Electrical Power and Energy Systems 133 (2021) 107231

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

# Stochastic expansion planning of battery energy storage for the interconnected distribution and data networks

Abdulraheem Hassan Alobaidi [a,*], Mahdi Khodayar [b], Ali Vafamehr [a], Harsha Gangammanavar [c], Mohammad E. Khodayar [a]

[a] Department of Electrical and Computer Engineering, Southern Methodist University, TX, USA
[b] Department of Computer Science, University of Tulsa, OK, USA
[c] Department of Engineering Management, Information, and Systems, Southern Methodist University, TX, USA

ARTICLE INFO

*Keywords:*
Battery energy storage
Data centers
Benders decomposition
Expansion planning

ABSTRACT

This paper presents a stochastic expansion planning framework to determine the installation time, location, and capacity of battery energy storage systems in the distribution networks with considerable penetration of photovoltaic generation and data centers. The presented framework aims to minimize the capital cost of the battery energy storage and the operation cost of the distribution network while ensuring the security of energy supply for the data centers that serve end-users in the data network as well as the reliability requirements of the distribution network. The proposed stochastic framework captures the interactions between the distribution network and data center operators considering limited shared information among these entities. Benders decomposition is used to capture the interactions between these autonomous operators in the electricity and data networks. The uncertainties associated with the electric demand, data center workload, solar PV generation, and the availability of the distribution branches are captured using Monte Carlo simulation. The representative scenarios are selected using a dissimilarity-based sparse subset selection algorithm. To evaluate the effectiveness of the proposed framework, numerical results are presented for a modified IEEE-34 bus distribution network with data centers and PV generation units.

## 1. Introduction

The significant annual increase in data center energy consumption impacts the long-term security and reliability of the distribution networks. The data center energy consumption that accounted for 1.8% of the US energy consumption in 2014, is increased by 4% annually in 2014–2020 [1]. The annual growth rate of the data center market is anticipated to be 8.5% in 2020–2027 [2]. The global concerns about greenhouse gas generation are shifting the electricity generation portfolio toward variable and uncertain renewable energy resources. Coordinated with such transitions and to mitigate the carbon footprints of data centers, considerable research efforts are dedicated to serving the data center electric demand with clean and renewable energy resources. Such efforts address the challenges associated with the variability and uncertainty in data center demand and renewable generation resources. The uncertainty in renewable generation resources would lead to voltage fluctuations and violations in the distribution feeders. Moreover, the renewable power injection could alter the direction of power flow in distribution feeders and cause protection system failures. The variability and uncertainty in the data center demand stemming from serving the end users' requests and workloads in the data network, could result in excessive voltage fluctuations and voltage drops in the distribution feeders. Therefore, coordination between the uncertain distributed energy supply and variable demand could help to improve the long-term operation of the distribution networks with such resources.

Battery energy storage (BES) offers several benefits to the distribution network including reducing the peak load at the main distribution feeder, mitigating the renewable generation curtailment, and improving the reliability and power quality. Considering the sustained growth in the data center demand, the long-term expansion planning of BES could help to achieve these objectives and postponing the potential feeder capacity expansions by balancing the generation and demand effectively. In this context, extensive research works were focused on the

* Corresponding author.
*E-mail addresses:* aalobaidi@smu.edu (A.H. Alobaidi), mahdi-khodayar@utulsa.edu (M. Khodayar), avafamehr@smu.edu (A. Vafamehr), harsha@smu.edu (H. Gangammanavar), mkhodayar@smu.edu (M.E. Khodayar).
*URL:* http://mailto:aalobaidi@smu.edu (A.H. Alobaidi).

https://doi.org/10.1016/j.ijepes.2021.107231
Received 16 December 2020; Received in revised form 12 April 2021; Accepted 20 May 2021
Available online 18 June 2021

**Nomenclature**

***Indices and sets***

| | |
|---|---|
| $\Omega_i$ | Set of candidate buses to install BES unit |
| $b$ | Index of distribution network branch |
| $c$ | Index of data center |
| $ch$ | Index of battery charging mode |
| $d$ | Index of representative day |
| $dc$ | Index of battery discharging mode |
| $e$ | Index of battery energy storage (BES) unit |
| $f$ | Index of distribution feeder |
| $g$ | Index of distributed generation |
| $i,j$ | Index of bus |
| $l$ | Index of load in the distribution network |
| $s$ | Index of scenario |
| $t$ | Index of hour |
| $v$ | Index of solar photovoltaic unit |
| $y$ | Index of year |

***Variables***

| | |
|---|---|
| $\mu_e^{i,y}$ | Binary variable for existence of BES unit $e$ at bus $i$ |
| $\pi_e^y, \pi_{c,s}^{t,d,y}$ | Dual variables |
| $E_{e,s}^{t,d,y}$ | Stored energy in the battery storage $e$ |
| $p_{b,s}^{t,d,y}$ | Real power flow in the distribution network branch $b$ |
| $p_{c,s}^{t,d,y}$ | Real power demand of data center $c$ |
| $p_{ch,e,s}^{t,d,y}$ | Real power of BES unit $e$ in charging mode |
| $p_{dc,e,s}^{t,d,y}$ | Real power of BES unit $e$ in discharging mode |
| $p_{f,s}^{t,d,y}$ | Purchased real power from the main feeder $f$ |
| $p_{g,s}^{t,d,y}$ | Real power dispatch of distributed generation unit $g$ |
| $p_{v,s}^{t,d,y}$ | Real power generation of photovoltaic unit $v$ |
| $q_{b,s}^{t,d,y}$ | Reactive power flow in the distribution network branch $b$ |
| $q_{e,s}^{t,d,y}$ | Reactive power output of BES unit $e$ |
| $q_{f,s}^{t,d,y}$ | Reactive power from the main feeder $f$ |
| $q_{g,s}^{t,d,y}$ | Reactive power dispatch of distributed generation unit $g$ |
| $q_{v,s}^{t,d,y}$ | Reactive power generation of photovoltaic unit $v$ |
| $U_{i,s}^{t,d,y}$ | Squared voltage at bus $i$ |
| $w_{c,s}^{t,d,y}$ | Workload processed by data center $c$ |
| $Y_e^{i,y}$ | Binary variable representing the installation decision for BES unit $e$ on bus $i$ |
| $Z_{(.),s}^{t,d,y}$ | Slack variables |
| $LS_{l,s}^{t,d,y}$ | Demand curtailment |

***Parameters***

| | |
|---|---|
| $\delta_f$ | Minimum power factor for feeder $f$ |
| $\Delta_s^{t,d,y}$ | Total workload received from the end-users |
| $\eta$ | Annual discount rate |
| $\lambda_g, \beta_g, \gamma_g$ | Coefficients in the quadratic cost function of distributed generation unit $g$ |
| $\phi_{b,s}^{t,d,y}$ | Availability of the distribution network branch $b$; 1 if a branch is available and 0 otherwise |
| $\rho_e$ | Operation cost of BES unit $e$ |
| $\rho_{f,s}^{t,d,y}$ | Hourly price of energy supplied by the main feeder $f$ |
| $\widetilde{P}_{v,s}^{t,d,y}$ | Forecasted real power of photovoltaic unit $v$ |
| $A_{(.)}$ | Bus-unit incidence matrix |
| $B$ | Bus-branch incidence matrix |
| $C_e$ | Investment cost of BES unit $e$ |
| $D_{(.)}$ | Bus-demand incidence matrix |
| $E_e^0$ | Initial stored energy in BES unit $e$ |
| $E_e^{max}$ | Maximum stored energy in BES unit $e$ |
| $E_e^{min}$ | Minimum stored energy in BES unit $e$ |
| $N_d$ | Number of days in representative $d$ |
| $p_c^{max}$ | Maximum real power consumption of data center $c$ |
| $p_e^{max}$ | Maximum real power dispatch of BES unit $e$ |
| $p_{l,s}^{t,d,y}$ | Real power load in the distribution network |
| $pr_s$ | Probability of scenario $s$ |
| $q_e^{max}$ | Maximum reactive power dispatch of BES unit $e$ |
| $q_{l,s}^{t,d,y}$ | Reactive power load in the distribution network |
| $r_{ij}^b$ | Resistance of branch $b$ connecting buses $i$ and $j$ |
| $T$ | Total number of hours |
| $T_d$ | Total number of representatives |
| $T_s$ | Total number of scenarios |
| $T_y$ | Total number of years |
| $V_i^{max}$ | Maximum acceptable bus voltage |
| $V_i^{min}$ | Minimum acceptable bus voltage |
| $w_c^{max}$ | Maximum workload in data center $c$ |
| $x_{ij}^b$ | Reactance of branch $b$ connecting buses $i$ and $j$ |
| $S_b^{max}$ | Maximum apparent power of distribution branch $b$ |
| $EENS_y^{max}$ | Acceptable annual expected energy not supplied |
| $VOLL$ | Value of lost load |

expansion planning of BES in distribution networks [3–17].

A planning framework for BES in the distribution network was proposed in [3] that captured the uncertainty in wind generation. A probabilistic optimal power flow was performed as a part of the planning framework and Tabu search with particle swarm optimization (PSO) was utilized as the solution methodology. In [4], optimal sizing and placement of BES were determined to minimize the power loss in the distribution network; however, the uncertainties in demand and renewable generation were not addressed. Here, the problem was formulated as a mixed-integer quadratically constrained quadratic programming, solved by the D-XEMS13 procedure in MATLAB. The authors in [5] proposed an algorithm based on the Benders decomposition to determine the location and capacity of the energy storage units. Using network reconfiguration, the proposed algorithm aimed to minimize the investment cost of energy storage, the cost of electricity, network loss, feeder overloading, and bus voltage deviations. The power flow constraints were approximated by second-order cones and the uncertainties in PV generation, price of electricity, and demand were considered using scenarios. A technique based on dual optimization was proposed in [6] to determine the capacity and location of distributed generation and BES units. The proposed technique sought to minimize the energy costs, greenhouse gas emissions, and real power losses and maintain the supply voltage within the acceptable limits. The proposed formulation captured the electric vehicle interconnection and Teaching Learning Based Optimization (TLBO) was adopted to solve the formulated problem. A probabilistic method for placement and sizing of energy storage units in the distribution network was proposed in [7] to improve the reliability of the energy supply. The uncertainties in demand and wind generation were considered and the optimal level of reliability was determined using load shedding in contingencies. The expansion planning of the energy storage in the distribution network was addressed in [8] to minimize the installation cost of the energy storage, the voltage deviations, network loss, and energy costs. The problem was formulated as a nonlinear programming problem and solved in two stages. The solution to the first stage problem yielded the location and size of the energy storage while the voltage deviation,

network loss, and energy cost were minimized in the second stage by solving an AC power flow problem. An expansion planning of distribution networks with energy storage systems (ESSs) was formulated as a multi-stage mixed-integer linear programming problem (MILP) in [9]. The expansion decisions included the installation of ESSs and the expansion and replacement of distribution lines, while the operation decisions were the scheduling of ESSs and the energy flow at the main substation. The formulated planning problem aimed to minimize the investment cost, the operation cost of ESS, the cost of the energy procured from the main substation, the curtailed load, and the annual outage penalties. The planning problem considered daily load scenarios to perform the economic dispatch and the extreme loading scenarios were used to check the network security and reliability of the planning decisions.

A non-parametric chance-constrained optimization was proposed in [10] for the expansion planning of ESSs in the distribution network. Here, the uncertainties in electric vehicle demands, residential loads, and renewable generation were taken into account using discrete empirical distributions. The expansion planning of ESS was formulated as a MILP problem in [11] where the uncertainties in wind and PV generations, the price of electricity at the main distribution feeder, the baseload, and the EV demand were represented by scenarios. The considered scenarios were selected using the k-means++ clustering approach. The expansion planning of distributed generation and BESs to maximize the payoff of the distribution network operator was presented in [12]. The problem was formulated as a mixed-integer nonlinear programming (MINLP) problem and solved using the PSO algorithm. A formulation for the expansion planning of BESs in the distribution network was presented in [13] to improve the utilization of wind power generation while minimizing the BES investment cost and distribution network operation cost. The uncertainties associated with the electricity demand, wind generation, and availability of micro-turbines were considered using scenarios. Chance-constrained programming was used to ensure the utilization of wind generation and a differential evolution algorithm was used to solve the proposed planning problem. The expansion planning of ESSs in a distribution network that leverages voltage sensitivity analysis and optimal power flow was presented in [14]. The location and capacity of ESSs were determined to prevent overvoltage and undervoltage incidents in the distribution network. The worst-case realization of the generation and demand profiles were considered using the historical data sets. The authors in [15] proposed a stochastic approach for the expansion planning of BES in a distribution network with conservative voltage reduction. The uncertainties in demand and renewable generation were addressed by developing scenarios. A hierarchical framework for locating and sizing the BESs was presented in [16]. The objective was to minimize the distribution network operation cost while maintaining the nodal voltages within acceptable limits. The location of the BESs was determined using voltage sensitivity analysis and the capacity of the BES was determined by solving a MINLP problem using natural aggregation algorithm. The uncertainties in the distributed generation outputs and demand were captured using scenarios. A stochastic expansion planning of ESSs and distributed generation resources was presented in [17] that captured the demand response to maximize social welfare. Here, the expansion planning problem was formulated as a MILP problem and the uncertainties in demand, wind speed, and solar radiation were addressed using scenarios.

The coordination among the energy storage and data center was addressed in the literature [18–22]. A day ahead resource planning for data centers in grid-connected microgrid with ESSs was addressed in [18]. The problem was formulated as a MILP problem to minimize the fuel cost and carbon footprint considering the delay-sensitive and delay-tolerant workloads in the data network. An algorithm using the Lyapunov optimization technique was proposed in [19] to balance the workloads among data centers with BES and minimize the real-time energy costs associated with processing the workloads. The formulated

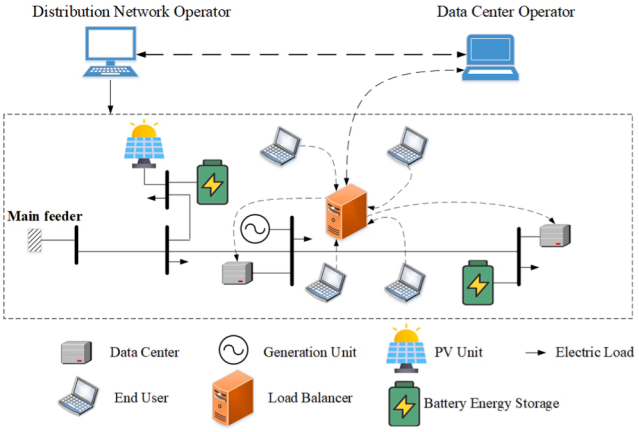


**Fig. 1.** Physical and control layouts of the distribution network with data centers.

stochastic programming problem addressed the uncertainty in the electricity prices and received workloads, to allocate the data center capacity and manage the battery energy flow. The energy management in data centers with energy storage considering carbon footprint offsets was discussed in [20]. The objective was to minimize the operation cost of the data center while satisfying the total carbon footprint requirement. A simulation-based capacity planning approach for energy storage in data centers was proposed in [21]. The power supply mix was characterized using the simulation models to quantify the capacity of energy storage. The expansion planning of the generation resources including the energy storage in the data center was presented in [22]. The proposed problem was formulated as a linear programming (LP) problem which minimized the investment and operation costs of the data center while satisfying the emission and service availability requirements.

Distributed algorithms were used to address the energy management in data centers [23–25]. Distributed data traffic routing was proposed in [23] to improve the energy consumption in the data center while avoiding congestion in the data switches. The coordinated energy management in co-location data centers was addressed in [24] using the alternating direction method of multipliers approach. The objective was to minimize the energy consumption and workload curtailment charges to ensure the quality of service provided to the servers. The energy management of data centers in grid-connected microgrid was addressed in [25]. A distributed algorithm was proposed to minimize the operation cost including the energy trade with the main grid, local generation cost, battery utilization cost, and workload distribution charges. The formulated stochastic programming problem captured the uncertainties in workloads, renewable energy resources, and energy prices.

The physical layout of the distribution network with the data centers is shown in Fig. 1. As shown in this figure, the distribution network and data centers are operated by two independent entities [22,24,26]. The data center operator (DCO) distributes the workloads received from the end-users among the data centers. The data centers process the workloads by using the electricity supplied by the distribution network. The distribution network is operated and managed by the distribution network operator (DNO). The decisions made by the DCO on allocating the workloads to the data centers will impact their electricity demand. The variations in data center demand modify the spatiotemporal demand profile in the distribution network and impact the economics and security of the distribution network. Similarly, the decisions made by the DNO impact the long-term security and reliability of energy supply to the data centers. Such decisions could further affect the efficiency and reliability of the services offered by the data centers to the end-users in the data network. As the information shared between DNO and DCO is limited, lacking coordination among these entities could lead to deficiency in energy supply to data centers and therefore shortage in the

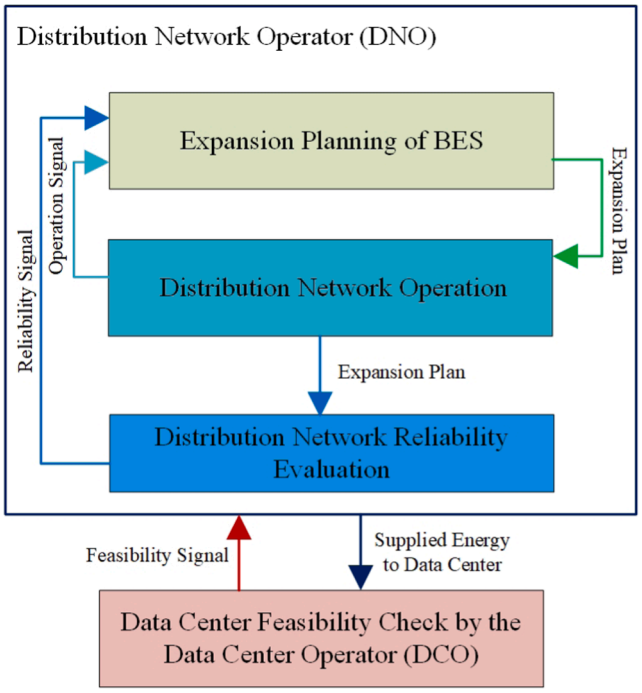


**Fig. 2.** The proposed stochastic framework for the BES expansion planning.

workload processing capacity, or increase in the electricity demand in the distribution network and eventually violation of the network constraints which could lead to demand curtailments. Therefore, effective coordination among DCO and DNO benefits the end-users in the data network by improving the quality of service provided by the data centers and improves the economics and security of the distribution network in the long-term operation. The expansion planning of the data center facilities considering the wind power generation in transmission network was addressed in [26]. The proposed algorithm aimed at minimizing the investment and operation costs of data centers and data routes. The uncertainty in the transmission network assets and the received requests are captured using scenarios. The interactions between the independent system operator (ISO) and DCO are captured using Benders decomposition and sharing the electricity price. In [27], the BES expansion planning in the distribution network with data center was addressed ignoring the reliability requirements of the distribution network, the operation cost of the BES units, and the uncertainty in the outages of the distribution branches. This research extends the earlier work in [27] by addressing these shortcomings.

In this paper, a stochastic framework for the expansion planning of BES in a distribution network that serves data centers is presented. The presented framework captures the interaction between DNO and DCO. As the information sharing among DNO and DCO is limited, the presented framework leverages Benders decomposition to formulate the expansion planning problem for the DNO and the long-term operation problem for the DCO. The uncertainties in demand, data center workload, the power output of PV generation, and the availability of the distribution branches are considered using scenarios. The proposed framework determines the expansion plans for BES units while ensuring energy security for the data centers and the reliability of the distribution network. The contributions of this work are summarized as follows:

- An expansion planning framework for BES units is proposed to determine the location, capacity, and installation time of the BES units in the distribution network with PV generation and data centers.
- The interactions among DNO and DCO are captured using the Bender decomposition technique to address the limited information shared among data center and distribution network operators.
- The uncertainties in electric demand, data center workload, the power output of PV generation units, and the outages of the distribution lines are considered in the long-term planning horizon. A dissimilarity-based sparse subset selection (DS3) algorithm is used to cluster the hourly data, select the most effective representative data, and determine the representative scenarios.

The paper is organized as follows, the problem formulation and solution methodology are presented in Sections 2 and 3, respectively. DS3 algorithm is presented in Section 4. A case study using the modified IEEE-34 bus distribution network is presented in Section 5. Finally, the conclusions are drawn in Section 6.

## 2. Problem formulation

Fig. 2 shows the structure of the proposed stochastic expansion planning problem for the BES in distribution networks that captures the interactions between the DCO and DNO to solve this problem. As shown in this figure, DNO solves the BES expansion planning problem (master problem) to determine the location, capacity, and time of BES installations. Here, the objective is to minimize the installation cost of the BES units. Once the initial expansion decisions are determined, DNO solves the distribution network operation sub-problem. The objective of this sub-problem is to minimize the expected operation cost of the distribution network while satisfying the network constraints. Here, DNO checks for the feasibility and optimality of the expansion decisions considering the imposed uncertainties in the long-term operation horizon. If the expansion decisions are not feasible or optimal in the distribution network, the operation signal in the form of a Benders cut is sent to the expansion planning problem to revise the expansion decisions. Once the expansion decisions are optimal and feasible for the distribution network, the distribution network reliability is evaluated by solving the distribution network reliability evaluation sub-problem. If the expansion plans do not satisfy the system reliability, a reliability signal (Benders cut) is sent to the expansion planning problem.

Once the expansion decisions satisfy the reliability constraint in the distribution network, the power supplied to the data center is passed to the data center feasibility sub-problem solved by the DCO. Similarly, the interaction between DNO and DCO is captured using the Benders decomposition technique. Here, DCO checks for the feasibility of serving the end-users' stochastic workloads given the energy allocated to the data centers by the DNO. If the solution provided by the DNO is infeasible in the data center sub-problem and the quality of service cannot be guaranteed, a feasibility Benders cut is generated and sent to the distribution operation sub-problem to change the operational decisions and update the allocated energy to the data centers. This process continues until the solution provided by the DNO is feasible for the DCO in the data network. The problem formulations for the master problem and sub-problems are discussed in detail in the following subsections.

### 2.1. Expansion planning problem – master problem (MP)

The master problem is formulated as a MILP problem (1)–(5). The objective function is shown in (1). The first term in (1) is the installation cost of the BES units which is formulated in (2) considering the annual discount rate. The second term in (1) is a positive variable that represents the total expected distribution network operation cost. The auxiliary variable ($\alpha$) is non-negative as shown in (3). It is assumed that only one BES unit is installed at each candidate bus in the planning horizon as enforced by (4) and (5).

$$min \quad Z_{lower} = IC + \alpha \tag{1}$$

s.t.

$$IC = \sum_{y=1}^{T_y} (1+\eta)^{1-y} \cdot \left( \sum_{i \in \Omega_i} \sum_{e} C_e \cdot \left( Y_e^{i,y} \right) \right) \tag{2}$$

$$\alpha \geqslant 0 \tag{3}$$

$$\sum_{y=1}^{T_y} \sum_{e} Y_e^{i,y} \leqslant 1 \ ; \forall i \in \Omega_i \tag{4}$$

$$\mu_e^{i,y} - \mu_e^{i,y-1} = Y_e^{i,y} \ ; \forall i \in \Omega_i \tag{5}$$

The solution to this problem ($\widehat{\mu}_e^{i,y}$) is passed to the distribution network operation problem (SP1) presented below.

### 2.2. Distribution network operation sub-problem (SP1)

This sub-problem is formulated as a linear programming (LP) problem (6)–(35). The objective function in SP1 is presented in (6). The distribution operation decisions minimize the expected operation cost of BES units, the expected cost of supplying energy from the upstream power network through the main distribution feeder, the expected operation cost of distributed generation assets, and the expected cost associated with the curtailed demand (*CU*) as shown in the first, second, third, and fourth terms of (6), respectively. The fifth term in (6) is the penalty associated with the mismatch in the nodal reactive power balance, where *M* is a large scalar. The expected operation cost of BES units, the expected cost of energy purchased from the main distribution feeder, and the expected operation cost of distributed generation units are computed by (7), (8), and (9), respectively. In this paper, the quadratic cost function in (9) is linearized using the piece-wise linearization technique. The expected cost associated with the demand curtailment is formulated as shown in (10) where *VOLL* is the value of lost load. The demand curtailment ($LS_{l,s}^{t,d,y}$) is limited by the total demand in the distribution network as shown in (11). The nodal real and reactive power balance equations are shown in (12) and (13), respectively. Here, the mismatch between the nodal reactive power generation and demand is captured by introducing the slack variables in (13). The mismatch in the nodal real power balance is handled by demand curtailment ($LS_{l,s}^{t,d,y}$) in (12). Here, it is assumed that the reactive power requirement of the data center is compensated by on-site capacitor banks, and therefore, the reactive power consumption of the data center is ignored. The slack variables are non-negative as presented in (14).

$$\begin{aligned} min \ \ S_1 = {} & OC_e + OC_f + OC_g + CU \\ & + \sum_{s=1}^{T_s} \sum_{y=1}^{T_y} \sum_{d=1}^{T_d} \sum_{t=1}^{T} M \cdot \left( Z_{1,s}^{t,d,y} + Z_{2,s}^{t,d,y} \right) \end{aligned} \tag{6}$$

$$OC_e = \sum_{s=1}^{T_s} pr_s \cdot \left( \sum_{y=1}^{T_y} (1+\eta)^{(1-y)} \cdot \left[ \sum_{d=1}^{T_d} N_d \times \sum_{t=1}^{T} \sum_{e} \left( \rho_e \cdot \left( p_{ch,e,s}^{t,d,y} + p_{dc,e,s}^{t,d,y} \right) \right) \right] \right) \tag{7}$$

$$OC_f = \sum_{s=1}^{T_s} pr_s \cdot \left( \sum_{y=1}^{T_y} (1+\eta)^{(1-y)} \cdot \left[ \sum_{d=1}^{T_d} N_d \times \sum_{t=1}^{T} \sum_{f} \left( \rho_{f,s}^{t,d,y} \cdot p_{f,s}^{t,d,y} \right) \right] \right) \tag{8}$$

$$OC_g = \sum_{s=1}^{T_s} pr_s \cdot \left( \sum_{y=1}^{T_y} (1+\eta)^{(1-y)} \cdot \left[ \sum_{d=1}^{T_d} N_d \times \sum_{t=1}^{T} \sum_{g} \left( \lambda_g \cdot \left( p_{g,s}^{t,d,y} \right)^2 + \beta_g \cdot p_{g,s}^{t,d,y} + \gamma_g \right) \right] \right) \tag{9}$$

$$CU = \sum_{s=1}^{T_s} pr_s \cdot \left( \sum_{y=1}^{T_y} (1+\eta)^{(1-y)} \cdot \left[ \sum_{d=1}^{T_d} N_d \times \sum_{t=1}^{T} \sum_{l} \left( VOLL \cdot \left( LS_{l,s}^{t,d,y} \right) \right) \right] \right) \tag{10}$$

$$0 \leqslant LS_{l,s}^{t,d,y} \leqslant p_{l,s}^{t,d,y} \tag{11}$$

$$\begin{aligned} & A_g \cdot p_{g,s}^{t,d,y} + A_v \cdot p_{v,s}^{t,d,y} + A_e \cdot \left( p_{dc,e,s}^{t,d,y} - p_{ch,e,s}^{t,d,y} \right) + \\ & A_f \cdot p_{f,b,s}^{t,d,y} = B \cdot p_{b,s}^{t,d,y} + D_l \cdot \left( p_{l,s}^{t,d,y} - LS_{l,s}^{t,d,y} \right) + D_c \cdot p_{c,s}^{t,d,y} \end{aligned} \tag{12}$$

$$\begin{aligned} & A_g \cdot q_{g,s}^{t,d,y} + A_v \cdot q_{v,s}^{t,d,y} + A_e \cdot q_{e,s}^{t,d,y} + A_f \cdot q_{f,b,s}^{t,d,y} + Z_{1,s}^{t,d,y} - Z_{2,s}^{t,d,y} \\ & \qquad = B \cdot q_{b,s}^{t,d,y} + D_l \cdot q_{l,s}^{t,d,y} \end{aligned} \tag{13}$$

$$Z_{1,s}^{t,d,y}, Z_{2,s}^{t,d,y} \geqslant 0 \tag{14}$$

The charging and discharging real power of the BES unit are limited by their maximum values as shown in (15) and (16), respectively. The charging and discharging modes of the BES unit are mutually exclusive as the operation cost of the BES unit is minimized in the objective function [28]. The reactive power output of the BES unit is limited by (17). The stored energy in the BES unit is limited by the maximum and minimum values shown in (18). The start and final stored energy in the BES unit are equal to the initial stored energy ($E_e^0$) as shown in (19). The hourly stored energy in the BES unit is constrained by (20) where $\tau_e^{ch}$ and $\tau_e^{dc}$ are the charging and discharging efficiencies, respectively. Here, the planning decision of installing BES units ($\mu_e^{i,y}$) is a continuous variable that is fixed to the solution obtained from the MP ($\widehat{\mu}_e^{i,y}$) as shown in (21).

$$0 \leqslant p_{ch,e,s}^{t,d,y} \leqslant \mu_e^{i,y} \cdot p_e^{max} \tag{15}$$

$$0 \leqslant p_{dc,e,s}^{t,d,y} \leqslant \mu_e^{i,y} \cdot p_e^{max} \tag{16}$$

$$-\mu_e^{i,y} \cdot q_e^{max} \leqslant q_{e,s}^{t,d,y} \leqslant \mu_e^{i,y} \cdot q_e^{max} \tag{17}$$

$$\mu_e^{i,y} \cdot E_e^{min} \leqslant E_{e,s}^{t,d,y} \leqslant \mu_e^{i,y} \cdot E_e^{max} \tag{18}$$

$$E_{e,s}^{1,t,y} = E_{e,s}^{T,d,y} = \mu_e^{i,y} \cdot E_e^0 \tag{19}$$

$$E_{e,s}^{t,d,y} = E_{e,s}^{t-1,d,y} + \left( \tau_e^{ch} \cdot p_{ch,e,s}^{t,d,y} - \frac{p_{dc,e,s}^{t,d,y}}{\tau_e^{dc}} \right) \tag{20}$$

$$\mu_e^{i,y} = \widehat{\mu}_e^{i,y} \quad : \pi_e^{i,y} \tag{21}$$

The real and reactive power dispatches of the distributed generation units are restricted by (22) and (23), respectively. The real power generation of the PV unit is restricted by the forecasted PV generation as shown in (24). The reactive power supply of the PV unit is limited by the PV inverters' capacity as shown in (25). The capacity of the main distribution feeder is restricted by (26) and (27), considering the minimum power factor ($\delta_f$) at the upstream network interconnection. The squared nodal voltage is constrained by the upper and lower bounds as shown in (28). The linearized distribution flow (DistFlow) model is used to represent the power flow in the distribution network branch as shown in (29) and (30) [29]. Here, the Big-M method is used to address the

availability of the distribution branch. The real and reactive power flows in a distribution network branch are affected by the availability of the branch as shown in (31) and (32), respectively. Once a distribution branch is not available, (29) and (30) are relaxed and the real and reactive power flow of the branch is equal to zero as enforced by (31) and (32). Hexagon approximation [30] is used to enforce the capacity limitation of the distribution branch in (33)–(35). The capacity of the distribution branch is limited by the maximum apparent power in (36). Eq. (36) is a circular constraint with a radius of the maximum apparent power ($S_b^{max}$). This quadratic constraint is approximated by an $n$ sided convex polygon with ($S_b^*$) calculated as (37). Using hexagon approximation i.e., $n = 6$, the circular constraint (36) is replaced with (33)–(35) as discussed in [30]. Similar constraints are used to enforce the apparent power capacity of the main distribution feeder.

$$p_g^{min} \leqslant p_{g,s}^{t,d,y} \leqslant p_g^{max} \tag{22}$$

$$-q_g^{max} \leqslant q_{g,s}^{t,d,y} \leqslant q_g^{max} \tag{23}$$

$$p_{v,s}^{t,d,y} \leqslant \widetilde{P}_{v,s}^{t,d,y} \tag{24}$$

$$-q_v^{max} \leqslant q_{v,s}^{t,d,y} \leqslant q_v^{max} \tag{25}$$

$$-p_f^{max} \leqslant p_{f,s}^{t,d,y} \leqslant p_f^{max} \tag{26}$$

$$-\tan\left(\delta_f\right)\cdot p_{f,s}^{t,d,y} \leqslant q_{f,s}^{t,d,y} \leqslant \tan\left(\delta_f\right)\cdot p_{f,s}^{t,d,y} \tag{27}$$

$$\left(V_i^{min}\right)^2 \leqslant U_{i,s}^{t,d,y} \leqslant \left(V_i^{max}\right)^2 \tag{28}$$

$$U_{i,s}^{t,d,y} - U_{j,s}^{t,d,y} \leqslant 2\left(r_{ij}^b\cdot p_{b,s}^{t,d,y} + x_{ij}^b\cdot q_{b,s}^{t,d,y}\right) + M\cdot\left(1-\phi_{b,s}^{t,d,y}\right) \tag{29}$$

$$U_{i,s}^{t,d,y} - U_{j,s}^{t,d,y} \geqslant 2\left(r_{ij}^b\cdot p_{b,s}^{t,d,y} + x_{ij}^b\cdot q_{b,s}^{t,d,y}\right) - M\cdot\left(1-\phi_{b,s}^{t,d,y}\right) \tag{30}$$

$$-M\cdot\phi_{b,s}^{t,d,y} \leqslant p_{b,s}^{t,d,y} \leqslant M\cdot\phi_{b,s}^{t,d,y} \tag{31}$$

$$-M\cdot\phi_{b,s}^{t,d,y} \leqslant q_{b,s}^{t,d,y} \leqslant M\cdot\phi_{b,s}^{t,d,y} \tag{32}$$

$$-\sqrt{3}\cdot\left(p_{b,s}^{t,d,y} + S_b^*\right) \leqslant q_{b,s}^{t,d,y} \leqslant -\sqrt{3}\cdot\left(p_{b,s}^{t,d,y} - S_b^*\right) \tag{33}$$

$$-\frac{\sqrt{3}}{2}\cdot S_b^* \leqslant q_{b,s}^{t,d,y} \leqslant \frac{\sqrt{3}}{2}\cdot S_b^* \tag{34}$$

$$\sqrt{3}\cdot\left(p_{b,s}^{t,d,y} - S_b^*\right) \leqslant q_{b,s}^{t,d,y} \leqslant \sqrt{3}\cdot\left(p_{b,s}^{t,d,y} + S_b^*\right) \tag{35}$$

$$\left(p_{b,s}^{t,d,y}\right)^2 + \left(q_{b,s}^{t,d,y}\right)^2 \leqslant \left(S_b^{max}\right)^2 \tag{36}$$

$$S_b^* = S_b^{max}\cdot\sqrt{\frac{\left(\frac{2\pi}{n}\right)}{\sin\left(\frac{2\pi}{n}\right)}} \tag{37}$$

Once SP1 is solved, the upper bound of the solution is calculated as in (38). If the lower bound ($\widehat{Z}_{lower}$) i.e., the solution to the MP, and the upper bound ($\widehat{Z}_{upper}$) satisfy $\left|\frac{\widehat{Z}_{upper}-\widehat{Z}_{lower}}{\widehat{Z}_{upper}+\widehat{Z}_{lower}}\right| \geqslant \epsilon$, Benders cut (39) is generated and sent to the MP. Otherwise, the distribution network reliability evaluation sub-problem (SP2) will be formulated and solved.

$$Z_{upper} = \widehat{IC} + \widehat{S}_1 \tag{38}$$

$$\alpha \geqslant \widehat{S}_1 + \sum_{y=1}^{T_y}\sum_e\sum_i \widehat{\pi}_e^{i,y}\cdot\left(\mu_e^{i,y} - \widehat{\mu}_e^{i,y}\right) \tag{39}$$

### 2.3. Distribution network reliability evaluation sub-problem (SP2)

To ensure that the BES expansion decisions satisfy the minimum reliability requirements of the distribution network, the distribution network reliability evaluation sub-problem is formulated as an LP problem. The objective is to minimize the annual demand curtailment [31,32]. The objective function (40) is subjected to the constraints (41) and (42) and (15)–(35). The reliability evaluation sub-problem is solved using the expansion planning decision ($\widehat{\mu}_e^{i,y}$). Here, the power supplied to the data centers in (41) is fixed to the solution obtained from the distribution network operation sub-problem (SP1) i.e., ($\widehat{p}_{c,s}^{t,d,y}$).

$$min\ S_{2,s}^{t,d,y} = \sum_l\left(LS_{l,s}^{t,d,y}\right) \tag{40}$$

s.t.

$$A_g\cdot p_{g,s}^{t,d,y} + A_v\cdot p_{v,s}^{t,d,y} + A_e\cdot\left(p_{dc,e,s}^{t,d,y} - p_{ch,e,s}^{t,d,y}\right) + A_f\cdot p_{f,b,s}^{t,d,y} = B\cdot p_{b,s}^{t,d,y} + D_l\cdot\left(p_{l,s}^{t,d,y} - LS_{l,s}^{t,d,y}\right) + D_c\cdot\widehat{p}_{c,s}^{t,d,y} \tag{41}$$

$$A_g\cdot q_{g,s}^{t,d,y} + A_v\cdot q_{v,s}^{t,d,y} + A_e\cdot q_{e,s}^{t,d,y} + A_f\cdot q_{f,b,s}^{t,d,y} = B\cdot q_{b,s}^{t,d,y} + D_l\cdot q_{l,s}^{t,d,y} \tag{42}$$

After solving the distribution network reliability sub-problem, the annual expected energy not supplied (EENS) is calculated using (43). If the annual EENS violates (44), the reliability cut (45) is sent to the MP. Otherwise, the power supplied to the data center ($\widehat{p}_{c,s}^{t,d,y}$) is passed to the data center feasibility sub-problem (SP3). Here, $\widehat{\pi}_e^{i,y}$ is the dual variable associated with the expansion decision of the BES units ($\mu_e^{i,y}$) in (21).

$$EENS_y = \sum_{s=1}^{T_s}\sum_{d=1}^{T_d}\sum_{t=1}^{T} pr_s\cdot N_d\cdot\widehat{S}_{2,s}^{t,d,y} \tag{43}$$

$$EENS_y \leqslant EENS_y^{max} \tag{44}$$

$$\sum_{s=1}^{T_s}\sum_{d=1}^{T_d}\sum_{t=1}^{T} pr_s\cdot N_d\cdot\widehat{S}_{2,s}^{t,d,y} + \sum_e\sum_i\widehat{\pi}_e^{i,y}\cdot\left(\mu_e^{i,y} - \widehat{\mu}_e^{i,y}\right) \leqslant EENS_y^{max} \tag{45}$$

### 2.4. Data center feasibility sub-problem (SP3)

The data center feasibility sub-problem (SP3) is formulated as an LP problem as shown in (46)–(51). The objective is to minimize the mismatch between the total workload collected from the end-users in the data network and the workload processed in the data centers as shown in (46). The workload in the data network represents the end-users' requests in the form of computing resources including processing, memory usage, and storage [33]. Here, the energy-intensive processing demand is considered as the workload. Positive slack variables are included in (47) to represent the mismatch between the received and processed workloads. The power consumed by a data center is associated with the workload processed as shown in (48). The capacity of a data center to process the workloads is limited by (49). Constraint (50) enforces the supplied energy to the data center to be equal to the solution obtained from SP1. The slack variables are non-negative as shown in (51).

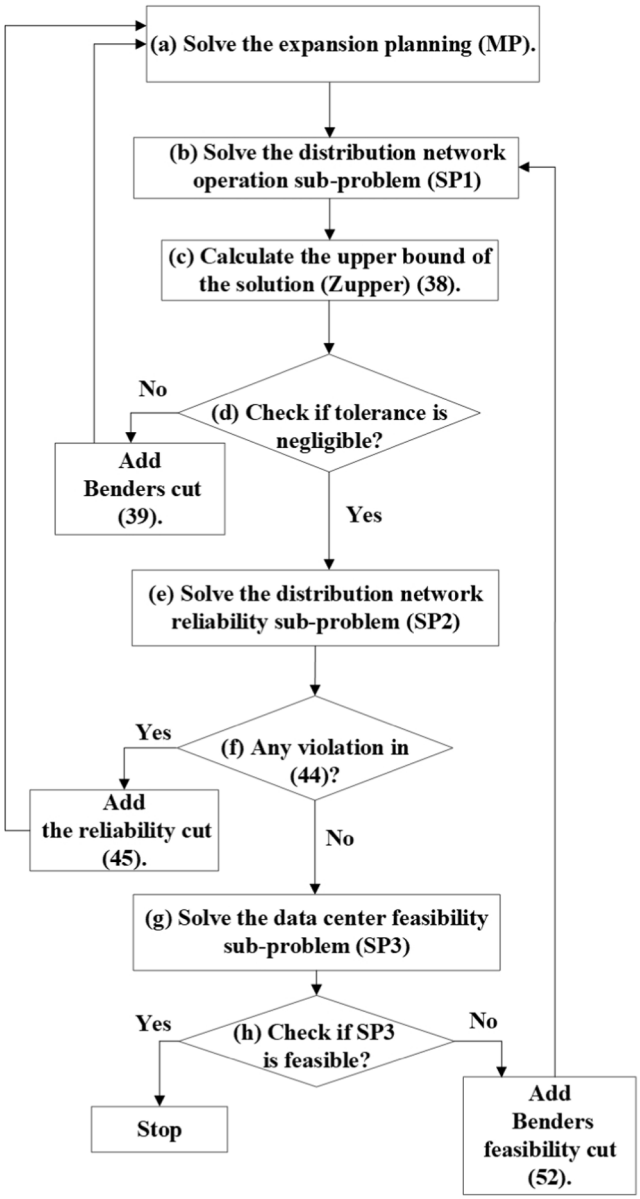


**Fig. 3.** The flowchart of the proposed stochastic planning framework.

$$min\ S_3 = \sum_{s=1}^{T_s}\sum_{y=1}^{T_y}\sum_{d=1}^{T_d}\sum_{t=1}^{T}\left(Z_{3,s}^{t,d,y} + Z_{4,s}^{t,d,y}\right) \tag{46}$$

s.t.

$$\sum_{c} w_{c,s}^{t,d,y} + Z_{3,s}^{t,d,y} - Z_{4,s}^{t,d,y} = \Delta_s^{t,d,y} \tag{47}$$

$$w_{c,s}^{t,d,y} \cdot p_c^{max} = w_c^{max} \cdot p_{c,s}^{t,d,y} \tag{48}$$

$$w_{c,s}^{t,d,y} \leqslant w_c^{max} \tag{49}$$

$$p_{c,s}^{t,d,y} = \widehat{p}_{c,s}^{t,d,y} \quad : \pi_{c,s}^{t,d,y} \tag{50}$$

$$Z_{3,s}^{t,d,y}, Z_{4,s}^{t,d,y} \geqslant 0 \tag{51}$$

In case of any mismatch in (46), Benders feasibility cut (52) is formed and added to SP1. Once the supplied energy to the data center is sufficient to process the workloads, i.e. $\widehat{S}_3 = 0$, the process stops.

$$\left(\widehat{Z}_{3,s}^{t,d,y} + \widehat{Z}_{4,s}^{t,d,y}\right) + \sum_{c} \widehat{\pi}_{c,s}^{t,d,y} \cdot \left(p_{c,s}^{t,d,y} - \widehat{p}_{c,s}^{t,d,y}\right) \leqslant 0 \tag{52}$$

## 3. Solution methodology

The flowchart of the proposed framework is shown in Fig. 3. The algorithm is composed of the following steps:

*Step (a)*: Solve the MP and determine the lower bound of the solution ($\widehat{Z}_{lower}$).

*Step (b)*: Using the solution of MP ($\widehat{\mu}_e^{i,y}$), solve the distribution network operation sub-problem (SP1). Go to step (c).

*Step (c)*: Calculate the upper bound of the solution $\widehat{Z}_{upper}$ using (38).

*Step (d)*: If $\left|\frac{\widehat{Z}_{upper} - \widehat{Z}_{lower}}{\widehat{Z}_{upper} + \widehat{Z}_{lower}}\right| \geqslant \in$, then add the Benders cut (39) to the MP and go to step (a). Otherwise, go to step (e).

*Step (e)*: Solve the distribution network reliability evaluation sub-problem (SP2) for the given expansion plan ($\widehat{\mu}_e^{i,y}$), calculate the annual EENS using (43), and go to step (f).

*Step (f)*: If there is any violation in the annual EENS requirement (44), add the reliability cut (45) to the MP and go to Step (a). Otherwise, go to Step (g).

*Step (g)*: Solve the data center feasibility sub-problem (SP3) for the given $\widehat{p}_{c,s}^{t,d,y}$ obtained from SP1.

*Step (h)*: If SP3 is feasible for the given $\widehat{p}_{c,s}^{t,d,y}$, i.e. $\widehat{S}_3 = 0$, then terminate the process. Otherwise, add the Benders feasibility cut (52) to SP1 and go to Step (b).

## 4. DS3 algorithm

The uncertainties associated with the hourly demand, hourly electricity prices of the utility grid, the hourly power generation of the solar PV units, the distribution branch outages, and the workloads received by the end-users in the data network are considered in this framework. To reduce the number of representative days and representative scenarios, the DS3 algorithm is used. This algorithm clusters the data points into a limited number of subsets. The subset is a set of data points that represents the original data set effectively. Each subset is formed using a dissimilarity metric which is the Euclidean distance among the data points [34]. The pairwise distances between the $n$ number of data points is used as a measure of dissimilarity as shown by matrix **A** in (53).

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix}_{n \times n} \tag{53}$$

where each entry $a_{j,k}$ in row $j$ indicates how well the $j$-th data point in the data set represents the $k$-th data point. Having a small dissimilarity $a_{j,k}$ indicates a better representation while having a large dissimilarity shows a strong statistical independence between the $j$-th and $k$-th data points. Matrix **X** with the size of $n \times n$ is defined in (54) to store the linear dependence coefficients of all data points. Here, each non-zero column $k$ indicates that the $k$-th data point is a representative.

$$\mathbf{X} = \begin{bmatrix} x_1^\top \\ \vdots \\ x_n^\top \end{bmatrix} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,n} \end{bmatrix}_{n \times n} \tag{54}$$

Each data point can be written as a linear combination of its corresponding representatives (i.e., corresponding non-zero columns). Each non-zero entry $x_{j,k} \in [0, 1]$ is the coefficient of the linear combination

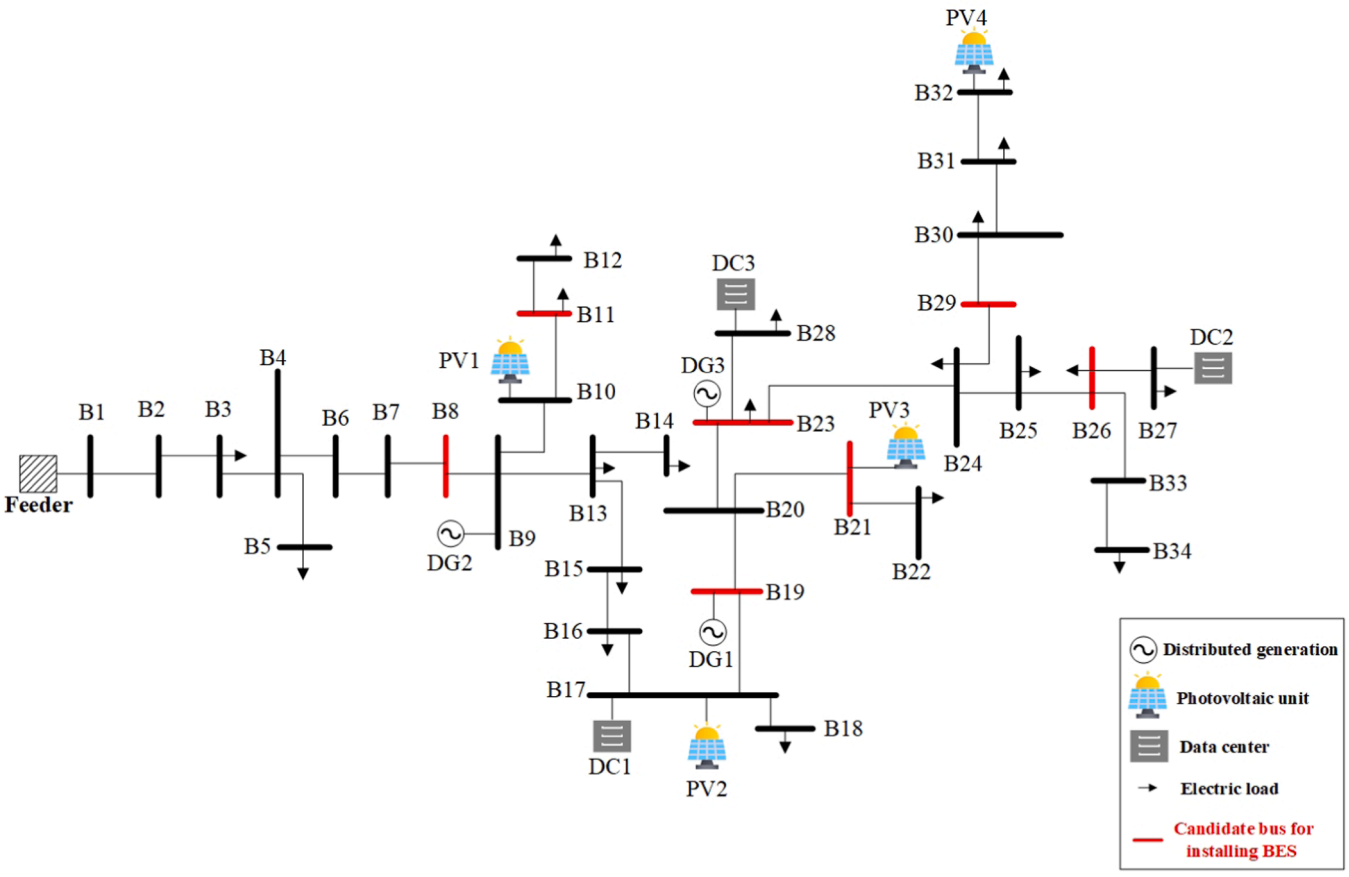


**Fig. 4.** The modified IEEE-34 bus distribution network with data centers and PV generation units.

corresponding to the $k$-th data point that represents the $j$-th data point. If $x_{j,k} = 0$, the $j$-th data point is not represented by the $k$-th data point. Matrix $\mathbf{X}$ is not only to find the representatives but also to cluster the data points according to the representatives. That is, one can assume a cluster corresponding to each representative with non-zero column in $\mathbf{X}$. To find the optimal representatives that best encode the data points corresponding to their clusters while minimizing the number of representatives, the DS3 algorithm solves the optimization problem (55), (56).

$$\min_{(x_{j,k})} \zeta . \sum_{j}^{n} \| x_j \|_2 + \sum_{k=1}^{n} \sum_{j=1}^{n} a_{j,k} . x_{j,k} \tag{55}$$

s.t.

$$\sum_{j=1}^{n} x_{j,k} = 1, \forall k \tag{56}$$

Here, $\| x_j \|_2$ is the $L2$-norm of the $j^{th}$ row of matrix $\mathbf{X}$. The objective function (55) has two terms. The first term represents the number of representative data points used in a linear combination to represent the $j$-th data point. Here, $\zeta$ is a regularization hyper-parameter that declines the number of representatives if increased. Having a small $\zeta$ (i.e., having a large number of representatives) would reduce the error in representing the original data and increase the process burden by handling more data samples. The expected distance between the representatives and the corresponding data points in the original data set is computed as in the second term. If the representatives provide a powerful data encoding with a small error, the second term decreases; hence, optimizing the total error in (55). The constraint (56) limits the number of data points represented by one representative (i.e., $k$-th representative). This constraint leads to a better dimensionality reduction while enhancing the encoding quality when the data points are written as a linear combination of representatives. Once the DS3 optimization is solved, the non-zero columns of the optimal solution matrix $\mathbf{X}$ are considered as the representatives that best show the statistical characteristics of the entire data points.

**Table 1**
Characteristics of distributed generation units.

| Unit | $P_{max}$ ($kW$) | $Q_{max}$ ($kVar$) | $\lambda_g$ ($/kWh^2$) | $\beta_g$ ($/kWh$) |
|---|---|---|---|---|
| DG1 | 300 | 150 | 0.00015 | 0.0025 |
| DG2 | 150 | 75 | 0.00028 | 0.038 |
| DG3 | 90 | 45 | 0.000778 | 0.00944 |

**Table 2**
Generation limits of PV units.

| Unit | Bus | $P_{max}$ (kW) | $Q_{max}$ (kVAR) |
|---|---|---|---|
| PV1 | 10 | 50 | 25 |
| PV2 | 17 | 300 | 150 |
| PV3 | 21 | 100 | 50 |
| PV4 | 32 | 250 | 125 |

## 5. Case study

The modified IEEE-34 bus distribution network shown in Fig. 4 is used to validate the proposed planning approach. The network is equipped with 3 distributed generation (DG) units, 4 solar PV generation units, and 3 data centers in the distribution network. The capacity of the distribution feeder is 4.4 MVA with a minimum power factor of 0.8. The

**Table 3**
Battery energy storage systems' characteristics.

| Battery | $P_{max}$ ($kW$) | $Q_{max}$ ($kVar$) | $E_{max}$ ($kWh$) | $C_e$ ($/kWh) | $\rho_e$ ($/kWh) |
|---|---|---|---|---|---|
| K1 | 100 | 60 | 400 | 450 | 0.18 |
| K2 | 50 | 35 | 200 | 420 | 0.17 |
| K3 | 25 | 15 | 100 | 380 | 0.15 |

**Table 4**
Number of days in each representative day of each year.

| Year | Representative | Number of Days |
|---|---|---|
| 1 | $R_{1,1}$ | 265 |
| | $R_{2,1}$ | 12 |
| | $R_{3,1}$ | 88 |
| 2 | $R_{1,2}$ | 131 |
| | $R_{2,2}$ | 82 |
| | $R_{3,2}$ | 152 |
| 3 | $R_{1,3}$ | 265 |
| | $R_{2,3}$ | 4 |
| | $R_{3,3}$ | 96 |
| 4 | $R_{1,4}$ | 232 |
| | $R_{2,4}$ | 20 |
| | $R_{3,4}$ | 113 |
| 5 | $R_{1,5}$ | 130 |
| | $R_{2,5}$ | 234 |
| | $R_{3,5}$ | 1 |
| 6 | $R_{1,6}$ | 136 |
| | $R_{2,6}$ | 224 |
| | $R_{3,6}$ | 5 |
| 7 | $R_{1,7}$ | 258 |
| | $R_{2,7}$ | 10 |
| | $R_{3,7}$ | 97 |
| 8 | $R_{1,8}$ | 216 |
| | $R_{2,8}$ | 11 |
| | $R_{3,8}$ | 138 |
| 9 | $R_{1,9}$ | 147 |
| | $R_{2,9}$ | 114 |
| | $R_{3,9}$ | 104 |
| 10 | $R_{1,10}$ | 179 |
| | $R_{2,10}$ | 81 |
| | $R_{3,10}$ | 105 |

characteristics of DG units and photovoltaic (PV) generation units are shown in Tables 1 and 2 respectively. The capacities of the data centers on buses B17, B27, and B28 are 100, 150, and 120 kW respectively. The characteristics of the BES units are shown in Table 3. Three types of Lithium-ion batteries are considered as the potential candidates. The characteristics of the BES units are obtained from [35] and their operation costs are calculated using the formulation in [36]. The charging and discharging efficiencies for the BES unit are 90%. As shown in Fig. 4, buses B8, B11, B19, B21, B23, B26, and B29 are considered as candidate buses to install the BES units. The minimum and the initial stored energy of all BES units are 15% and 50% of the maximum energy capacity, respectively.

The time horizon for the expansion planning is 10 years, the annual discount rate is 10%, and the annual growth of demand in the distribution network and data network is 5% [27]. The acceptable annual EENS is zero, and the convergence tolerance ($\in$) is 0.01%. The value of lost load is 10 $/kWh. In the first year, the solar PV penetration which is defined as the ratio of the total PV generation to the total demand in the distribution network is 17.21%, and the maximum demand of the data center is 20.49% of the peak demand in the distribution network. The following cases are considered:

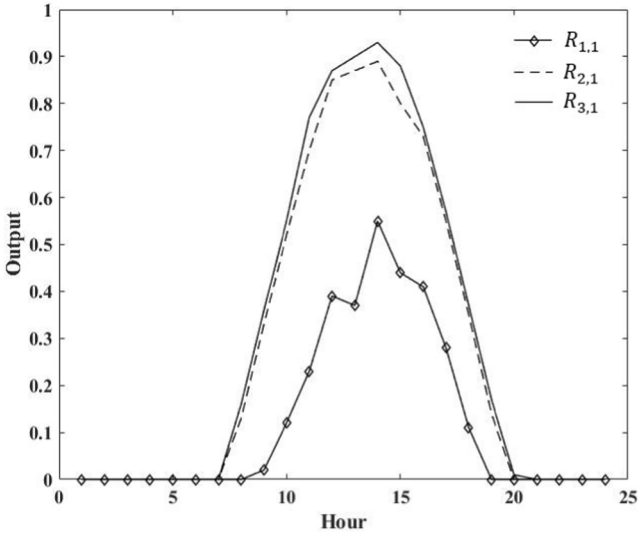


**Fig. 5.** Normalized PV generation for the representative days 1–3 ($R_{1,1}$, $R_{2,1}$, $R_{3,1}$) in the first year.

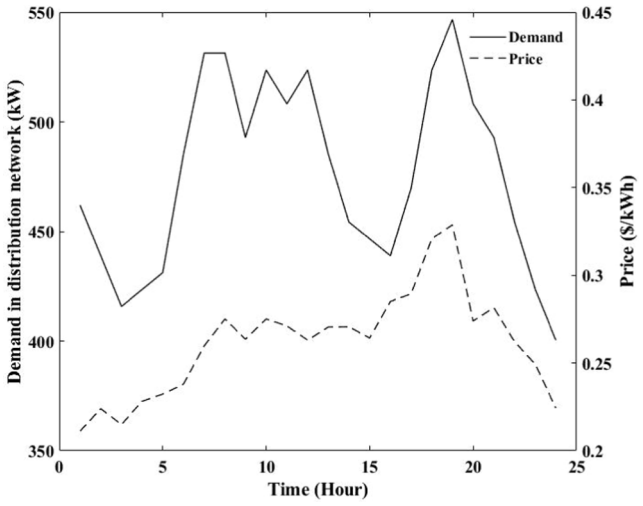


**Fig. 6.** Total hourly demand in the distribution network and hourly energy price in the first representative day of the first year.

Case 1 – Expansion planning of the BES units considering the forecasted demand, electricity prices, solar PV generation, and workload in the data network.
Case 2 – Expansion planning of the BES units considering the forecasted values with contingencies in the distribution network.
Case 3 – Expansion planning of the BES units with uncertainty in the forecasted demand, electricity prices, solar PV generation, distribution branch outages, and workload in the data network.
Case 4 – The impact of EENS on the expansion planning of the BES units.

### *5.1. Case 1 – Expansion planning of the BES units considering the forecasted demand, electricity prices, solar PV generation, and workload in the data network*

In this case, the distribution network with data centers is considered under normal operating conditions with no outages in the system. By applying the DS3 algorithm, the original data set which consists of 4 ×

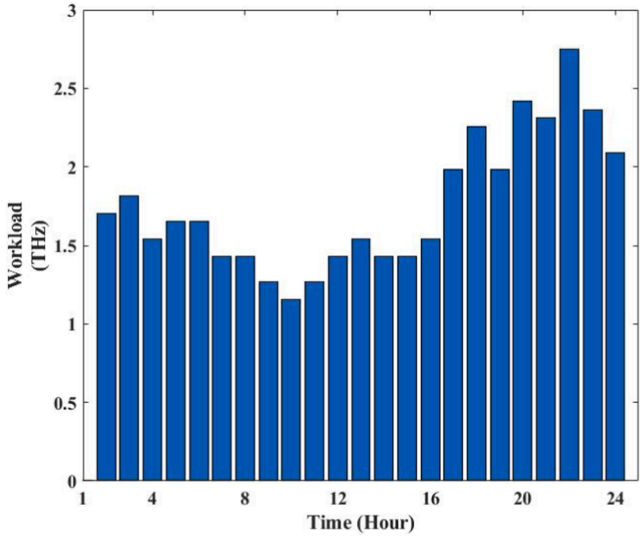


**Fig. 7.** The total workload in the first representative day of the first year.

**Table 5**
The expansion decision for BES in Case 1.

| Bus | Year | Storage type |
|---|---|---|
| B23 | 7 | K2 |
| B26 | 3 | K2 |
| B29 | 4 | K2 |

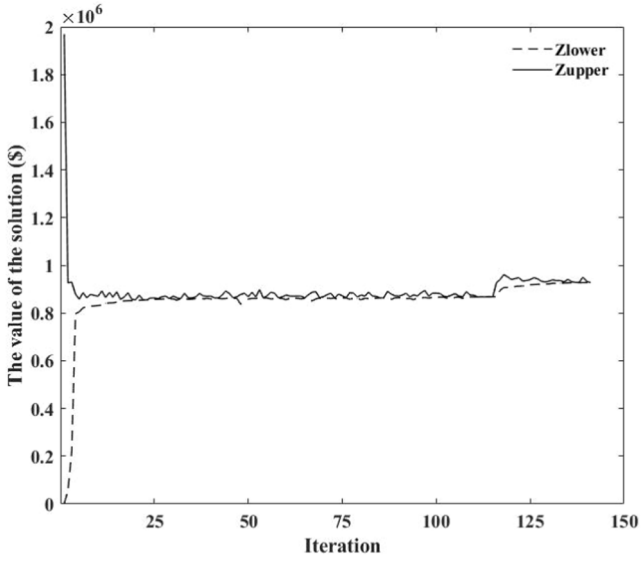


**Fig. 8.** The upper and lower bounds of the solution at each iteration in Case 1.

8760 = 35040 data points are clustered into 3 representative data subsets consisting of $3 \times 24 \times 4 = 288$ data points in each year. Table 4 shows the representative days and the number of days in each year ($R_{d,y}$). The normalized PV outputs for the three representative days in the first year are shown in Fig. 5. Here, the maximum output for a PV unit is determined by multiplying the nominal capacity of each PV unit by the normalized PV output. Similarly, the demand profile and the hourly energy prices on the first representative day of the first year ($R_{1,1}$) are shown in Fig. 6. The workload in the data network for the first representative day of the first year is shown in Fig. 7.

Table 5 shows the installed BES in Case 1. As shown in this table, the

**Table 6**
Outages in the distribution lines.

| Line | Year | Representatives | Hour |
|---|---|---|---|
| B23-B24 | 5 | R1 | 13:00 |
| B20-B23 | 6 | R2 | 16:00 |
| B16-B17 | 7 | R3 | 18:00 |
| B25-B26 | 4 | R2 | 12:00 |

**Table 7**
The expansion decisions for BES in Case 2.

| Bus | Year | Storage type |
|---|---|---|
| B11 | 7 | K3 |
| B26 | 3 | K1 |
| B29 | 6 | K3 |

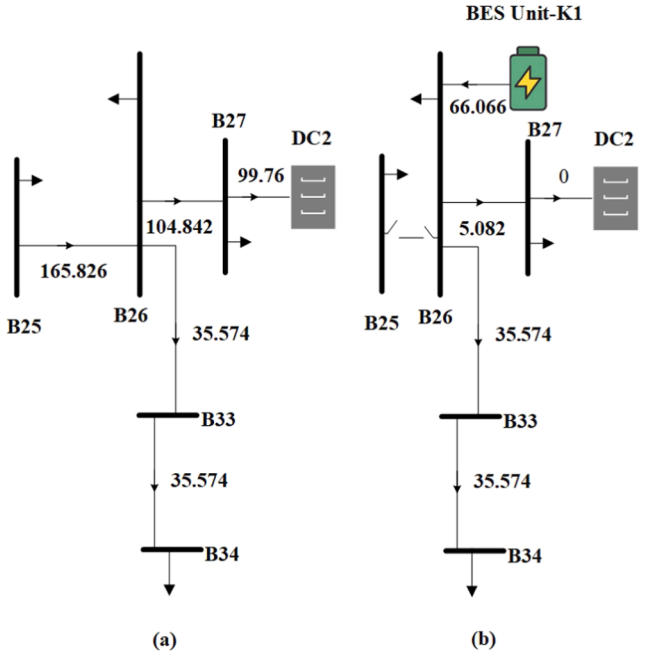


**Fig. 9.** Power flow in the distribution network. (a) Case 1 (b) Case 2.

expansion planning decision is to install three BES units type K2 with total capacity of 600 kWh. The location and installation time of the BES units are on buses B23, B26, and B29 in the seventh, third, and fourth year, respectively. The installation and operation costs of BES units are $163,588.86 and $5,958.50, respectively. The cost of purchased energy from the main feeder and the operation cost of distributed generation units are $417,642.92 and $340,354.47, respectively. The total planning cost for this case is $927,544.75. Fig. 8 shows the upper and lower bounds of the solution in each iteration. As shown in this figure, the upper and lower bounds of the solution converge as the number of iterations increases. The increase in the lower and upper bounds at 116th iteration is because of the feasibility cut passed from the data center feasibility sub-problem (SP3).

### *5.2. Case 2 – Expansion planning of the BES units considering the forecasted values with contingencies in the distribution network*

In this case, the impacts of outages in the distribution network on the expansion plans of the BES units are addressed. The outages in the branches between buses B23-B24, B20-B23, B16-B17, and B25-B26 are considered in the representative days of Case 1 as shown in Table 6. The expansion plans for the BES units are shown in Table 7. Compared to

**Table 8**
Probability of scenarios.

| Scenario | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Probability (%) | 28.2 | 33.6 | 34.6 | 0.4 | 3.2 |

Case 1, three BES units are installed with total capacity of 600 kWh which is the same in Case 1. However, the location, year, and type of BES units are changed because of the branch outages in the distribution network. The planning decision is to install one BES unit type K1 at Bus B26 in the third year, and two BES unit type K3 at Bus B11 and B29 in the seventh and sixth year, respectively.

The installed BES units will help to serve the demand once the distribution branches are on outage. The BES unit installed on bus B29 along with the PV generation at Bus B32 would supply the demand on buses B30, B31, and B32 during the outage of branch B20-B23. In Case 1, at hour 16:00 in the second representative day of the sixth year, the demand on buses B30, B31, and B32 along with the demand on buses B23-B34 in the distribution network are supplied by branch B20-B23 that carries 168.566 kW, and by the distributed generations on bus B23 as well as the PV generation units on bus B21 and B32. During the same period in Case 2, branch B20-B23 is on outage, and the demand on buses B30, B31, and B32 are supplied by the BES unit type K3 on bus B29 along with PV generation unit on bus B32.

Furthermore, the outage in the distribution network would lead to changes in the workload process by the data centers. Fig. 9a. shows the power flow in a section of the distribution network at hour 12:00 on the second representative day of the fourth year in Case 1, while Fig. 9b. shows the power flow in the same period in Case 2 with branch B25-B26 on outage. As shown in Fig. 9a, the branch B25-B26 supplies the demand in the distribution network on buses B26, B27, and B34 as well as the demand of the data center DC2. The total power consumption of data center DC2 on bus B27 is 99.76 kW. However, in Fig. 9b, the power consumption of DC2 is zero and the workloads are redistributed among data centers DC1 and DC3. Here, because of the outage on branch B25-B26, the BES unit is installed on Bus B26 to merely serve the electric demand on buses B26, B27, and B34; and DC1 and DC3 serve the received workloads and increase their demand by 99.76 kW.

The total installation cost of the BES units is $176,186.68 in this case which is higher than that in Case 1 by $12,597.82. The operation cost of BES units is increased to $14,816.52. The operation cost of distributed generation is increased to $350,307.34 in this case because of the increase in the power dispatch of DG1 and DG3 during the branch outages. The cost of the purchased power from the main feeder is reduced to $415,336.94. The total cost in this case is $956,647.48 which is $29,102.73 higher than that in Case 1. In Case 2, the branch outages lead to installing different types of BES units and increase the dispatch of the distributed generation units.

### *5.3. Case 3 – Expansion planning of the BES units with uncertainty in the forecasted demand, electricity prices, solar PV generation, distribution branch outages, and workload in data network*

The stochastic solution for the BES expansion planning is presented in this case. The uncertainties in the distribution network's demand, the PV generation, electricity prices, the workloads received by the data centers, and the outages of the distribution branches are considered. The forecast errors of the uncertain parameters are represented by the Gaussian probability distribution functions for which the mean values are the forecasted data used in Cases 1 and 2, and the standard deviations are 3% of the mean values. The mean time to failure (MTTF) and the mean time to repair (MTTR) for distribution branches B23-B24, B20-B23, B16-B17, and B25-B26 are 38,400 h and 15 h, respectively. Monte-Carlo simulation is used to generate 500 scenarios. Here, 8 matrices each with (500 × 87600) data points are generated to represent 8 uncertain parameters i.e. demand, electricity price, workload for the data centers, PV generation, and the outages in four distribution lines in 10 years. DS3 algorithm [34] is used to reduce the number of scenarios to 5 with associated probabilities shown in Table 8. Similarly, for each year in each scenario, 3 representative days are clustered using DS3 algorithm. Fig. 10 shows the number of days in each representative day in each year ($R_{d,y}$) for each scenario. After applying the DS3 algorithm, three representative days for each year are obtained in each scenario. Considering 10 years in the planning horizon, 30 representative days are considered in each scenario as shown in Fig. 10. Fig. 11 shows the normalized PV power generation in each scenario for the first

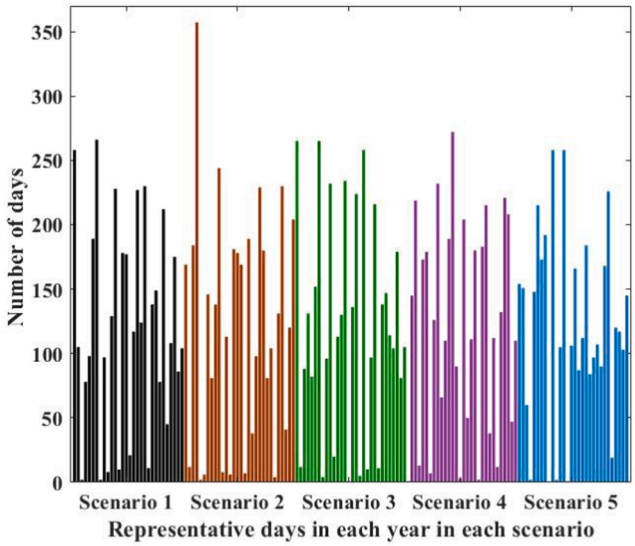


**Fig. 10.** Number of days in the representative days in each year ($R_{d,y}$) for each scenario.

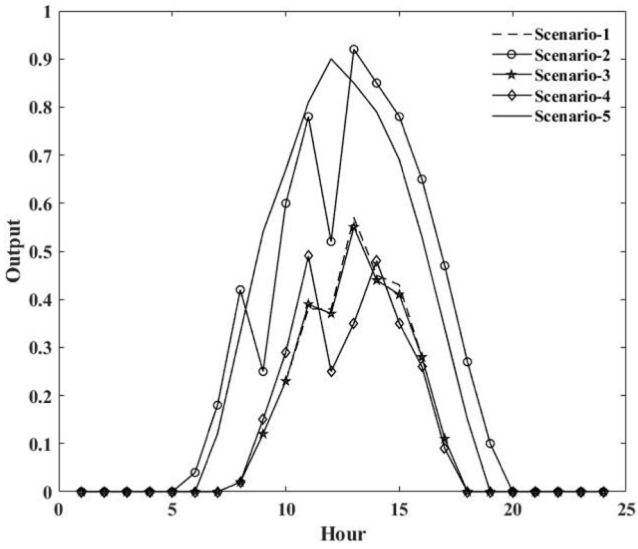


**Fig. 11.** Normalized PV generation in each scenario for the first representative day in the first year ($R_{1,1}$) in Case 3.

**Table 9**
The expansion decisions for BES in Case 3.

| Bus | Year | Storage type |
|---|---|---|
| B23 | 8 | K2 |
| B26 | 4 | K1 |
| B29 | 5 | K2 |

**Table 10**
Summary of the planning solutions of all cases.

| # | Case 1 | Case 2 | Case 3 |
|---|---|---|---|
| Total BES capacity (kWh) | 600 kWh | 600 kWh | 800 kWh |
| Total investment cost ($) | $163,588.86 | $176,186.68 | $214,286.43 |
| Total operation cost ($) | $763,955.89 | $780,460.8 | $770,458.81 |
| Total planning cost ($) | $927,544.75 | $956,647.48 | $984,745.24 |

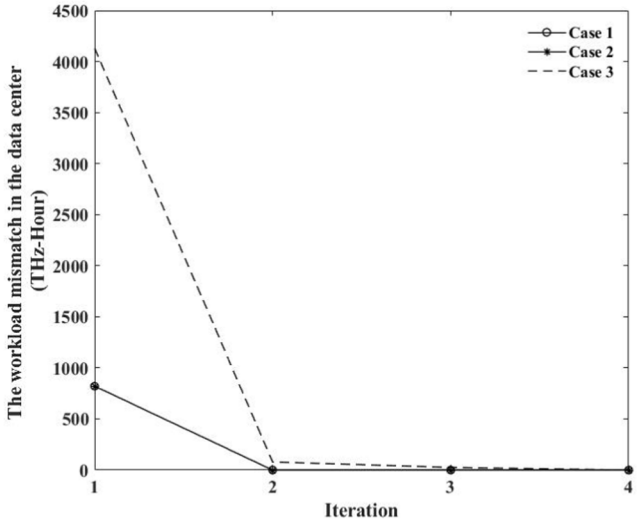


**Fig. 12.** The mismatch between the required and served workloads by the data centers at each iteration in Cases 1–3.

**Table 11**
The expansion decisions in Case 3 with the total acceptable EENS of 524.32 MWh.

| Bus | Year | Storage type |
|---|---|---|
| B23 | 7 | K3 |
| B26 | 3 | K2 |
| B29 | 4 | K2 |

representative day in the first year ($R_{1,1}$).

Table 9 shows the installation plan for the BES units. Compared to Case 1, an additional 200 kWh capacity of the BES units is installed. The expansion decision is installing one BES unit type K1 on bus B26 in the fourth year and two BES units type K2 on bus B29 and B23 in the fifth and eighth year, respectively. Compared to Case 1, the installation cost of BES units is increased to $214,286.43. Furthermore, the total expected operation costs of the distributed generation and BES units are decreased to $338,424.69 and $3,831.28, respectively. The expected cost of the power supplied from the distribution feeder is $428,202.84 which is $10,559.92 higher than that in Case 1. Table 10 summarizes the solution outcomes in all cases. The total expansion and expected operation cost, in this case, is $984,745.24 which is 6.17% and 2.94% higher than the total expansion planning costs in Cases 1 and 2, respectively. Fig. 12 shows the mismatch between the total required and served workloads by the data centers in each iteration for Cases 1–3. As shown in this figure, the mismatch reaches zero after two interactions (iterations) between the DNO and DCO in Cases 1 and 2. The workload mismatch in the data centers converges to zero after four interactions between the DNO and DCO in Case 3.

### 5.4. The impact of EENS on the expansion planning of BES

In this section, the impact of EENS on the expansion planning of BES units is evaluated in Case 3. Here, the acceptable EENS in the first year is 41.686 MWh which is 1% of the total expected demand in the first year. The acceptable EENS increases by 5% annually. Therefore, the acceptable total EENS for the planning horizon is 524.32 MWh. The value of lost load is the same as in previous cases. Table 11 shows the expansion plans for the BES units in this case. The expansion planning decision is to install 500 kWh BES units, which is 300 kWh lower than the installed capacity once the acceptable total EENS is zero. Consequently, the investment cost for BES units is reduced to $139,983.58 with the total EENS of 4063.451 kWh. The penalty associated with the curtailed demand is $15,876.55 and total operation cost is $802,510.48. The total expansion planning cost is $958,370.61 which is 2.68% lower than that in Case 3 with zero acceptable total EENS.

**Table 12**
The impact of VOLL on the investment, operation, and total planning costs.

| VOLL ($/kWh) | Total EENS (kWh) | Investment Cost ($) | Operation Cost ($) | Demand Curtailment Cost ($) | Total Planning Cost ($) |
|---|---|---|---|---|---|
| 50 | 0 | 214,286 | 770,458 | 0 | 984,745 |
| 40 | 482 | 202,775 | 774,474 | 7,441 | 984,691 |
| 30 | 2390 | 163,588 | 788,395 | 27,650 | 979,634 |
| 20 | 2390 | 163,588 | 788,395 | 18,433 | 970,417 |
| 10 | 4063 | 139,983 | 802,510 | 15,876 | 958,370 |

The total EENS, the investment cost for the BES units, the total operation and planning costs as well as the cost associated with the curtailed demand for Case 3 are shown in Table 12 with the changes in the VOLL from $10/kWh to $50/kWh. As shown in this table, as the VOLL increases, the total EENS is reduced because of the increase in the penalty associated with the curtailed demand. Consequently, the operation cost is reduced. Moreover, the increase in the VOLL will increase the installed capacity of the BES units to reduce the total EENS. Consequently, the investment cost and the total planning cost are increased with the increase in the VOLL.

## 6. Conclusion

A scenario based expansion planning framework is proposed in this paper for optimal sizing and locating the BES units in the distribution network with data center facilities. The planning algorithm seeks to minimize the BES installation cost, the operation cost of the distribution network while providing sufficient power supply to the data centers and ensuring the reliability of the distribution network. The proposed algorithm captures the interactions between DNO and DCO using Bender decomposition technique. Monte Carlo simulation is utilized to generate a large number of scenarios based on the probability distribution functions representing the forecast errors in uncertain parameters. DS3 clustering technique is used to perform scenario reduction and select the effective scenarios. The modified IEEE-34 bus distribution network with three data centers is used to validate the effectiveness of the proposed algorithm. The numerical results show that incorporating the uncertainty in the availability of the distribution lines, electricity demand and prices, workloads in the data network, and generation of PV units will lead to an increase in the installed BES capacity in the distribution network. Moreover, the outages in the distribution branches lead to changes in the location and types of the BES units and consequently, changes in the expansion planning costs and resource allocation strategies in data centers to process the received workloads. The impact of EENS on the expansion planning decisions is evaluated. It is shown that the total expansion planning cost is reduced as the EENS increases. Moreover, as the VOLL increases, the total EENS decreases, and the investment cost and the total planning cost increase.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

## References


[1] The Power of Efficiency to Cut Data Center Energy Waste. Online. https://www.nrdc.org/experts/pierre-delforge/power-efficiency-cut-data-center-energy-waste.
[2] Online. https://www.grandviewresearch.com/industry-analysis/data-center-construction-market.
[3] Sedghi M, Ahmadian A, Aliakbar-Golkar M. Optimal storage planning in active distribution network considering uncertainty of wind power distributed generation. IEEE Trans. Power Syst. 2016;31:304–16.
[4] Lazzeroni P, Repetto M. Optimal planning of battery systems for power losses reduction in distribution grids. Electric Power Systems Research 2019;167:94–112.
[5] Nick M, Cherkaoui R, Paolone M. Optimal planning of distributed energy storage systems in active distribution networks embedding grid reconfiguration. IEEE Trans. Power Syst. 2018;33:1577–90.
[6] Khalid M, Akram U, Shafiq S. Optimal planning of multiple distributed generating units and storage in active distribution networks. IEEE Access 2018;6:55234–44.
[7] Awad AS, El-Fouly TH, Salama MM. Optimal ess allocation and load shedding for improving distribution system reliability. IEEE Transactions on Smart Grid 2014;5:2339–49.
[8] M. Nick, M. Hohmann, R. Cherkaoui, M. Paolone, Optimal location and sizing of distributed storage systems in active distribution networks, in: 2013 IEEE Grenoble Conference, IEEE, 2013, pp. 1–6.
[9] Shen X, Shahidehpour M, Han Y, Zhu S, Zheng J. Expansion planning of active distribution networks with centralized and distributed energy storage systems. IEEE Transactions on Sustainable Energy 2017;8:126–34.
[10] Akhavan-Hejazi H, Mohsenian-Rad H. Energy storage planning in active distribution grids: A chance-constrained optimization with non-parametric probability functions. IEEE Transactions on Smart Grid 2018;9:1972–85.
[11] de Quevedo PM, Munoz-Delgado G, Contreras J. Impact of electric vehicles on the expansion planning of distribution systems considering renewable energy, storage, and charging stations. IEEE Transactions on Smart Grid 2019;10:794–804.
[12] Saboori H, Hemmati R. Maximizing disco profit in active distribution networks by optimal planning of energy storage systems and distributed generators. Renew. Sustain. Energy Rev. 2017;71:365–72.
[13] Zhang Y, Dong ZY, Luo F, Zheng Y, Meng K, Wong KP. Optimal allocation of battery energy storage systems in distribution networks with high wind power penetration. IET Renew. Power Gener. 2016;10:1105–13.
[14] Giannitrapani A, Paoletti S, Vicino A, Zarrilli D. Optimal allocation of energy storage systems for voltage control in lv distribution networks. IEEE Transactions on Smart Grid 2017;8:2859–70.
[15] Zhang Y, Ren S, Dong ZY, Xu Y, Meng K, Zheng Y. Optimal placement of battery energy storage in distribution networks considering conservation voltage reduction and stochastic load composition. IET Generation, Transmission & Distribution 2017;11:3862–70.
[16] X. Qian, S. Zhang, J. Liu, Y. Zheng, W. Liu, Hierarchical optimal planning of battery energy storage systems in radial distribution networks, in: 2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2), IEEE, 2019, pp. 84–89.
[17] Asensio M, de Quevedo PM, Mu noz-Delgado G, Contreras J. Joint distribution network and renewable energy expansion planning considering demand response and energy storage—part i: Stochastic programming model. IEEE Transactions on Smart Grid 2018;9:655–66.
[18] P. Wang, L. Xie, Y. Lu, Z. Ding, Day-ahead emission-aware resource planning for data center considering energy storage and batch workloads, in: 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), IEEE, 2017, pp. 1–6.
[19] Guo Y, Fang Y. Electricity cost saving strategy in data centers by using energy storage. IEEE Trans. Parallel Distrib. Syst. 2013;24:1149–60.
[20] C. Ren, D. Wang, B. Urgaonkar, A. Sivasubramaniam, Carbon-aware energy capacity planning for datacenters, in: 2012 IEEE 20th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, IEEE, 2012, pp. 391–400.
[21] D. Gmach, J. Rolia, C. Bash, Y. Chen, T. Christian, A. Shah, R. Sharma, Z. Wang, Capacity planning and power management to exploit sustainable energy, in: 2010 International Conference on Network and Service Management, IEEE, 2010, pp. 96–103.
[22] F. Kong, X. Liu, Greenplanning: Optimal energy source selection and capacity planning for green datacenters, in: 2016 ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS), IEEE, 2016, pp. 1–10.
[23] L. Zhou, L.N. Bhuyan, K. Ramakrishnan, Dream: Distributed energy-aware traffic management for data center networks, in: Proceedings of the Tenth ACM International Conference on Future Energy Systems, 2019, pp. 273–284.
[24] Y. Guo, M. Pan, Coordinated energy management for colocation data centers in smart grids, in: 2015 IEEE International Conference on Smart Grid Communications (SmartGridComm), IEEE, 2015, pp. 840–845.
[25] Yu L, Jiang T, Zou Y. Distributed real-time energy management in data center microgrids. IEEE Transactions on Smart Grid 2018;9:3748–62.
[26] Vafamehr A, Khodayar ME, Manshadi SD, Ahmad I, Lin J. A framework for expansion planning of data centers in electricity and data networks under uncertainty. IEEE Transactions on Smart Grid 2019;10:305–16.
[27] A. Vafamehr, Operation and planning of data centers in electricity networks, Ph.D. thesis, 2019.
[28] Kargarian A, Hug G, Mohammadi J. A multi-time scale co-optimization method for sizing of energy storage and fast-ramping generation. IEEE Transactions on Sustainable Energy 2016;7:1351–61.
[29] Baran ME, Wu FF. Network reconfiguration in distribution systems for loss reduction and load balancing. IEEE Power Engineering Review 1989;9:101–2.
[30] Ahmadi H, Martı JR. Linear current flow equations with application to distribution systems reconfiguration. IEEE Trans. Power Syst. 2014;30:2073–80.
[31] Zhang X, Shahidehpour M, Alabdulwahab AS, Abusorrah A. Security-constrained co-optimization planning of electricity and natural gas transportation infrastructures. IEEE Trans. Power Syst. 2015;30:2984–93.
[32] Khodaei A, Shahidehpour M. Microgrid-based co-optimization of generation and transmission planning in power systems. IEEE transactions on power systems 2012;28:1582–90.
[33] V.S. Shekhawat, A. Gautam, A. Thakrar, Datacenter workload classification and characterization: An empirical approach, in: 2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS), IEEE, 2018, pp. 1–7.
[34] Elhamifar E, Sapiro G, Sastry SS. Dissimilarity-based sparse subset selection. IEEE transactions on pattern analysis and machine intelligence 2015;38:2182–97.
[35] K. Mongird, V.V. Viswanathan, P.J. Balducci, M.J.E. Alam, V. Fotedar, V.S. Koritarov, B. Hadjerioua, Energy storage technology and cost characterization report, Technical Report, Pacific Northwest National Lab. (PNNL), Richland, WA (United States), 2019.
[36] T.A. Nguyen, Optimization in microgrid design and energy management, Ph.D. thesis, 2014.