Data-driven Adaptive Operation of Soft Open Points in Active Distribution Networks

Abstract—The integration of soft open point (SOP) effectively improves the flexibility of active distribution networks (ADNs). However, in practical operation, accurate network parameters are difficult to obtain and the operation state changes rapidly with distributed generators (DGs). With the development of information technologies, massive operation data can be acquired in ADNs. How to utilize multi-source data has become the key to realize the intelligent operation of ADNs. This paper proposes a data-driven operation strategy of SOP based on model-free adaptive control (MFAC). First, considering the inaccurate parameters and frequent change of operation states, a data-driven framework is formulated for the real-time operation of SOP. Then, the operation strategies of multiple SOPs are further improved with inter-area coordination. The results of case studies show that driven by the measurement data, the potential benefits of SOPs are explored to adaptively respond to system state changes and improve the operational performance of ADNs.

Index Terms—active distribution network (ADN), soft open point (SOP), data-driven, adaptive control, coordinated operation.

<div align="center">NOMENCLATURE</div>

Sets

$\mathcal{N}$	Set of all nodes in ADN
$\mathcal{N}_{\mathrm{SOP}}$	Set of connected nodes of SOPs
Γ	Set of boundary branches
T	Set of time periods

Indices

t	Index of time periods
i	Index of nodes
l	Index of boundary branches
m, n	Index of boundary nodes

Variables

$Y[t], \widetilde{Y}[t]$	Measurement/estimation of variables to be optimized
$U[t], \widetilde{U}[t]$	Measurement/estimation of node voltage at time t
$X[t]$	Strategy of SOP at time t
$\Phi[t]$	Pseudo-Jacobi matrix at time t
$\tilde{P}_0[t]$	Estimated active power provided by the substation at time t
$P_i[t]$	Injected active power of node i at time t
$P^{\mathrm{ref}}[t]$	Total net load of ADN at time t
c_i	Cluster center index of node i
δ_i	Local density of node i
d_i	Distance between node i and other nodes
s_{ij}	Pseudo electrical distance between the node i and node j
ΔU_i^*	Measurement of voltage amplitude variation of node i
$\Delta P_j^*, \Delta Q_j^*$	Change of active and reactive power injection of node j
$\hat{P}_{a,l}[t], \hat{Q}_{a,l}[t]$	Boundary active and reactive power transmission of sub-area a at time t
$\hat{U}_{a,m}[t], \hat{U}_{a,n}[t]$	Voltage of boundary nodes m and n in sub-area a at time t
$U_{a,m}^{\mathrm{ref}}[t], U_{a,n}^{\mathrm{ref}}[t]$	Voltage reference of boundary nodes m and n in sub-area a at time t

$\bar{U}_{a,m}[t], \bar{U}_{a,n}[t]$	Estimated voltage of boundary node m and n in sub-area a at time t
$\boldsymbol{X}_a[t], \boldsymbol{X}_b[t]$	Strategy of SOP in upstream area a/downstream area b at time t
$\boldsymbol{X}'_a[t], \boldsymbol{X}'_b[t]$	Auxiliary variables of the upstream area a/downstream area b at time t
$P_{b,l}[t], Q_{b,l}[t]$	Boundary decision variables of the downstream area b at time t

Parameters

$\mu, \eta, \lambda, \rho, \omega$	Weight factors
γ	Clustering parameter
$\boldsymbol{U}^{\mathrm{ref}}$	Voltage reference of each node in ADN
N_{A}	Number of divided sub-areas
N_{C}	Number of cluster centers
N_n	Number of all nodes
N_{T}	Number of total time periods
$\overline{X}, \underline{X}$	Maximum and minimum limits of $\boldsymbol{X}$
ΔT	Optimization horizon
Δt	Control horizon
ε	Predefined tolerance

I. Introduction

With the flexible inter-connectivity and high controllability, soft open point (SOP) has become a critical equipment for the structural evolution of active distribution networks (ADNs) [1]. By accurately and rapidly regulating the power flow between connected feeders [2], SOP can effectively cope with the problem of voltage violation caused by the high shares of distributed generators (DGs) [3]. However, under practical and complex operation environments, accurate models and parameters of distribution networks are still difficult to obtain [4] and the day-ahead optimized strategy of SOP is challenged by frequent changes of operation state and uncertainties from DG [5]. The model-based operation strategies are more appropriate for the system analysis but bring significant challenges to the real-time operation and optimization of SOP.

Extensive research on the optimal operation of SOP has been carried out in ADNs. In [6], a multi-objective SOP operation optimization model was established to realize network loss reduction and voltage deviation mitigation. A two-stage robust operation strategy of SOP was proposed in [7] to cope with the uncertainty of PV generation. To improve the efficiency of SOP strategy, a combined decentralized and local voltage control strategy of SOP is proposed in [8]. It can achieve near-global optimization while reducing the burden of communication. However, the control strategies mentioned above rely on the accurate parameters and physical mechanism models of distribution networks. A more intelligent and adaptive regulation mode of SOP is essential in the absence of accurate parameters.

The rapid development of information and communication technologies (ICT) has promoted the information level of ADNs [9]. The massive amount of heterogeneous operation data can be measured and acquired in distribution networks [10]. It contains important information such as the characteristics of user behavior and the trend of system state [11]. How to fully utilize the operational data has become the key to improve the intelligent operation level of ADNs.

Data-driven control gradually draws wide attention, which is the approach that based on the measurement or historical data to realize the operation and optimal control of ADNs, without the requirement of accurate mechanism model [12]. It only utilizes the operational data to statistically describe and model the input and output relationships of a complex system. Data-driven control can effectively address the absence of accurate models and parameters of mechanism modeling and has strong applicability to the systems with strong nonlinearity. Thus, the data-driven control can provide auxiliary analysis support for the model-based operation strategies.

Considering the data sources, there are mainly two forms of data-driven control approach, including machine learning-based control and iteratively adaptive control. For simulating the dynamic process of ADN, the machine learning-based control uses a large amount of offline historical data to train neural networks and matches online data to realize the operation optimization. However, it is time consuming to complete the training process and hardly adapts to the system state changes. For example, if the network topology changes, the previous trained model may be not effective and needs to be retrained. The other form is iteratively adaptive control, which fully utilizes the operational data to model the input and output relationships of ADN operation. The iterative interaction process with ADNs is conducted to realize the real-time control of SOP. As an accurate mechanism model is not required, data-driven control can effectively address the absence of accurate parameters of mechanism modelling and has strong applicability to systems with features of nonlinearity and uncertainty.

Preliminary research has been carried out on the data-driven operation of distribution networks. In terms of the analysis framework, a research framework of big data was proposed in [13] for the intelligent operation of power grids. The authors in [14] summarized and provided a prospect of data-driven analysis of power system security assessment. As for the efficient operation, a data-driven approach was used in [15] to linearize power flow calculations. The authors in [16] further implemented local control of DG based on data-driven methods to reduce dependence on topology parameters. Regarding the prediction of source-load coordination, short-term photovoltaic prediction was performed based on massive data in [17]. A data-driven approach based on artificial intelligence was further proposed in [18] to realize the risk assessment in distribution networks. Thus, it is promising to excavate the multi-source data to improve the adaptive ability of SOP and to facilitate the intelligent operation of ADNs in complex environments.

Model-free adaptive control (MFAC) is a typical data-driven approach [19], of which the basic framework has been proposed in [20]. Under some practical assumptions, the convergence and stability of MFAC can be guaranteed in the multi-input multi-output nonlinear systems [21]. Ref. [22] introduced the idea of MFAC in the design of power system stabilizers to suppress low-frequency oscillations. By using MFAC, the coordinated power control issue of interconnected microgrids was effectively solved in [23]. Compared with the conventional data-driven approach, the MFAC approach has several merits to make it suitable for the adaptive operation of SOPs. First, MFAC only depends on the measurement data of the controlled system [24], which could be easily implemented with a small computational burden for its simplicity. Massive historical data and training process are not required in MFAC, which are necessary for artificial intelligence-based methods. In addition, MFAC can effectively adapt to the topology reconfiguration and operation state changes in ADNs.

Facing the parameter inaccuracy and state change of ADN under complex operation environments, this paper proposes an adaptive real-time operation strategy of SOP based on MFAC. Through the interaction of measurement information with ADNs, the data-driven model of SOP regulation is established by using real-time data, which effectively adapts to the state changes and improves the operational performance of ADNs.

The main contributions are summarized as follows:

1) Considering the lack of accurate parameters and frequent change of operation state, a data-driven framework for SOP operation is formulated based on MFAC. Based on the measurement data, the potential benefits of SOPs are explored to adaptively respond to system state changes and improve the operational performance of ADN in complex uncertain environments.

2) Oriented for the decentralized control, the operation strategies of multiple SOPs are further improved with inter-area coordination. First, the controllable areas of multiple SOPs are divided dynamically and the adaptive operation model of SOP is established based on the measurement information in each area. Then, the operation strategies of SOPs are further improved by inter-area coordination to maintain the voltage within the limits. By fully excavating the real-time data for SOP operation in ADNs, the voltage profile is effectively improved while network power losses are minimized simultaneously.

The remainder of this paper is organized as follows. Section II describes the data-driven adaptive operation control of SOP, and the principle of MFAC with SOP is elaborated. In Section III, the data-driven coordinated operation with multiple SOPs is proposed. The case studies and effectiveness verification are given in Section IV. The conclusions are stated in Section V.

II. DATA-DRIVEN ADAPTIVE OPERATION OF SOP

Facing the problem of parameter inaccuracy and state change of ADN under complex operation environments, a data-driven framework of SOP operation is proposed based on MFAC to dynamically realize the adaptive control in ADN.

A. Data-driven Adaptive Control

The essence of data-driven adaptive operation of SOP is to establish the relationship between the objectives of controllable area and SOP strategy, which can be abstractly expressed as:

$$Y = f(X) \tag{1}$$

where f represents the unsolved function relationship. Y denotes the variables in objectives, such as node voltage and network power loss. X denotes the operation strategy of SOPs, including active power transmission and reactive power outputs, shown as follows.

$$X = [P^{\text{sop}}, Q_1^{\text{sop}}, Q_2^{\text{sop}}]^{\text{T}} \tag{2}$$

Assuming that the partial derivative of f is continuous, the dynamic linearization model is established at each time instant:

$$\widetilde{Y}[t + \Delta t] = Y[t] + \mathbf{\Phi}^{\text{T}}[t] * (X[t] - X[t - \Delta t]) \tag{3}$$

where $\widetilde{Y}[t + \Delta t]$ is the estimation of variables to be optimized at time $t + \Delta t$. $Y[t]$ is the measurement of variables to be optimized at time t. $\mathbf{\Phi}[t]$ is the pseudo-Jacobi matrix (PJM) which implies the sensitivity between objective and control strategy of SOP. The estimation function of $\mathbf{\Phi}[t]$ is represented as follows:

$$\widehat{\mathbf{\Phi}}[t] = \text{argmin} \left(\left\| U[t] - \widetilde{U}[t] \right\|_2^2 + \mu \left\| \widehat{\mathbf{\Phi}}[t] - \widehat{\mathbf{\Phi}}[t - \Delta t] \right\|_2^2 \right) \tag{4}$$

where $\widehat{\mathbf{\Phi}}[t]$ is the estimation of $\mathbf{\Phi}[t]$. The iterative expression for solving $\widehat{\mathbf{\Phi}}[t]$ by gradient descent method is shown as follows:

$$\widehat{\mathbf{\Phi}}[t] = \widehat{\mathbf{\Phi}}[t - \Delta t] + \left(\Delta Y[t] - \widehat{\mathbf{\Phi}}[t - \Delta t] \Delta X[t - \Delta t] \right) \frac{\eta \Delta X[t - \Delta t]}{\mu + \|\Delta X[t - \Delta t]\|_2^2} \tag{5.a}$$

$$\widehat{\mathbf{\Phi}}[t] = \widehat{\mathbf{\Phi}}[1], \text{ if } \widehat{\mathbf{\Phi}}[t] \leq \epsilon \text{ or } \Delta X[t - \Delta t] \leq \epsilon \text{ or sign}(\widehat{\mathbf{\Phi}}[t]) \neq \text{sign}(\widehat{\mathbf{\Phi}}[1]) \tag{5.b}$$

$$\widehat{\boldsymbol{\Phi}}_{ij}[1] = S_{ij} = \partial \Delta U_i^* / \partial \Delta P_j^* + \partial \Delta U_i^* / \partial \Delta Q_j^* \; i \in \mathcal{N}, j \in \mathcal{N}_{\text{SOP}} \qquad (5.c)$$

where $\Delta \boldsymbol{U}[t]$ denotes $\boldsymbol{U}[t] - \boldsymbol{U}[t-\Delta t]$. $\Delta \boldsymbol{X}[t-\Delta t]$ denotes $\boldsymbol{X}[t-\Delta t] - \boldsymbol{X}[t-2\Delta t]$. ϵ is a small positive constant. $\widehat{\boldsymbol{\Phi}}[\mathbf{1}]$ is the initial value of $\widehat{\boldsymbol{\Phi}}[t]$. The Eq. (5.b) is introduced to make Eq. (5.a) have stronger parameter tracking ability.

The initial state of ADN may not be optimal, but will be within the physical operation constraints, which is guaranteed by the regulation of distribution system operator (DSO). The initial state of ADN is represented by the initial value of PJM $\widehat{\boldsymbol{\Phi}}[1]$, which will influence the convergence speed and steady-state control results. If the initial value of PJM $\widehat{\boldsymbol{\Phi}}[1]$ is well obtained, the proposed control method can adaptively adjust the operation strategies of SOPs with less steps. The optimal objective can be achieved by iteration. In the iteration process, the physical operation constraints of ADN are always maintained within a reasonable range. Thus, to reduce the iteration number and alleviate the impact on ADN, an appropriate initial value of the PJM $\widehat{\boldsymbol{\Phi}}[1]$ should be selected. The sensitivities between objective and SOP regulation are used to initialize $\widehat{\boldsymbol{\Phi}}[1]$. The calculation of sensitivities is shown in Eq. (5.c).

From the perspective of physical essence, the PJM $\widehat{\boldsymbol{\Phi}}[t]$ is similar to the sensitivity between the objective and SOP regulation. Compared with the sensitivity-based method, the proposed data-driven method has some merits. 1) To realize the optimization of control objectives, the value of $\widehat{\boldsymbol{\Phi}}[t]$ is adaptively updated by iteration. It facilitates the obtaining of the fast convergence speed and improved control performance. 2) As various control objectives can be incorporated, the proposed method has better scalability in the application scenarios.

The objective function of adaptive control is formulated as follows, in which the minimum voltage deviation, network power losses, and the change of SOP output can be incorporated.

$$J(\boldsymbol{X}[t]) = \min \left(\left\| \boldsymbol{Y}^{\text{ref}} - \widetilde{\boldsymbol{Y}}[t+\Delta t] \right\|_2^2 + \lambda \| \boldsymbol{X}[t] - \boldsymbol{X}[t-\Delta t] \|_2^2 \right) \qquad (6)$$

Through equations (1)-(6), a data-driven operation model of SOP is established by using measurement data. Without models and accurate parameters of ADN, the input-output relationship of SOP operation is depicted dynamically, which can adaptively respond to system state changes. The data-driven framework of real-time operation of SOPs is depicted in Fig. 1. Note that the real-time operation of SOP means an intra-day online control process with the control horizon Δt, instead of the traditional day-ahead operation optimization.
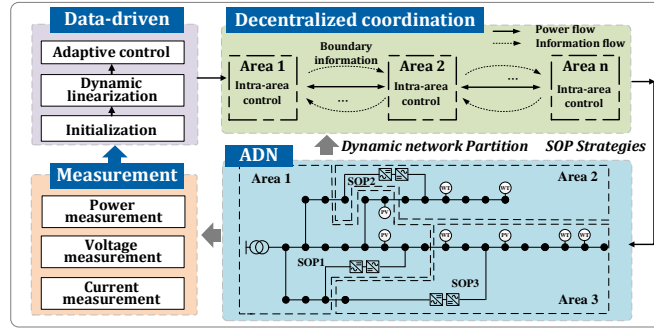

Fig. 1. Data-driven framework of the SOP operation.

Remark: The proposed data-driven method mainly includes two kinds of parameters: fixed parameters (μ, η, λ, ρ) and time-varying parameters ($\widehat{\boldsymbol{\Phi}}[t]$).

1) The fixed parameters generally have predefined ranges to guarantee the convergence and stability, namely $\eta \in (0,2]$, $\mu > 0$, $\lambda > 0$ and $\rho \in (0,1]$ provided in [12]. The values of fixed parameters influence the convergence speed, and improper fixed parameters will affect the convergence of data-driven SOP strategies. In this paper, the fixed parameters λ, ρ, η and μ are set as 1.0 based on the empirical values provided in [19].

2) The time-varying parameters, namely PJM $\widehat{\boldsymbol{\Phi}}[t]$ will affect the convergence value of data-driven SOP strategies. The analysis of PJM $\widehat{\boldsymbol{\Phi}}[t]$ is elaborated in two aspects: the initial value and the time varying values. First, the sensitivity of the initial value of $\widehat{\boldsymbol{\Phi}}[t]$ has been analyzed above. The sensitivities between objective and SOP regulation can be used to initialize the PJM. Then, to adapt to the variation of system states, the time-varying parameter $\widehat{\boldsymbol{\Phi}}[t]$ estimated and updated by Eq. (5) based on the real-time measurement data in each iteration.

Note that as $\widehat{\boldsymbol{\Phi}}[t]$ is iteratively estimated and updated by the data-driven method, the control performance can be guaranteed. Thus, to avoid the impact of the inappropriate parameters, the fixed parameters can be set within the predefined ranges based on the practical experiences of DSO.

B. Control Performance

1) Convergence and stability analysis

Assumption (1): The controlled system satisfies the Lipschitz condition.

Assumption (2): The partial derivative of $f_i(\cdots)$, $i = 1, \cdots, N$, is continuous for the input $\boldsymbol{X}$.

As proved in Ref. [20], if Assumptions (1) and (2) are satisfied in the non-linear control systems, the convergence and stability of the data-driven control method can be guaranteed. a) The voltage deviation is converged, namely $\lim_{t \to \infty} \| \Delta \boldsymbol{e}[t] \|_v = 0$ exists, where

where $P^{\text{ref}}[t]$ denotes the total net load of ADN at time t, which can be calculated by the sum of injected power of each node. $\tilde{P}_0[t + \Delta t]$ is the estimated active power provided by the substation at time $t + \Delta t$.

Then by the variable expansion, namely $Y^{\text{ref}} = [U^{\text{ref}}, P^{\text{ref}}]$ and $\tilde{Y}[t + \Delta t] = [\tilde{U}[t + \Delta t], \tilde{P}_0[t + \Delta t]]$, the term of network power loss reduction $\left\| P^{\text{ref}} - \tilde{P}_0[t + \Delta t] \right\|_2^2$ can also be integrated into the formulation of data-driven model. Then Eq. (11) is reformulated as Eq. (6).

Based on Eq. (3), the iterative expression for solving $X[t]$ by gradient descent method is shown as follows.

$$X[t] = X[t - \Delta t] + P_r\left(Y^{\text{ref}} - Y[t]\right)\frac{\rho \Phi^{\text{T}}[t]}{\lambda + \left\| \hat{\Phi}[t] \right\|_2^2} \tag{12}$$

where λ and ρ are weighting factors, which are determined by DSO. $\Phi[t]$ can be solved by Eq. (4)-(5). P_r indicates the projection operator to limit the SOP output within the feasible range Θ, which is defined in Eq. (13.a). P_r is expressed as Eq. (13.b).

$$\Theta = \begin{cases} (P^{\text{sop}}[t])^2 + (Q_1^{\text{sop}}[t])^2 \leq S_n^2 \\ (P^{\text{sop}}[t])^2 + (Q_2^{\text{sop}}[t])^2 \leq S_n^2 \end{cases} \tag{13.a}$$

$$P_r = \begin{cases} \left(I - \frac{\nabla \Theta_t \nabla \Theta_t^{\text{T}}}{\nabla \Theta_t^{\text{T}} \nabla \Theta_t} \right), \text{if } X[t] \in \Theta \\ 1, \qquad\qquad \text{if } X[t] \notin \Theta \end{cases} \tag{13.b}$$

where $\nabla \Theta_t$ represents the norm vector of the tangent plane of Θ.

To further consider the range of active transmission power and reactive power of SOP [26], the control saturation range is set by the maximum and minimum limits of SOP outputs.

$$X[t] = \overline{X}, \text{if } X[t] \geq \overline{X} \tag{14.a}$$

$$X[t] = \underline{X}, \text{if } X[t] \leq \underline{X} \tag{14.b}$$

where $\overline{X}$ and $\underline{X}$ denote the limits of $X[t]$. The convergence condition of intra-area MFAC is defined as follows.

$$R_1 = \max\left\{ \|X[t] - X[t - \Delta t]\|_2, \|Y^{\text{ref}} - Y[t]\|_2 / N_n \right\} \leq \varepsilon \tag{15}$$

Combined with analysis in Section II, the data-driven adaptive operation model of SOP in each area a can be summarized:

$$\min f_a = J(X_a[t])$$
$$\text{st. } (4) - (6), (12) - (14) \tag{16}$$

Thus, the above model established based on the real-time measurement data can be effectively solved by the serial iterations with the guaranteed convergence [24].

Note that if the measurement error has an upper bound, the steady-state control error of system is proved to be bounded in [19]. As the measuring data used in the proposed approach are the preprocessed data rather than the raw data, the measurement error generally has an upper bound. The impact of measurement error is further analyzed based in case studies.

B. Inter-area Coordinated Operation of Multiple SOPs

With the boundary information interaction, the control performance of SOPs can be improved through the inter-area coordination. It helps to solve problems in the area without controllable resource and fulfill the requirement of near-global optimization while reducing the calculation burden.

1) Boundary information

Based on the network partition in Section III.A, the overlaps between two neighboring sub-areas are generally the boundary branch and its two end nodes. Then the boundary information is defined as the power transmission of boundary branch and voltage of boundary nodes. Γ denotes the set of boundary branch. As for $\forall l \in \Gamma$, the end nodes of l are named as m and n. Define that node m belongs to the upstream area a. Node n belongs to downstream area b. The boundary information between area a and area b is described as follows:

$$\hat{P}_{a,l}[t] = \hat{P}_{b,l}[t], \hat{Q}_{a,l}[t] = \hat{Q}_{b,l}[t] \tag{17.a}$$

$$r_p = \left\| \left(\hat{P}_{a,l}[t] - \hat{P}_{b,l}[t], \hat{Q}_{a,l}[t] - \hat{Q}_{b,l}[t] \right)^{\text{T}} \right\|_\infty \tag{17.b}$$

where $\hat{P}_{a,l}[t]$ and $\hat{P}_{b,l}[t]$ denote active power transmission between area a and area b at time t. $\hat{Q}_{a,l}[t]$ and $\hat{Q}_{b,l}[t]$ denote reactive power transmission between two areas at time t. r_p denotes the residual of boundary power transmission.

$$\hat{U}_{a,m}[t] = \hat{U}_{b,m}[t], \hat{U}_{a,n}[t] = \hat{U}_{b,n}[t] \tag{18.a}$$

$$r_v = \left\| \left(\hat{U}_{a,m}[t] - \hat{U}_{b,m}[t], \hat{U}_{a,n}[t] - \hat{U}_{b,n}[t] \right)^{\text{T}} \right\|_\infty \tag{18.b}$$

where $\hat{U}_{a,m}[t]$ and $\hat{U}_{b,m}[t]$ denote boundary voltage of node m between areas a and b at time t. $\hat{U}_{a,n}[t]$ and $\hat{U}_{b,n}[t]$ denote boundary voltage of node n between areas a and b at time t. r_v denotes the residual of boundary voltage.

2) Inter-area coordinated voltage control

The consistency algorithm is used to coordinate the strategies of SOPs among areas, in which the upstream area is defined as

the area adjacent to substations. The inter-area coordination process is described as follows.

a) The upstream area receives the boundary information transmitted by the downstream area, and conducts the intra-area operation of SOP to optimize voltage of boundary nodes.

The objective function of upstream area is updated as:

$$J'(\boldsymbol{X}_a[t], \boldsymbol{X}'_a[t]) = \min (J(\boldsymbol{X}_a[t]) + \left\| U^{\mathrm{ref}}_{a,n} - \tilde{U}_{a,n}[t+\Delta t] \right\|^2_2$$
$$+ \lambda \left\| \boldsymbol{X}'_a[t] - [\hat{\boldsymbol{P}}_{b,l}[t-\Delta t], \hat{\boldsymbol{Q}}_{b,l}[t-\Delta t]]^{\mathrm{T}} \right\|^2_2) \tag{19}$$

where $\boldsymbol{X}'_a[t]$ is the auxiliary variable denoting boundary power transmission. $U^{\mathrm{ref}}_{a,n}$ is the voltage reference of boundary node n in area a. The reference values of boundary voltage $U^{\mathrm{ref}}_{a,m}$ and $U^{\mathrm{ref}}_{a,n}$ in objective function are updated by boundary information.

$$U^{\mathrm{ref}}_{a,m} = \hat{U}_{b,m}[t-\Delta t], \ U^{\mathrm{ref}}_{a,n} = \hat{U}_{b,n}[t-\Delta t] \tag{20}$$

b) Then, the upstream area transmits the boundary information to downstream area, including boundary power costs $\hat{P}_{a,l}[t]$, $\hat{Q}_{a,l}[t]$ and boundary voltage references $\hat{U}_{a,m}[t]$, $\hat{U}_{a,n}[t]$.

c) The downstream area receives the boundary information transmitted by upstream area, and conducts the intra-area operation of SOP to optimize the power transmission of boundary branch. The related objective function is updated as:

$$J'(\boldsymbol{X}_b[t], \boldsymbol{X}'_b[t]) = \min (J(\boldsymbol{X}_b[t]) + \left\| U^{\mathrm{ref}}_{b,m} - \tilde{U}_{b,m}[t+\Delta t] \right\|^2_2$$
$$+ \lambda \left\| \boldsymbol{X}'_b[t] - [\hat{\boldsymbol{P}}_{a,l}[t-\Delta t], \hat{\boldsymbol{Q}}_{a,l}[t-\Delta t]]^{\mathrm{T}} \right\|^2_2) \tag{21}$$

where $U^{\mathrm{ref}}_{b,m}$ is the voltage reference of boundary node m in area b. $\boldsymbol{X}'_b[t]$ denotes the extended variables of downstream area b, shown as follows:

$$\boldsymbol{X}'_b[t] = [P_{b,l}[t], Q_{b,l}[t]]^{\mathrm{T}} \tag{22}$$

where $P_{b,l}[t]$ and $Q_{b,l}[t]$ are the boundary power transmission of downstream area b at time t. The reference values of boundary node voltage $U^{\mathrm{ref}}_{b,m}$ and $U^{\mathrm{ref}}_{b,n}$ in objective function are updated as:

$$U^{\mathrm{ref}}_{b,m} = \hat{U}_{a,m}[t-\Delta t], U^{\mathrm{ref}}_{b,n} = \hat{U}_{a,n}[t-\Delta t] \tag{23}$$

The downstream area obtains the solution of $\boldsymbol{X}'_b[t]$, which assigns as the boundary power references $\hat{P}_{b,l}[t]$, $\hat{Q}_{b,l}[t]$.

d) Then the downstream area transmits the boundary information to upstream area, including boundary power references $\hat{P}_{b,l}[t]$, $\hat{Q}_{b,l}[t]$ and boundary voltage costs $\hat{U}_{b,m}[t]$, $\hat{U}_{b,n}[t]$.

e) Repeat the above process until the difference between the boundary information of adjacent areas meets the accuracy requirements, which is defined as follows.

$$R_2 = \left\| [r_p, r_v]^{\mathrm{T}} \right\|_\infty \le \varepsilon \tag{24}$$

In summary, the required data for the coordinated operation of multiple SOPs is shown in Table I. Based on the boundary information, each area conducts intra-area adaptive control of SOP and exchanges the updated boundary information. The strategies of SOPs are coordinated to achieve better performance with lighter calculation burden.

TABLE I
REQUIRED DATA OF INTER-AREA COORDINATED CONTROL

Intra-area Data	Data from Upstream to Downstream	Data from Downstream to Upstream
Node voltage measurement $U[t]$, node injected power $P_i[t]$	Boundary node voltage $(\hat{U}_{a,m}[t], \hat{U}_{a,n}[t])$, boundary power flow $(\hat{P}_{a,l}[t], \hat{Q}_{a,l}[t])$	Boundary node voltage $(\hat{U}_{b,m}[t], \hat{U}_{b,n}[t])$, boundary power flow $(\hat{P}_{b,l}[t], \hat{Q}_{b,l}[t])$

C. Implementation of Data-driven Strategy of SOPs

The flow chart of the MFAC-based operation of SOPs is shown in Fig. 2.

First, on an hour-level time scale, the controllable area is determined with the center of each SOP. On a second-level time scale, by using the real-time voltage measurement, the data-driven model of SOP operation is established in each area, which can adaptively respond to the frequent system state changes. Then the strategies of multiple SOPs are further improved by inter-area coordination to improve the operational performance of ADNs. Note that as for the area without SOP, the boundary information is also measured and transmitted to the adjacent areas. The inter-area coordination will facilitate to solve the operation problems of this area.

Thus, by establishing the data-driven operation model based on MFAC, the problems of lacking accurate parameters and the frequent change of operation state can be effectively avoided. Real-time measurement data and interaction between SOP and distribution networks are used to realize the adaptive operation control of SOPs, which can effectively alleviate the voltage violation and reduce power losses.
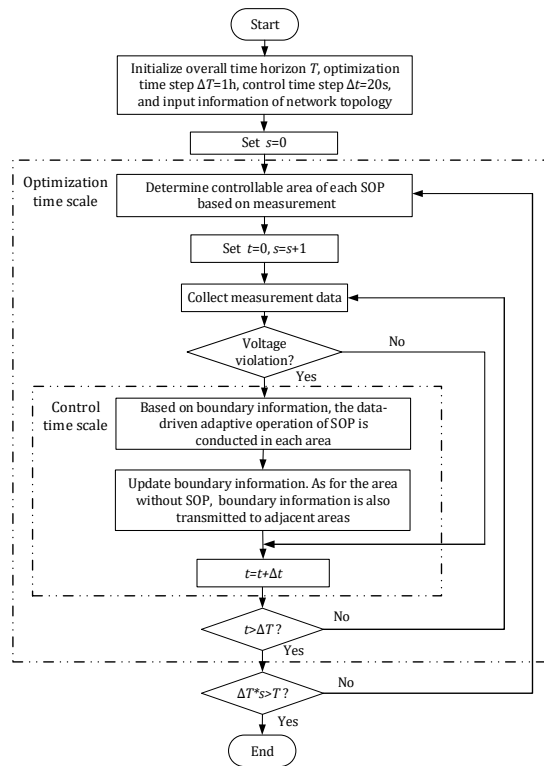
Fig. 2. Flow chart of the adaptive operation of SOPs.

IV. CASE STUDIES AND ANALYSIS

In this section, the effectiveness of the proposed data-driven adaptive operation control of SOP is verified on the modified IEEE 33-node system. The proposed model is implemented in the MATLAB R2014a. The numerical experiments were carried out on a computer with an Intel Xeon CPU E5-1620 processor running at 3.70 GHz and 32 GB of RAM.

A. Modified IEEE 33-node System

The modified IEEE 33-node distribution system including a substation and 32 branches, of which the rated voltage level is 12.66 kV. Total active power and reactive power demands are 3715.0 kW and 2300.0kvar. The structure of the test case is shown in Fig. 3.
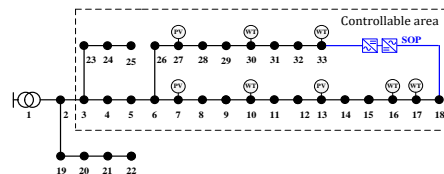

Fig. 3. Structure of the modified IEEE 33-node system.

To fully consider the high penetration of DGs on power losses and voltage deviation, three photovoltaics (PVs) and five wind turbines (WTs) are integrated into the networks, whose active power reaches almost 100% of peak demand. PVs are installed at nodes 7, 13, and 27, whose capacity is 500 kWp, 300 kWp, 400 kWp. WTs are installed at nodes 10, 16, 17, 30, and 33, respectively. The capacities of WTs are 300 kVA, 200 kVA, 200 kVA, 200 kVA, 300 kVA, respectively. One group of SOP with a capability of 1000 kVA is installed between node 18 and 33.
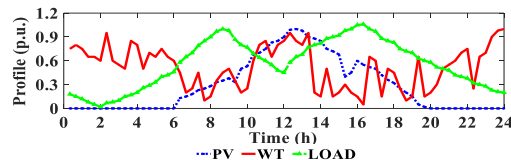
Fig. 4. Operation curves of DG and load.

The daily DGs and loads operation curves are shown in Fig. 4, in which 5-minute is taken as a unit period. The upper and lower limits of the voltage range are set to 1.06 p.u. and 0.94 p.u. The weighting factors λ, ρ, η and μ are set as 1.0. The predetermined tolerance ε is set as 0.01 to ensure the convergence. Set the duration of optimization horizon ΔT as 1 hour and the control horizon Δt as 20 seconds, which is also adopted as the sampling time interval of voltage measurement.

B. Data-driven Adaptive Operation of SOP

1) MFAC-based operation control of SOP

To verify the effectiveness of the data-driven operation control of SOP based on MFAC, two scenarios are adopted.

Scenario I: There is no optimization strategy conducted on SOP, and the initial operation state of ADN is obtained.

Scenario II: The proposed data-driven adaptive operation of SOP is conducted.

The controllable area of SOP is shown in Fig. 3. Taking 10:00 am as an example, the proposed data-driven strategy of SOP is conducted and the voltage profiles of dynamic process are shown in Fig. 5. The red dotted line represents the voltage profile in Scenario I. The blue dashed line shows the converged voltage profile in Scenario II and the grey area denotes the dynamic process of voltage profile in Scenario II.
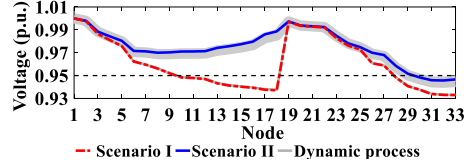

Fig. 5. Voltage profiles of adaptive control process.

It can be seen that due to the optimization of objective function in each iteration, the voltage profiles are maintained within the limits during the dynamic process. To further ensure that node voltage does not exceed the limits during adaptive control, a predictive mechanism is added to the data-driven operation control. When SOP strategy $X[t]$ is obtained, $\tilde{U}[t + \Delta t]$ can be estimated based on Eq. (3). If voltage violation occurs, $X[t]$ will be reset as $(X[t] + X[t - \Delta t])/2$.

The calculation efficiency of MFAC approach is determined by the calculation time in each iteration and iteration number of convergence. As simple algebraic operation is required for MFAC approach, calculation burden of each control time step is light. In addition, MFAC can ensure the convergence within limited iterations [19].

Fig. 6(a) represents the MFAC-based operation strategy of SOP at 10:00 am in Scenario II. Fig. 6(b) verifies that node voltages effectively obtain convergence. Fig. 6(c) shows the convergence of residuals in Eq. (15). The residual decreases rapidly at the beginning until it converges to the predetermined accuracy ε, and then decreases at a slower rate within 10 iterations.

(a) Operation strategy of SOP

(b) Voltage profiles of node 18 and node 33

(c) Convergence of the residual
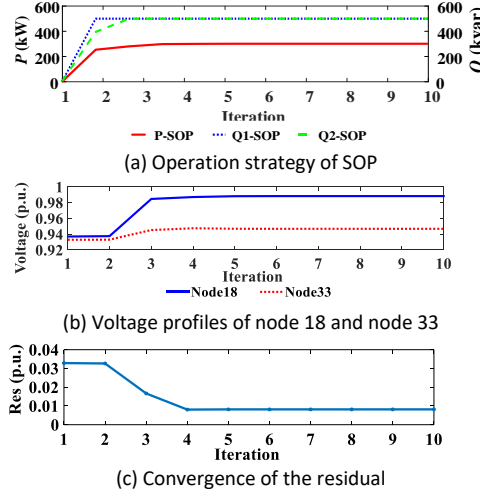
Fig. 6. Adaptive control process of SOP at 10:00 am.

To analyze the impact of the initial value of the PJM $\hat{\Phi}[1]$ on the data-driven SOP control, two PJMs $\hat{\Phi}_1[1]$ and $\hat{\Phi}_2[1]$ with different initial values are adopted. $\hat{\Phi}_1[1]$ is initialized by the sensitivities between objective and SOP regulation and $\hat{\Phi}_2[1]$ is set to $\mathbf{1}$, in which $\mathbf{1}$ represents the all-ones matrix. Taking 10:00 am as an example, the data-driven voltage control results of node 18 are shown in Fig. 7.

Fig. 7. Data-driven control performance with different initial PJMs.

It can be seen that the sensitivities between objective and SOP regulation can appropriately initialize the initial values of $\hat{\Phi}[1]$, which has a satisfactory convergence performance.

To analyze the sensitivity of parameters, the data-driven control has been conducted with different fixed parameter values.

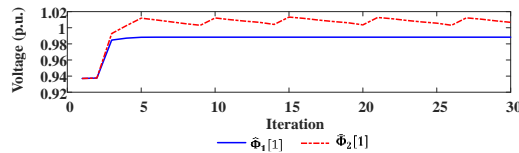When each fixed parameter is changed, the others remain as 1. Taking the control performance with different λ and ρ as examples, voltages of node 18 at 10:00 am are shown in Fig. 8.

(a) Control performance with different values of parameter λ

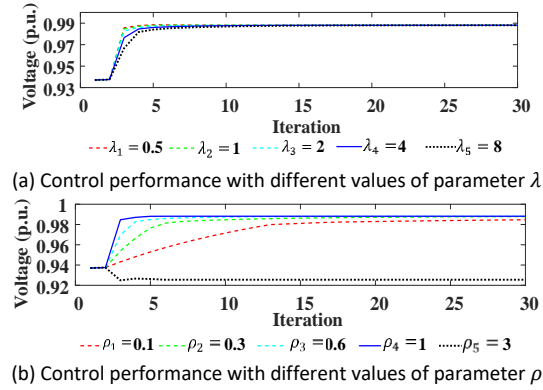(b) Control performance with different values of parameter ρ

Fig. 8. Sensitivity analysis of different parameter values.

As shown in Fig. 8, control results are not sensitive to the fixed parameters within the predefined ranges. Conversely, control results may be unstable or not convergent with the parameters outside the predefined ranges, which is consistent with the conclusion of convergence analysis. Besides, convergence rate is obviously influenced by λ and ρ. When ρ increases or λ decreases, the convergence rate increases.

2) Adaptability to changes of operation state

The adaptability of the data-driven operation of SOP is verified with the changes of system state, including the network reconfiguration and the fluctuation of DG outputs.

It is assumed that a fault occurs on the branch between nodes 7 and 8, and the local topology change consequently occurs at 10:30 am. The branch between nodes 7 and 8 is disconnected, and the tie switch between nodes 12 and 22 is closed. The voltage profiles before and after the reconfiguration are shown in Fig. 9(a). The active power transmission of SOP before and after the reconfiguration is shown in Fig. 9(b). From Fig. 9, it can be seen that the MFAC-based operation control of SOP can adaptively respond to the reconfiguration of network topology.

(a) Voltage profile
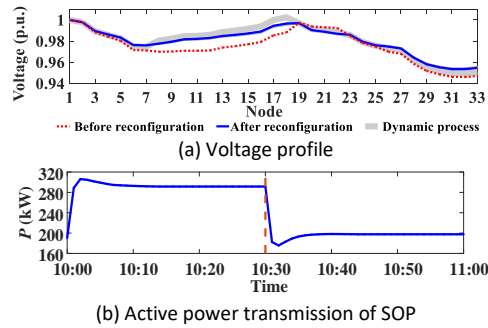

(b) Active power transmission of SOP

Fig. 9. Dynamic process with network reconfiguration.

Besides, the proposed data-driven approach can also adapt to the economic network reconfiguration. With the topology change, the control range of SOP will correspondingly change. First, the sensitivities between objectives and SOP outputs should be recalculated to determine the new control range of SOP. Then, the time-varying parameter $\mathbf{\Phi}[t]$ is calculated by Eq. (5) and the control strategy of SOP is accordingly updated to adapt to the topology change.

To further demonstrate the adaptability of the proposed data-driven approach, economic network reconfiguration is assumed to be conducted at 10:30 am. The tie switches between nodes 12 and 22, nodes 25 and 29 are closed, and branch switches between node 10 and 11, as well as nodes 27 and 28 are disconnected. The voltage profiles before and after reconfiguration are shown in Fig. 10(a). The active power transmission of SOP before and after reconfiguration is shown in Fig. 10(b).

(a) Voltage profile

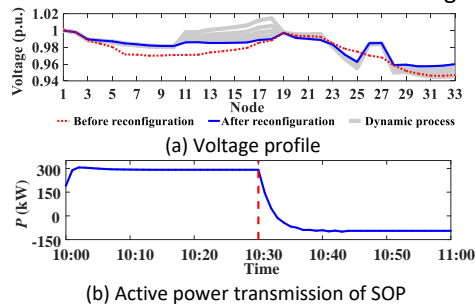
(b) Active power transmission of SOP

Fig. 10. Dynamic process with network reconfiguration.

As shown in Fig. 9 and Fig. 10, the proposed SOP control approach can adapt to the possible topology changes, including local topology change and economic network reconfiguration.

Considering the DG fluctuation from 10:00 am to 10:30 am, the voltage changes of node 18 and 33 with DG fluctuation are shown in Fig. 11(a). The dynamic process of active power transmission of SOP is shown in Fig. 11(b). As the response time of SOP strategy is limited by the number of iterations of MFAC, 5-minute is taken as the sampling interval of DG fluctuation. As shown in Fig. 11, the data-driven method can quickly respond to the fluctuation of DG by adaptively updating the control strategy of SOP.

(a) Voltage profiles of node 18 and 33
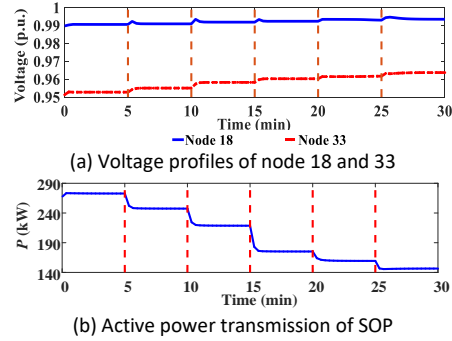
(b) Active power transmission of SOP

Fig. 11. Dynamic process with DG fluctuation.

Thus, the proposed adaptive operation strategy of SOP can effectively respond to the rapid change of system state.

3) Real-time simulation verification

With the capability of high-accuracy simulating the dynamic behaviors in ADNs, real-time simulation provide a critical means to test the effectiveness of control strategies. To further illustrate the effectiveness of the data-driven approach, the real-time simulations have been conducted based on the data-driven operation strategies. As a widely used real-time digital simulator, RT-LAB is adopted to perform the real-time simulations of ADNs. The simulation data will be sent to the SOP controller for the data-driven optimization of operation strategies. In each iteration of data-driven control step, the voltage is measured and updated to the controller. Then the strategy of SOP is calculated and transmitted to the simulator. Note that the optimization time step ΔT of SOP operation is limited by the number of iterations of MFAC as well as the communication time. The optimization time step is set as 5 seconds during the simulation process. The control time step of SOP operation Δt is set as 1 second, which is also adopted as the sampling time interval of voltage measurement. The simulation time step of ADNs is set as 50us in RT-LAB.

Taking node 18 at 10:00 am as an example, voltage amplitude and phase-A voltage waveform at node 18 are shown in Fig. 12(a) and Fig. 12(b). And the current of branch between nodes 17 and 18 are shown in Fig. 12(c). It can be seen that the voltage profile of node 18 is maintained within the limits during the dynamic process and converges within 5 iterations. As the real-time output of DG is varied, the fluctuation of steady-state voltage still exits.

The SOP operation is realized through a dynamic adjustment process in the proposed data-driven approach. As the change rate of SOP output is constrained in each control step, the impact on ADNs caused by SOP variation is effectively mitigated.

(a) Voltage amplitude at node 18

(b) Phase-A voltage at node 18

(c) Phase-A current of branch between nodes 17 and 18
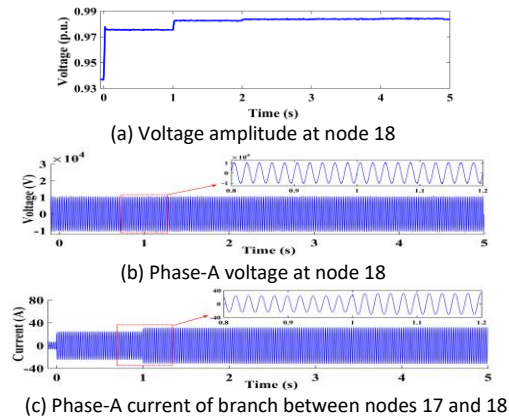
Fig. 12. Real-time simulation results.

C. Coordinated Operation with Multiple SOPs

1) Operation of multiple SOPs with inter-area coordination

The inter-area coordination is further adopted to improve the control performance of SOPs. In the intra-area optimization, each area only needs the intra-area data, instead of data of the entire network. In the inter-area coordination, the reduced calculation burden and the near-optimal effect can be obtained with a small amount of boundary information interaction.

Fig.11. Network partition result at 10:00 am with SOPs.

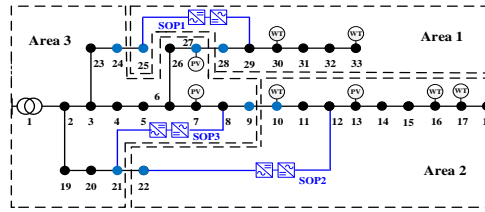In order to elaborate the coordinated operation with multiple SOPs three groups of SOPs with a capability of 1000 kVA are integrated as shown in Fig. 13. Based on the network partition described in section III.A, ADN can be divided into three sub-areas with SOPs. In each sub-area, there is an SOP that plays a major control role. Due to the fluctuation of DGs and loads in different time periods, the node voltages vary from time to time. The controllable areas of SOPs can be updated based on voltage changes of nodes. For example, the result of network partition at 10:00 am is shown in Fig. 13, in which the boundary nodes are marked with blue.
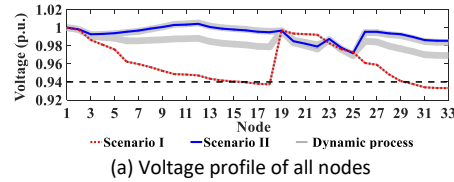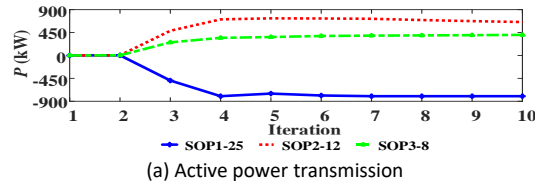

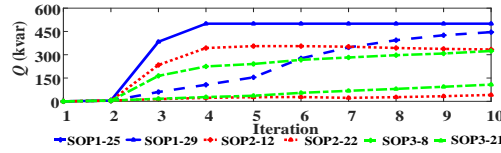(a) Voltage profile of all nodes
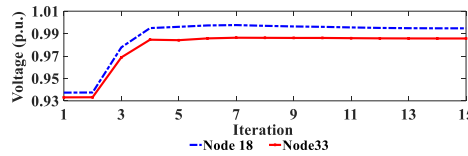
(b) Voltage profiles of node18 and 33

Fig. 14. Voltage profiles at 10:00 am.

(a) Active power transmission

(b) Reactive power compensation

Fig. 15. Iterative process of data-driven control of SOP.

To improve the optimization level of ADN in each sub-area, the data-driven operation of SOPs described in section III.B is implemented. Taking 10:00 am as an example, the voltage profiles of all nodes are shown in Fig. 14(a). The dynamic voltage process of nodes 18 and 33 is shown in Fig. 14(b). The adaptive process of data-driven control of SOP is shown in Fig. 15.

TABLE II
DATA EXCHANGE OF INTER-AREA COORDINATION IN EACH ITERATION

Number of data from downstream area to upstream area		Number of data from upstream area to downstream area	
Area 1→Area 3	8	Area 3→Area 1	8
Area 2→Area 3	8	Area 3→Area 2	8

The required interaction data among areas in each iteration is listed in Table II. In each iteration, each sub-area conducts intra-area adaptive control and exchanges boundary information with the neighbor sub-area. As it takes 12 iterations to reach the convergence, the total amount of interaction data is (8×4)×12.

The daily data-driven operation strategy of SOPs is shown in Fig. 16, which can rapidly respond to the voltage fluctuations caused by DG outputs.
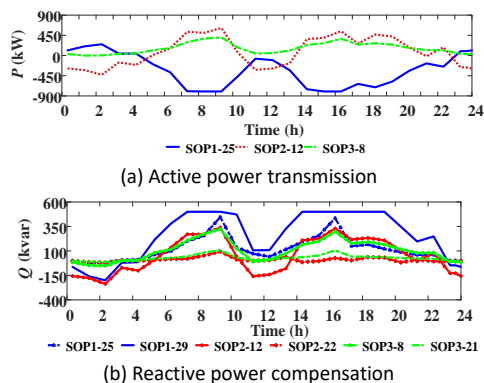
Fig. 16. Daily data-driven operation strategy of SOPs.

2) Operation result analysis

The daily operation results of the two scenarios are shown in Table II. Compared with the results in Scenario I, the proposed data-driven operation strategy of SOPs in Scenario II can effectively mitigate the voltage deviation of ADN and reduce power losses of whole networks. To verify the control performance of data-driven strategy of SOPs, Scenario III is further adopted with the centralized operation control of SOPs, in which the objective is the same with that of Scenario II.

TABLE III
OPERATION RESULTS OF THE THREE SCENARIOS

Scenario	Minimum voltage (p.u.)	Maximum voltage (p.u.)	Index of voltage deviation (p.u.)	Power losses (kWh)
I	0.9330	1.0313	0.0172	1179.7000
II	0.9710	1.0065	0.0035	865.2918
III	0.9689	1.0017	0.0054	839.5731

The voltage deviation index of ADN is defined in Eq. (25). In Scenario II it is reduced by 79.43% compared to Scenario I. The total power loss is reduced by 26.65%, which is a considerable improvement in economic efficiency. It can be also seen that the proposed control strategies have a similar performance with centralized strategies in Scenario III, which have the best utilization of SOP outputs.

$$\text{Index}_1 = \sum_{s \in T} \sum_{i \in \mathcal{N}} \left| U_i^{\text{ref}} - U_i[s] \right| / (N_n * N_{\text{T}}) \qquad (25)$$

Fig. 17(a) shows the daily active power losses for Scenarios I and II. Fig. 17(b) shows the converged voltage profiles in Scenarios I and II. It can be seen that the high penetration of DGs leads to the frequent voltage fluctuation and severe voltage deviation in Scenario I. By applying the proposed data-driven operation control of SOPs in Scenario II, SOPs adjust the active power transmission and the reactive power in each area rapidly and accurately. As shown in Fig. 17(b), the voltage deviation of the system is effectively mitigated and the voltage profiles are maintained within the desired voltage range.
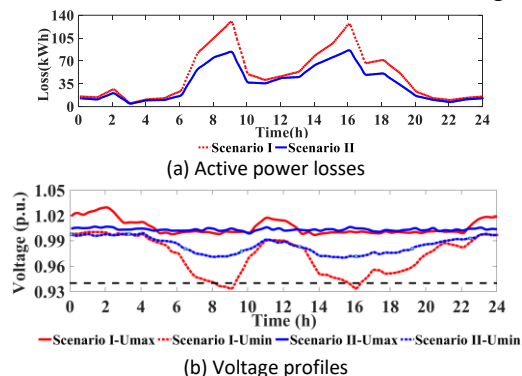
Fig. 17. Daily operation results of Scenarios I and II.

D. Modified IEEE 123-node System

The modified IEEE 123-node system is adopted to verify the scalability of the proposed method on the large-scale ADNs. The detailed parameters can refer to [26]. Three WTs and six PVs are integrated into the networks. WTs are installed at nodes 28, 92 and 108, respectively. The capacities of the WTs are 1000 kVA. PVs are installed at nodes 33, 42, 86, 97, 111 and 116, whose capacities are 300 kWp. Two groups of SOP are installed between nodes 55 and 95, as well as nodes 117 and 123. The parameters of the above regulation devices are set to the same value as the IEEE-33 node system.

To verify the effectiveness of the data-driven SOP operation, the same three scenarios in subsection IV.B are also adopted in the modified IEEE 123-node system.

TABLE IV

OPERATION RESULTS OF THE THREE SCENARIOS

Scenario	Minimum voltage (p.u.)	Maximum voltage (p.u.)	Index of voltage deviation (p.u.)	Power losses (kWh)
I	0.9129	1.0276	0.0298	1439.0
II	0.9600	1.0104	0.0071	1035.8
III	0.9584	1.0143	0.0058	1011.5

The daily operation results of the three scenarios are shown in Table IV. Similar to the conclusions in subsection IV.B, the proposed data-driven method is scalable to large-scale ADNs.

To verify the robustness of the proposed data-driven method, the impact of measurement error is further analyzed. Considering 3% voltage measurement error [27], 100 scenarios are randomly generated by the Monte Carlo method to conduct data-driven control of SOPs. The voltage profiles with and without measurement error of node 18 at 10:00 am are shown in Fig. 18. It can be seen that the control strategy of SOP can also effectively converge with the measurement error in an acceptable range. The fluctuation of control results caused by the measurement error is within 1%. If the measurement error is larger, it will have a greater impact on control results as shown in Fig. 18. However, devices with higher measurement accuracy such as PMU are widely used in ADNs, whose measurement error is generally less than 0.2% [27]. Thus, its measurement accuracy is sufficient to satisfy the data-driven method proposed in this paper. Even if there is a larger measurement error, the redundant measurement of multiple time periods can be used to form an overdetermined equation for optimal control.
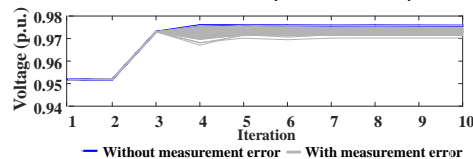

Fig. 18. Dynamic process of voltage profiles with measurement error.

TABLE V

COMPARISON OF THE COMPUTATIONAL EFFICIENCY

Scenario	Calculation time of SOP strategies for 10:00 am (s)	Calculation time of SOP strategies for the whole day (s)
II	0.0569	1.7784
III	0.7880	17.5965

Taking the control process at 10:00 am and of the whole day as examples, the comparison of computational efficiency with the centralized method is shown in Table V. The results show that the proposed data-driven approach can effectively reduce the computational time while obtain the near-optimal solution.

Thus, the proposed operation strategies of multiple SOPs with inter-area coordination can effectively reduce the impact of high penetration of DG and mitigate the voltage violation in ADN. Besides, the proposed operation strategy can also effectively adapt to the complex operation environments.

V. CONCLUSIONS

This paper proposes a data-driven adaptive operation strategy of SOP, which can effectively adapt to the topology reconfiguration and frequent operation state changes. Then, oriented for the decentralized control, the operation strategies of multiple SOPs are further coordinated with inter-area coordination. The results shows that based on the measurement data, the data-driven operation strategy of SOPs can adaptively respond to system state changes and improve the operational performance of ADNs, including the mitigation of voltage deviation and reduction of power losses.

There are also some limitations of the data-driven control method in practical applications. First, the overall measurement data of the controlled region is needed to conduct the data-driven control of SOPs. Reducing the number of required measurements can be further investigated, which is expected to further lighten the communication burdens and improve adaptability. The measurement accuracy will also affect control performance. Second, the interactive iteration process between SOP and ADNs is used to realize the adaptive operation control of SOPs, which may cause sustained disturbances on distribution networks during the control process. To avoid disturbances, the digital twin of ADNs can be further used to interact with data-driven control model of SOP until the converged operation strategy is obtained.

VI. REFERENCES

[1] Bloemink J, Green T. Benefits of distribution-level power electronics for supporting distributed generation growth. *IEEE Trans. Power Deliv.*, vol. 28, no. 2, pp. 911-919, 2013.
[2] A. Mohammad, A. Jaafre, G. Mokryani, "Planning and operation of LV distribution networks: a comprehensive review," *IET Energy Syst. Integr.*, vol. 1, no. 2, pp. 133-146, 2019.
[3] C. Lai, Y. Jia, M. McCulloch, *et al*, "Daily clearness index profiles cluster analysis for photovoltaic system", *IEEE Trans. Ind. Inform.*, vol. 13, no. 5, pp. 2322-2332, 2017.
[4] J. Li, Z. Xu, J. Zhao, *et al*, "Distributed online voltage control in active distribution networks considering PV curtailment," *IEEE Trans. Ind. Inform.*, vol. 15, no. 10, pp. 5519-5530, 2019.
[5] W. Cao, J. Wu, N. Jenkins, "Operating principle of soft open points for electrical distribution network operation," *Appl. Energy*, vol. 164, pp. 245-257, 2016.
[6] C. Long, J. Wu, L. Thomas, "Optimal operation of soft open points in medium voltage electrical distribution networks with distributed generation," *Appl. Energy*, vol. 184, pp. 427-437, 2016.
[7] H. Ji, C. Wang, P. Li, *et al*, "Robust operation of soft open points in active distribution net-works with high penetration of photovoltaic integration," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 280-289, 2019.

[8]   P. Li, H. Ji, H. Yu, *et al*, "Combined decentralized and local voltage control strategy of soft open points in active distribution networks," *Appl. Energy*, vol. 241, pp. 613-624, 2019.

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

[10] *R. Zhang, T. Jiang, G. Li, et al,* "Day-ahead scheduling of multi-carrier energy systems with multi-type energy storages and wind power*," CSEE J. Power Energy Syst.*, vol. 4, no. 3, pp. 283-292, 2018.

[11] J. Zhang, C. Y. Chung and L. Guan, "Noise effect and noise-assisted ensemble regression in power system online sensitivity identification," *IEEE Trans. Ind. Inform.*, vol. 13, no. 5, pp. 2302-2310, 2017.

[12] 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.

[13] C. Tu, X. He, Z. Shuai, *et al*, "Big data issues in smart grid- A review," *Renew. Sust. Energ. Rev.*, vol. 79, pp. 1099-1107, 2017.

[14] T. Huang, Q. Guo, H. Sun, *et al*, "A deep spatial-temporal data-driven approach considering microclimates for power system security assessment," *Appl. Energy*, vol. 237, pp. 36-48, 2019.

[15] Y. Liu, Y. Zhang, Y. Wang, *et al*, "Data-driven power flow linearization: a regression approach," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2569-2580, 2019.

[16] S. Karagiannopoulos, P. Aristidou, G. Hug, "Data-driven local control design for active distribution grids using off-line optimal power flow and machine learning techniques," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6461-6471, 2019.

[17] H. Chao, W. Long, L. Lai, "Data-driven short-term solar irradiance forecasting based on information of neighboring sites," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9918-9927, 2018.

[18] X. Shi, R. Qiu, T. Mi, *et al,* "Adversarial feature learning of online monitoring data for operational risk assessment in distribution networks," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 975-985, 2020.

[19] 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.

[20] 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.

[21] H. Zhang, J. Zhou, Q. Sun, *et al*, "Data-driven control for interlinked AC/DC microgrids via model-free adaptive control and dual-droop control," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 557-571, 2015.

[22] X. Bu, Z. Hou, H. Zhang, *et al*, "Data-driven multi agent systems consensus tracking using model free adaptive control," *IEEE Trans. Neur. Net. Lear. Syst.*, vol. 9, no. 5, pp. 1514-1524, 2018.

[23] Y. Guo, Z. Hou, S. Liu, *et al*, "Data-driven model-free adaptive predictive control for a class of MIMO nonlinear discrete-time systems with stability analysis," *IEEE Access*, vol. 7, pp. 102852-102866, 2019.

[24] Y. Zhao, Z. Yuan, C. Lu, *et al*, "Improved model-free adaptive wide-area coordination damping controller for multiple-input-multiple-output power systems," *IET Gener. Transm. Distrib.*, vol. 10, no. 13, pp. 3264-3275, 2016.

[25] A. Rodriguez, A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, pp. 1492-1496, 2014.

[26] P. Li, H. Ji, C. Wang, *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.

[27] M. Asprou, E. Kyriakides, "Identification and estimation of erroneous transmission line parameters using PMU measurements," *IEEE Trans. Power Deliv.*, vol. 32, no. 6, pp. 2510-2519, 2017.