Data-Driven Distributionally Robust Unbalanced Operation of Distribution Networks with High Penetration of Photovoltaic Generation and Electric Vehicles

Mohammad Ramin Feizi[a,*], Mohammad Khodayar[a] and Jiayong Li[a,b]

[a]Department of Electrical and Computer Engineering, Southern Methodist University, Dallas, TX, United States
[b]College of Electrical and Information Engineering, Hunan University, Changsha 410082, China

ARTICLE INFO

Keywords:
unbalanced distribution system operation, solar photovoltaic generation, electric vehicle, distributionally robust optimization, uncertainties

Abstract

The uncertainties associated with the large-scale integration of electric vehicles (EV) and renewable energy resources introduce new challenges to operating the unbalanced three-phase distribution networks. In this paper, a data-driven distributionally robust optimization framework is proposed for the operation of the distribution network considering the uncertainties associated with the interconnected EV fleets and solar photovoltaic (PV) generation. The proposed framework leverages the column-and-constraint generation (C&CG) approach to minimize the operation cost considering the worst-case probability distributions of PV generation, the available energy in EV fleets, the arrival and departure times of EV fleets as well as their minimum and maximum energy capacities. The proposed approach is applied to the modified three-phase unbalanced IEEE 34-bus and IEEE 123-bus networks. To evaluate the performance of the proposed distributionally robust optimization framework, the results are compared to those procured by solving scenario-based stochastic programming and robust optimization problems. Furthermore, the impact of vehicle-to-grid capability on the operation of the distribution network is investigated and the in-sample and out-of-sample performances of the proposed framework are evaluated.

1. Introduction

The worldwide electric vehicle registration is increased by 41% in 2020 as the vehicle sales are decreased by 16% in 2020 due to the global pandemic [1]. As the EV sales grow, the share of gasoline vehicles' sales is estimated to decline from 94% in 2019 to 81% in 2050 [2]. Incentives for EVs including sales tax reductions and exemptions as well as subsidies for the charging assets are aimed to promote the transition toward sustainable mobility, reduce greenhouse gas generation, and improve local air quality in urban areas [3]. The increase in the number and capacity of EV interconnection will increase the EV charging demand in the distribution feeder. The EV charging load increases the residential demand peak and network loss and further increases the risk of overloading, voltage drops, and power quality degradation [4, 5]. Capacity expansions, installing energy storage, and regulating the EV charging profile are among the efforts to mitigate the adverse effects of large-scale integration of EVs in the distribution networks [6].

Solar photovoltaic (PV) generation is the prominent distributed renewable energy resource (DER) that could be installed close to the customers. The reduction in the capital cost through the project lifetime, the changes in state regulation and policies to promote PV technology (e.g., 100% renewable generation in California by 2045), and the tax incentives [2] are expected to further promote this technology and increase its installed capacity in the distribution networks. The growth in capacity of the uncertain and variable PV generation may lead to excessive voltage fluctuations, voltage violations, and unbalanced loading and increase the vulnerability of the distribution networks to overloading and protection malfunctions. The vehicle-to-grid (V2G) technology with considerable power injection capacity coupled with PV generation in the distribution feeders could further increase the probability of reverse power flow and voltage variations. Coordinating the charging and discharging schedule of EV clusters with DERs including PV generation and energy storage systems (ESS) could decrease the risk of such undesirable conditions. Providing services such as peak shaving, valley filling, reduction in load curtailment by EVs further improve the energy economics and reliability in the distribution networks [7], [8]. In [9], the coordination between EVs and DERs is formulated and solved by a differential evolution optimization algorithm to mitigate the unbalance among phases, improve the voltage profile, and decrease the distribution system loss. In [10] a hybrid optimization is proposed to reduce the EV operation cost. The scheduling of EV charging stations is coordinated with PV and battery storage systems using the real-time empirical data.

The uncertainties associated with the operation and interconnection of EVs including the uncertainties in the daily energy consumption, arrival and departure times, maximum and minimum available energy, and the number of interconnected EVs, introduce challenges to the operation of distribution networks with EVs. Stochastic programming (SP) and robust optimization (RO) problem formulations are used in the literature to address the uncertainties associated with the integration of EVs. In [11] a two-stage SP problem is

*This document is the results of the research project funded by the National Science Foundation Grant ECCS-1710923.

*Corresponding author
mfeizi@smu.edu (.M.R. Feizi); mkhodayar@smu.edu (M. Khodayar); j-y.li@connect.polyu.hk (J. Li)
ORCID(s): 0000-0002-7270-6518 (.M.R. Feizi)

Nomenclature

Parameters

$\bar{p}_{(.),k,t}$	Maximum real power.
$\bar{q}_{(.),k,t}$	Maximum reactive power.
$\bar{S}_{l,k}$	Maximum complex power of branch l on phase k.
η_e^{ch}, η_e^{dc}	Battery charging and discharging efficiency.
$\overline{p}_{i,k,g}$	Maximum power at segment g of the generation cost curve.
$\phi_{l,k}$	Availability of phase k on branch l.
$\rho_{f,t}$	Price of electricity of feeder.
$\underline{V}, \overline{V}$	Squared lower and upper limits of nodal voltage.
$AD_{d,b}$	Element of demand-bus incidence matrix.
$AE_{e,b}$	Element of energy storage-bus incidence matrix.
$AI_{i,b}$	Element of unit-bus incidence matrix.
$AL_{l,b}$	Element of line-bus incidence matrix.
$AN_{f,b}$	Element of feeder-bus incidence matrix.
$AV_{v,b}$	Element of PV-bus incidence matrix.
cc	Cost of load curtailment.
$I_{e,k,t}^{arv}$	Binary parameter for interconnection of EV cluster.
$I_{e,k,t}^{dep}$	Binary parameter for interconnection of EV cluster.
PF_f	Power factor at distribution feeder f.
$w_{i,g}$	Marginal cost of the unit i in segment g.
$\hat{E}_{e,k,t}^{arv}$	Forecasted available energy at arrival time.
$\hat{E}_{e,k,t}^{dep}$	Forecasted available energy at departure time.
$\hat{E}_{e,k,t}^{max}$	Forecasted maximum available energy for EV cluster.
$\hat{E}_{e,k,t}^{min}$	Forecasted minimum available energy for EV cluster.
$\hat{p}_{e,k,t}^{ch,max,m}$	Maximum charging power for EV cluster e.
$\hat{p}_{e,k,t}^{dc,max,m}$	Maximum discharging power for EV cluster e.
$I_{e,k,t}$	Binary parameter for interconnection of EV cluster.

Sets and Indecies

$\mathbb{C}_s$	Set of the closed switches in spanning tree s.
$\mathbb{O}_s$	Set of the open switches in spanning tree s.
$\mathbb{S}$	Set of all spanning trees.
b	Index of bus.
d	Index of demand.
e	Index of electric vehicle cluster.
f	Index of distribution feeder.
i	Index of distributed generation.
k	Index of phase.
l	Index of distribution branch.
m, n	Index of sample.
s	Index of spanning tree.
v	Index of PV generation unit.

Variables

$\mathbf{U}_{\mathbf{b,t}}^{\mathbf{m}}$	Squared voltage vector on bus b.
$E_{e,k,t}^{m}$	Energy capacity of EV cluster.
$p_{(.),k,t}^{m}$	Real power of a unit.
$p_{e,k,t}^{ch,m}$	Real charging power for electric vehicle cluster e.
$p_{e,k,t}^{dc,m}$	Real discharging power for electric vehicle cluster e.
$ps_{i,k,g,t}^{m}$	Power in segment g of the cost curve for a distributed generation.
$q_{(.),k,t}^{m}$	Reactive power of a unit.
$y_{l,t}$	Binary variable representing the status of the distribution branch, 1 if connected, and 0 otherwise.
$S_{(.),k,t}^{m}$	Apparent power on phase k.

formulated to determine the location and capacity of public EV charging stations considering the uncertainties in EVs' arrival and departure times and their charging patterns, as well as the preferred walking distance to the charge station. The objective function is to maximize the availability and access of EVs to the charging stations. In [12] the capacity of EV chargers, solar PV panels, and battery energy storage in the EV charging station is determined considering the uncertainty associated with the EV mobility and charging demand, PV generation, and electricity price. The problem is formulated as a SP problem using scenarios. In [13], the dispatch of EV charging stations equipped with PV generation and energy storage systems is determined using a stochastic dynamic programming approach that captures the uncertainty in electricity price, PV output, as well as electricity tariffs. In [14] a RO problem is formulated to determine the charging and discharging schedule of EV aggregators, considering the uncertainty in the available energy of EVs that is affected by mobility patterns and driver behaviors. The objective is to minimize the operation cost of the system including the operation cost of EV aggregators. In [15], the energy cost and voltage deviation are minimized while capturing the uncertainty in charging power and energy capacity, and the volatility in real and reactive demand. A RO problem is formulated to determine the real and reactive power dispatch of EVs and Benders' decomposition technique is used to solve the problem.

While in the SP approach, the probability distribution functions of the uncertain parameters are known to generate scenarios, formulating a RO problem eliminates the need for this assumption. In the RO problem, the uncertainty is defined as a set and the worst-case realization of uncertainty is considered. In most cases, the estimated probability distribution function for uncertain variables may not be accurate. Moreover, to accurately represent the uncertainties, a large number of scenarios is required which further increases the computation burden of solving the SP problems. Compared to the solution procured by SP, the worst-case realization of the uncertainty in the RO formulation provides a more conservative solution and by regulating the budget

of uncertainty, the conservativeness of the solution is adjusted [16]. In [16, 17] a hybrid stochastic and robust optimization approach is proposed to reduce the computation burden of the SP problem by considering the uncertainty sets. Distributionally robust optimization (DRO) is introduced in [18, 19, 20] to capture the worst-case realization of the probability distribution of the uncertain variables. In [18], a decomposition framework is proposed to solve a two-stage data-driven transmission expansion planning problem. The L_1 norm is used to construct the ambiguity set using the empirical data, that represents the uncertainty in wind generation. In [19] a data-driven risk-averse stochastic unit commitment problem formulation is proposed where L_1 and L_∞ norms are used to construct the ambiguity set of probability distributions for wind generation. The formulated problem is solved using Bender's decomposition technique. In [20] a DRO framework for the unit commitment problem is proposed where the Wasserstein metric is used as a measure to construct the ambiguity set and the solution to the formulated problem is compared to that of SP and RO problems. In [21] a two-stage DRO problem with the Kullback-Divergence metric is presented for the planning of EV charging stations and renewable resources in the power network. Here, the location of the charging stations is determined in the first stage and the capacities of renewable generation and energy storage are optimized in the second stage.

In this paper, a Wasserstein metric-based DRO problem is formulated and solved for the unbalanced distribution network operation considering the uncertainties associated with EVs and PV generation. The formulated problem is a two-stage optimization problem where the topology of the radial distribution network is determined in the first stage problem, and the distribution network operation cost is minimized considering the worst-case probability distribution of the uncertainties in the second stage problem. The contributions of this paper are summarized as follows:

- A data-driven DRO problem is proposed for the unbalanced operation of the distribution networks considering the network switching, and the performance of the procured solution is compared with the solutions to the SP and RO counterparts.
- The uncertainties associated with forecasted PV generation as well as the EVs i.e. the available energy at arrival and departure times, and the maximum and minimum available energy and power for EVs were considered.
- The V2G capability of EVs is addressed and the in-sample and out-of-sample performance of DRO and SP problems are compared.

The rest of the paper is organized as follows. In section 2, a two-stage data-driven DRO problem is formulated using the constructed ambiguity set. The ambiguity set is formed using the empirical data and the Wasserstein metric is used as a distance measure. In section 3, the solution framework for the proposed problem is presented. C&CG technique is used to solve this problem. The numerical results and conclusion are presented in sections 4 and 5 respectively.

2. Problem Formulation

In this section, first, the formed ambiguity set for the probability distribution of uncertain variables is presented and later, the data-driven DRO problem is formulated using the procured ambiguity set.

2.1. Forming the ambiguity set

The probability distributions of the forecasted PV generation, the maximum and minimum capacity of EV clusters, the maximum power of EV clusters, and the available energy at arrival and departure times are considered to be uncertain. The historical data of PV generation and EV clusters' characteristics are used to construct an empirical probability distribution. Using the empirical data on the maximum and minimum available energy ($\hat{E}^{max}_{e,k,t}$, $\hat{E}^{min}_{e,k,t}$), maximum power consumption ($\hat{p}^{max}_{e,k,t}$), and the available energy at arrival and departure times ($\hat{E}^{arv}_{e,k,t}$, $\hat{E}^{dep}_{e,k,t}$) and the forecasted PV generation ($\hat{p}_{v,k,t}$), the empirical probability distribution for each variable is formed. For a set of empirical data of size N, the probability distribution of empirical data is $\hat{\Lambda} = (\hat{\theta}_1, ..., \hat{\theta}_n)$ where $\hat{\theta}_n = 1/N$. Based on the empirical probability distributions the unknown probability distributions of uncertain variables associated with PVs and EV clusters are procured with the confidence level β. Here, the Wasserstein metric shown in (1), is used to quantify the distance $\mathcal{D}_v(\Lambda, \hat{\Lambda})$ between the unknown probability distribution and the empirical probability distribution.

$$\mathcal{D}_v(\Lambda, \hat{\Lambda}) := \inf_{\pi} \sum_{|v|} \sum_{|\hat{v}|} \pi^{v,\hat{v}} \cdot \left\| v - \hat{v} \right\|_1 \tag{1}$$

Here, v and $\hat{v}$ are the vectors of the uncertain variables with unknown and empirical probability distributions, respectively. Furthermore, $|v|$ and $|\hat{v}|$ are the sizes of vectors v and $\hat{v}$ respectively. Furthermore, $\pi^{v,\hat{v}}$ represents the joint probability distribution of random variables v and $\hat{v}$. The ambiguity set is defined as $\mathbb{D} = \left\{ \mathcal{D}_v(\Lambda, \hat{\Lambda}) \leq \alpha \right\}$ where α is the maximum distance between the two probability distributions defined as $\alpha = D\sqrt{\frac{2}{N} \ln \frac{1}{(1-\beta)}}$ [20]. The ambiguity set is formed by (2)-(5) where $m = 1, ..., N$ and $n = 1, ..., N$ are the indices of the samples from unknown and empirical probability distributions, and $\pi^{m,n}$ is the joint probability of θ^m (unknown marginal probability of sample m) and $\hat{\theta}^n$ (marginal probability of empirical data sample n). Here, the constraints (3) and (4) are enforcing the relationship between the joint probability distribution $\pi^{n,m}$ and the marginal probability distributions of unknown and empirical data samples. Constraint (5) shows the property of the unknown probability distribution.

$$\sum_{n} \sum_{m} \pi^{m,n} \left| v^m - v^n \right| \leq \alpha \tag{2}$$

$$\sum_{n} \pi^{m,n} = \theta^{m}, \qquad \forall m \tag{3}$$

$$\sum_{m} \pi^{m,n} = \hat{\theta}^{n}, \qquad \forall n \tag{4}$$

$$\sum_{m} \theta^{m} = 1 \tag{5}$$

2.2. Formulation of Data-driven DRO Problem

In this section, a two-stage data-driven optimization problem is formulated for the unbalanced operation of the distribution network. In the first stage, the topology of the distribution network is determined by procuring the state of the switches in the network. In the second-stage problem, the worst-case realization of the probability distribution of the uncertain variables is determined within a bounded Wasserstein distance from the empirical probability distribution. The rest of the decision variables include the real and reactive power dispatch of DERs and the distribution feeder.

The states of the switches in the first-stage problem maintain the radial topology of the distribution network. Here, the radial distribution system is represented as a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ denotes the set of vertices (buses) with K vertices and $\mathbf{E}$ denotes the set of edges (lines) with M edges where $K - 1 = M$. The spanning trees of the distribution network include all possible radial topologies. The following algorithm is developed to find the switching state of the distribution lines that ensures the radial operation of the distribution network:

Step 1: In a distribution network graph G with K vertices and M edges, find all combinations of edges with $K - 1$ number of edges.

Step 2: Check if all K vertices are present in the selected combinations in Step 1. If so, return these combinations as spanning trees, otherwise, discard it.

Step 3: The spanning trees form a set S in which in each spanning tree $s \in S$, the state of a switchable line that is open i.e., $l \in \mathbb{O}_s$ is 0, i.e., $y_{l,t} = 0$. Similarly, the state of the switchable line that is closed i.e., $l \in \mathbb{C}_s$ is 1, i.e., $y_{l,t} = 1$. As one of the spanning trees should be selected, constraint (8) is enforced. The selected spanning tree is formed by a certain combination of closed and open switches as shown in (8).

$$\begin{aligned}\min_{y_{l,t}} \mathbf{0} + \max_{\Lambda \in \mathbb{D}} \min_{\mathbf{p},\mathbf{q},\mathbf{U}} E_{\theta}\Bigg(& \sum_{f}\sum_{k}\sum_{t} \rho_{f,t} \cdot p^{m}_{f,k,t} + \\ & \sum_{b}\sum_{k}\sum_{t}\sum_{g}\sum_{i} AI_{i,b} \cdot w_{i,g} \cdot ps^{m}_{i,k,g,t} + \\ & \sum_{b}\sum_{k}\sum_{t} cc \cdot AD_{d,b} \cdot \left(\overline{p}_{d,k,t} - p^{m}_{d,k,t}\right)\Bigg)\end{aligned} \tag{6}$$

Subject to: (2)-(5),

$$\sum_{l=1}^{K-1} y_{l,t} = K - 1 \tag{7}$$

$$\sum_{s \in \mathbb{S}} \left(\prod_{l \in \mathbb{C}_s} y_{l,t} \cdot \prod_{l \in \mathbb{O}_s} \left(1 - y_{l,t}\right) \right) = 1 \tag{8}$$

$$\begin{aligned}\sum_{l} AL_{l,b} \cdot p^{m}_{l,k,t} + \sum_{i} AI_{i,b} \cdot p^{m}_{i,k,t} + \sum_{v} AV_{v,b} \cdot \hat{p}^{m}_{v,k,t} + \\ \sum_{f} AN_{f,b} \cdot p^{m}_{f,k,t} + \sum_{e} AE_{e,b} \cdot (p^{dc,m}_{e,k,t} - p^{ch,m}_{e,k,t}) \\ = \sum_{d} AD_{d,b} \cdot p^{m}_{d,k,t}\end{aligned} \tag{9}$$

$$\begin{aligned}\sum_{l} AL_{l,b} \cdot q^{m}_{l,k,t} + \sum_{i} AI_{i,b} \cdot q^{m}_{i,k,t} + \sum_{f} AN_{f,b} \cdot q^{m}_{f,k,t} \\ = \sum_{d} AD_{d,b} \cdot q^{m}_{d,k,t}\end{aligned} \tag{10}$$

$$ps^{m}_{i,k,g,t} \leq \overline{p}_{i,k,g} \tag{11}$$

$$p^{m}_{i,k,t} = \sum_{g} ps^{m}_{i,k,g,t} \tag{12}$$

$$AL_{l,b} \cdot \mathbf{U}^{m}_{b,t} + 2(\widetilde{\mathbf{R}}_{l} \cdot \mathbf{p}^{m}_{l,t} + \widetilde{\mathbf{X}}_{l} \cdot \mathbf{q}^{m}_{l,t}) \leq M \cdot (1 - \boldsymbol{\phi}_{l}) \tag{13}$$

$$AL_{l,b} \cdot \mathbf{U}^{m}_{b,t} + 2(\widetilde{\mathbf{R}}_{l} \cdot \mathbf{p}^{m}_{l,t} + \widetilde{\mathbf{X}}_{l} \cdot \mathbf{q}^{m}_{l,t}) \geq -M \cdot (1 - \boldsymbol{\phi}_{l}) \tag{14}$$

$$\underline{V} \leq U^{m}_{b,k,t} \leq \overline{V} \tag{15}$$

$$p^{m}_{l,k,t} \geq -\phi_{l,k} \cdot y_{l,t} \cdot \bar{S}_{l,k} \cdot \cos(\frac{\pi}{8}) \tag{16}$$

$$p^{m}_{l,k,t} \leq \phi_{l,k} \cdot y_{l,t} \cdot \bar{S}_{l,k} \cdot \cos(\frac{\pi}{8}) \tag{17}$$

$$q^{m}_{l,k,t} \geq -\phi_{l,k} \cdot y_{l,t} \cdot \bar{S}_{l,k} \cdot \cos(\frac{\pi}{8}) \tag{18}$$

$$q^{m}_{l,k,t} \leq \phi_{l,k} \cdot y_{l,t} \cdot \bar{S}_{l,k} \cdot \cos(\frac{\pi}{8}) \tag{19}$$

$$p^{m}_{l,k,t} + q^{m}_{l,k,t} \geq -\phi_{l,k} \cdot y_{l,t} \cdot \bar{S}_{l,k} \cdot (\sin(\frac{\pi}{8}) + \cos(\frac{\pi}{8})) \tag{20}$$

$$p^{m}_{l,k,t} + q^{m}_{l,k,t} \leq \phi_{l,k} \cdot y_{l,t} \cdot \bar{S}_{l,k} \cdot (\sin(\frac{\pi}{8}) + \cos(\frac{\pi}{8})) \tag{21}$$

$$p^{m}_{l,k,t} - q^{m}_{l,k,t} \geq -\phi_{l,k} \cdot y_{l,t} \cdot \bar{S}_{l,k} \cdot (\sin(\frac{\pi}{8}) + \cos(\frac{\pi}{8})) \tag{22}$$

$$p^{m}_{l,k,t} - q^{m}_{l,k,t} \leq \phi_{l,k} \cdot y_{l,t} \cdot \bar{S}_{l,k} \cdot (\sin(\frac{\pi}{8}) + \cos(\frac{\pi}{8})) \tag{23}$$

$$0 \leq p^{m}_{i,k,t} \leq \bar{p}_{i,k} \tag{24}$$

$$-\bar{q}_{i,k} \leq q^{m}_{i,k,t} \leq \bar{q}_{i,k} \tag{25}$$

$$p^{m}_{d,k,t} \leq \bar{p}_{d,k,t} \tag{26}$$

$$q^{m}_{f,k,t} \leq tan\left(\cos^{-1} PF_f\right) \cdot p^{m}_{f,k,t} \tag{27}$$

$$q^{m}_{f,k,t} \geq -tan\left(\cos^{-1} PF_f\right) \cdot p^{m}_{f,k,t} \tag{28}$$

$$\hat{E}^{min,m}_{e,k,t} - M \cdot (1 - I_{e,k,t}) \leq E^{m}_{e,k,t} \tag{29}$$

$$E^{m}_{e,k,t} \leq \hat{E}^{max,m}_{e,k,t} + M \cdot (1 - I_{e,k,t}) \tag{30}$$

$$E^{m}_{e,k,t} \leq \hat{E}^{arv,m}_{e,k} + M \cdot \left(1 - I^{arv}_{e,k,t}\right) \tag{31}$$

$$E^{m}_{e,k,t} \geq \hat{E}^{arv,m}_{e,k} - M \cdot \left(1 - I^{arv}_{e,k,t}\right) \tag{32}$$

$$E^{m}_{e,k,t} \leq \hat{E}^{dep,m}_{e,k} + M \cdot \left(1 - I^{dep}_{e,k,t}\right) \tag{33}$$

$$E^{m}_{e,k,t} \geq \hat{E}^{dep,m}_{e,k} - M \cdot \left(1 - I^{dep}_{e,k,t}\right) \tag{34}$$

$$E^{m}_{e,k,t} - E^{m}_{e,k,t-1} - \eta^{ch}_{e} p^{ch,m}_{e,k,t} + p^{dc,m}_{e,k,t}/\eta^{dc}_{e} \leq M \cdot \left(1 - I_{e,k,t}\right) \tag{35}$$

$$E^{m}_{e,k,t} - E^{m}_{e,k,t-1} - \eta^{ch}_{e} p^{ch,m}_{e,k,t} + p^{dc,m}_{e,k,t}/\eta^{dc}_{e} \geq -M \cdot \left(1 - I_{e,k,t}\right) \tag{36}$$

$$0 \leq p^{ch,m}_{e,k,t} \leq I_{e,k,t} \cdot \hat{p}^{ch,max,m}_{e,k,t} \tag{37}$$

$$0 \leq p^{dc,m}_{e,k,t} \leq I_{e,k,t} \cdot \hat{p}^{dc,max,m}_{e,k,t} \tag{38}$$

$$Y_t \leq (y_{l,t})_{l \in \mathbb{O}_s} \tag{39}$$

$$Y_t \leq (y_{l,t})_{l \in \mathbb{C}_s} \tag{40}$$

$$Y_t \geq (y_{l,t})_{l \in \mathbb{C}_s} + (y_{l,t})_{l \in \mathbb{O}_s} - 1 \tag{41}$$

The objective function presented in (6), is the operation cost of DERs, the cost of supplying electricity from the distribution feeder and the penalty associated with the curtailed demand. To maintain the radial topology, (7) enforces the number of branches (edges) to be equal to $K - 1$ and (8) would lead to selecting one set of switchable lines that form a spanning tree in the set $\mathbb{S}$. The nonlinear terms appear in (8) are linearized using (39)-(41). Here, Y_t is binary variable, (39) and (40) ensure that Y_t is zero if $(y_{l,t})_{l \in \mathbb{O}_s}$ or $(y_{l,t})_{l \in \mathbb{C}_s}$ is zero. The last constraint (40) ensures that Y_t is 1 if both $(y_{l,t})_{l \in \mathbb{O}_s}$ and $(y_{l,t})_{l \in \mathbb{O}_s}$ are 1.

Here, (9) and (10) show the nodal real and reactive power balance in the distribution network. The operation cost of the distributed generation (DG) unit is formulated as a quadratic function of the output power which is further linearized using a piece-wise linearization technique in (6), (11) and (12). Constraints (13) and (14), present the real and reactive power flow in distribution branch l, where, $\phi_{l,k} \in \boldsymbol{\phi}_{\mathbf{l}}$, $U^{m}_{b,k,t} \in \mathbf{U}^{\mathbf{m}}_{\mathbf{b,t}}$. The nodal voltage is limited by lower and upper bounds as shown in (15). The voltage unbalance between the phases is small and the loss in the distribution branches is neglected. The real and reactive power flows satisfy a circular constraint where the squared real power flow and the squared reactive power flow on each branch are limited by the squared apparent power capacity of the branch. The circular constraint is further linearized by a convex polygon (16)-(23) using the methods proposed in [22, 23]. As the number of sides increases, the approximated polygon is a tighter representation of the circular constraint. Here, an octagon is used where the radius of the approximated enclosing circle of the octagon is $\bar{S}_{l,k} = S^{max}_{l,k} \cdot \sqrt{(2\pi/8)/(\sin(2\pi/8))}$. Similar constraints could be written for the distribution feeder given the apparent power capacity of the feeder. Constraints (24) and (25) are limiting the real and reactive power for a DG unit by its capacity. The served demand is less than the total demand as enforced by (26). For distribution feeder f, the reactive power on each phase is limited by the acceptable power factor of the feeder as shown in (27)-(28). The limits on the available energy in an EV cluster e are shown in (29) and (30). As shown in these constraints, the limits on the available energy are enforced once the vehicle clusters are connected to the network, i.e. $I_{e,k,t} = 1$. The available energy at arrival time is limited by (31) and (32), where these constraints are enforced once the vehicle clusters are connected to the network at the arrival time, i.e. $I^{arv}_{e,k,t} = 1$. Similarly, the available energy at departure time is limited by (33) and (34), where (33) and (34) are enforced once the vehicle clusters are connected to the network at the departure time, i.e. $I^{dep}_{e,k,t} = 1$. Here, $I_{e,k,t} = 1$ for all hours between the arrival and departure times. The relationship between the available energy and the power dispatch of each EV cluster is enforced by (35) and (36). Constraint (37) and (38) enforce the limitation of charging and discharging power for each EV cluster. The data-driven DRO problem is formulated as (6)-(38), (2)-(5) where the object function is reformulated as (42).

$$\max_{\Lambda \in \mathbb{D}} \min_{\mathbf{p,q,U}} \sum_{m} \theta^{m} \cdot \Bigg(\sum_{f} \sum_{k} \sum_{t} \rho_{f,t} \cdot p^{m}_{f,k,t} + \sum_{b} \sum_{k} \sum_{t} \sum_{g} \sum_{i} AI_{i,b} \cdot w_{i,g} \cdot ps^{m}_{i,g,k,t} + \sum_{b} \sum_{k} \sum_{t} cc \cdot AD_{d,b} \cdot \left(\bar{p}_{d,k,t} - p^{m}_{d,k,t}\right) \Bigg) \tag{42}$$

3. Solution Methodology

To solve the problem presented in (6)-(38), (2)-(5), the following solution algorithm based on C&CG technique is presented. Here, the abstract form of the problem is formu-

lated as (43)-(47).

$$\min_{\boldsymbol{y}} \ \mathbf{0} + \max_{\Lambda \in \mathbb{D}} \min_{\boldsymbol{x}} \ \boldsymbol{b}^{\top}\boldsymbol{x} \tag{43}$$

subject to:

$$\boldsymbol{A}\boldsymbol{y} \leq \boldsymbol{d} \tag{44}$$

$$\boldsymbol{F}\boldsymbol{x} - \boldsymbol{E}\boldsymbol{y} \leq \boldsymbol{k} \tag{45}$$

$$\boldsymbol{G}\boldsymbol{x} - \boldsymbol{K}\boldsymbol{y} = \boldsymbol{g} \tag{46}$$

$$\boldsymbol{J}\boldsymbol{x} = \boldsymbol{h} \tag{47}$$

The proposed algorithm is presented in the following steps:
Step 1) Initialization: Initialize the upper bound $UB = \infty$, lower bound $LB = -\infty$, $\epsilon \leq 10^{-3}$ and the iteration counter $w = 0$.
Step 2) Solve the master problem: Solve the master problem for each iteration as follows:

$$\min_{\boldsymbol{y}} \ \mathbf{0} + \eta \tag{48}$$

subject to:

$$\boldsymbol{A}\boldsymbol{y} \leq \boldsymbol{d} \tag{49}$$

$$\boldsymbol{\eta} \geq \boldsymbol{\theta}^{*}\mathbf{b}^{\top}\boldsymbol{x}^{(l)}, \quad \forall l \leq w \tag{50}$$

$$\boldsymbol{F}\boldsymbol{x}^{(l)} - \boldsymbol{E}\boldsymbol{y} \leq \boldsymbol{k}, \quad \forall l \leq w \tag{51}$$

$$\boldsymbol{G}\boldsymbol{x}^{(l)} - \boldsymbol{K}\boldsymbol{y} = \boldsymbol{g}, \quad \forall l \leq w \tag{52}$$

$$\boldsymbol{J}\boldsymbol{x}^{(l)} = \boldsymbol{h}, \quad \forall l \leq w \tag{53}$$

Update the LB as the following equation (54):

$$LB = \hat{\eta}_{w+1} \tag{54}$$

Step 3) Solve the subproblem: The objective function of the subproblem and it's constraints are given as follows;

$$\max_{\Lambda \in \mathbb{D}} \min_{\boldsymbol{x}} \ \boldsymbol{b}^{\top}\boldsymbol{x} \tag{55}$$

subject to:

$$\boldsymbol{F}\boldsymbol{x} \leq \boldsymbol{k} + \boldsymbol{E}\boldsymbol{y}^{*} \tag{56}$$

$$\boldsymbol{G}\boldsymbol{x} = \boldsymbol{g} + \boldsymbol{K}\boldsymbol{y}^{*} \tag{57}$$

$$\boldsymbol{J}\boldsymbol{x} = \boldsymbol{h} \tag{58}$$

constraints (2)-(5)

Then, the subproblem presented in (55)-(58) is reformulated using it's dual form in (59)-(60):

$$\boldsymbol{Q}(\boldsymbol{y}) = \max_{\Lambda \in \mathbb{D}, \gamma, \lambda, \mu} (\boldsymbol{k} + \boldsymbol{E}\boldsymbol{y}^{*})^{\top}\boldsymbol{\gamma} + (\boldsymbol{g} + \boldsymbol{K}\boldsymbol{y}^{*})^{\top}\lambda + \boldsymbol{h}^{\top}\boldsymbol{\mu} \tag{59}$$

subject to:

$$\boldsymbol{F}^{\top}\boldsymbol{\gamma} + \boldsymbol{G}^{\top}\lambda + \boldsymbol{J}^{\top}\boldsymbol{\mu} \leq \boldsymbol{b} \tag{60}$$

constraints (2)-(5)

Where, γ, λ and μ are the dual variables for constraints (56), (57) and (58) respectively. Then update the UB using (61) and go to *Step 4*.

$$UB = min\{UB, \boldsymbol{Q}(\boldsymbol{y}^{*(w+1)})\} \tag{61}$$

Step 4) Check the convergence criterion: If $UB - LB \leq \epsilon$ end the process, otherwise proceed to step 5.
step 5) Generate columns and constraints: Add (50)-(53) to the master problem, set $w = w + 1$ and go back to step 2.

4. Numerical Results

In this section, the modified IEEE 34-bus and IEEE 123-bus distribution systems are considered. The simulations are carried out on a server with dual 14 Core Intel Xeon 2.6 GHz and 380 GB of memory and CPLEX 12.8 is used as a solver. Historical sample data are considered for PV generation and EV cluster interconnections. The EV cluster data include minimum and maximum energy capacity of EVs, the available energy at arrival and departure times, and the maximum power of EV clusters. The arrival and departure times are considered as parameters. Using a Gaussian probability distribution, sample data for PV generation are generated with the mean value equal to the forecasted PV generation and the standard deviation of 0.1 of the mean value. For EVs clusters, the sample data of available energy at the arrival and departure times, the maximum power output of EV clusters as well as the maximum and minimum available energy of EV clusters are generated using a Gaussian probability distribution with a mean equal to the forecasted values and the standard deviation equal to 0.1 of the mean value. The maximum charging/discharging power for single-phase and three-phase EVs are 7.4 kW and 22.2 kW, respectively. The charging and discharging efficiencies for an EV cluster are 90%. The load curtailment cost is 20 \$/kWh. The available energy at the arrival time is 20% of the maximum energy capacity and the available energy at the departure time is 95% of the maximum energy capacity of the EV clusters. For the sake of simplicity, the arrival and departure times for all EV clusters are 9:00 and 18:00 respectively.

Table 1
Dispatchable DG units' characteristics

DG	Bus	P^{min} (kW)	P^{max} (kW)	Q^{min} (kVAR)	Q^{max} (kVAR)
1	4	0	150	-75	75
2	6	0	100	-50	50
3	24	0	50	-25	25

4.1. IEEE 34-bus system

The IEEE 34-bus distribution network supplies three-phase and single-phase loads where the peak real demands on phase A, phase B, and phase C are 606 kW, 584 kW, and 579 kW respectively. Similarly, the peak reactive demands on these phases are 357, 344, and 343 kVar respectively. Three 3-phase PV generation units are connected to the distribution network, where each has the maximum forecasted output power of 126.92 kW. The maximum total PV generation is 21.52% of the total peak demand. Fig 1 shows the diagram of the modified IEEE 34-bus distribution network and the interconnection of PVs, DGs, and EV clusters. Here, the switchable lines are shown as dashed lines. The characteristics of the dispatchable DG units are shown in Table 1. Table 2 shows the operation cost of DG1-DG3 for each segment of the linearized cost curve. The hourly electricity price is considered for the feeder. As shown in Fig. 1, four EV clusters are considered. EV cluster 1 has a three-phase interconnection, while EV clusters 2, 3, and 4 are connected to the grid on phases A, B, and C, respectively. Table 3 shows the characteristics of the EV clusters.

The following cases are considered:

Case 1 – Deterministic operation of the distribution network
Case 2 – Scenario-based stochastic operation of the distribution network
Case 3 – Robust operation of the distribution network
Case 4 – Data-driven distributionally robust operation of the distribution network

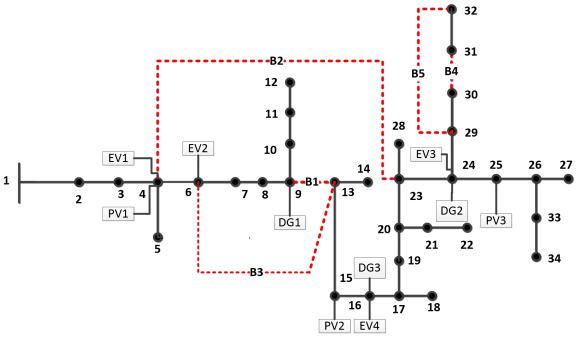


Figure 1: Modified IEEE 34-bus distribution network

4.1.1. Case 1 - Deterministic operation of the distribution network

In this case, the deterministic solutions of the distribution network operation problem with and without EV interconnections are investigated. The operation cost of the distribution network without EV interconnection is $4530.19 and the operation cost increases to $5336.82 when EV clusters are interconnected to the distribution network. Fig. 2 shows the dispatch of feeder, PVs, DGs and charging and discharging power of EVs on phase A. The total energy provided by the feeder and DGs is 11860 kWh. Tables 4 and 5 show the state of the switchable lines in the network without and with EV interconnection respectively.

Table 2
The marginal cost of each segment for DGs ($/kWh)

DG	$w_{i,1}$	$w_{i,2}$	$w_{i,3}$	$w_{i,4}$
1	0.18	0.22	0.26	0.30
2	0.17	0.24	0.31	0.38
3	0.15	0.25	0.35	0.45

Table 3
Characteristics of EV clusters in the IEEE 34-bus system

EV	Number of vehicles	Bus	P^{max} (kW)	E^{min} (kWh)	E^{max} (kWh)
1	40	4	888	360	3600
2	25	6	185	75	750
3	40	24	296	120	1200
4	30	16	222	90	900

Table 4
The status of switchable branches in Case 1 without EV interconnection

Switchable branch	Hours (24 hours)
B1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B2	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B3	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B4	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B5	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

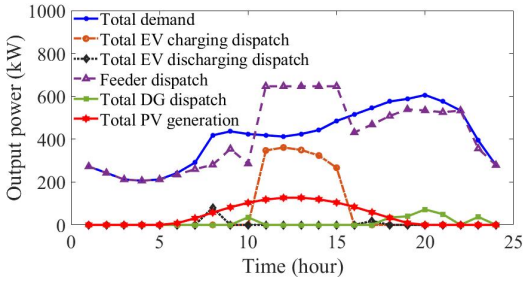


Figure 2: The dispatch of feeder, DGs, PVs, and the charging and discharging power of EVs on phase A in Case 1.

4.1.2. Case 2 – Scenario-based stochastic operation of the distribution network

In this case, the uncertainty associated with the PV generation and EV clusters are considered using 40 scenarios. The expected operation cost of the distribution network is increased to $5364.58 compared to the deterministic solution

Table 5
The status of switchable branches in Case 1 with EV interconnection

Switchable branches	Hours (24 hours)
B1	0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B2	1 1 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B3	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B4	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B5	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Table 6
The status of switchable branches in Case 2

Switchable branch	Hours (24 hours)
B1	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B2	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B3	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B4	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B5	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

as a result of the imposed uncertainties in the operation horizon. The switching states of the switchable lines are shown in Table 6. As shown in this table, the branch connecting buses 4 and 23 is closed and the branch connecting buses 9 to 13 is opened. The expected dispatch of feeder, PVs, DGs and charging and discharging power of EVs on phase A is shown in Fig. 3. Here, the total expected energy provided by the feeder and DGs is 11885 kWh which is increased compare to that in Case 1.

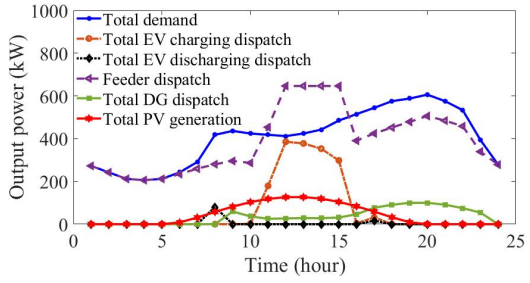


Figure 3: The expected dispatch of feeder, DGs, PVs, and charging and discharging power of EVs on phase A in Case 2.

4.1.3. Case 3 – Robust operation of the distribution network

In this section, a two-stage robust optimization problem is formulated and solved using the C&CG algorithm. The non-linear terms in the sub-problem are linearized using McCormick envelopes. The uncertainty sets for the PV generation and EV cluster characteristics are formed using the upper and lower bounds of the empirical data. Fig. 4, shows the worst-case dispatch of feeder, PVs, DGs, and charging and discharging power of EVs on phase A. In this case, the total energy output of DGs and feeder is increased to 12392 kWh, which is larger than those in Cases 1 and 2. The operation cost of the distribution network is \$11834.53 which is more than those in Cases 1 and 2. As expected, the solution to the robust optimization problem is more conservative compared to the solution to the SP and deterministic problems. The switching states of the switchable branches in the distribution system are the same as those in Case 2, except for hours 5 and 7, where branch 1 is closed and branches 2 and 3 are opened. The total demand curtailment is 348.3650 kWh which is more than those in Cases 1 and 2.

Table 7
The status of switchable branches in Case 4

Switchable branches	Hours (24 hours)
B1	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B2	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B3	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B4	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0
B5	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1

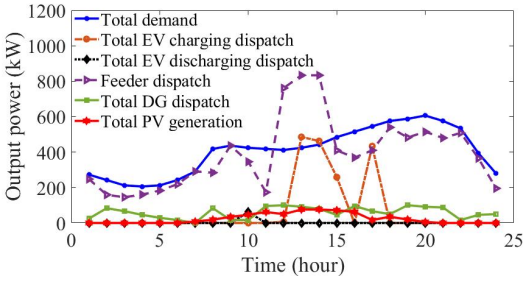


Figure 4: The worst-case dispatch of feeder, DGs, PVs, and charging and discharging of EVs in Case 3 on phase A.

4.1.4. Case 4 – Data-driven distributionally robust operation of the distribution network

In this section, the worst-case probability distributions of uncertain variables are procured given the 40 sample data for each variable. Fig. 5, shows the expected dispatch of feeder, PVs, DGs, and charging and discharging power of EVs for phase A, considering the worst-case probability of each scenario. The total expected energy output of feeder and DGs in this case is 12117 kWh which is less than that in Case 3 and more than that in Cases 1 and 2. The switching states of distribution branches are given in Table 7. Here, without switchable branches, the maximum voltage deviation during the operation horizon on phases A, B, and C, is 0.052; however, with switching capability, such deviation will decrease to 0.016. In this case, the impacts of the size of the historical data and selected confidence level, V2G capability in EV clusters, and the out-of-sample performance of DRO are investigated in the IEEE-34 distribution network.

a) The impacts of sample size and confidence level

Table 8, shows the impact of the number of samples on the total in-sample operation cost of the distribution network in this case. The results are compared to those presented in

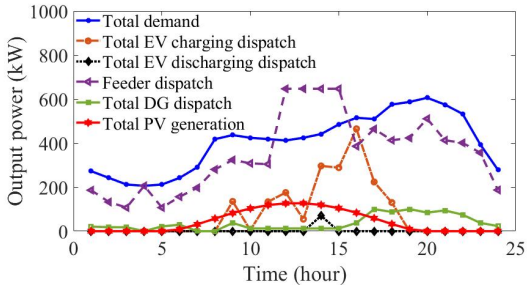


Figure 5: The expected dispatch of feeder, DGs, PVs, and charging and discharging of EVs on phase A in Case 4, considering the worst-case probability distribution of uncertain variables.

Case 2. As shown in this table the expected operation cost of the distribution network as the solution to the data-driven DRO problem, converges to the solution to the SP problem as the size of sample data increases. As shown in Table 8, the solution to the formulated DRO problem i.e., the expected operation cost of the distribution network, for different sizes of sample data, is greater than the solution to the SP problem. This shows that solving the DRO problem provides a more conservative solution compared to the SP solutions. Furthermore, it is shown in this case that the solution to the SP problem i.e., the expected operation cost of the distribution network, increases as the number of samples increases; however, the solution to the DRO problem decreases as the number of samples increases. The solution to the DRO problem is less conservative compared to the solution to the RO problem. Here, the expected operation cost of the distribution network is \$8019.10 that is increased to \$11834.53 when the distribution network operation problem is solved as a RO problem. Table 9 shows the solution time with different sizes of sample data in Cases 2 and 4. As shown in this table, the solution time for solving the DRO problem (Case 4) is higher than that for the SP problem (Case 2). Furthermore, as the number of samples increases the solution time increases.

The total demand curtailments on all phases are shown in Table 10. The total demand curtailments on phases A, B, and C are increased by 41.638 kWh, 87.714 kWh, and 27.534 kWh in this case compared to those in Case 1. Furthermore, the demand curtailments on phases A, B, and C are increased by 40.608 kWh, 84.6 kWh, and 26.80 kWh in this case compared to those in Case 2. The demand curtailments on all phases, in this case, are less than those in Case 3.

As the size of the sample data increases, the Wasserstein radius α decreases; however, the Wasserstein radius is also affected by the confidence level β. As the confidence level increases the Wasserstein radius increases. Table 11, shows the impact of the confidence level on the solution of the DRO problem (i.e., the expected operation cost of the distribution network) for 40 data samples. As the confidence level increases, the increase in the Wasserstein radius will lead to considering the probability distributions of uncertain variables that are further away from the probability distribution of the empirical data which results in the higher expected operation costs.

Table 8
The expected operation cost (\$) in Cases 2 and 4 with different sizes of sample data

Case	Size of sample data				
	20	40	60	80	100
Case 2	5544.47	5364.58	5331.19	5307.90	5320.40
Case 4	8019.10	7922.62	7915.80	7924.46	7907.78

Table 9
Solution times for Cases 2 and 4 with different sizes of sample data

Case	Size of sample data				
	20	40	60	80	100
Case 2	00:00:32	00:00:52	00:01:27	00:01:59	00:02:32
Case 4	00:03:52	01:42:58	24:42:57	31:19:39	33:47:54

Table 10
The total curtailment on each phase in all cases

Cases	Total Curtailment (kWh)		
	Phase A	Phase B	Phase C
Case 1	0	0	0
Case 2	1.03	3.114	0.731
Case 3	97.254	194.332	56.779
Case 4	41.638	87.714	27.534

Table 11
The expected operation cost in Case 4 with different confidence levels

Confidence level	Wasserstein radius	Cost (\$)
0.99	5.758	7922.62
0.95	4.644	7920.70
0.9	4.072	7914.41
0.8	3.404	7908.74
0.7	2.944	7896.95
0.6	2.568	7847.02
0.5	2.234	7815.14

b) The impacts of V2G capability in EV clusters

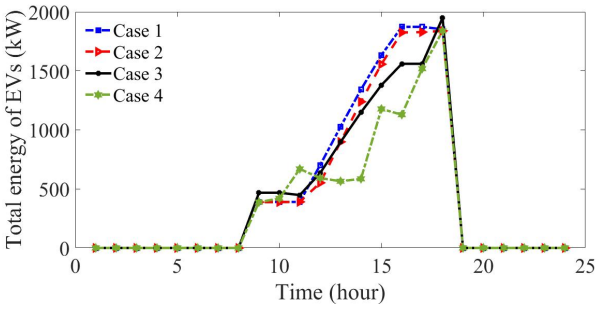
In this section the impacts of the V2G capability of the EVs on the operation cost are investigated. Table 12 presents the operation cost for all cases with and without V2G capability for EV clusters. As shown in this table, the operation cost will decrease as EVs are equipped with the V2G capability. Fig 6 shows the total available energy in EV clusters with V2G capability.

c) Out-of-sample performance

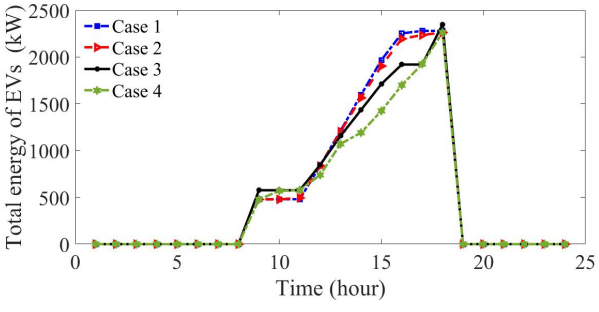
In order to evaluate the out of sample performance of SP and DRO problems, the dispatch of the distribution feeder is considered as a first-stage decision which is not dependent on samples and scenarios. The out-of-sample performance of SP and DRO solutions are shown in Table 13. Here, the out-of-sample performance of DRO solution is better than that for SP as the operation cost of the distribution network using DRO formulation is less than that using SP for out-

Table 12
Distribution network operation cost with and without the V2G capability

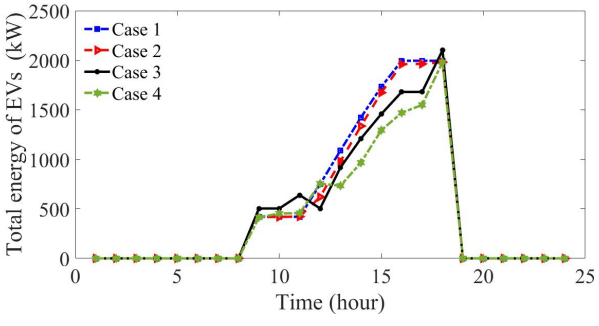
Cases	Cost ($)	
	With V2G	Without V2G
Case 1	5336.82	5378.49
Case 2	5364.58	5405.30
Case 3	10670.90	11834.53
Case 4	7922.618	7964.053



(a)



(b)



(c)

Figure 6: The total available energy in EV clusters in Cases 1 and 3 and the total expected available energy in Cases 2 and 4 connected to (a) phase A, (b) phase B and (c) phase C in

of-sample EV and PV generation data. Fig. 7 demonstrates the distribution of out-of-sample cost for 40 sample data in Cases 2 and 4. As shown in this figure, the probability distribution of out-of-sample operation cost using DRO is concentrated than the solution to the SP problem. This implies that the solution to DRO problem is more robust compared to the solution to the SP problem.

Table 13
The out-of-sample expected operation costs ($) in Cases 2 and 4 with different sizes of sample data

Cases	Size of sample data			
	40	60	80	100
Case 2	5094.03	5096.25	5090.45	5098.52
Case 4	5015.10	4914.99	5008.95	5040.27

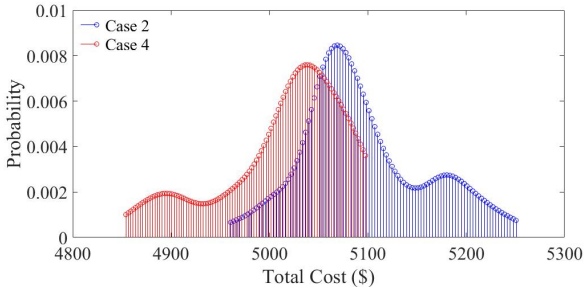


Figure 7: Out-of-sample distribution of operation cost in Case 2 and Case 4.

4.2. IEEE 123-bus system

The modified IEEE 123-bus test system with switchable lines, is shown in Fig. 8. The peak real and reactive demands on phase A are 1420 kW and 775 kVar, the peak real and reactive demands on phase B are 915 kW and 515 kVar, and the peak real and reactive demands on phase C are 1155 kW and 635 kVar, respectively. Table 14 shows the characteristics of three-phase DG units. The marginal costs at each segment of the DG units' cost curves are shown in Table 15 and the hourly price of electricity at the distribution feeder is similar to that for the IEEE 34-bus distribution system. The characteristics of the EV clusters are given in Table 16. The maximum total solar PV generation is 36.1% of the total peak demand. The total PV generation and demand profiles are shown in Fig. 9. The maximum total EV demand is 46.6% of the total peak demand. Similar to the IEEE 34-bus network four case studies are considered.

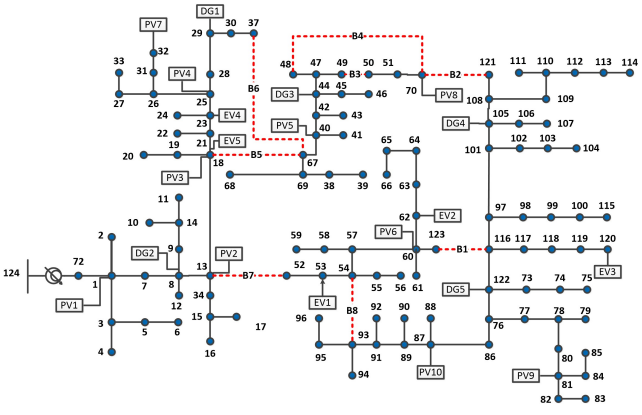


Figure 8: The modified IEEE 123-bus system.

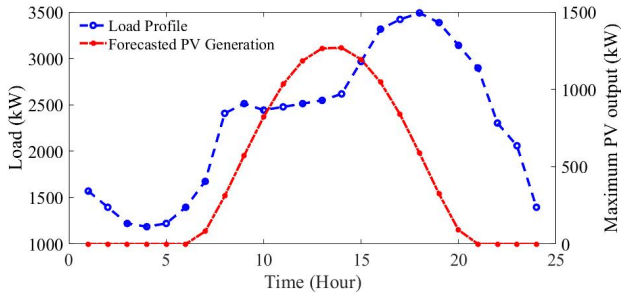


Figure 9: PV generation and demand profiles for IEEE 123-bus system

Table 14
Characteristics of DG units in the IEEE 123-bus system

DG	Bus	P^{min} (kW)	P^{max} (kW)	Q^{min} (kVAR)	Q^{max} (kVAR)
1	29	0	200	-100	100
2	8	0	200	-100	100
3	44	0	100	-60	60
4	105	0	120	-60	60

Table 15
The marginal cost of each segment of DGs in the IEEE 123-bus system ($/kWh)

DG	$w_{i,1}$	$w_{i,2}$	$w_{i,3}$	$w_{i,4}$
1	0.18	0.22	0.26	0.30
2	0.17	0.24	0.3	0.38
3	0.15	0.25	0.35	0.45
4	0.15	0.25	0.35	0.45

Table 16
Characteristics of EV fleets in IEEE 123-bus system

EV	Number of vehicles	Bus	P^{max} (kW)	E^{min} (kWh)	E^{max} (kWh)
1	75	18	555	225	2250
2	60	23	444	180	1800
3	90	62	666	270	270
4	150	53	1110	450	4500
5	60	123	444	180	1800

4.2.1. Case 1 – Deterministic operation of the distribution network

In this case, once the EV interconnection is ignored the operation cost of the distribution network is $8183.75. The interconnection of EVs to the distribution network will increase the operation cost to $9545.05. Table 17 shows the switching states of the switchable branches considering the EV interconnection.

4.2.2. Case 2 – Scenario-based stochastic operation of the distribution network

In this case, the expected operation cost for 40 scenarios is $9576.46, and the switching states of the switchable branches are shown in Table 18. Similar to the previous case study, the expected operation cost is higher compared to operation cost in Case 1 (solution to the deterministic problem).

Table 17
The status of switchable branches with EV interconnection in Case 1 for the IEEE 123-bus system

Switchable branches	Hours (24 hours)
B1	0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B2	1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B3	0 1 0 0 0 0 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1
B4	1 0 1 1 1 1 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
B5	0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0
B6	1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 1 1 1
B7	0 1 0 1 1 1 1 1 0 1 1 1 0 1 0 0 1 0 1 0 0 1 1 1
B8	1 0 1 0 0 0 0 0 1 0 0 0 1 0 1 1 0 1 0 1 1 0 0 0

Table 18
The status of switchable branches in Case 2 for the IEEE 123-bus system

Switchable branches	Hours (24 hours)
B1	0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B2	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B3	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B4	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B5	0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B6	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B7	0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B8	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Table 19
The status of switchable branches in Case 3 for the IEEE 123-bus system

Switchable branches	Hours (24 hours)
B1	0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B2	1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B3	0 1 0 0 0 0 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1
B4	1 0 1 1 1 1 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
B5	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
B6	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
B7	0 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0
B8	1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1

4.2.3. Case 3 – Robust operation of the distribution network

In the RO problem formulation, the uncertainty sets for solar PV generation and EVs are determined using the upper and lower values of the empirical data. The total operation cost is $24067.32, which is higher than the operation costs in Cases 1 and 2. The states of the switchable branches are given in Table 19.

Table 20
Total demand curtailments on all phases in the IEEE 123-bus system

Cases	Total Curtailment (kWh)		
	Phase A	Phase B	Phase C
Case 1	0	0	0
Case 2	0	0	0
Case 3	567.919	0	172.346
Case 4	388.970	14.867	11.528

Table 21
The status of the switchable branches in Case 4 for the IEEE 123-bus system

Switchable branches	Hours (24 hours)
B1	1 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1
B2	0 1 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 0 1 1 1 1 0
B3	1 1 1 0 0 1 1 1 0 0 1 0 0 1 1 1 0 0 1 1 1 0 1 1
B4	0 0 0 1 1 0 0 0 1 1 0 1 1 0 0 0 1 1 0 0 0 1 0 0
B5	0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0 1 0 0
B6	1 1 1 1 1 1 1 1 1 1 0 0 1 1 1 1 1 0 1 1 1 0 1 1
B7	1 1 1 0 1 0 1 1 1 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1
B8	0 0 0 0 0 1 0 0 0 0 1 1 1 0 1 1 1 1 0 1 1 1 1 0

4.2.4. Case 4 – Data-Driven Distributionally robust operation of the distribution network

In this case, the expected operation cost of the distribution network considering the worst-case probability distribution of the uncertain variables with 40 data samples, is \$17305.01. The expected operation cost of the distribution network in this case (the solution to the DRO problem) is higher than those in Cases 1 and 2, and less than that in Case 3. The total demand curtailments in Cases 1-4 are shown in Table 20. As shown in this table, the total demand curtailments on all phases in Cases 1 and 2 are zero. The total demand curtailment in Case 3 is the highest and the total demand curtailment in Case 4 is higher than those in Cases 1 and 2. Table 21, demonstrates the sates of the switchable branches in this case and Table 22 shows the expected operation cost of the distribution network as the number of data samples increases from 20 to 100. As shown in Table 22, with the increase in the number of samples, the value of the objective function of the DRO problem is decreasing and gets closer to that of the SP problem. Moreover, the solution to the SP problem increases with the increase in the number of sample data. However, similar to the previous case study, the value of the objective function in the DRO problem i.e., the expected operation cost of the distribution network decreases as the number of data samples increases.

Considering 40 data samples, the expected operation cost of the distribution network with V2G capability for the interconnected EVs is \$17305.01. The expected operation cost increases to \$17612.28 once the EV clusters do not have the V2G capability. The total operation cost of the distribution network with and without V2G capability of EVs in all cases are shown in Table 23.

Table 22
The expected operation cost of the IEEE 123-bus system with different numbers of empirical data samples

Cases	Size of Historical Data				
	20	40	60	80	100
Case 2	14591	14706	14675	14767	14901
Case 4	17310	17305	17298	17290	17288

Table 23
The operation cost of the IEEE 123-bus system with and without V2G capability in EVs

Cases	Distribution network operation cost (\$)	
	With V2G capability	Without V2G capability
Case 1	9545.05	9879.51
Case 2	14705.52	15000.06
Case 3	24067.32	27475.95
Case 4	17305.01	17612.28

5. Conclusion

In this paper, the data-driven distributionally robust operation of the distribution network with high penetration of PV generation and EV interconnection is evaluated. The proposed formulation captures the uncertainty in PV generation and the characteristics of the interconnected EV clusters. A solution algorithm based on the C&CG approach is used to solve the proposed formulation. It is shown that the expected operation cost of the distribution network using the proposed DRO formulation converges to the solution of the SP problem as the number of data samples increases. Moreover, increasing the confidence level will increase the Wasserstein radius and therefore, increase the operation of the distribution network. The solution to the DRO problem is compared to those of SP and RO problems. It is shown that the operation cost of the distribution network is highest once the operation problem is formulated as a RO problem. The solution to the distributionally robust operation of the distribution network is less conservative compared to the solution to the RO problem and therefore, the operation cost of the distribution network using the DRO formulation is less than that using the RO formulation. The demand curtailment of the distribution network is compared to the SP and RO solutions. It is shown that the demand curtailment is the highest, once the distribution network operation problem is formulated as a RO problem. Furthermore, the out-of-sample performance of the DRO problem is compared to that of the SP problem. It is shown that the distributionally robust operation of the distribution network would lead to a more conservative solution once it is exposed to unseen uncertainties. Therefore the out-of-sample performance of the DRO solution is better than that for the SP. The impact of V2G on the operation cost of the distribution network is further investigated. It is shown that the V2G capability of EV clusters will reduce the operation cost of the distribution network.

References

[1] IEA, Global ev outlook 2021: Accelerating ambitions despite pandemic (2021).

[2] U. EIA, Annual energy outlook 2020: with projections to 2050 (2020).

[3] M. S. Kumar, S. T. Revankar, Development scheme and key technology of an electric vehicle: An overview, Renewable and Sustainable Energy Reviews 70 (2017) 1266–1285.

[4] A. Dubey, S. Santoso, Electric vehicle charging on residential distribution systems: Impacts and mitigations, IEEE Access 3 (2015) 1871–1893.

[5] N. B. Arias, S. Hashemi, P. B. Andersen, C. Træholt, R. Romero, Distribution system services provided by electric vehicles: recent status, challenges, and future prospects, IEEE Transactions on Intelligent Transportation Systems (2019).

[6] K. Clement-Nyns, E. Haesen, J. Driesen, The impact of charging plug-in hybrid electric vehicles on a residential distribution grid, IEEE Transactions on power systems 25 (2009) 371–380.

[7] Z. Ma, N. Yang, S. Zou, Y. Shao, Charging coordination of plug-in electric vehicles in distribution networks with capacity constrained feeder lines, IEEE Transactions on Control Systems Technology 26 (2017) 1917–1924.

[8] E. Akhavan-Rezai, M. F. Shaaban, E. F. El-Saadany, F. Karray, Managing demand for plug-in electric vehicles in unbalanced lv systems with photovoltaics, IEEE Transactions on Industrial Informatics 13 (2017) 1057–1067.

[9] H. Lu, J. Hossain, M. R. Islam, L. Li, Multiobjective optimization technique for mitigating unbalance and improving voltage considering higher penetration of electric vehicles and distributed generation, IEEE Systems Journal (2020).

[10] K. Chaudhari, A. Ukil, K. N. Kumar, U. Manandhar, S. K. Kollimalla, Hybrid optimization for economic deployment of ess in pv-integrated ev charging stations, IEEE Transactions on Industrial Informatics 14 (2017) 106–116.

[11] S. Faridimehr, S. Venkatachalam, R. B. Chinnam, A stochastic programming approach for electric vehicle charging network design, IEEE Transactions on Intelligent Transportation Systems 20 (2018) 1870–1882.

[12] B. Zhang, Q. Yan, M. Kezunovic, Placement of ev charging stations integrated with pv generation and battery storage, in: 2017 Twelfth International Conference on Ecological Vehicles and Renewable Energies (EVER), IEEE, 2017, pp. 1–7.

[13] Y.-T. Liao, C.-N. Lu, Dispatch of ev charging station energy resources for sustainable mobility, IEEE Transactions on Transportation Electrification 1 (2015) 86–93.

[14] X. Bai, W. Qiao, Robust optimization for bidirectional dispatch coordination of large-scale v2g, IEEE Transactions on Smart Grid 6 (2015) 1944–1954.

[15] S. Pirouzi, J. Aghaei, M. A. Latify, G. R. Yousefi, G. Mokryani, A robust optimization approach for active and reactive power management in smart distribution networks using electric vehicles, IEEE Systems Journal 12 (2017) 2699–2710.

[16] S. Minniti, A. Haque, N. Paterakis, P. Nguyen, A hybrid robust-stochastic approach for the day-ahead scheduling of an ev aggregator, in: 2019 IEEE Milan PowerTech, IEEE, 2019, pp. 1–6.

[17] A. Akbari-Dibavar, K. Zare, S. Nojavan, A hybrid stochastic-robust optimization approach for energy storage arbitrage in day-ahead and real-time markets, Sustainable Cities and Society 49 (2019) 101600.

[18] A. Bagheri, J. Wang, C. Zhao, Data-driven stochastic transmission expansion planning, IEEE Transactions on Power Systems 32 (2016) 3461–3470.

[19] C. Zhao, Y. Guan, Data-driven stochastic unit commitment for integrating wind generation, IEEE Transactions on Power Systems 31 (2015) 2587–2596.

[20] R. Zhu, H. Wei, X. Bai, Wasserstein metric based distributionally robust approximate framework for unit commitment, IEEE Transactions on Power Systems 34 (2019) 2991–3001.

[21] R. Xie, W. Wei, M. E. Khodayar, J. Wang, S. Mei, Planning fully renewable powered charging stations on highways: A data-driven robust optimization approach, IEEE transactions on transportation electrification 4 (2018) 817–830.

[22] X. Chen, W. Wu, B. Zhang, Robust restoration method for active distribution networks, IEEE Transactions on Power Systems 31 (2015) 4005–4015.

[23] H. Ahmadi, J. R. Martı, Linear current flow equations with application to distribution systems reconfiguration, IEEE Transactions on Power Systems 30 (2014) 2073–2080.