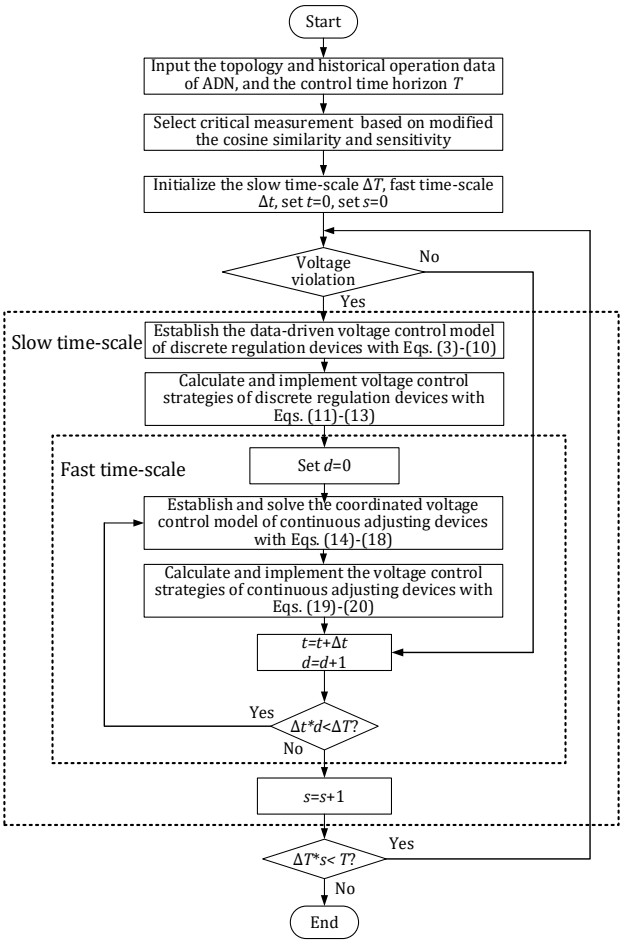


Fig. 3. Flowchart of data-driven coordinated voltage control.

Thus, by establishing the data-driven model of coordinated voltage control on multiple time-scales, the nodal voltage can be maintained within the desired range. In addition, by selecting the critical measurements, the requirement of total measurement amounts is reduced in the data-driven model.

IV. Case Studies and Analysis

To verify the advantages of the proposed data-driven coordinated voltage control approach on multiple time-scales, two standard IEEE test cases are adopted. The case studies are performed in MATLAB R2016b, which is installed on a computer with an Intel(R) Core(TM) i7-6700HQ CPU@2.60GHz and 16 GB of RAM.

A. *Modified IEEE 33-node System*

The topology of the tested 33-node distribution system is illustrated in Fig. 4, including a substation and 33 nodes. A ten-tap step OLTC with 1% voltage adjustment per tap is connected to node 1. And the rated voltage level and total power demands of the tested system can refer to Ref. [35].

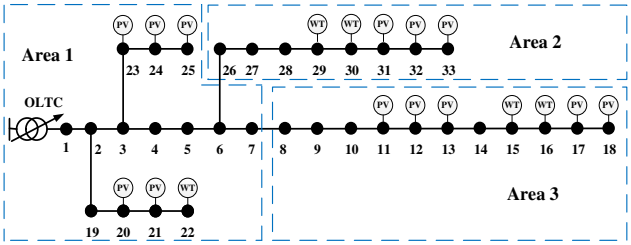


Fig. 4. Topology of the tested 33-node system.

Accounting for the high penetration of DGs, 13 units of PV are installed at nodes 11, 12, 17, 18, 20, 21, 23, 24, 25, 31, 32, and 33, of which capacities are 100 kWp each. 3 units of WT are installed at nodes 13, 15, and 16, of which capacities are 500 kVA each. 3 units of WT are installed at nodes 22, 29, and 30, of which capacities are 100 kVA each.

B. *Selection of Critical Measurements*

Typical scenarios are clustered based on the annual historical data of ADN, as shown in Fig. 5.

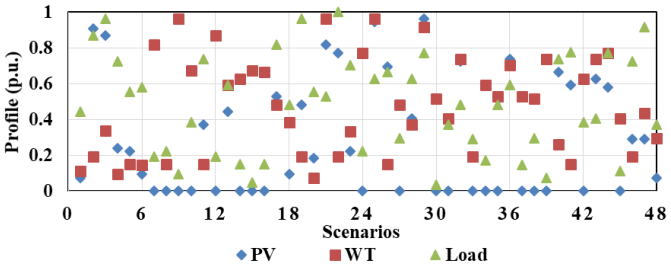
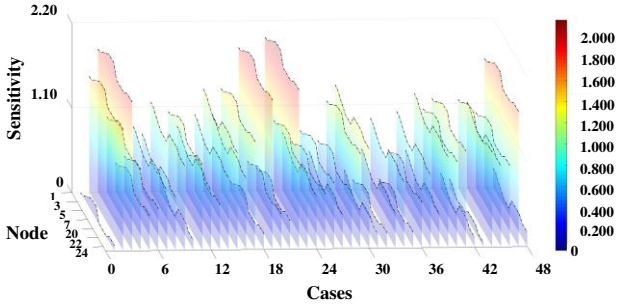
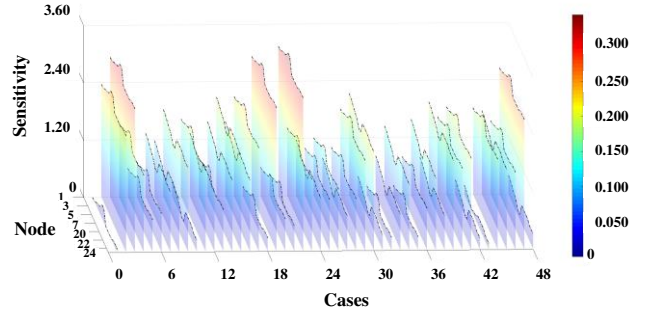


Fig. 5. 48 typical scenarios of ADN.

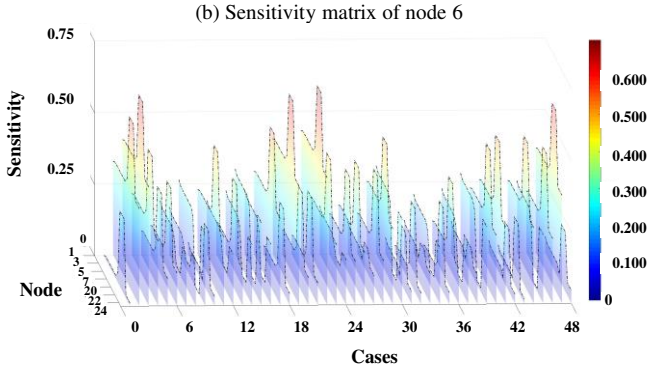
Then the voltage-to-power sensitivities of nodes in each typical scenario can be obtained based on the historical data. Subsequently, the characteristic graph of each node can be obtained under typical scenarios. Taking nodes 5, 6, 19, and 20 in area 1 as examples, the sensitivity characteristic graphs are shown in Fig. 6.



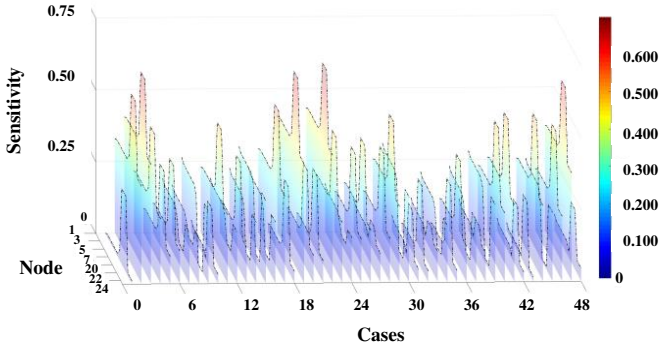
(a) Sensitivity matrix of node 5



(b) Sensitivity matrix of node 6



(c) Sensitivity matrix of node 19



(c) Sensitivity matrix of node 20

Fig. 6. Graphs of the sensitivity matrixes.

On the basis of the sensitivity matrix mentioned in Section III.A, the modified cosine similarity is calculated to quantify the similarity between nodes. Table I shows the results of the modified cosine similarity calculation.

It is illustrated from Fig. 6 and Table I that the similarity between nodes 5 and 6 is higher than the similarity between nodes 5 and 19, or nodes 5 and 20. Assuming the threshold ε is 0.99, nodes 5 and 6 can be clustered into one category, while nodes 19 and 20 belong to different categories due to lower similarities. Then in a category, the most representative nodes are recommended as the critical measurements. The result of critical measurements selection is shown in Fig. 7.

TABLE I
SIMILARITY RESULTS BETWEEN NODES

Nodes	5-6	5-19	5-20	6-19	6-20	19-20
Similarity	0.9989	0.9509	0.7064	0.9487	0.7059	0.8488

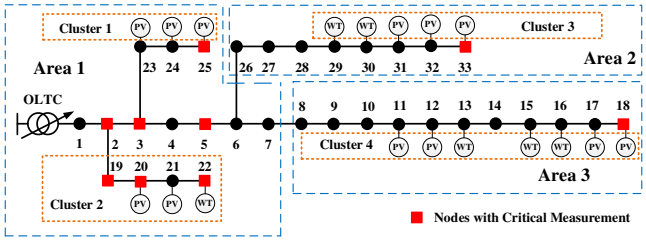


Fig. 7. Results of critical measurements selection.

C. Analysis of Voltage Control Results

1) Data-driven coordinated voltage control

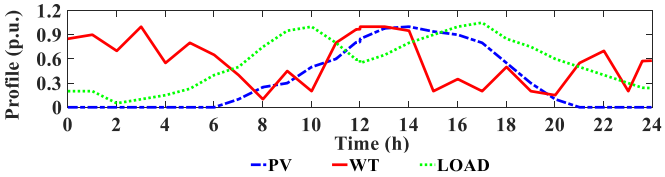


Fig. 8. Daily curves of power loads and DGs.

Fig. 8 shows the typical daily curves of power loads and DGs, which are used to analyse the voltage control results. The time resolution of the fast time-scale is set as 0.5-minute. [0.95, 1.05] p.u. is set as the limited operational range of voltage profiles. The weight factors are all set as 1.0. Based on the operational data, the data-driven control method needs to select an identical optimal reference for both upper and lower violation of nodal voltage. Thus, the voltage reference $\boldsymbol{U}^{\mathrm{ref}}$ is set as $\mathbf{1}$.

Considering that frequent switching actions of OLTC increase security risks to ADN, set the duration of slow time-scale ΔT as 4 hours. The prediction information of the hourly operation curves is obtained based on Fig. 8. To obtain prediction information with higher accuracy, a graph neural network can be utilized in future work [36]. Considering the frequent fluctuation of DGs, set the fast time-scale Δt and the sampling interval of voltage measurements as 0.5-minute.

Then three scenarios are utilized to demonstrate the advantages of the proposed data-driven coordinated voltage control with critical measurements. The description of the studied scenarios is shown in Table II.

TABLE II
DESCRIPTION OF THE THREE SCENARIOS

Scenario	Description
I	The initial state of ADN is obtained without the regulation of nodal voltage.
II	The proposed data-driven coordinated operation strategies of DGs and OLTC are adopted.
III	The uncoordinated voltage control strategies are adopted.

In Scenario III, the discrete and continuous regulation devices are not coordinated on multiple time-scales. OLTC is regulated only based on real-time information without the prediction information of DGs. In addition, the cooperation between multiple DG clusters is not taken into consideration.

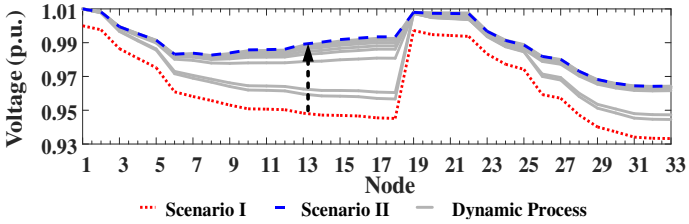
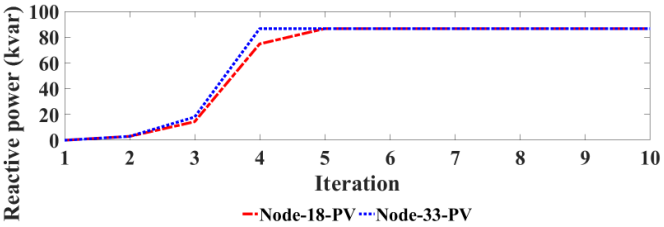
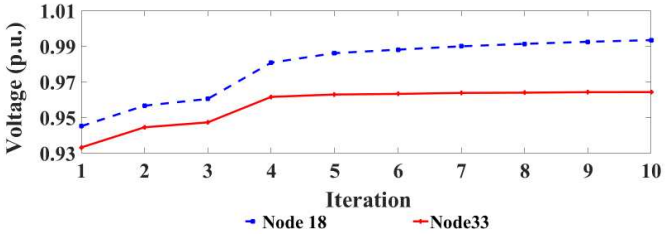


Fig. 9. Voltage control performance at 10:00 am.

The proposed data-driven coordinated voltage control with critical measurements is adopted for mitigating the voltage violation of ADN. The operation state of ADN from 10:00 am to 10:05 am is taken as an example. Fig. 9 illustrates the voltage control performance in which the grey areas indicate the voltage adjustment process of Scenario II.



(a) Dynamic process of PVs' reactive power outputs at nodes 18 and 33



(b) Voltage adjustment process of nodes 18 and 33

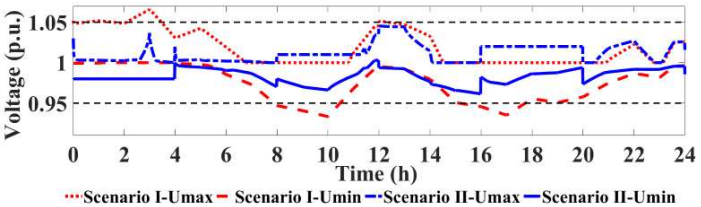
Fig. 10. Dynamic adjustment process from 10:00 am to 10:05 am.

The dynamic process of reactive power outputs of PVs is shown in Fig. 10(a). The voltage adjustment process of nodes 18 and 33 is shown in Fig. 10(b). DGs are utilized to regulate nodal voltage by generating reactive power to mitigate the voltage deviation. As for the computation efficiency, the proposed data-driven coordinated voltage control with critical measurements relieves the computation burden as it only requires a simple algebraic operation. Also, convergence can be ensured within limited iterations [25].

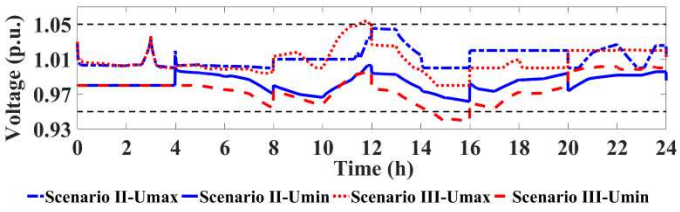
2) Analysis of voltage control performance

To test the daily voltage control performance, the proposed data-driven coordinated voltage control with critical measurements is carried out on the whole day.

The daily voltage control performances of the three scenarios are illustrated in Fig. 11.



(a) Comparison of voltage control performances of Scenarios I and II



(b) Comparison of voltage control performances of Scenarios II and III

Fig. 11. Comparison of voltage control performances of three scenarios.

Fig. 11(a) compares the voltage control performances of Scenarios I and II. In Scenario I, the highly penetrated DGs creates serious voltage violation. However, in Scenario II, by implementing the proposed coordinated voltage control approach, the OLTC regulates voltage on a slow time scale. While DGs are coordinated on a fast time scale to suppress the voltage fluctuation. Consequently, the nodal voltage of ADN is regulated and maintained within a reasonable operation range. However, the multiple regulation devices are operated uncoordinatedly in Scenario III, which fails to mitigate the voltage violation, as illustrated in Fig. 11(b).

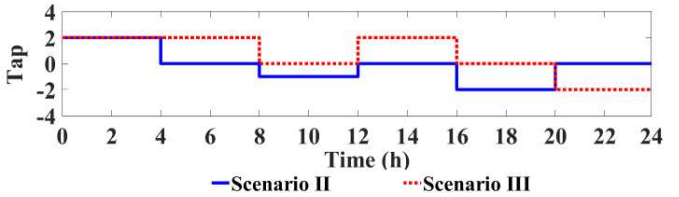


Fig. 12. Daily adjustment strategies of OLTC.

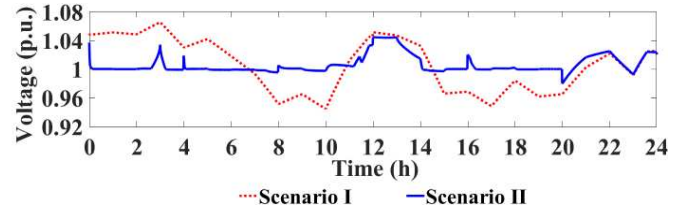


Fig. 13. Daily voltage control performance of node 18.

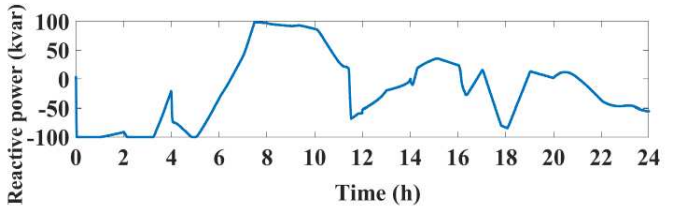


Fig. 14. Daily operation strategies of DG at node 18

The daily adjustment strategies of OLTC in Scenarios II and III are represented in Fig. 12. The daily voltage profiles of node 18 are shown in Fig. 13. Daily operation strategies of DG at node 18 are shown in Fig. 14. It can be seen from Figs. 12, 13, and 14 that the adjustment strategies of OLTC and DGs are coordinated to maintain the voltage within the intended limitation. From 0:00 am to 6:00 am, the power generated from DGs is beyond the power demands. The OLTC adjusts the tap step, while DGs coordinately absorb the reactive power to regulate the voltages. Conversely, from 6:00 am to 12:00 am, DGs cannot satisfy the high power demand. The OLTC adjusts the tap step to increase nodal voltages. Meanwhile, DGs generate reactive power to alleviate voltage deviation.

In addition, on the slow time scale, the OLTC changes the tap step based on the prediction information, to prevent further exacerbation of voltage profiles. At each operation time of OLTC, the voltage fluctuation appears and DGs are coordinated to suppress the voltage fluctuation.

To quantify the voltage control effect, the voltage deviation index (VDI) of the three scenarios is stated in Table III, which can be defined in Eq. (27).

TABLE III
VOLTAGE CONTROL PERFORMANCES OF THE THREE SCENARIOS

Scenario	Min. voltage (p.u.)	Max. voltage (p.u.)	VDI	Number of tap movement
I	0.9332	1.0658	0.0179	0
II	0.9615	1.0454	0.0087	5
III	0.9374	1.0540	0.0152	4

Table III illustrates that the improvement of voltage profiles is much more obvious in Scenario II compared to Scenarios I and III. In Scenario II, the VDI is diminished by 51.40% compared with Scenario I, while decreased by 42.76% than the uncoordinated voltage control in Scenario III.

$$VDI = \frac{\sum_{t=1}^{T}\sum_{i=1}^{N}|U_{t,i}-1|}{T\cdot N} \tag{27}$$

To evaluate the suppression of voltage fluctuations, the daily voltage flicker index (VFI) can be calculated according to Ref. [37]. The definition of the VFI is shown in Eq. (28).

$$VFI = \sum_{t\in T}\left|U_{t,i}-U_{t-\Delta t,i}\right|/(T\cdot N)\ i\in\mathcal{N} \tag{28}$$

In Scenario II, the VFI is decreased by 30.32% compared with Scenario I, namely from 1.88e-3 to 1.31e-3.

The computational efficiency of the proposed data-driven voltage control with critical measurements in Scenario II is 0.029 second for each iteration, which is demonstrated to be suitable for real-time voltage control.

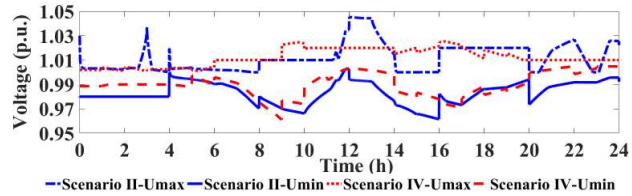
3) Comparison with existing approaches

To demonstrate the effectiveness and advantage of the proposed method, another two existing voltage control methods are studied and compared. The former one is the model-based centralized optimization approach, which requires accurate physical models and parameters of ADNs. The latter one is a deep reinforcement learning-based approach, which may need the retraining process under topology changes of ADNs.

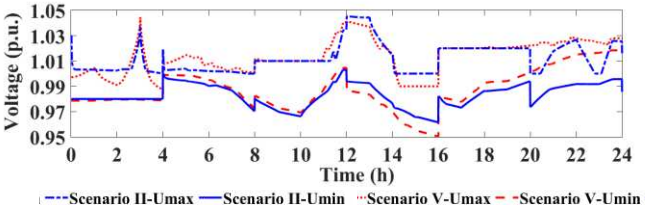
a) Ref. [35] proposed a mixed-integer second-order cone programming model for voltage control in ADNs with accurate network parameters. Based on Ref. [35], the model-based centralized control approach is adopted in Scenario IV, which has the optimal voltage control performance.

b) Deep deterministic policy gradient network (DDPG) is a typical deep reinforcement learning-based algorithm with two-layer actor-critic networks, that is suitable for large-scale control problems with deterministic and continuous actions [38]. Based on the training process, the operational strategies can be determined by DDPG without the interaction with ADNs. To compare the control performance of the proposed data-driven coordinated voltage control, a DDPG-based voltage control approach is further adopted in Scenario V.

The comparisons of voltage control performance in Scenarios II, IV, and V are illustrated in Fig. 15 and Table IV.



(a) Comparison of voltage control performances of Scenarios II and IV



(b) Comparison of voltage control performances of Scenarios II and V

Fig. 15. Comparison of voltage control performances of three scenarios.

TABLE IV
VOLTAGE CONTROL PERFORMANCES OF THE THREE SCENARIOS

Scenario	Min. voltage (p.u.)	Max. voltage (p.u.)	VDI	Number of tap movement
II	0.9615	1.0454	0.0087	5
IV	0.9611	1.0254	0.0076	4
V	0.9503	1.0446	0.0130	4

As shown in Table IV, it is illustrated that the proposed data-driven method has similar control performance to the model-based centralized control approach and the DDPG-based voltage control approach. All three methods can effectively mitigate the voltage deviation and maintain voltages within the desired ranges. Considering the influence of hyperparameters, the results of DDPG approach can be further improved in the future work.

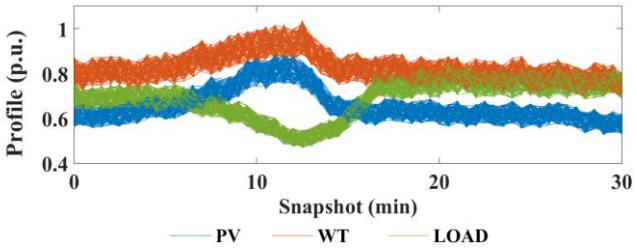
To assess the optimization effect of the data-driven method in Scenario II, the average optimization rate (AOR) is defined as follows [39].

$$AOR = (1 - |(\gamma_{\mathrm{d}} - \gamma_{\mathrm{c}})/(\gamma_{\mathrm{c}} - \gamma_{\mathrm{O}})|) * 100\% \quad (29)$$

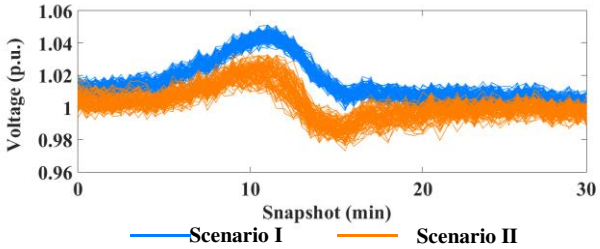
where γ_{d} denotes the voltage deviation index of the proposed data-driven voltage control approach. γ_{c} is the voltage deviation index of the model-based centralized control approach. While γ_{O} denotes the voltage deviation index of the initial operational state of ADN. The AOR of the proposed data-driven method is 89.32%. It means the proposed data-driven voltage control method reaches 89.32% of the optimal solution without relying on accurate physical models.

4) Adaptability to the uncertainty of ADN

The uncertainties of ADN generally include rapid fluctuations of DGs and loads, as well as the topology changes. When the DG outputs fluctuate, the measurements of ADN will change. Based on the real-time measurements, the proposed data-driven control model is updated correspondingly. Then, the data-driven control model can rapidly adjust the operational strategies of controlled devices to adapt to the uncertainty of DGs and loads. The rapid fluctuations of loads and DGs from 11:30 am to 12:00 am are considered to test the adaptability of the proposed data-driven voltage control method. Considering 10% random disturbance of loads and DGs, 100 scenarios are randomly generated by Monte Carlo method to conduct the proposed data-driven control.



(a) Random disturbance of loads and DGs



(b) Voltage control performances of node 18 of Scenarios I and II

Fig. 16. The dynamic process with DG fluctuation.

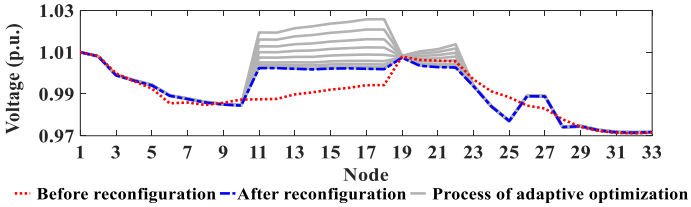
TABLE V
VOLTAGE CONTROL PERFORMANCES WITH DG UNCERTAINTY

Scenario	Expected min. voltage (p.u.)	Expected max. voltage (p.u.)	VDI
I	0.9647	1.0466	0.0103
II	0.9697	1.0276	0.0087
IV	0.9725	1.0462	0.0088

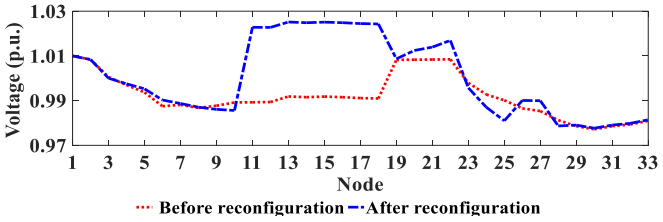
The random disturbance of loads and DGs is shown in Fig. 16(a). The voltage profiles of node 18 in Scenarios I and II are shown in Fig. 16(b). The voltage control performance of the whole ADN is illustrated in Table V. It can be seen from Fig. 16 and Table V that the proposed data-driven method can respond to the random fluctuations of DG and loads and mitigate the voltage deviation. However, the control performance of model-based approach is deteriorated under DG uncertainty.

In addition, to test the adaptability of the proposed data-driven method to the topology change of ADN, an economic network reconfiguration is further considered. The economic network reconfiguration is presumed to be carried out at 10:30 am. The tie switches between nodes 12 and 22, as well as nodes 25 and 29, are closed, and branch switches between nodes 10 and 11, as well as nodes 27 and 28, are disconnected.

Fig. 17 illustrates the voltage control performances during the network reconfiguration process. It can be seen from Fig. 17(a) that the proposed data-driven method can dynamically adjust voltage profiles to adapt to the new topology.



(a) Voltage control performances of Scenario II



(b) Voltage control performances of Scenario IV

Fig. 17. The adaptive process to the topology changes.

D. *Modified IEEE 123-node System*

The modified IEEE 123-node distribution system is adopted to verify the scalability of the proposed data-driven method on ADNs with three-phase unbalanced conditions. Fig. 18 shows the topology of the test system. A ten-tap step OLTC with 1% voltage adjustment per tap is connected to node 1. To fully consider the impact of asymmetric access of DGs, six PVs with a capacity of 1000 kWp and three WTs with a capacity of 1000 kVA are integrated into the distribution networks. The locations of DGs are listed in Table VI. The same Scenarios I and II in Section IV.C are also carried out in the test case.

The voltage control performance of node 78 is taken as an example to illustrate the control effects on the three-phase ADNs. The daily voltage control performance is demonstrated in Table VII.

TABLE VI
INSTALLATION PARAMETERS OF DGS

Parameter	PV						WT		
Location	33	42	59	77	86	91	28	51	76
Phase	A	B	B	C	C	A	A	C	B

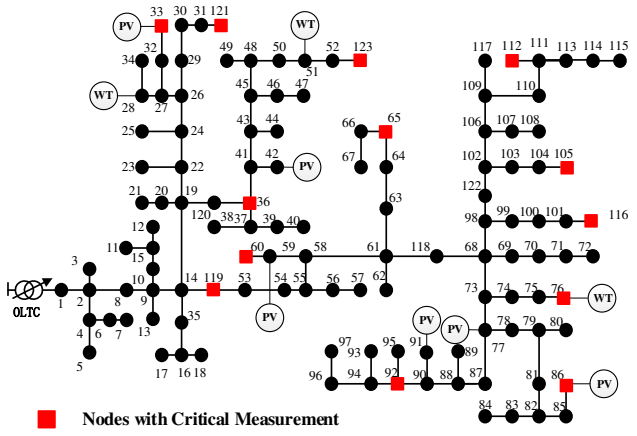
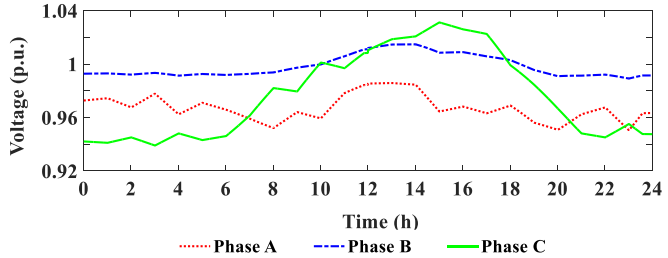
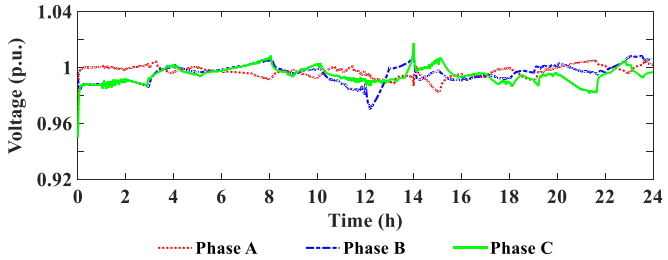


Fig. 18. Topology of the tested 123-node system.



(a) Voltage control performances of node 78 in Scenario I



(b) Voltage control performances of node 78 in Scenario II

Fig. 19. Voltage profiles of the unbalanced three-phase ADN.

TABLE VII
VOLTAGE CONTROL PERFORMANCES OF THE THREE-PHASE CASE

Phase	Min. voltage (p.u.)		Max. voltage (p.u.)		*VDI*	
	Scenario I	Scenario II	Scenario I	Scenario II	Scenario I	Scenario II
Phase A	0.9385	0.9503	1.0573	1.0375	0.0270	0.0084
Phase B	0.9719	0.9526	1.0850	1.0484	0.0146	0.0068
Phase C	0.9301	0.9504	1.0724	1.0463	0.0226	0.0085

It can be seen from Table VII, in Scenario II, the *VDI* of three phases are diminished by 68.89%, 53.42% and 62.39% respectively, compared with Scenario I. Thus, the proposed data-driven voltage control method can effectively mitigate voltage deviation from the nominal value in each phase, as shown in Fig. 19. The scalability of the proposed data-driven voltage control method is verified.

Thus, the proposed data-driven coordinated voltage control approach with critical measurements can effectively alleviate the impact of DG integration and suppress the serious voltage violation in ADNs. In addition, the proposed voltage control approach is also capable of responding to the frequent topology and state variations of ADNs.

V. CONCLUSIONS

A data-driven coordinated voltage control method is presented for the coordination of OLTC and DG inverters on multiple time-scales to maintain voltages within the desired range. By utilizing the real-time measurement data, a data-driven coordinated voltage control model is established, in which discrete and continuous regulation devices are considered and coordinated on multiple time-scales. In addition, a method of critical measurement selection method is proposed to guarantee the voltage control performance under the partial measurements in practical ADNs. Comprehensive case studies with different scenarios are conducted to verify the effectiveness of the proposed data-driven method. The results demonstrate that the proposed method can effectively suppress the voltage fluctuation on multiple time-scales and rapidly adapts to the frequent variations in ADNs.

Several research directions are worth to be investigated in the future. A combination of data-driven model and physical model can be researched for better control performance. Then, the time-series features of the energy storage system should be further taken into account. In addition, a graph neural network can be further considered for a better prediction accuracy.

APPENDIX

Proof of Theorem A1:

Based on the general expression of nonlinear systems in Ref. [24], a brief proof of Theorem A1 is provided as follows:

The discrete data model of ADN can be transformed from Eq. (1), which is formulated as Eq. (A.1).

$$\boldsymbol{y}[t+\Delta t]=f\begin{pmatrix}\boldsymbol{y}[t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t\big],\boldsymbol{x}[t],\\ \cdots,\boldsymbol{x}[k-n_x\Delta t]\end{pmatrix} \tag{A.1a}$$

$$f(\cdot)=(f_1(\cdot),\cdots,f_m(\cdot))^{\mathrm{T}}\in\textstyle\prod_{n_y+n_x+2}\mathbf{R}^m \tag{A.1b}$$

where t denotes the indices of instants. $\boldsymbol{y}\in\mathbf{R}^m$, denotes the measurements vector of ADN such as the voltage amplitude of each node. $\boldsymbol{x}\in\mathbf{R}^m$. n_y and n_x are two unknown constants.

$$\Delta\boldsymbol{y}[t+\Delta t]=\boldsymbol{\Phi}[t]*\Delta\boldsymbol{x}[t] \tag{A.2a}$$

$$\Delta\boldsymbol{y}[t+\Delta t]=\boldsymbol{y}[t+\Delta t]-\boldsymbol{y}[t] \tag{A.2b}$$

$$\Delta\boldsymbol{x}[t]=\boldsymbol{x}[t]-\boldsymbol{x}[t-\Delta t] \tag{A.2c}$$

Then Eq. (A.3) can be obtained by Combining Eq. (A.1) and Eq. (A.2).

$$\begin{aligned}&\Delta\boldsymbol{y}[t+\Delta t]=\boldsymbol{y}[t+\Delta t]-\boldsymbol{y}[t]\\&=f\big(\boldsymbol{y}[t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t\big],\boldsymbol{x}[t],\cdots,\boldsymbol{x}[k-n_x\Delta t]\big)\\&-f\big(\boldsymbol{y}[t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t\big],\boldsymbol{x}[t-\Delta t],\boldsymbol{x}[t-\Delta t],\cdots,\boldsymbol{x}[k-n_x\Delta t]\big)\\&+f\big(\boldsymbol{y}[t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t\big],\boldsymbol{x}[t-\Delta t],\boldsymbol{x}[t-\Delta t],\cdots,\boldsymbol{x}[k-n_x\Delta t]\big)\\&-f\begin{pmatrix}\boldsymbol{y}[t-\Delta t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t-\Delta t\big],\boldsymbol{x}[t-\Delta t],\\ \cdots,\boldsymbol{x}[k-n_x\Delta t-\Delta t]\end{pmatrix}\end{aligned} \tag{A.3}$$

Based on the Cauchy mean value theorem, Eq. (A.4) can be further obtained.

$$\begin{aligned}&f\big(\boldsymbol{y}[t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t\big],\boldsymbol{x}[t],\cdots,\boldsymbol{x}[k-n_x\Delta t]\big)\\&-f\begin{pmatrix}\boldsymbol{y}[t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t\big],\boldsymbol{x}[t-\Delta t],\boldsymbol{x}[t-\Delta t],\\ \cdots,\boldsymbol{x}[k-n_x\Delta t]\end{pmatrix}\\&=(\partial f(\cdot)/\partial\boldsymbol{x}[t])\Delta\boldsymbol{x}[t]\\&\big(\boldsymbol{y}[t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t\big],\boldsymbol{x}[t-\Delta t],\boldsymbol{x}[t-\Delta t],\cdots,\boldsymbol{x}[k-n_x\Delta t]\big)^{\mathrm{T}}\\&\le\partial f(\cdot)/\partial\boldsymbol{x}[t]\le\big(\boldsymbol{y}[t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t\big],\boldsymbol{x}[t],\cdots,\boldsymbol{x}[k-n_x\Delta t]\big)^{'}\end{aligned} \tag{A.4}$$

where $\partial f(\cdot)/\partial\boldsymbol{x}[t]$ is the partial derivative value of $f(\cdot)$ with respect to $\boldsymbol{x}[t]$.

Define a time-varying variable $\boldsymbol{B}[t]$ as follows:

$$\begin{aligned}\boldsymbol{B}[t]=&f\begin{pmatrix}\boldsymbol{y}[t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t\big],\boldsymbol{x}[t-\Delta t],\boldsymbol{x}[t-\Delta t],\\ \cdots,\boldsymbol{x}[k-n_x\Delta t]\end{pmatrix}\\&-f\begin{pmatrix}\boldsymbol{y}[t-\Delta t],\cdots,\boldsymbol{y}\big[t-n_y\Delta t-\Delta t\big],\boldsymbol{x}[t-\Delta t],\\ \cdots,\boldsymbol{x}[k-n_x\Delta t-\Delta t]\end{pmatrix}\end{aligned} \tag{A.5}$$

Introducing Eq. (A.3) and Eq. (A.5) into Eq. (A.2), Eq. (A.6) can be obtained.

$$\Delta\boldsymbol{y}[t+\Delta t]=\frac{\partial f(\cdot)}{\partial\boldsymbol{x}[t]}\Delta\boldsymbol{x}[t]+\boldsymbol{B}[t] \tag{A.6}$$

Introducing a variable $\boldsymbol{\varpi}[t]$, in each control instant t, the functional relationship between $\boldsymbol{B}[t]$ and $\Delta\boldsymbol{x}[t]$ yields:

$$\boldsymbol{B}[t]=\boldsymbol{\varpi}[t]\Delta\boldsymbol{x}[t] \tag{A.7}$$

When $\Delta\boldsymbol{x}[t]\neq 0$, there is a unique solution of $\boldsymbol{\varpi}[t]$, such that Eq. (A.7) holds.

$$\boldsymbol{\Phi}[t]=\frac{\partial f(\cdot)}{\partial\boldsymbol{x}[t]}+\boldsymbol{\varpi}[t] \tag{A.8}$$

By defining $\boldsymbol{\Phi}[t]$ as Eq. (A.8), Eq. (A.2) can be obtained. Thus, theorem A1 is proved.

REFERENCES

[1] H. Gao, J. Li and B. Xu, "Principle and implementation of current differential protection in distribution networks with high penetration of DGs," *IEEE Trans. Power Deliv.*, vol. 32, no. 1, pp. 565-574, 2017.

[2] K. Mahmoud, M. Abdel-Nasser and M. Lehtonen, "Low-computational voltage-assessment approach considering fine-resolution simulations for distribution systems with photovoltaics," *IEEE Syst. J.*, vol. 15, no. 4, pp. 5321-5331, 2021.

[3] X. Chen, M. B. Mcelroy, Q. Wu, Y. Shu and Y. Xue, "Transition towards higher penetration of renewables: an overview of interlinked technical, environmental and socio-economic challenges," *J. Mod. Power Syst. Clean Energy*, vol.7, no. 1, pp.1-8, 2019.

[4] Erdiwansyah, Mahidin, H. Husin, Nasaruddin, M. Zaki and Muhibbuddin, "A critical review of the integration of renewable energy sources with various technologies," *Prot. Control Mod. Power Syst.*, vol. 6, no. 1, pp. 37-54, 2021.

[5] F. Ding and B. Mather, "On distributed PV hosting capacity estimation, sensitivity study, and improvement," *IEEE Trans. Sustain. Energy*, vol. 8, no. 3, pp. 1010-1020, 2017.

[6] M. M. Salem, N. I. Elkalashy, Y. Atia and T. A. Kawady, "Modified inverter control of distributed generation for enhanced relaying coordination in distribution networks," *IEEE Trans. Power Deliv.*, vol. 32, no. 1, pp. 78-87, 2017.

[7] H. Zhang, J. Chen, J. Yan, X. Song, R. Shibasaki, and J. Yan, "Urban power load profiles under ageing transition integrated with future EVs charging," *Advances in Applied Energy*, Vol. 1, pp. 100007, 2021.

[8] K. Mahmoud and M. Lehtonen, "Comprehensive analytical expressions for assessing and maximizing technical benefits of photovoltaics to distribution systems," *IEEE Trans. Smart Grid*, vol. 12, no. 6, pp. 4938-4949, 2021.

[9] Z. Tang, D. J. Hill, and T. Liu, "Fast distributed reactive power control for voltage regulation in distribution networks," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 802-805. 2019.

[10] Z. Tang, D. J. Hill, T. Liu and H. Ma, "Hierarchical voltage control of weak subtransmission networks with high penetration of wind power," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 187-197, 2018.

[11] X. Zhu, J. Wang, N. Lu, N. Samaan, R. Huang and X. Ke, "A hierarchical VLSM-based demand response strategy for coordinative voltage control between transmission and distribution systems," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 4838-4847, 2019.

[12] X. Liu *et al.*, "Multi-stage voltage support optimization for microgrids with multiple distributed generation units," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 141-156, 2021.

[13] X. Sun and J. Qiu, "Hierarchical voltage control strategy in distribution networks considering customized charging navigation of electric vehicles," *IEEE Trans. Smart Grid*, vol. 12, no. 6, pp. 4752-4764, 2021.

[14] S. Salih and P. Chen, "On coordinated control of OLTC and reactive power compensation for voltage regulation in distribution systems with wind power," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4026-4035, 2016.

[15] G. Kryonidis, C. Demoulias and G. Papagiannis, "A two-stage solution to the bi-objective optimal voltage regulation problem," *IEEE Trans. Sustain. Energy*, vol. 11, no. 2, pp. 928-937, 2020.

[16] K. M. Muttaqi, A. D. T. Le, M. Negnevitsky and G. Ledwich, "A coordinated voltage control approach for coordination of OLTC, voltage regulator, and DG to regulate voltage in a distribution feeder," *IEEE Trans. Ind. Appl.*, vol. 51, no. 2, pp. 1239-1248, 2015.

[17] Y. Li, P. Zhang, M. Althoff and M. Yue, "Distributed formal analysis for power networks with deep integration of distributed energy resources," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 5147-5156, 2019.

[18] Y. Gao, B. Foggo, and N. Yu, "A physically inspired data-driven model for electricity theft detection with smart meter data," *IEEE Trans. Ind. Inform.*, vol. 15, no. 9, pp. 5076-5088, 2019.

[19] L. Bai *et al.*, "A data-driven network optimization approach to coordinated control of distributed photovoltaic systems and smart buildings in distribution systems," *IET Energy Syst. Integr.*, Early Access, 2021.

[20] J. Zhang, Z. Chen, C. He, Z. Jiang and L. Guan, "Data-driven-based optimization for power system var-voltage sequential control," *IEEE Trans. Ind. Inform.*, vol. 15, no. 4, pp. 2136-2145, 2019.

[21] B. Foggo and N. Yu, "Improving supervised phase identification through the theory of information losses," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2337-2346, 2020.
[22] W. Wang, N. Yu, Y. Gao and J. Shi, "Safe off-policy deep reinforcement learning algorithm for Volt-VAR control in power distribution systems," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3008-3018, 2020.
[23] Z. Hou, S. Jin, "Data-driven model-free adaptive control for a class of MIMO nonlinear discrete-time system," *IEEE Trans. Neural Networks*, vol. 22, no. 12, pp. 2173-2188, 2011.
[24] Z. Hou, S. Jin, "A novel data-driven control approach for a class of discrete-time nonlinear systems," *IEEE Trans. Control Syst. Technol.*, vol. 19, no. 6, pp. 1549-1558, 2011.
[25] Z. Hou, Y. Zhu, "Controller-dynamic-linearization-based model free adaptive control for discrete-time nonlinear systems," *IEEE Trans. Ind. Inform.*, vol. 9, no. 4, pp. 2301-2309, 2013.
[26] Y. Huo *et al.*, "Data-driven adaptive operation of soft open points in active distribution networks" *IEEE Trans. Ind. Inform.*, vol. 17, no. 12, pp. 8230-8242, 2021.
[27] H. Lu, *et al*, "Network simplification-based cluster coordinated optimization method for distributed PVs with inadequate measurement," *IEEE Access*, vol. 8, pp. 65283-65293, 2020.
[28] K. Alzaareer, M. Saad, H. Mehrjerdi, D. Asber and S. Lefebvre, "Development of new identification method for global group of controls for online coordinated voltage control in active distribution networks," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 3921-3931, 2020.
[29] T. Xu, W. Wu, W. Zheng, H. Sun and L. Wang, "Fully distributed quasi-Newton multi-area dynamic economic dispatch method for active distribution networks," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4253-4263, 2018.
[30] J. E. Sarmiento, C. A. Alvez, B. de Nadai N., A. C. Zambroni de Souza, E. M. Carreno and P. F. Ribeiro, "A complex-valued three-phase load flow for radial networks: high-performance and low-voltage solution capability," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3241-3249, 2019.
[31] A. Sirisha and A. Pradhan, "Cosine similarity based directional comparison scheme for subcycle transmission line protection," *IEEE Trans. Power Deliv.*, vol. 35, no. 5, pp. 2159-2167, 2020.
[32] A. Rodriguez, A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, pp. 1492-1496, 2014.
[33] H. Su, *et al.*, "Novel voltage-to-power sensitivity estimation for phasor measurement unit-unobservable distribution networks based on network equivalent," *Appl. Energy*, vol. 250, pp. 302-312, 2019.
[34] B. Wang, C. Zhang, C. Li, P. Li, Z. Y. Dong and J. Lu, "Hybrid interval-robust adaptive battery energy storage system dispatch with SoC interval management for unbalanced microgrids," *IEEE Trans. Sustain. Energy*, vol. 13, no. 1, pp. 44-55, 2022.
[35] P. Li, *et al.*, "Coordinated control method of voltage and reactive power for active distribution networks based on soft open point," *IEEE Trans. Sustain. Energy*, vol. 8, no. 4, pp. 1430-1442, 2017.
[36] C. Li, Z. Dong, G. Chen, B. Zhou, J. Zhang and X. Yu, "Data-driven planning of electric vehicle charging infrastructure: a case study of Sydney, Australia," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3289-3304, 2021.
[37] K. Mahmoud, M. Lehtonen, "Three-level control strategy for minimizing voltage deviation and flicker in PV-rich distribution systems", *Int. J. Elec. Power & Energy Sys.*, vol. 120, pp. 105997, 2020.
[38] P. Li, *et al.*, "Deep reinforcement learning-based adaptive voltage control of active distribution networks with multi-terminal soft open point," *Int. J. Elec. Power*, vol. 141 no.8, pp. 108138, 2022.
[39] D. Cao, J. Zhao, W. Hu, F. Ding, Q. Huang and Z. Chen, "Attention enabled multi-agent DRL for decentralized Volt-VAR control of active distribution system using PV inverters and SVCs," *IEEE Trans. Sustain. Energy*, vol. 12, no. 3, pp. 1582-1592, 2021.

Yanda Huo received the B.S. degree in electrical engineering from Tianjin University, Tianjin, China, in 2016. He is currently working toward the Ph.D. degree in electrical engineering with Tianjin University, Tianjin, China.

His current research interest is data-driven control of distribution networks.

Peng Li (Senior Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Tianjin University, Tianjin, China, in 2004 and 2010, respectively.

He is currently a Professor with the School of Electrical and Information Engineering, Tianjin University. His current research interests include operation and planning of active distribution networks, modelling and transient simulation of power systems. Prof. Li is an associate Editor of IEEE TRANSACTION ON SUSTAINABLE ENERGY, CSEE Journal of Power and Energy Systems, Sustainable Energy Technologies and Assessments, and IET Renewable Power Generation.

Haoran Ji (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Tianjin University, Tianjin, China, in 2014 and 2019, respectively.

From 2019 to 2021, he was a Postdoctoral Research with Tianjin University. He is currently an Associate Professor in Tianjin University. His research interests include distributed generation systems and optimal operation of distribution networks. He was supported by China National Postdoctoral Program for Innovative Talents in 2019.

Hao Yu (Member, IEEE) received the B.S. and Ph.D. degree in electrical engineering from Tianjin University, Tianjin, China, in 2010 and 2015, respectively. He is currently an Associate Professor with the School of Electrical and Information Engineering, Tianjin University.

His current research interests include the operation analysis and optimization of active distribution networks and integrated energy systems. He is the assistant editor IET Energy Systems Integration.

Jinyue Yan received the Ph.D. degree from the Royal Institute of Technology KTH, Stockholm, Sweden, in 1991.

He is currently a Chair Professor of Energy Engineering at Mälardalen University, Sweden. Prof. Yan is an Academician of European Academy of Sciences and Arts, and serves as the advisory expert to the UN, EU, & ADB. His research interests include advanced energy systems, renewable energy, advanced power generation, climate change mitigation technologies and related environment and policy etc. Prof. Yan published about 400 papers including papers in Science, Nature Energy, Nature Climate & Nature Communications and hold 10+ patents with about 15000+ citations and H-index 62.

Prof. Yan is the Editor-in-Chief of Advances in Applied Energy & Editor-in-Chief of Handbook of Clean Energy Systems. He has led research platform (Future Energy Profile) with funding of over 80 million Euro by Swedish Knowledge Foundation and industrial partners.

Jianzhong Wu (Member, IEEE) received the Ph.D. from Tianjin University, Tianjin, China, in 2004. From 2004 to 2006, he was at Tianjin University, where he is an Associate Professor. From 2006 to 2008, he was a Research Fellow at the University of Manchester, Manchester, U.K. He is currently a Professor with the Cardiff School of Engineering, Institute of Energy, London, U.K.

Prof. Wu is the Editor-in-Chief of Applied Energy. His current research interests include energy infrastructure and smart grids.

Chengshan Wang (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Tianjin University, Tianjin, China, in 1991.

He is currently a Professor with the School of Electrical and Information Engineering, Tianjin University. Prof. Wang is a Member of Chinese Academy of Engineering. His research interests include distribution system analysis and planning, distributed generation system and microgrid.

Prof. Wang is the Editor-in-Chief of IET Energy Systems Integration. He is the Director of the Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin, China.