Electrical Power and Energy Systems 117 (2020) 105701

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

A short-term decision-making model for a price-maker distribution company in wholesale and retail electricity markets considering demand response and real-time pricing

Farzad Hassanzadeh Moghimi, Taghi Barforoushi*

Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran

ARTICLE INFO

Keywords:
Bi-level model
Demand response
Energy storage system (ESS)
Real-time pricing (RTP)
Smart grid
Strategic DisCo

ABSTRACT

Adopting strategic behavior in the wholesale electricity market can revolutionize the performance of the Distribution Company (DisCo) in the operation of the distribution network. Besides, the demand response program in real-time pricing (RTP) environment, which is practicable within the smart distribution network, has a significant positive impact on the strategic behavior of this company. In this paper, a new framework is proposed to develop the sell and purchase strategies for a strategic distribution company in the energy and retail markets. The DisCo in this paper is the owner and operator of the distribution system that can affect the price of the energy due to the ownership of conventional and energy storage systems (ESSs). The uncertainty associated with the demands within the distribution network is considered by sets of scenarios. Also, the elasticity of the demands, which pertains to the retail price, is taken into account in the demand response program. Retail energy prices are also determined for customers under the RTP scheme. The problem is modeled in the form of a bi-level optimization problem, the upper-level of which includes maximizing the profit from energy sales to consumers under the RTP scheme and managing the production of distributed generation (DG) units, amount of charging or discharging of the storage system and deciding about LC program. And the lower-level is a market-clearing of the wholesale market aiming at maximizing social welfare. The upper network (sub-transmission network) is modeled in the form of a DC load flow equations to consider the impact of power transmission limits. By converting the bi-level problem to MPEC model and linearizing it using the dual theory and KKT conditions to a MILP model, finally, the MILP model is solved using the GAMS software. The performance of the proposed model is shown in two case studies, based on the IEEE-33BUS distribution network, and the 3-bus and 14-bus sub-transmission networks. Simulations investigate the impact of the strategic behavior of the DisCo on the financial and technical aspects of its operations. Simulations also indicate that the proposed model is an appropriate tool for analyzing the performance of the strategic DisCo in the wholesale and retail energy markets. Also, using the proposed approach can result in increasing of DisCo's profit, while declining in wholesale prices and load interruptions.

1. Introduction

1.1. Aim

In recent decades, due to the need for the power industry to enter the competitive environment, various structural changes have been created. The purpose of creating these structural changes was to achieve competition in production and distribution (retailing). In this regard, one of the companies that have emerged in the new competitive environment is electricity retailers. Retailers take part in the wholesale electricity market on behalf of small consumers. It acquires energy at a certain price from the wholesale markets and sells the energy to consumers by a specific pricing scheme. Also, due to the variety in the structures of competitive electricity markets, in some of the structures, operation of the distribution network, as well as, retailing can be the responsibility of distribution companies [1]. Both the role of operation and retailing, make it possible for distribution companies to use DG units and storage resources within the distribution network to meet part of the customer's consumption, reducing the retail risk caused by uncertainties [2–4].

The expansion of distribution networks coupled with the high volume of customer coverage, especially in the metropolitan areas, has led

* Corresponding author.
E-mail address: Barforoshi@nit.ac.ir (T. Barforoushi).

https://doi.org/10.1016/j.ijepes.2019.105701
Received 2 July 2019; Received in revised form 21 October 2019; Accepted 9 November 2019

Nomenclature

Indices and Sets

$h, \hat{h}$, H index and set of time periods
n, m, N index and set of buses in the upper network
$i, \hat{i}$, I index and set of buses in the distribution network
ω, Ω index and set of scenarios
q, D^{other} index and set of other demands in upper network
k, K index and set of bidding blocks
d, D index and set of DisCo in the upper network
j, J index and set of power suppliers/GENCO in the upper network
b,B index and set of offering blocks
ess, ESS index and set of the ESS
g, G index and set of DGs
t, T index and set of types of consumers
$line$, LINE index and set of distribution feeders

Parameters and Constants

P_ω probability of scenario ω
$\lambda_{i,t}^{LC}$ VOLL of consumer type t at bus i ($/MWh)
α_g, β_g ,γ_g parameters for the cost function of DG g
g_{line}, b_{line} r_{line} ,x_{line} parameters for distribution feeder $line$
S^{Grid} main grid power limit in the distribution network (MVA)
S^{line} distribution feeder capacity
$\overline{V}$ $\underline{V}$ max. and min. limits for voltage (pu)
$\overline{P^{DG}}$ $\underline{P}^{DG}$ max. and min. active power limits for DG g (KW)
$\overline{Q^{DG}}$ $\underline{Q}^{DG}$ max. and min. reactive power limits for DG g (KW)
R_g^{up} R_g^{dn} ramp rate for DG g (KW)
SUR_g SDR_g startup and shutdown ramp rate for DG g (KW)
η^{charge} $\eta^{discharge}$ charging and discharging efficiency of ESS ESS
E_{ess}^{max} E_{ess}^{min} max. and min. level of ESS ESS
$P_{ess}^{maximum\ \ charge\ \ rate}$ $P_{ess}^{maximum\ \ discharge\ \ rate}$ max. charging and discharging rate of ESS
$\overline{Q_i}$ $\underline{Q}_i$ max. and min. limits of reactive compensation on the bus I (MVAR)
$P_{i,t,\omega,h}^{D\ \ initial}$ initial demand of consumer type t at bus i in hour h in scenario ω (MW)
$Q_{i,t,\omega,h}^{D\ \ initial}$ initial demand of consumer type t at bus i in hour h in scenario ω (MVAR)
$E^t_{h,\hat{h}}$ elasticity matrix of consumer behavior
$\lambda_t^{flat\ \ tarrif}$ flat tariff of retailing energy of consumer type t ($/MWh)
$\overline{\lambda_t^{service}}$ max. allowed service price for consumer type t ($/MWh)
$\lambda_{t,\omega,h}^{average}$ max. allowed average service price offered to consumer type t in hour h in scenario ω ($/MWh)
β^{cap} price cap for Disco's bidding ($/MWh)
$\alpha_{h,b,\omega,j}^{supplier}$ price offered by supplier/GENCO j of block b in hour h in scenario ω
$\beta_{h,q,k}^{other}$ price bid by other demand q by block k in hour h
$X_{n,m}$ reactance of line between bus n and m
F_{nm} capacity of the line between bus n and m
$P_{j,b}^{max}$ max limit of block b of supplier/GENCO j (MW)
$D_{h,d,k}^{disco^{max}}$ max limit of block k of DisCo d in hour h (MW)
$D_{h,q,k}^{other}$ max cap. of block k of other demand q in hour h (MW)

Variables

$S_{g,\omega,h}$ commitment state of DG g in hour h in scenario ω
$P_{i,t,\omega,h}^{D}$ active demand of consumer type t at bus i in hour h in scenario ω after demand response (MW)
$Q_{i,t,\omega,h}^{D}$ reactive demand of consumer type t at bus i in hour h in scenario ω after demand response (MW)
$Q_{i,\omega,h}^{shunt}$ reactive power compensation at bus i in hour h scenario ω (MVAR)
$P_{line,\omega,h}^{loss}$ active loss power in feeder $line$ in hour h in scenario ω (MW)
$Q_{line,\omega,h}^{loss}$ reactive loss power in feeder $line$ in hour h in scenario ω (MVAR)
$Q_{line,\omega,h}^{feeder\ \ receive}$ $Q_{line,\omega,h}^{feeder\ \ send}$ send and receive reactive power of feeder $line$ in hour h in scenario ω (MVAR)
$P_{line,\omega,h}^{feeder\ \ receive}$ $P_{line,\omega,h}^{feeder\ \ send}$ send and receive active power of feeder $line$ in hour h in scenario ω (MW)
$V_{i,\omega,h}$ voltage amplitude at bus i in hour h in scenario ω (pu)
$P_{g,h,\omega}^{DG}$ active power generation of DG g in hour h in scenario ω (MW)
$Q_{g,\omega,h}^{DG}$ reactive power generation of DG g in hour h in scenario ω (MVAR)
$P_{\omega,h,i}^{Grid}$ active purchased power form upper network at bus i in hour h in scenario ω (MW)
$Q_{\omega,h,i}^{Grid}$ reactive purchased power form upper network at bus i in hour h in scenario ω (MW)
$P_{i,t,\omega,h}^{LC}$ active power of consumer type t reduced by LC at bus i in hour h in scenario ω (MW)
$Q_{i,t,\omega,h}^{LC}$ reactive power of consumer type t reduced by LC at bus i in hour h in scenario ω (MVAR)
$P_{i,\omega,h}^{D}$ active power of consumer type t after demand response and LC in hour h in scenario ω (MW/MVAR)
$Q_{i,\omega,h}^{D}$ reactive power of consumer type t after demand response and LC in hour h in scenario ω (MVAR)
$\lambda_{\omega,h}^{Active}$ price of purchasing energy from the upper network in hour h in scenario ω ($/MWh)
$\lambda_{t,\omega,h}^{service}$ Service price offered to consumer type t in hour h in scenario ω ($/MWh)
$\lambda_{t,\omega,h}$ retailing sell price for consumer type t in hour h in scenario ω ($/MWh)
$\gamma_{ess,h}^{ESS}$ commitment state of ESS ESS in hour h
$P_{ess,\omega,h}^{discharge}$ $P_{ess,\omega,h}^{charge}$ power charged/discharged in/from ESS ESS in hour h in scenario ω (Kw)
$E_{ess,\omega,h}^{ESS}$ ESS ESS level in hour h in scenario ω (Kw)
$\beta_{h,d,k,\omega}^{disco}$ price bid by block k of disco d in hour h in scenario ω ($/MWh)
$\theta_{h,n,\omega}$ voltage angel of bus n in hour h in scenario ω
$D_{h,q,k,\omega}^{other}$ power purchased by block k of other demand q in hour h in scenario ω (MW)
$D_{h,d,k,\omega}^{disco}$ power purchased by block k of DisCo d in hour h in scenario ω (MW)
$P_{h,j,b,\omega}^{supplier}$ power supplied by block b of supplier/GENCO j in hour h in scenario ω (MW)
$LMP_{h,n,\omega}$ locational marginal price at bus n in hour h in scenario ω ($/MWh)

to a significant increase in electricity consumption. These changes in consumption along with the DG units and ESS have turned the DisCo into a strategic player in the wholesale electricity market. It will be possible for the DisCo to act as a price-maker in the wholesale market. Besides, there are many improvements in computer and communication technologies, allow the DisCo to take advantage of the RTP pricing scheme, as well as more efficient operation and control. TOU or CCP, as pricing schemes, do not provide good control in the distribution network. On the other hand, the RTP scheme, despite its cost and complexity, can provide optimal demand-side control, and also has a positive effect on the profitability of distribution companies [3,5]. Hence, the RTP scheme would be the best choice for DisCos. Therefore, the problem lies with a strategic DisCo, as a buyer in the wholesale electricity market, determining the purchasing strategy, and selling to elastic customers at the retail level, under RTP scheme. In some researches, the DisCo's problem has been described as a single-level problem [6], however, those models do not sufficiently represent DisCos' strategic behavior. Thus, the DisCo needs a comprehensive decision-making model to determine both purchase strategy in the wholesale market, and selling strategy to consumers in the retail level, along with the optimal operation of its resources. Here, a new framework based on the leader-follower game, is proposed in the form of bi-level optimization problem. In other words, the impact of DisCo's strategic behavior on the wholesale market is modeled by a bi-level optimization problem. RTP is utilized for better demand-side control, while in many references this pricing method has not been used. Uncertainty in demands consumption is modeled through a set of scenarios. The objective of the upper-level problem is maximizing the DisCo's profit by considering the constraints of the distribution network while the lower-level problem involves maximizing social welfare by considering its constraints.

1.2. Literature review

In recent years, the performance of distribution companies in the electricity markets has attracted the attention of many researchers. In [2], the decision framework is presented in the form of a bi-level problem for analyzing the strategic behavior of the DisCo in the energy reserve markets. The upper level involves making decisions for the optimal operation of distributed energy sources (DERs) and bidding in the energy and reserve markets by the aim to minimize the operation cost. The lower level of the problem involves the market clearing in the energy and reserve markets. By neglecting the distribution network constraints, [2] focused on interactions between the DisCo and both markets above to reduce the cost of operation by changing the energy price in the market. In [7], the strategic DisCo's decision problem is presented in the form of a bi-level optimization problem in which the upper-level consist of cost minimization, while the lower-level problem, in two stages, includes market clearing and interaction with existing MGs. The distribution company also utilizes renewable generation units and uses the CVaR to manage uncertainties. Investigations are focused on the impact of MGs behavior on the performance of the DisCo and energy market. The company relies on existing aggregators to buy demand response and uses the flat tariffs scheme to sell energy to customers. In [8], the trading strategy of a proactive DisCo that has a bi-directional interaction with the real-time market is presented as a bi-level optimization problem, in which the amount of power purchased from the Day-ahead market and the price of the day-ahead market, as well as the retailing prices, are predefined. At the upper-level problem, trading strategy of the DisCo in the real-time market is considered with the objective of optimizing its profit by considering the DR program, while in the lower-level problem, the operation cost of upper-network is minimized. In [9], similar to [2,8], the problem of the DisCo is expressed as a bi-level problem. However, the upper-level problem includes the competition between distribution companies to buy energy from the day-ahead market, which is modeled using the matrix game, and at the lower-level problem, the output of DG units and the exact amount of loads are determined. In [10], a two-stage bi-level model is presented for pricing, dispatching of resources, and the purchasing of the required energy for a retailer in a smart grid. In the first stage, which consists of two levels, the retailer is modeled as the leader and the consumer as the follower to determine consumption level and the retail price. According to the results of the first stage, the retailer determines the decision to dispatch the energy storage resources as well as the amount of purchasing energy from the energy market. At this point, the retailer is participating in the market as a price taker. The bi-level model, at the first stage, is handled by the Steckelberg game and, at the second stage, is done by robust optimization method. Results show that the distribution company uses more storage resources against market price fluctuations. In [11], a bi-level trading model of DisCo is presented at the distribution network level in which the company participates in both DA and real-time markets. It is assumed that the market price and the number of purchases will be constant at the beginning of the planning period. The company's trading in the planning period includes interaction with the real-time market and with aggregators. The interaction with the aggregator is expressed as a bi-level optimization problem that aims at maximizing the profits of the proactive DisCo and maximizing the profits of the aggregators at the lower level. In [12], a bi-level model is presented to implement the trading strategy of an active DisCo. The upper-level includes maximizing the profits of the DisCo, and the lower-level, in two stages, consists of maximizing social welfare in the energy market and minimizing the operation cost in the real-time market. Despite considering the impact of DisCo on the energy market, the selling price to customers is considered as a parameter. In [13], a bi-level model is presented for expressing the strategic behavior of DisCo in the energy market. The upper-level includes maximizing the profit of DisCo, and the lower-level of which includes social welfare maximization in the energy market. The company interacts with existing aggregators to meet its required demand response. Although the strategic behavior of the distribution company is included, the distribution network is not considered. Results indicate that increasing the resources of DisCo increases its impact on the energy market.

Although many studies focused on the impact of DisCos on the market through bi-level problem, some of them, expressed the behavior of the DisCo in the form of a single-level model. In [6], a single-level non-linear model is presented for optimizing the performance of a DisCo interacting with a private MGs, aiming at minimizing the system cost (i.e. distribution network and micro-grid). Results show that optimum interaction makes the DisCo and MGs act more economic. In [14], a nonlinear stochastic model of a DisCo in the DA and RT markets is presented in the presence of storage resources and time-varying pricing scheme. The first stage involves maximizing the DisCo's profit, while the second stage involves minimizing the costs of the inaccurate estimation from the RT market. To solve the nonlinear model, the TABU search algorithm was used. Results indicate that the presence of the storage resources along with the time-varying pricing scheme, have a positive effect on the DisCos profit. Similarly, in [15], a model of the DisCo for participating in the DA market is presented in two stages, aiming at minimizing the total operation costs. In [16], a short-term two-stage decision model is presented for a DisCo, in which both DA and RT markets, as well as DG units and Load Curtailment (LC), are considered, with the objective of minimizing the operation costs. In [17], a nonlinear short-term stochastic model is presented for a DisCo in the presence of voltage-sensitive loads. The objective is maximizing the DisCo's profit, which is solved using the TABU search algorithm. In [18], the DisCo's problem is solved with the objective of maximizing the profit, along with focusing on the demand response program and the TOU pricing scheme. Results show that using the model increases profit, and reduce the risk of DisCo. It also technically corrects the voltage profile and reduces network losses. In [3], similar to [18], by focusing on the demand response program in the RTP scheme, a model has been presented to determine the optimal selling and purchasing strategies from the wholesale market for a price-taker DisCo, considering the uncertainty of elastic demands and wholesale market price. In addition, the DisCo owns a number of DG units. Results show that the presented model had a positive effect on voltage profile and network reliability. In [19], a model is presented for optimizing the bidding strategy of DisCo for participating in pool market, considering bilateral contracts and DR program. DR programs work more efficient in markets with high degree of uncertainty. RTP also increases demands' responsiveness, as well as increasing market efficiency.

The advantages of the RTP scheme are peak shaving and postponing the investments in generation expansion [5]. Some of references have focused on RTP and DR. In [20], an optimal demand management strategy is presented using a DR program and RTP scheme in a smart

Table 1
Relevant features of previous researches.

Reference	Optimization model	Role of DisCo in the market	Leader	Follower	Pricing Method for retailing level	Demand Response program	Approach	Other features
[3]	Single-level	Price-taker	–	–	RTP	RTP	MILP	–
[6]	Single-level	Price-taker	–	–	–	–	Non-Linear (ARO)	–
[14]	Single-level	Price-taker	–	–	Time varying pricing	–	Non-Linear (TABU search algorithm)	–
[15]	Single-level	Price-taker	–	–	Hourly Predefined Prices	LC	NLP	DC load flow for DN
[16]	Single-level	Price-taker	–	–	–	LC	MILP	Focused on Minimization of costs
[17]	Single-level	Price-taker	–	–	RTP	RTP/Elastic loads	Non-Linear (TABU search algorithm)	Voltage-sensitive loads are considered
[18]	Single-level	Price-taker	–	–	TOU	TOU	MILP	–
[19]	Single-level	Price-taker	–	–	Various methods	Various methods	Genetic algorithm	–
[9]	Bi-level	–	–	–	Predefined Prices	–	value iteration-based reinforcement learning algorithm	Matrix game bi-level modeling of competition of DisCos In energy market
[4]	Bi-level	Price-taker	DisCo	MGs	Flat tariff	–	hierarchical two-level algorithm	Decentralized optimization
[10]	Bi-level	Price-taker	*Retailer	Customers	RTP	RTP	RO	–
[11]	Bi-level	Price-taker	DisCo	Aggregators	–	Aggregator based DR	MPPDC (Non-Linear)	Predefined wholesale Prices and amount of purchase
[21]	Bi-level	Price-taker	*Retailer	Customers	Dynamic pricing	DR Market	MILP	DN is not considered
[24]	Bi-level	Price-taker	DisCo	MGs	Flat tariff	–	MILP	–
[25]	Bi-level	Price-taker	DisCo	MGs	–	–	MILP	–
[26]	Bi-level	Price-taker	DisCo	MGs	–	–	MILP	DN is not considered
[27]	Bi-level	Price-taker	DisCo	MGs	–	–	MILP	Reactive power in DN is not considered
[28]	Bi-level	Price-taker	Retailer	MGs	–	–	Non-Linear (MOPSO)	DN & customers are not considered
[29]	Bi-level	Price-taker	DisCo	EV Parking Lots	Various methods	Various methods	MILP	Predefined retailing Prices in all methods
[30]	Bi-level	Price-taker	LSE (load-serving entity)	Aggregators	Dynamic pricing	Aggregator based DR	MILP	–
[33]	Bi-level	Price-taker	DG owners	DisCo	–	–	NLP	DN is not considered
[2]	Bi-level	Price-maker	DisCo	DAM & RM Participants	–	LC	NLP	DN is not considered
[7]	Bi-level	Price-maker	DisCo	DAM Participants & MGs	Flat tariff	Interaction with Aggregator	MILP	No elastic loads
[8]	Bi-level	Price-maker	DisCo	RM Participants	Hourly Predefined Prices	Elastic loads	MPPDC(Non-Linear)	Predefined retailing Prices & DAMC results
[12]	Bi-level	Price-maker	DisCo	DAM Participants	Hourly Predefined Prices	–	NLP	–
[13]	Bi-level	Price-maker	DisCo	DAM Participants	Hourly Predefined Prices	Interaction with Aggregator	MILP	DN is not considered
[31]	Bi-level	Price-taker/Price-maker	Aggregator	RTM participants	–	LC/DDL/flexible loads	MILP	Price-taker in DAM Price-maker in RTM
[35]	Bi-level	Price-maker	DisCo	DAM Participants	–	LC	NLP	DN is not considered
[36]	Bi-level	Price-maker	DisCo	DAM Participants	Flat tariff	LC	NLP	
This paper	**Bi-level**	**Price-maker**	**DisCo**	**DAM Participants**	**RTP**	**RTP/Elastic demands/LC**	**MILP**	**Consideration of EES & TL limits**

grid, which considers the cost of electrical energy and the production of renewable units as well as the individual customers' consumption. Besides, the model, which is solved using genetic algorithm, optimizes the consumption pattern of household appliances and electric vehicles. In [21], the impact of a retailer on consumers' behavior in the implementation of DR program under the RTP scheme has been studied by using a bi-level model. The retailer is modeled as the leader, and consumers respond is modeled as followers. Results show that RTP scheme has a positive impact on the reduction of peak consumption level and the retailer's profit. However, consumers' payments have increased compared to flat tariffs and TOU pricing scheme. Also, network constraints were not considered. Also, the results of the demand management program in [22], a comparison between flat tariff and time-varying pricing scheme, shows that the time-varying pricing scheme will shift demand and reduce peak consumption. In [23], the scheduling model of a smart distribution network is presented, in which RTP scheme is used. Aiming at maximizing the profit, optimal retailing price, operation of DG, energy storage resources, and LC are scheduled. The IGDT bi-level method is also used to model the risk. Results show that by applying the model, demand profiles and the voltage is smoother and the peak consumption is reduced.

The DisCo within the Active Distribution Network (AND) can interact with other utilities and exchange power. Several studies discuss how DisCo interacts with existing MGs in the network. In [4], a bi-level decentralized optimization model is presented for optimal operation of an AND, which aims to maximize the DisCo' profit and the MG's profit, considering the network security constraint. The model is solved using a heuristic algorithm. In [24], a bi-level model is presented for a DisCo's interacting with MGs, in the energy and local reserve market. At the upper-level, the DisCo maximizes its profit by determining the amount of power purchased from the markets, as well as the volume and the price of energy exchanged with MGs, considering the wholesale market prices. At the lower-level, MGs decide on DG units and Interruptible Loads (ILs), according to the price signals of the DisCo in the local market, to minimize the cost of energy. The results show that the DisCo, as a strategic participant in the local market, tries to adjust the price of energy, while MGs have more participation in the reserve market. In a framework similar to the model in [24], the decision problem in [25] is expressed as a bi-level model. Results show that the resources of MG have a significant effect on local market prices. In [26], in addition to the concepts and models mentioned in [24,25], a bi-level model is presented in which two different pricing methods are suggested for interacting with MGs. In the first method, pricing is different for MGs. However, uniform pricing is considered for MGs in the second method. Results show that in the first method, the power purchased from the Disco by MG and the Disco's profit has increased, although, in the second method, the cost of MG has decreased. In [27], a two-stage bi-level model is presented for power management of the distribution network. At the upper level, the energy acquisition cost of the distribution network is minimized so that the best decision is made by forecasting energy prices in the energy market as well as wind power generation. At the lower level, the costs of MGs are minimized. However, reactive power is not considered in the distribution network. Besides, the DisCo does not directly sell energy to consumers. In [28], a bi-level model is presented for the interaction between the retailer and the existing MGs in the distribution network. The upper-level includes the profit maximization of the retailer and the lower-level consists of minimizing the operation cost of the MG. Although the retailer and the MGs are acting in the distribution network, the model has not considered the effect of distribution network. Additionally, the retailer does not sell directly to end-users, and only the interaction with the micro-network has been considered. Furthermore, the proposed model is nonlinear and solved with MOPSO algorithm. In [29], a bi-level model is presented for optimal planning of smart distribution network with EV parking lots. The company utilizes renewable energy resources and uses energy storage systems to control uncertainties. The upper-level involves reducing the cost of purchasing power from the wholesale market, and the lower-level includes the maximization of the parking lots' profit. Different types of DR programs are considered, however, the prices in the wholesale market and selling prices are considered as parameters.

In [30], a short-term decision model for a Load Serving Entity (LSE), which interacts with existing aggregators, is presented in the presence of renewable energy, traditional generation, and energy storage. A dynamic pricing scheme is used to facilitate the LSE's interaction with the demand response aggregators. The model is expressed as a bi-level optimization problem by modeling the LSE as leader, and aggregators as followers. At the upper-level problem, the LSE's profit, and at the lower level problem, the pay-off function of aggregators are maximized. According to the results, the pricing method has increased LES's profit and reduced LC actions. In [31], a bi-level model is presented for DR aggregator. The aggregator utilizes various DR sources, such as LC program, elastic demands, and DDLs (duration differentiated loads). At the upper level, the objective is to maximize the aggregator's profit. The aggregator is assumed to be price-taker in the DA market, however, it is considered as price maker in the real-time market. At lower level, the objective is minimizing the cost of energy imbalances from the energy market. In [32], a bi-level model is presented to investigate the hierarchical market structure between the distribution and transmission networks. The upper-level aims to maximize the market participant's payoff function and the lower-level minimizes the operation cost of the distribution network. However, the locational marginal prices (LMPs) are not endogenously generated. To optimize contract pricing between DISCO and DG owners, a bi-level model is presented in [33], in which the upper-level objective is to maximize the DG owners' profit while neglecting the physical constraints, the lower level considers the DISCO network constraints. To avoid the non-convexity of constraints, the paper only concerns active power and voltage magnitudes as decision variables. In [34], a single-level energy acquisition model has been presented, in which DisCo acquires the required energy by various resources, such as the wholesale market, DG units, and independent DG units in the medium voltage network, and LC actions. The proposed model aims to minimize the cost of acquiring energy. The model in [35] extended the model in [34] as a bi-level optimization problem for describing the dynamic behavior of other distribution companies. The objective of the upper-level is minimizing the cost of purchasing power and dispatch of DG units, and at the lower-level, the market-clearing problem is solved to determine the price of energy. In [36], a similar model to [35] is presented in which the network is modeled by the AC power distribution equations.

The reviewed papers have weaknesses that can affect the distribution company's decisions. In this respect, in [6,14–19], the DisCo's decision problem is expressed by a single-level model. By considering the DisCo's position in energy markets as a strategic player, these models cannot accurately reflect the behavior of a DisCo. In models presented by a number of studies (e.g. [27–29]), the behavior of DisCo has been modeled by bi-level optimization problem, but impacts on some players in the distribution network (MGs or EV parking lots, etc.) have been considered. In some researches, the pricing scheme was not used, such that customers of the distribution network were ignored (e.g. [2,25,26]), or the retailing price was predetermined (e.g. [8,12,13]). Also, in some models, the distribution network was neglected in the DisCos' decision problem. Table 1 shows the characteristics of previous research along with the specifications of this paper.

1.3. Contributions

Strategic DisCos need a comprehensive model to determine how to purchase from the wholesale market and propose the retail prices of electricity to consumers in uncertain circumstances. With the help of this model, they can project their own performances in short periods and provide a more accurate investigation into the factors influencing their behavior. According to previous studies, the behavior of the strategic distribution companies with the ownership of DGs and ESSs has not been considered. In other words, in the models presented in previous studies, the role of the resources of DisCos, such as DG units, ESSs, and the DR program, have been less highlighted. Also, the impact of an increase in the price of fuel for DGs, and transmission limits of the upper network, at which the DisCo acquire its required energy, has not been studied. Besides, in some studies such as [2,32,33] the physical constraints of the transmission lines have not been considered. In [14–19], as well as [34], the model mentioned as a single-level optimization problem, neglecting the influence of company on the energy market. In addition, in some studies, DisCo has an influence on aggregators or MGs, and interact with them for its profit or have an influence on consumers for more DR (e.g. in [4,10,11,24–26,31]). In some studies, RTP and LC program, as well as elastic demands, as alternatives for DR program, are not considered (e.g. in [3,35,36]). Therefore, in this paper, a new framework is presented for optimizing the strategy of a price-maker DisCo, for acquiring energy from the wholesale market and obtaining hourly retail prices offered to consumers, based on the RTP scheme, considering the uncertainty in customers' demand. The proposed framework is a bi-level problem. At the upper-level, the DisCo maximizes its profit and determines its strategy pertaining to the optimal operation of the distribution network and its resources. Bidding curves for the wholesale market are constructed at upper level, as well. The DisCo can influence the wholesale market and reduce the cost of energy in its favor by its optimal bidding curves. The distribution network is modeled in the form of an AC load flow equations. At the lower level, the objective is to maximize social welfare, which determines the amount of power purchased from the wholesale market by market participants and LMPs. The upper-network is modeled in the form of a DC load flow equations. The impacts of the consumers' elasticity, as well as the increase in the fuel prices on the decision of the DisCo, are also studied. The strategic DisCo can reduce the cost of energy acquisition from the market and increase its profit by adopting appropriate behavior in the wholesale market, which leads to a reduction in prices. The proposed model is formulated as an MPEC problem, which is solved after the linearization and conversion to the MILP problem, using the GAMS software. According to the contribution of this paper, the main points are as follows:

- Proposing a new decision-making model for a strategic DisCo for acquiring energy from the wholesale market and retailing in the retail market, aiming to lead to a reduction in wholesale prices and an increase in profit of strategic DisCo.
- Considering the elasticity of consumers, while ascertaining the hourly retail prices according to the RTP scheme.
- Considering Energy storage system (ESS) and the LC program, as alternatives for the DR program, which can facilitate having more impact on the energy market by DisCo.
- Investigating the impact of transmission limits of the upper network, and operation cost of DGs on DisCo's decisions.
- Converting the bi-level problem, into a MPEC model, and then to a MILP, using dual theory and KKT conditions, as well as linearization methods.

1.4. Paper organization

The rest of the paper is organized as follows. In Section 2, the proposed framework is presented. In Section 3, DisCo's decision-making model and its constraints are presented. In Section 4, a comprehensive study of this model, and the results of this study are presented and also, in Section 5, the relevant conclusions of the study are presented.

2. The proposed framework

Fig. 1 shows the proposed framework. It should be mentioned that the framework is designed to help the DisCo to optimally manage its own distributed energy resources, including DG units and ESSs, and determine the strategy for purchasing energy from the wholesale market and determining the retail prices to customers. As it is illustrated in Fig. 1, the framework is presented in the form of a bi-level problem, in which the objective of the upper-level is to maximize the DisCo's profit within the given planning period. The constraints of the upper-level problem are network constraints, constraints on DG units and ESSs, pricing constraints and DR and constraints of the binding curves to the wholesale market. The retail price for consumers, the production of DGs, charging or discharging of the ESS, the proposed bidding curves to the wholesale market are among the most important variables of the upper-level problem. Some data, shown in Fig. 1, are

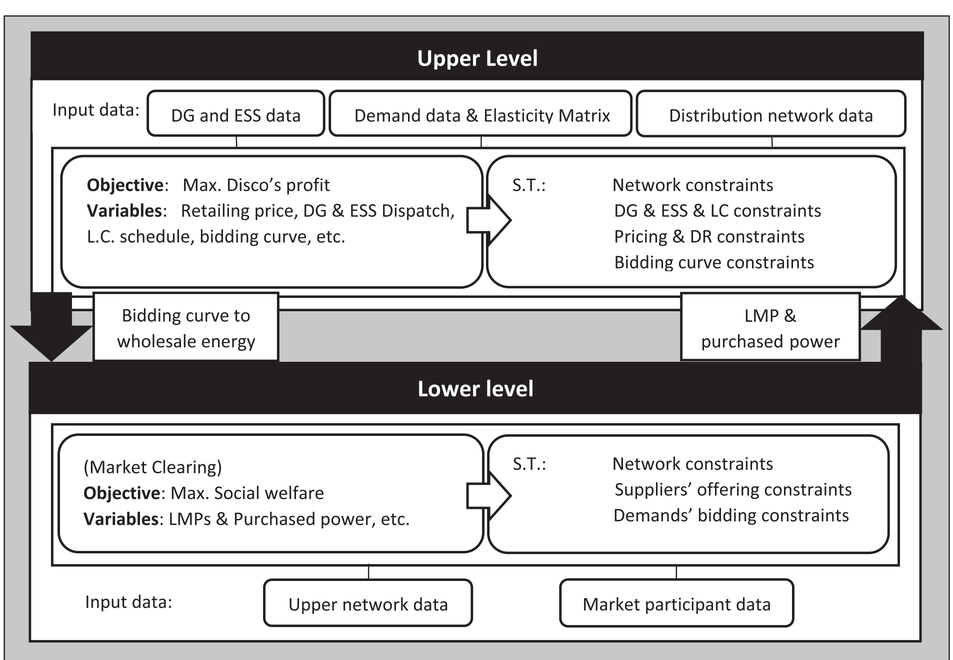

Fig. 1. The proposed bi-level framework for the interaction between the strategic DisCo and the wholesale electricity market.

needed for solving of the upper-level problem. This data includes information on DGs and ESS, data of consumers' demand and their elasticity coefficient factors and also distribution network. The retail price offered to customers, the production of DGs and the storage level of the ESS and the amount of LC program usage (if needed) are determined in the planning period. Also, the performance of the DR program, which is modeled base on the elasticity coefficient of the consumers, is determined at the upper level for each hour. The nonlinear constraints associated with equations in the AC distribution system are replaced by an appropriate approximation, and other nonlinear terms of the network constraints are approximated by piecewise approximation method. The lower-level problem, i.e. market-clearing, includes the maximization of the social welfare, to determine the cost of energy at each bus in the upper network and also, the amount of power purchased by the DisCo in each hour. The constraints of the market-clearing problem are upper network constraints, which is modeled as a DC load flow equations, and the constraints of the supply curves of suppliers as well as bidding curves of consumers. At the lower level, there is no uncertainty regarding suppliers and demands at the upper network. As it is mentioned in the next sections, the bi-level model can be converted to the MPEC model, and subsequently to the MILP model by using duality theory and KKT conditions. The proposed MILP model can be solved using the GAMS software.

3. Model formulation

3.1. Bi-level model

The short-term decision-making model of the strategic DisCo is expressed in the form of a bi-level problem, whose upper-level problem is to maximize the DisCo's profit, modeled based on formulations in [3], and the lower-level problem includes maximizing the social welfare. Equation (1) is the upper-level objective function and (43) is the lower-level objective function. Additionally, the dual variables associated with the lower-level problem, are expressed along with each constraint. The primal variables for the lower-level problem are { $\theta_{h,n,\omega}$, $D^{other}_{h,q,k,\omega}$, $D^{disco}_{h,d,k,\omega}$, $P^{supplier}_{h,j,b,\omega}$ }, while their dual variables are { $LMP_{h,n,\omega}$, $\xi_{h,n,m,\omega}$, $\delta^{1}_{h,\omega}$, :$\eta^{other\quad min}_{h,q,k,\omega}$, $\eta^{other\quad max}_{h,q,k,\omega}$, $\eta^{disco\quad min}_{h,d,k,\omega}$, $\eta^{disco\quad max}_{h,d,k,\omega}$, $\mu^{min}_{h,j,b,\omega}$, $\mu^{max}_{h,j,b,\omega}$}. The primal variables of upper-level problem are{ $\beta^{disco}_{h,d,k,\omega}$, $\gamma^{BSS}_{bss,h}$, $P^{discharge}_{bss,\omega,h}$, $P^{charge}_{bss,\omega,h}$, $E^{Bss}_{bss,\omega,h}$, $S_{g,\omega,h}$, $Q^{D}_{i,t,\omega,h}$, $P^{D}_{i,t,\omega,h}$, $Q^{shunt}_{i,\omega,h}$, $Q^{loss}_{line,w,h}$, $P^{loss}_{line,w,h}$, $Q^{feeder\quad receive}_{line,\omega,h}$, $Q^{feeder\quad send}_{line,\omega,h}$, $P^{feeder\quad receive}_{line,\omega,h}$, $P^{feeder\quad send}_{line,\omega,h}$, $V_{i,\omega,h}$, $Q^{DG}_{g,\omega,h}$, $P^{DG}_{g,h,\omega}$, $Q^{Grid}_{\omega,h,i}$, $P^{Grid}_{\omega,h,i}$, $Q^{LC}_{i,t,\omega,h}$, $P^{LC}_{i,t,\omega,h}$, $Q^{D}_{i,\omega,h}$, $P^{D}_{i,t,\omega,h}$, $\lambda^{Active}_{\omega,h}$, $\lambda^{service}_{t,\omega,h}$, $\lambda_{t,\omega,h}$ }. Note that {$\beta^{disco}_{h,d,k,\omega}$ }, which pertains to bidding decision, is a variable within the upper-level problem while it is a parameter within the lower-level problem.

3.1.1. Upper-level problem

The upper-level problem includes maximizing the DisCo's profit by considering the operating and the distribution network constraints as well as ascertaining the optimum bidding price for participation in the wholesale market, which is represented by (1)–(41).

3.1.1.1. Upper-level objective function. The objective function of the upper-level problem, which is the DisCo's profits, is represented by (1). The DisCo can procure its energy needs through its participation in the wholesale market and the purchase from upper-network or through its production as well as using the LC program and also can sell the energy to consumers within the RTP scheme. Hence, the objective function of the upper-level problem consists of four terms. The first term is the revenue from retailing of electricity to the customers in RTP environment. Regarding the consumption, the amount of energy demand can vary within the hours according to the demand response program. The company can use its LC program if needed. The second term of (1) illustrates the cost of the LC program. The third term of (1) is the cost of supplying energy from the upper-network, which is connected to the distribution network. The fourth term of (1) is the generation cost of the DG units which are owned by the DisCo. The nonlinear cost function of DGs is linearized by using the piecewise approximation method [37]. Note that because of the small size of the DG units within the distribution network, no cost has been considered as start-up and shut-down cost, as well as the cost of reactive power purchased from upper-network.

3.1.1.2. Upper-level problem constraints. ***Network Equations:*** Constraints (2)–(9) represent the power flow equations of the distribution network. Due to the nonlinearity of equations (2)–(5), the linear approximation is represented in equations (6)–(9). This linear approximation has been used with respect to the equations in [3] and the linearization methods in [38,39]. Active and reactive power balance constraints on each bus are represented by (10)–(13). It should be noted that the distribution network and upper network are connected by a substation in the bus 1 and the purchasing power can be delivered to the distribution network only from this point. Hence, the power purchased from the market is considered merely in the equations of bus 1. Equations (14) and (15) show the consumers' demand after applying DR and LC programs. Equations (16) and (17) represent the power loss within the distribution network, which are nonlinear, and can be linearized by using the methods in [3,38]. (18) and (19) represent the main substation capacity limit of the distribution system and distribution feeder capacity limits, respectively. Since both of these equations are nonlinear, they are approximated with the piecewise approximation method. (20) and (21) represent the bus voltage constraints. Also, the voltage of bus 1 in the distribution network is constant.

DG Constraints: (22) and (23) show the operation constraints for DGs operation. The binary variable $S_{g,\omega,h}$ specifies the commitment state of the DG units, which is determined by the optimality of the DisCo's problem. Constraints (24) and (25) ensure that the dynamic constraints of DG operation, such as ramp-up and ramp-down constraints, as well as startup-ramp and shut down-ramp, are met [37]. Also, for simplicity minimum up/down times constraints of DG units are neglected.

ESS Constraints: According to [17,40], constraints (26)–(29) represent the behavior of ESS. Constraint (26) shows the amount of energy stored in the ESS. At the first hour, the ESS is assumed to have a certain amount of energy stored. After the first hour, the ESS behavior is determined according to the constraints (27) and (29), which, respectively, indicate the capacity of ESS, and the charge and discharge limits, as well as controlling the charge and discharge of the ESS.

Shunt Compensators' Constraints: The compensation of the required reactive power and the voltage regulation in the distribution system, as well as the delivery of energy with proper power quality, are the responsibilities of the DisCo. Therefore, in order to cope with these responsibilities, the DisCo should use the shunt reactive power compensators within the distribution system. Constraint (30) defines the range of performance and the compensation level of these compensators.

The LC Program Constraints: Constraint (31) ensures that the amount of LC is positive and less than the customer's power consumption. Additionally, (32) ensures that the power factor will maintain at the same level after the LC program.

Demand Response: The constraints (33) and (34) express the demand response model used in this study. The demand response is modeled based on the elasticity matrix so that the loads react to price changes based on their elasticity. The elasticity matrix is a 24 × 24 matrix that includes all hours of the planning period. The elements on the main diagonal of this matrix indicate self-elasticity. Self-elasticity expresses changes in consumption at a specific hour relative to the price changes at that hour. Also, non-diagonal elements of the elasticity matrix represent the cross-elasticity, which expresses the changes in consumption at specific hour relative to the price changes at other hours. According to [41], consumers can be classified according to their behavior patterns into various models such as anticipating/postponing/flexible/inflexible/optimizing consumers, which each of them has a specific pattern of elasticity matrix.

Retailing Price Constraints: The constraint (35) illustrates the energy sale price to customers, which is derived from the sum of the cost of purchasing energy from the upper network and the service cost. The first part of this price represents the impact of the price of the wholesale market on retail prices. The second part is determined by the constraints (36) and (37) and the operating conditions.

Bidding Constraints: Constraint (38) limits the price bids to a maximum value to participate in the wholesale market. Also, the constraint (39) guarantees the downward trend of the bidding curve offered by the DisCo.

Purchasing power: Constraint (40) indicates that the amount of power purchased from the upper network is equal to the accepted amount after the market-clearing, and (41) indicates that the price of bus 1 should be considered as the price of purchasing energy for the DisCo.

3.1.2. Lower-level problem

The lower-level problem includes maximizing the social-welfare by considering the upper network constraints as well as bidding and offering constraints of the market participants [42], which are represented by (42)–(48).

3.1.2.1. Lower-level objective function. The lower-level objective function is shown by (42), which consists of three terms. The first term denotes the offering curve of suppliers in the wholesale market and the second and the third part denote the bidding curves of consumers in the market.

3.1.2.2. Lower-level problem constraints. ***Network and participants constraints:*** Constraints (43) represent the power balance for each bus of the upper network. The constraints on the supplier's offer for sale and the purchase of consumers in the upper network are shown in by (44)–(46). The power lines' limits in the upper network are shown in (47). The constraint on the voltage angle of the reference bus in the upper network is shown in (48).

$$\text{Max.} \sum_{\omega\in\Omega}\sum_{i\in I}\sum_{t\in T}\sum_{h\in H} P_\omega \lambda_{t,\omega,h} P^{D}_{i,t,\omega,h} - \sum_{\omega\in\Omega}\sum_{i\in I}\sum_{t\in T}\sum_{h\in H} P_\omega \lambda^{LC}_{i,t} P^{LC}_{i,t,\omega,h} - \sum_{\omega\in\Omega}\sum_{h\in H} P_\omega \lambda^{Active}_{\omega,h} P^{Grid}_{\omega,h,i=1} - \sum_{\omega\in\Omega}\sum_{g\in G}\sum_{h\in H} P_\omega [\alpha_g + \beta_g P^{DG}_{g,h,\omega} + \gamma_g P^{DG^2}_{g,h,\omega}] \tag{1}$$

Subject to :

$$P^{feeder\quad send}_{line,\omega,h} = g_{line} V^2{}_{i,\omega,h} - g_{line} V_{i,\omega,h} V_{\hat{A}\, i,\omega,h} \cos\left(\delta_{i,\omega,h} - \delta_{\hat{A}\, i,\omega,h}\right) - b_{line} V_{i,\omega,h} V_{\hat{A}\, i,\omega,h} \sin\left(\delta_{i,\omega,h} - \delta_{\hat{A}\, i,\omega,h}\right); \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{2}$$

$$P^{feeder\quad receive}_{line,\omega,h} = -g_{line} V^2{}_{i,\omega,h} + g_{line} V_{i,\omega,h} V_{\hat{A}\, i,\omega,h} \cos(\delta_{\hat{A}\, i,\omega,h} - \delta_{i,\omega,h}) + b_{line} V_{i,\omega,h} V_{\hat{A}\, i,\omega,h} \sin\left(\delta_{\hat{A}\, i,\omega,h} - \delta_{i,\omega,h}\right); \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{3}$$

$$Q^{feeder\quad send}_{line,\omega,h} = -b_{line} V^2{}_{i,\omega,h} - g_{line} V_{i,\omega,h} V_{\hat{A}\, i,\omega,h} \sin(\delta_{i,\omega,h} - \delta_{\hat{A}\, i,\omega,h}) + b_{line} V_{i,\omega,h} V_{\hat{A}\, i,\omega,h} \cos\left(\delta_{i,\omega,h} - \delta_{\hat{A}\, i,\omega,h}\right); \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{4}$$

$$Q^{feeder\quad receive}_{line,\omega,h} = b_{line} V^2{}_{i,\omega,h} + g_{line} V_{i,\omega,h} V_{\hat{A}\, i,\omega,h} \sin(\delta_{\hat{A}\, i,\omega,h} - \delta_{i,\omega,h}) - b_{line} V_{i,\omega,h} V_{\hat{A}\, i,\omega,h} \cos\left(\delta_{\hat{A}\, i,\omega,h} - \delta_{i,\omega,h}\right); \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{5}$$

$$P^{feeder\quad send}_{line,\omega,h} = g_{line} V_{i,\omega,h} - g_{line} V_{\hat{A}\, i,\omega,h} - b_{line} \delta_{i,\omega,h} - b_{line} \delta_{\hat{A}\, i,\omega,h} + \frac{1}{2} P^{feeder\quad loss}_{line,\omega,h}; \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{6}$$

$$P^{feeder\quad receive}_{line,\omega,h} = g_{line} V_{i,\omega,h} - g_{line} V_{\hat{A}\, i,\omega,h} - b_{line} \delta_{i,\omega,h} - b_{line} \delta_{\hat{A}\, i,\omega,h} - \frac{1}{2} P^{feeder\quad loss}_{line,\omega,h}; \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{7}$$

$$Q^{feeder\quad send}_{line,\omega,h} = -b_{line} V_{i,\omega,h} + b_{line} V_{\hat{A}\, i,\omega,h} - g_{line} \delta_{i,\omega,h} + g_{line} \delta_{\hat{A}\, i,\omega,h} + \frac{1}{2} P^{feeder\quad loss}_{line,\omega,h}; \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{8}$$

$$Q^{feeder\quad receive}_{line,\omega,h} = -b_{line} V_{i,\omega,h} + b_{line} V_{\hat{A}\, i,\omega,h} - g_{line} \delta_{i,\omega,h} + g_{line} \delta_{\hat{A}\, i,\omega,h} - \frac{1}{2} P^{feeder\quad loss}_{line,\omega,h}; \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{9}$$

$$P^{Grid}_{\omega,h,i} + \sum_{g\in G} P^{DG}_{g,\omega,h} - \sum_{line\in LINE} P^{feeder\quad send}_{line,\omega,h} + \sum_{line\in LINE} P^{feeder\quad receive}_{line,\omega,h} = P^{D}_{i,\omega,h}; \; i = 1\; \forall\, \omega \in \Omega, \forall\, h \in H \tag{10}$$

$$Q^{Grid}_{\omega,h,i} + \sum_{g\in G} Q^{DG}_{g,\omega,h} - \sum_{line\in LINE} Q^{feeder\quad send}_{line,\omega,h} + \sum_{line\in LINE} Q^{feeder\quad receive}_{line,\omega,h} = Q^{D}_{i,\omega,h}; \; i = 1\; \forall\, \omega \in \Omega, \forall\, h \in H \tag{11}$$

$$\sum_{g\in G} P^{DG}_{g,\omega,h} - \sum_{line\in LINE} P^{feeder\quad send}_{line,\omega,h} + \sum_{line\in LINE} P^{feeder\quad receive}_{line,\omega,h} + \sum_{bss\in BSS} \eta^{discharge} P^{discharge}_{bss,\omega,h} = P^{D}_{i,\omega,h} + \sum_{bss\in BSS} P^{charge}_{bss,\omega,h}; \forall\, i \neq 1, \forall\, \omega \in \Omega, \forall\, h \in H \tag{12}$$

$$\sum_{g\in G} Q^{DG}_{g,\omega,h} - \sum_{line\in LINE} Q^{feeder\quad send}_{line,\omega,h} + \sum_{line\in LINE} Q^{feeder\quad receive}_{line,\omega,h} = Q^{D}_{i,\omega,h} - Q^{shunt}_{i,\omega,h}; \forall\, i \neq 1, \forall\, \omega \in \Omega, \forall\, h \in H \tag{13}$$

$$P^{D}_{i,\omega,h} = \sum_{t\in T} (P^{D}_{i,t,\omega,h} - P^{LC}_{i,t,\omega,h}); \forall\, i \in I, \forall\, \omega \in \Omega, \forall\, h \in H \tag{14}$$

$$Q^{D}_{i,\omega,h} = \sum_{t\in T} (Q^{D}_{i,t,\omega,h} - Q^{LC}_{i,t,\omega,h}); \forall\, i \in I, \forall\, \omega \in \Omega, \forall\, h \in H \tag{15}$$

$$P^{loss}_{line,w,h} = r_{line} \frac{P^{feeder\quad send^2}_{line,\omega,h} + Q^{feeder\quad send^2}_{line,\omega,h}}{V^2}; \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{16}$$

$$Q^{loss}_{line,w,h} = x_{line} \frac{P^{feeder\quad send^2}_{line,\omega,h} + Q^{feeder\quad send^2}_{line,\omega,h}}{V^2}; \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{17}$$

$$P^{Grid^2}_{\omega,h,i} + Q^{Grid^2}_{\omega,h,i} \leq \bar{S}^{Grid}; \; i = 1\; \forall\, \omega \in \Omega, \forall\, h \in H \tag{18}$$

$$P^{feeder\quad send^2}_{line,\omega,h} + Q^{feeder\quad send^2}_{line,\omega,h} \leq \bar{S}^{line}; \forall \quad line \quad \in \quad LINE, \forall\, \omega \in \Omega, \forall\, h \in H \tag{19}$$

$$\underline{V} \leq V_{i,\omega,h} \leq \overline{V}; \forall\, i \in I, \forall\, \omega \in \Omega, \forall\, h \in H \tag{20}$$

$$V_{i,\omega,h} = Constant;\ i = 1\ \forall\ \omega \in \Omega, \forall\ h \in H \tag{21}$$

$$\underline{P}^{DG} S_{g,\omega,h} \le P^{DG}_{g,\omega,h} \le \overline{P^{DG}} S_{g,\omega,h};\ \forall\ g \in G, \forall\ \omega \in \Omega, \forall\ h \in H \tag{22}$$

$$\underline{Q}^{DG} S_{g,\omega,h} \le Q^{DG}_{g,\omega,h} \le \overline{Q^{DG}} S_{g,\omega,h};\ \forall\ g \in G, \forall\ \omega \in \Omega, \forall\ h \in H \tag{23}$$

$$P^{DG}_{g,\omega,h+1} - P^{DG}_{g,\omega,h} \le R^{up}_g S_{g,\omega,h} + SUR_g [S_{g,\omega,h+1} - S_{g,\omega,h}] + \overline{P^{DG}} [1 - S_{g,\omega,h+1}];\ \forall\ g \in G, \forall\ \omega \in \Omega, \forall\ h \in H \tag{24}$$

$$P^{DG}_{g,\omega,h} - P^{DG}_{g,\omega,h+1} \le R^{dn}_g S_{g,\omega,h+1} + SDR_g [S_{g,\omega,h} - S_{g,\omega,h+1}] + \overline{P^{DG}} [1 - S_{g,\omega,h}];\ \forall\ g \in G, \forall\ \omega \in \Omega, \forall\ h \in H \tag{25}$$

$$E^{Ess}_{ess,\omega,h} = E^{Ess}_{ess,\omega,h-1} + \eta^{charge} P^{charge}_{ess,\omega,h} - P^{discharge}_{ess,\omega,h};\ h \ne 1, \forall\ ess \in ESS, \forall\ \omega \in \Omega \tag{26}$$

$$E^{min}_{ess} \le E^{ESS}_{ess,\omega,h} \le E^{max}_{ess};\ \forall\ ess \in ESS, \forall\ \omega \in \Omega, \forall\ h \in H \tag{27}$$

$$0 \le P^{charge}_{ess,\omega,h} \le P^{maximum\ \ charge\ \ rate}_{ess} \gamma^{ESS}_{ess,h};\ \forall\ ess \in ESS, \forall\ \omega \in \Omega, \forall\ h \in H \tag{28}$$

$$0 \le P^{discharge}_{ess,\omega,h} \le P^{maximum\ \ discharge\ \ rate}_{ess} (1 - \gamma^{ESS}_{ess,h});\ \forall\ ess \in ESS, \forall\ \omega \in \Omega, \forall\ h \in H \tag{29}$$

$$\underline{Q}_i \le Q^{shunt}_{i,\omega,h} \le \overline{Q_i};\ \forall\ i \in I, \forall\ \omega \in \Omega, \forall\ h \in H \tag{30}$$

$$0 \le P^{LC}_{i,t,\omega,h} \le P^{D}_{i,t,\omega,h};\ \forall\ i \in I, \forall\ t \in T, \forall\ \omega \in \Omega, \forall\ h \in H \tag{31}$$

$$P^{LC}_{i,t,\omega,h} Q^{Dinitial}_{i,t,\omega,h} - Q^{LC}_{i,t,\omega,h} P^{Dinitial}_{i,t,\omega,h} = 0;\ \forall\ i \in I, \forall\ t \in T, \forall\ \omega \in \Omega, \forall\ h \in H \tag{32}$$

$$P^{D}_{i,t,\omega,h} = p^{Dinitial}_{i,t,\omega,h} \left(1 + \sum_{\hat{h} \in H} E^{t}_{h,\hat{h}} \frac{\lambda_{t,\omega,\hat{h}} - \lambda^{flat\ \ tarrif}_{t}}{\lambda^{flat\ \ tarrif}_{t}} \right);\ \forall\ i \in I, \forall\ t \in T, \forall\ \omega \in \Omega, \forall\ h \in H \tag{33}$$

$$Q^{D}_{i,t,\omega,h} = Q^{Dinitial}_{i,t,\omega,h} \left(1 + \sum_{\hat{h} \in H} E^{t}_{h,\hat{h}} \frac{\lambda_{t,\omega,\hat{h}} - \lambda^{flat\ \ tarrif}_{t}}{\lambda^{flat\ \ tarrif}_{t}} \right);\ \forall\ i \in I, \forall\ t \in T, \forall\ \omega \in \Omega, \forall\ h \in H \tag{34}$$

$$\lambda_{t,\omega,h} = \lambda^{Active}_{\omega,h} + \lambda^{service}_{t,\omega,h};\ \forall\ t \in T, \forall\ \omega \in \Omega, \forall\ h \in H \tag{35}$$

$$\lambda^{service}_{t,\omega,h} \le \overline{\lambda^{service}_{t}};\ \forall\ t \in T, \forall\ \omega \in \Omega, \forall\ h \in H \tag{36}$$

$$\sum_{i \in I} \sum_{h \in H} \lambda^{service}_{t,\omega,h} p^{Dinitial}_{i,t,\omega,h} \le \lambda^{average}_{t,\omega,h} \sum_{i \in I} \sum_{h \in H} p^{Dinitial}_{i,t,\omega,h};\ \forall\ i \in I, \forall\ t \in T, \forall\ \omega \in \Omega, \forall\ h \in H \tag{37}$$

$$0 \le \beta^{disco}_{h,d,k,\omega} \le \beta^{cap};\ \forall\ h \in H, \forall\ d \in D, \forall\ k \in \mathrm{K}, \forall\ \omega \in \Omega \tag{38}$$

$$\beta^{disco}_{h,d,k,\omega} \ge \beta^{disco}_{h,d,k+1,\omega};\ \forall\ h \in H, \forall\ d \in D, \forall\ k \in \mathrm{K}, \forall\ \omega \in \Omega \tag{39}$$

$$\sum_{k \in K} D^{disco}_{h,d,k,\omega} = P^{Grid}_{\omega,h,i};\ i = 1, d = 1, \forall\ h \in H, \forall\ \omega \in \Omega \tag{40}$$

$$\lambda^{Active}_{\omega,h} = LMP_{h,n,\omega};\ n = 1, \forall\ h \in H, \forall\ \omega \in \Omega \tag{41}$$

Where, $D^{disco}_{h,d,k,\omega}$ and $LMP_{h,n,\omega} \in$ arg {

$$\text{Min.} \sum_{j \in J} \sum_{b \in B} \alpha^{supplier}_{h,b,\omega,j} P^{supplier}_{h,j,b,\omega} - \sum_{d \in D^{disco}} \sum_{k \in K} \beta^{disco}_{h,d,k,\omega} D^{disco}_{h,d,k,\omega} - \sum_{q \in D^{other}} \sum_{k \in K} \beta^{other}_{h,q,k} D^{other}_{h,q,k,\omega} \tag{42}$$

Subject to :

$$\sum_{d \in D^{disco}} \sum_{k \in K} D^{disco}_{h,d,k,\omega} + \sum_{q \in D^{other}} \sum_{k \in K} D^{other}_{h,q,k,\omega} - \sum_{j \in J} \sum_{b \in B} P^{supplier}_{h,j,b,\omega} + \sum_{m \in N} \frac{1}{X_{n,m}} (\theta_{h,n,\omega} - \theta_{h,m,\omega}) = 0:\ LMP_{h,n,\omega};\ \forall\ n \tag{43}$$

$$0 \le P^{supplier}_{h,j,b,\omega} \le P^{max}_{j,b}:\ \mu^{min}_{h,j,b,\omega}, \mu^{max}_{h,j,b,\omega};\ \forall\ j \in J, \forall\ b \tag{44}$$

$$0 \le D^{disco}_{h,d,k,\omega} \le D^{disco^{max}}_{hd,k}:\ \eta^{disco\ \ min}_{h,d,k,\omega}, \eta^{disco\ \ max}_{h,d,k,\omega};\ \forall\ d \in D^{disco}, \forall\ k \tag{45}$$

$$0 \le D^{other}_{h,q,k,\omega} \le D^{other}_{hq,k}:\ \eta^{other\ \ min}_{h,q,k,\omega}, \eta^{other\ \ max}_{h,q,k,\omega};\ \forall\ q \in D^{other}, \forall\ k \tag{46}$$

$$\sum_{m \in N} \frac{1}{X_{n,m}} (\theta_{h,n,\omega} - \theta_{h,m,\omega}) \le F_{nm}:\ \xi_{h,n,m,\omega};\ \forall\ n, \forall\ m \in N \tag{47}$$

$$\theta_{h,n=1,\omega} = 0:\ \delta^{1}_{h,\omega}\};\ \forall\ h, \forall\ \omega \tag{48}$$

3.2. MPEC

One of the suitable solutions to deal with the bi-level optimization problems is to replace the lower-level problems with their Karush Kuhn-Tucker (KKT) conditions. Note that $\{\beta^{disco}_{h,d,k,\omega}\}$, which pertains to bidding decision, is a variable of the upper-level problem, while it is a parameter within the lower-level problem. Therefore, the lower-level problem ((38)–(44)) is linear and so is convex. So it can be replaced with its KKT conditions. Equality constraints including (43) and (48) are defined as $h(x) = 0$, and inequality constraints including (44)–(47) defined as $g(x) \le 0$. Then the Lagrangian function can be defined as $L = F(x)$, where $F(x)$ and x are objective function and primal variable set of the lower-level problem, respectively. So, equations, which are obtained by derivation of $F(x)$, are described by (51)–(54). Additionally, primal, dual, and complementary constraints are described by (55)–(61). This replacement renders a mathematical program with equilibrium constraints (MPEC), which is illustrated below

$$\text{Max.}\quad (1) \tag{49}$$

Subject to

$$(2 - 41), (43), (48) \tag{50}$$

$$\alpha^{supplier}_{h,b,\omega,j} - LMP_{h,n,\omega} + \mu^{max}_{h,j,b,\omega} - \mu^{min}_{h,j,b,\omega} = 0;\ \forall\ h \in H, \forall\ j \in J, \forall\ b \in B, \forall\ \omega \in \Omega \tag{51}$$

$$-\beta^{disco}_{h,d,k,\omega} + LMP_{h,n,\omega} + \eta^{disco\ \ max}_{h,d,k,\omega} - \eta^{disco\ \ min}_{h,d,k,\omega} = 0;\ \forall\ h \in H, \forall\ d \in D^{disco}, \forall\ k \in K, \forall\ \omega \in \Omega \tag{52}$$

$$-\beta^{other}_{h,q,k} + LMP_{h,n,\omega} + \eta^{other\ \ max}_{h,q,k,\omega} - \eta^{other\ \ min}_{h,q,k,\omega} = 0;\ \forall\ h \in H, \forall\ d \in D^{other}, \forall\ k \in K, \forall\ \omega \in \Omega \tag{53}$$

$$\sum_{m \in N} \frac{1}{X_{n,m}} (LMP_{h,n,\omega} - LMP_{h,m,\omega}) + \sum_{m \in N} \frac{1}{X_{n,m}} (\xi_{h,n,m,\omega} - \xi_{h,m,n,\omega}) + (\delta^{-1}_{h,\omega})_{n=1} = 0;\ \forall\ h \in H, \forall\ n \in N, \omega \in \Omega \tag{54}$$

$$0 \le P^{supplier}_{h,j,b,\omega} \perp \mu^{min}_{h,j,b,\omega} \ge 0;\ \forall\ h \in H, \forall\ j \in J, \forall\ b \in B, \forall\ \omega \in \Omega \tag{55}$$

$$0 \le (P^{max}_{j,b} - P^{supplier}_{h,j,b,\omega}) \perp \mu^{max}_{h,j,b,\omega} \ge 0;\ \forall\ h \in H, \forall\ j \in J, \forall\ b \in B, \forall\ \omega \in \Omega \tag{56}$$

$$0 \le D^{disco}_{h,d,k,\omega} \perp \eta^{discomin}_{h,d,k,\omega} \ge 0;\ \forall\ h \in H, \forall\ d \in D^{disco}, \forall\ k \in K, \forall\ \omega \in \Omega \tag{57}$$

$$0 \le (D^{disco^{max}}_{hd,k} - D^{disco}_{h,d,k,\omega}) \perp \eta^{disco\ \ max}_{h,d,k,\omega} \ge 0;\ \forall\ h \in H, \forall\ d \in D^{disco}, \forall\ k \in K, \forall\ \omega \in \Omega \tag{58}$$

$$0 \le D^{other}_{h,q,k,\omega} \perp \eta^{other\ \ min}_{h,q,k,\omega} \ge 0;\ \forall\ h \in H, \forall\ q \in D^{other}, \forall\ k \in K, \forall\ \omega \in \Omega \tag{59}$$

$$0 \le (D^{other^{max}}_{hq,k} - D^{other}_{h,q,k,\omega}) \perp \eta^{other\ \ max}_{h,q,k,\omega} \ge 0;\ \forall\ h \in H,\ \forall\ q \in D^{other},\ \forall\ k \in K,\ \forall\ \omega \in \Omega \tag{60}$$

$$0 \le \left[F_{nm} - \sum_{m \in N} \frac{1}{X_{n,m}}(\theta_{h,n,\omega} - \theta_{h,m,\omega})\right] \perp \xi_{h,n,m,\omega} \ge 0;\ \forall\ h \in H,\ \forall\ n,\ \forall\ m \in N,\ \forall\ \omega \in \Omega \tag{61}$$

In which (49) and (50) contain the upper-level problem and its constraints, as well as equation constraints in the lower-level problem.

3.2.1. MPEC linearization

In the aforementioned MPEC problem, there are two types of non-linear terms:

- Term $\sum_{k \in K} \sum_{d \in D^{disco}} D^{disco}_{h,d,k,\omega} LMP_{h,n,\omega}$ in the (1), which is equal with the term $\lambda^{Active}_{\omega,h} P^{Grid}_{\omega,h,i}$ according to the constraints (40) and (41), is nonlinear, and should be linearized. Therefore, this term can be linearized, and also can be replaced with the nonlinear term of the objective function, by using strong duality theorem, and KKT conditions ([43,44]). The strong duality theorem says that if a problem is convex (continuous and linear), the objective functions of the primal and dual problems have the same value at the optimum. Thus by applying the strong duality theorem to the lower-level problem, we have :

$$\sum_{j \in J} \sum_{b \in B} \alpha^{supplier}_{h,b,\omega,j} P^{supplier}_{h,j,b,\omega} - \sum_{d \in D^{disco}} \sum_{k \in K} \beta^{disco}_{h,d,k,\omega} D^{disco}_{h,d,k,\omega} - \sum_{q \in D^{other}} \sum_{k \in K} \beta^{other}_{h,q,k} D^{other}_{h,q,k,\omega} = - \sum_{d \in D^{disco}} \sum_{k \in K} \eta^{disco\ \ max}_{h,d,k,\omega} D^{disco^{max}}_{hd,k} + \sum_{j \in J} \sum_{b \in B} \mu^{max}_{h,j,b,\omega} P^{max}_{j,b} - \sum_{q \in D^{other}} \sum_{k \in K} \eta^{other\ \ max}_{h,q,k,\omega} D^{other^{max}}_{hq,k} + \sum_{n} \sum_{m \in N} \xi_{h,n,m,\omega} F_{nm} \tag{62}$$

To obtain $D^{disco}_{h,d,k,\omega} LMP_{h,n,\omega}$, (58) can be used. Hence :

$$\eta^{disco\ \ max}_{h,d,k,\omega} (D^{disco}_{h,d,k,\omega} - D^{disco^{max}}_{hd,k}) = 0 \rightarrow \eta^{disco\ \ max}_{h,d,k,\omega} (D^{disco}_{h,d,k,\omega}) = \eta^{disco\ \ max}_{h,d,k,\omega} (D^{disco^{max}}_{hd,k}) \tag{63}$$

Then, By substituting (63) in (62), we will have:

$$\sum_{j \in J} \sum_{b \in B} \alpha^{supplier}_{h,b,\omega,j} P^{supplier}_{h,j,b,\omega} + \sum_{d \in D^{disco}} \sum_{k \in K} (\eta^{disco\ \ max}_{h,d,k,\omega} - \beta^{disco}_{h,d,k,\omega}) D^{disco}_{h,d,k,\omega} - \sum_{q \in D^{other}} \sum_{k \in K} \beta^{other}_{h,q,k} D^{other}_{h,q,k,\omega} = + \sum_{j \in J} \sum_{b \in B} \mu^{max}_{h,j,b,\omega} P^{max}_{j,b} - \sum_{q \in D^{other}} \sum_{k \in K} \eta^{other\ \ max}_{h,q,k,\omega} D^{other^{max}}_{hq,k} + \sum_{n} \sum_{m \in N} \xi_{h,n,m,\omega} F_{nm} \tag{64}$$

Also, we have from (52) :

$$-\beta^{disco}_{h,d,k,\omega} + LMP_{h,n,\omega} + \eta^{disco\ \ max}_{h,d,k,\omega} - \eta^{disco\ \ min}_{h,d,k,\omega} = 0$$

By substituting the above equation in (64) we will have:

$$\sum_{j \in J} \sum_{b \in B} \alpha^{supplier}_{h,b,\omega,j} P^{supplier}_{h,j,b,\omega} + \sum_{d \in D^{disco}} \sum_{k \in K} (\eta^{disco\ \ min}_{h,d,k,\omega} - LMP_{h,n,\omega}) D^{disco}_{h,d,k,\omega} - \sum_{q \in D^{other}} \sum_{k \in K} \beta^{other}_{h,q,k} D^{other}_{h,q,k,\omega} = + \sum_{j \in J} \sum_{b \in B} \mu^{max}_{h,j,b,\omega} P^{max}_{j,b} - \sum_{q \in D^{other}} \sum_{k \in K} \eta^{other\ \ max}_{h,q,k,\omega} D^{other^{max}}_{hq,k} + \sum_{n} \sum_{m \in N} \xi_{h,n,m,\omega} F_{nm} \tag{65}$$

Considering (57), we have :

$$\eta^{disco\ \ min}_{h,d,k,\omega} (-D^{disco}_{h,d,k,\omega}) = 0 \tag{66}$$

Then, the converted linear term is:

$$\sum_{k \in K} \sum_{d \in D^{disco}} D^{disco}_{h,d,k,\omega} LMP_{h,n,\omega} = \sum_{j \in J} \sum_{b \in B} \alpha^{supplier}_{h,b,\omega,j} P^{supplier}_{h,j,b,\omega} - \sum_{q \in D^{other}} \sum_{k \in K} \beta^{other}_{h,q,k} D^{other}_{h,q,k,\omega} + \sum_{j \in J} \sum_{b \in B} \mu^{max}_{h,j,b,\omega} P^{max}_{j,b} - \sum_{q \in D^{other}} \sum_{k \in K} \eta^{other\ \ max}_{h,q,k,\omega} D^{other^{max}}_{hq,k} + \sum_{n} \sum_{m \in N} \xi_{h,n,m,\omega} F_{nm};\ \forall\ h \in H,\ \omega \in \Omega \tag{67}$$

- Complementary constraints determine that the inner product of the primal inequality constraints and their respective dual variables is zero. Constraints (55)–(61) are nonlinear and should be linearized. So, they can be written as a series of mixed-integer linear conditions using the method in [42–44].according to the method, each complementary condition $0 \le a \perp b \ge 0$ can be converted to linear equivalent set of mixed-integer linear conditions $0 \le a$ and $0 \le b$ and $a \le uM$ and $b \le (1-u)M$, where u is an auxiliary binary variable and M is a big positive constant. Condition (68)–(75), (76)–(83), (84)–(91) and (92)–(95) pertains to constraint of suppliers ((55)–(56)), DisCo ((57)–(58)), other demands ((59)–(60)) and upper-network TL ((60)), respectively.

$$0 \le P^{supplier}_{h,j,b,\omega} \tag{68}$$

$$P^{supplier}_{h,j,b,\omega} \le \mathrm{M_1 U_{h,j,b,\omega1}} \tag{69}$$

$$0 \le \mu^{min}_{h,j,b,\omega} \tag{70}$$

$$\mu^{min}_{h,j,b,\omega} \le \mathrm{M_2(1 - U_{h,j,b,\omega1})} \tag{71}$$

$$0 \le (P^{max}_{j,b} - P^{supplier}_{h,j,b,\omega}) \tag{72}$$

$$(P^{max}_{j,b} - P^{supplier}_{h,j,b,\omega}) \le \mathrm{M_3 U_{h,j,b,\omega2}} \tag{73}$$

$$0 \le \mu^{max}_{h,j,b,\omega} \tag{74}$$

$$\mu^{max}_{h,j,b,\omega} \le \mathrm{M_4(1 - U_{h,j,b,\omega2})} \tag{75}$$

$$0 \le D^{disco}_{h,d,k,\omega} \tag{76}$$

$$D^{disco}_{h,d,k,\omega} \le \mathrm{M_5 U_{h,d,k,\omega3}} \tag{77}$$

$$0 \le \eta^{disco\ \ min}_{h,d,k,\omega} \tag{78}$$

$$\eta^{disco\ \ min}_{h,d,k,\omega} \le \mathrm{M_6(1 - U_{h,d,k,\omega3})} \tag{79}$$

$$0 \le (D^{disco^{max}}_{hd,k} - D^{disco}_{h,d,k,\omega}) \tag{80}$$

$$(D^{disco^{max}}_{hd,k} - D^{disco}_{h,d,k,\omega}) \le \mathrm{M_7 U_{h,d,k,\omega4}} \tag{81}$$

$$0 \le \eta^{disco\ \ max}_{h,d,k,\omega} \tag{82}$$

$$\eta^{disco\ \ max}_{h,d,k,\omega} \le \mathrm{M_8(1 - U_{h,d,k,\omega4})} \tag{83}$$

$$0 \le D^{other}_{h,q,k,\omega} \tag{84}$$

$$D^{other}_{h,q,k,\omega} \le \mathrm{M_9 U_{h,q,k,\omega5}} \tag{85}$$

$$0 \le \eta^{other\ \ min}_{h,q,k,\omega} \tag{86}$$

$$\eta^{other\ \ min}_{h,q,k,\omega} \le \mathrm{M_{10}(1 - U_{h,q,k,\omega5})} \tag{87}$$

$$0 \le (D^{other^{max}}_{hq,k} - D^{other}_{h,q,k,\omega}) \tag{88}$$

$$(D^{other^{max}}_{hq,k} - D^{other}_{h,q,k,\omega}) \le \mathrm{M_{11} U_{h,q,k,\omega6}} \tag{89}$$

$$0 \le \eta_{h,q,k,\omega}^{other\ \ max} \tag{90}$$

$$\eta_{h,q,k,\omega}^{other\ \ max} \le \mathrm{M}_{12}(1 - \mathrm{U}_{h,q,k,\omega 6}) \tag{91}$$

$$0 \le \left[F_{nm} - \sum_{m \in N} \frac{1}{X_{n,m}}(\theta_{h,n,\omega} - \theta_{h,m,\omega})\right] \tag{92}$$

$$\left[F_{nm} - \sum_{m \in N} \frac{1}{X_{n,m}}(\theta_{h,n,\omega} - \theta_{h,m,\omega})\right] \le \mathrm{M}_{13}\mathrm{U}_{h,n,m,\omega 7} \tag{93}$$

$$0 \le \xi_{h,n,m,\omega} \tag{94}$$

$$\xi_{h,n,m,\omega} \le \mathrm{M}_{14}(1 - \mathrm{U}_{h,n,m,\omega 7}) \tag{95}$$

$$\mathrm{U}_{h,j,b,\omega 1}, \mathrm{U}_{h,j,b,\omega 2}, \mathrm{U}_{h,d,k,\omega 3}, \mathrm{U}_{h,d,k,\omega 4}, \mathrm{U}_{h,q,k,\omega 5}, \mathrm{U}_{h,q,k,\omega 6}, \mathrm{U}_{h,n,m,\omega 7} \in \{0, 1\} \tag{96}$$

3.3. MILP

Finally, the original bi-level optimization problem can be recast by the following MILP:

$$\begin{aligned}\text{Max.} &\sum_{\omega \in \Omega} \sum_{i \in I} \sum_{t \in T} \sum_{h \in H} P_\omega \lambda_{t,\omega,h} P_{i,t,\omega,h}^{D} - \sum_{\omega \in \Omega} \sum_{i \in I} \sum_{t \in T} \sum_{h \in H} P_\omega \lambda_{i,t}^{LC} P_{i,t,\omega,h}^{LC} - \\ &\sum_{\omega \in \Omega} \sum_{h \in H} P_\omega \Big(\sum_{j \in J} \sum_{b \in B} \alpha_{h,b,\omega,j}^{supplier} P_{h,j,b,\omega}^{supplier} - \sum_{q \in D^{other}} \sum_{k \in K} \beta_{h,q,k}^{other} D_{h,q,k,\omega}^{other} \\ &+ \sum_{j \in J} \sum_{b \in B} \mu_{h,j,b,\omega}^{max} P_{j,b}^{max} - \sum_{q \in D^{other}} \sum_{k \in K} \eta_{h,q,k,\omega}^{other\ \ max} D_{hq,k}^{other^{max}} \\ &+ \sum_{n} \sum_{m \in N} \xi_{h,n,m,\omega} F_{nm} \Big) - \sum_{\omega \in \Omega} \sum_{g \in G} \sum_{h \in H} P_\omega [\alpha_g + \beta_g P_{g,h,\omega}^{DG} + \gamma_g P_{g,h,\omega}^{DG^2}]\end{aligned} \tag{97}$$

Subject to:

$$(2) - (37), (39), (44), (68) - (96) \tag{98}$$

$$\alpha_{h,b,\omega,j}^{supplier} - LMP_{h,n,\omega} + \mu_{h,j,b,\omega}^{max} - \mu_{h,j,b,\omega}^{min} = 0; \forall\, h \in H, \forall\, j \in J, \forall\, b \in B, \forall\, \omega \in \Omega \tag{99}$$

$$-\beta_{h,d,k,\omega}^{disco} + LMP_{h,n,\omega} + \eta_{h,d,k,\omega}^{disco\ \ max} - \eta_{h,d,k,\omega}^{disco\ \ min} = 0; \forall\, h \in H, \forall\, d \in D^{disco}, \forall\, k \in K, \forall\, \omega \in \Omega \tag{100}$$

$$-\beta_{h,q,k}^{other} + LMP_{h,n,\omega} + \eta_{h,q,k,\omega}^{other\ \ max} - \eta_{h,q,k,\omega}^{other\ \ min} = 0; \forall\, h \in H, \forall\, d \in D^{other}, \forall\, k \in K, \forall\, \omega \in \Omega \tag{101}$$

$$\sum_{m \in N} \frac{1}{X_{n,m}}(LMP_{h,n,\omega} - LMP_{h,m,\omega}) + \sum_{m \in N} \frac{1}{X_{n,m}}(\xi_{h,n,m,\omega} - \xi_{h,m,n,\omega}) + (\delta_{h,\omega}^{-1})_{n=1} = 0; \forall\, h \in H, \forall\, n \in N, \omega \in \Omega \tag{102}$$

4. Numerical studies

The proposed model is studied on two test systems. This section presents the results of their simulation and analysis.

4.1. Case study 1

4.1.1. Data

A system composed of the IEEE 33-BUS distribution network and the 3-bus sub-transmission network [45] is considered as case study 1. The base voltage in this network is assumed to be 12.6 kV. Additionally, three DG units and one ESS are installed in this network. Fig. 2 illustrates the single-line diagram of the network the characteristics of the DG units and the ESS are shown in Tables 2 and 3, respectively.

It is assumed that the ESS is connected to the distribution network in bus10. The uncertainty in customer consumption is illustrated by producing 1000 sets of scenarios and reducing them to 8 sets of scenarios [46]. Fig. 3 illustrates the expected load in the distribution network. The load coefficient values are derived from [3]. Self-elasticity coefficients are considered −0.2 [41], which express changes in consumption at a specific hour with respect to the price changes at that same hour.

As shown in Fig. 3, the upper network is a three-bus network connected to the distribution network at bus 1. As previously mentioned, DCLF equations are used to model the upper network. The price of active power at each bus within the upper network is obtained by the market-clearing stage. Note that no price is considered for reactive power. Also, it is assumed that the DisCo has sufficient knowledge of the energy market and other participants due to the fact that this

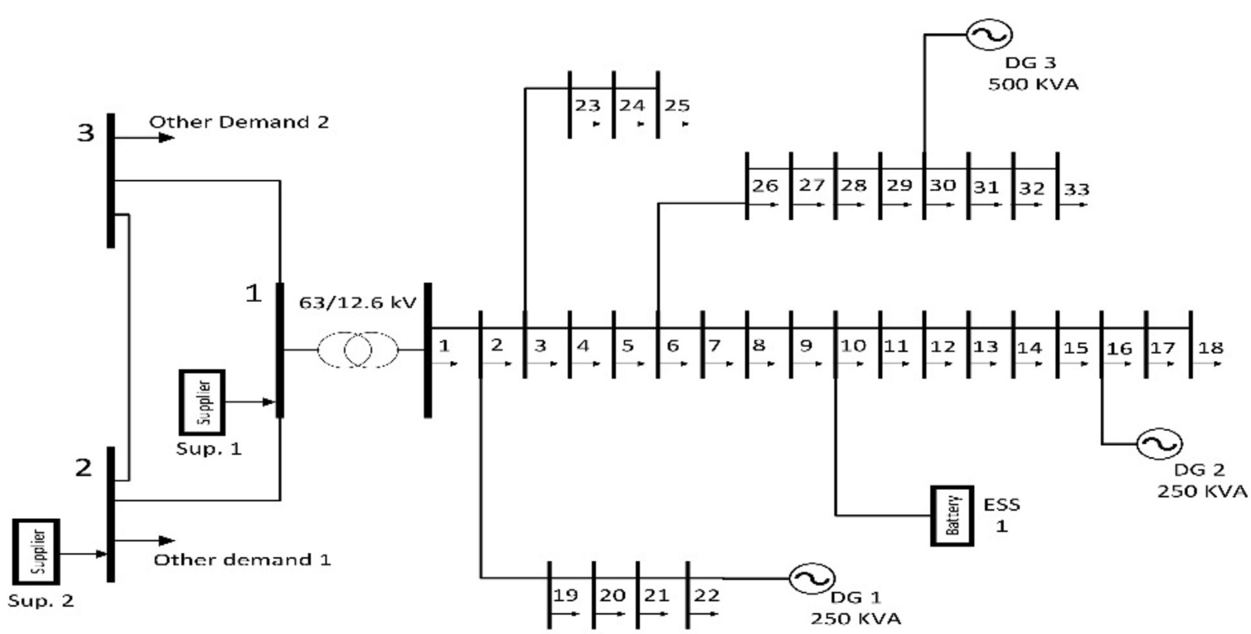

Fig. 2. IEEE 33-BUS distribution network and the 3-bus sub-transmission network.

Table 2
Characteristics of DG units.

DG number	Bus number	P^{max}(Kw)	P^{min}(Kw)	Q^{max}(Kvar)	Q^{min}(Kvar)	Ramp up(Kw/min)	Ramp down(Kw/min)
1	16	250	0	100	−100	250	250
2	22	250	0	100	−100	250	250
3	30	500	0	100	−100	500	500
DG number	SU Ramp(Kw/min)	SD Ramp(Kw/min)		Case	α	β	γ
1	250	250		**#1 & #2 & #3**	0	30	0.1
2	250	250		**#4 & #5 (costly DG)**	10	50	1
3	500	500					

Table 3
Characteristics of the ESS.

ESS Number	Bus number	P^{chmax}(Kw)	P^{dismax}(Kw)	η^{ch}	η^{dis}	Initial level (kw)	Capacity (Kw)
1	10	300	300	0.9	0.9	500	1500

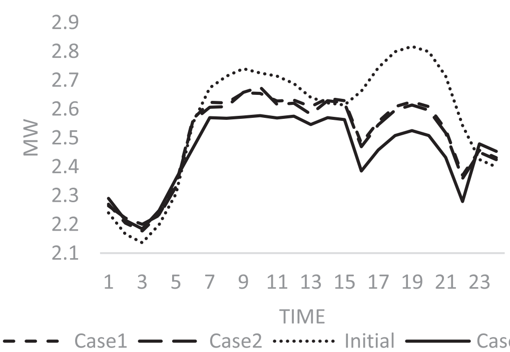

Fig. 3. Expected demand in the distribution system (case study 1).

company participated in the market for a long time and it is almost possible for it to estimate the data of other participants with the available historical data and using method mentioned in [47]. Hence, from the Disco's perspective, it is assumed that two suppliers with capacities of 3.5 and 5 megawatts are located at bus 1 and bus 2, respectively. The loads for bus 2 and bus 3 are 1.8 and 2.2 megawatts, respectively. Tables 4 and 5 illustrate the characteristics of suppliers and demands participating in the wholesale market.

It is assumed that the service cost ($\lambda_{t,\omega,h}^{service}$) should be less than 16 $/MWh, and the average offered price to customers should be less than 8 $/MWh [3]. Also, $\lambda_t^{flat\quad tarrif}$ is considered as $ 65/MWh. VOLL the distribution network is also assumed to be 1000 $/MWh.

4.1.2. Results

It is worth noting that the initial model of this paper is based on [3]. However, since data of the test system in [3] was not completely available, the simulation was performed on the IEEE-33bus for validating the proposed model. The bi-level model was then developed in this paper to study the strategic behavior of the DisCo in the wholesale market. In order to investigate the impacts of different parameters on the performance of the strategic DisCo, various study cases are defined according to Table 6. The simulation results for the five cases are shown in Table 7 and Figs. 3–8.

4.1.2.1. Base case. The strategic behavior and impact of the DisCo that tries to influence the wholesale market are studied in the base case. In the base case, uncertainty in the demand of distribution network customers is not considered. The only uncertainty taken into account here is that the DisCo acts as a strategic player in the wholesale market. As can be seen from Table 7, in this case, the DisCo uses the full capacity of DG units to meet the needs of consumers. The energy produced by DG units during the planning period is equal to 24 MW. In other words, the DisCo plays its strategic role through the purchase of energy from the wholesale market. The total energy required by consumers is 69.92 MWh, at which 45.92 MWh are purchased from the wholesale market. The cost of purchasing energy from the wholesale market and the operation cost of DG units are 2066.73 $ and 720.9 $, respectively. The expected revenue and profit of the DisCo in the planning period are 3984.98 $ and 1197.34 $, respectively. Furthermore, no LC is required, however, when the DisCo acted as the price taker in the wholesale market, a certain amount of LC was required in the network. This is due to the company's influence on the wholesale market and a significant decrease in market prices.

4.1.2.2. Case 1. In this case, the uncertainty in the demand of consumers of the distribution network is taken into account. This uncertainty is modeled by producing 1000 scenarios and reducing them to 8 scenarios [39]. This uncertainty has led to a reduction in the expected demand. Table 7 shows that the total energy consumption has decreased by 10.9% with respect to the base case. The energy purchased from the wholesale market has decreased by 16.59% compared to the base case, while the total production of DG units has been used. The expected profit of DisCo has also dropped by 7.88% compared to the base case. As shown in Fig. 3, demand response has reduced demand in some hours, and on the other hand, has increased demand in the other hours. For instance, compared with the base case, at hours 8 and 19, demand has decreased by 3.42% and 6.81%, respectively, while at hour 3, the demand has increased by 1.88%. Therefore, load shifting and peak load reduction can be considered as impacts of demand response and RTP.

Fig. 4 shows the price of electricity on bus 1 during the planning period. Since there is no uncertainty in the wholesale market, and also

Table 4
Demands' characteristics in the wholesale market.

Demands	Bus number	Cap. B1 (Mw)	Cap. B2 (Mw)	Cap. B3 (Mw)	Bid Price B1 ($/MWh)	Bid Price B2 ($/MWh)	Bid Price B3 ($/MWh)
1	2	1	0.6	0.2	60	45	20
2	3	1	0.7	0.5	60	30	10
Strategic DisCo	1	1.2	1	0.8			

Table 5
Suppliers' characteristics in the wholesale market.

Power suppliers	Bus number	Cap. B1 (Mw)	Cap. B2 (Mw)	Cap. B3 (Mw)	Offer Price B1 ($/MWh)	Offer Price B2 ($/MWh)	Offer Price B3 ($/MWh)
1	1	1	2	0.5	10	35	60
2	2	1	2.5	0.5	25	45	79

Table 6
Introducing case studies.

case	Uncertainty in demand	ESS	Upper network limits	DGs with high fuel cost	Low elasticity of demands
Base case					
# 1	✓				
# 2	✓	✓			
# 3	✓	✓	✓		
#4	✓	✓		✓	
#5	✓	✓		✓	✓

the DisCo has market power, the company tries to control the market price. Since the power required by the DisCo from the upper network is in a particular block, the market price in the relevant hours is also fixed. Note that when the DisCo was considered as a price taker in the wholesale market, the price is under the control of other consumers. Fig. 5 shows the expected real-time price profile, offered to customers over the planning period. It should be noted that the price profile is influenced by the amount of consumption as well as the wholesale market price. During peak hours, customers are offered a higher price. Also, the low price offered to the customer, at the off-peak hours, has led to an increase and a shift in demand.

Fig. 6 shows the total system losses over the planning period. It should be noted that the maximum power loss is 0.1 MW (i.e., 3.62%). Due to the shunt compensation provided by the distribution network, reactive power delivered by the upper network has been decreased. Thus, power losses and voltage profile have been controlled. In this case, the minimum voltage is obtained 12.5 kV, which is within acceptable level.

4.1.2.3. Case 2. In this case, in addition to considering uncertainty in demand, the ESS is considered within the distribution network. As shown in Fig. 3, the demand profile is similar to case 1. Also, according to Fig. 5, the demand is reduced during the hours when the price is high, and vice versa. Therefore, peak demand has fallen and, consequently, the retail price has been adjusted. Fig. 4 shows that the market price has fallen in some hours compared to the case 1. Compared with case 1, at hour 3, the market price has dropped from 43.4$/MWh to 39.3 $/MWh, (i.e. 9.45%). This drop in prices also directly affected retail prices (Fig. 5). Therefore, the ESS, by creating more flexibility, has increased the power of the DisCo in the wholesale market. According to Table 7, DisCo's revenue decreased by 0.4% with respect to the case 1, because of declining in sales. Demand reduction due to the DR program has reduced the DisCo's sales, and therefore has

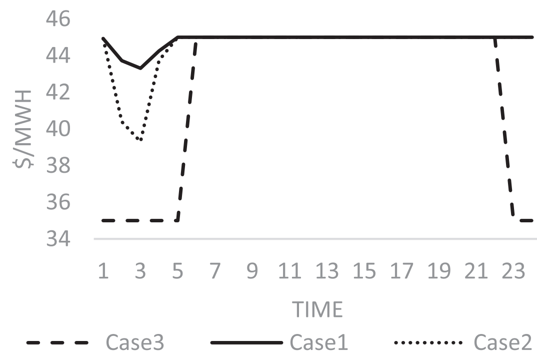

Fig. 4. The expected price of energy on bus1 in the sub-transmission network.

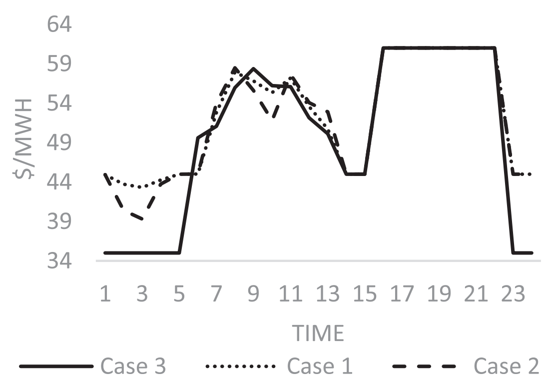

Fig. 5. Expected retail price presented to customers.

reduced its revenue. However, due to the storage system, the power purchased from the upper network has decreased, which ultimately led to an increase in the profit of DisCo. Fig. 6 shows the system losses for case 1, case 2, and case 3. It can be seen that the distribution network losses, in this case, have decreased compared to case 1. The maximum system loss is reduced by 5% with respect to case 1. Fig. 7 shows the power purchased from the upper network and the power produced by the DG units. As shown in Fig. 7 (and also Table 7), the purchase of energy from the wholesale market is reduced compared to case 1. The ESS with charging in low-price hours, as well as discharging during high-price hours, reduced the total energy purchased from the upper. According to Table 7, the purchased energy has fallen by 0.65 MWh compared to case 1, and also the profit of DisCo has increased by 1.9%.

4.1.2.4. Case 3. In case3, in addition to the parameters considered in

Table 7
Simulation results for different cases.

Case	Expected revenue ($)	Expected profit ($)	Expected grid purchase cost ($)	Expected LC cost ($)	Expected DG generation cost ($)	Expected generated energy (MWh)	Expected purchased energy (MWh)	Total expected energy (purchase & generation) (MWh)
Base case	3984.98	1197.34	2066.73	0	720.9	24	45.92	69.92
#1	3543.11	1103.01	1719.19	0	720.9	24	38.30	62.30
#2	3526.72	1125.34	1680.48	0	720.9	24	37.65	61.65
#3	3390.73	1065.23	1676.24	0	649.25	21.61	39.36	60.98
#4	3550.76	670.86	2879.89	0	0	0	63.99	63.99
#5	4651.54	1828.52	2823.01	0	0	0	62.73	62.73

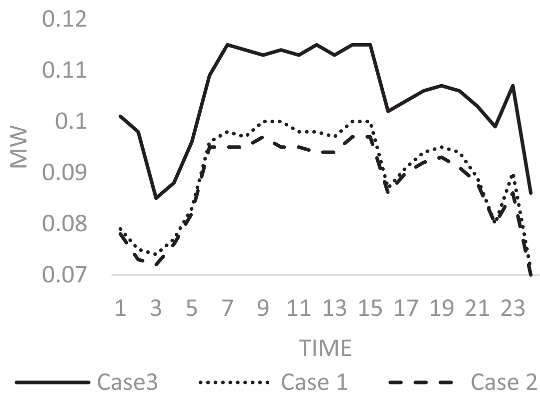

Fig. 6. Expected losses of the distribution system.

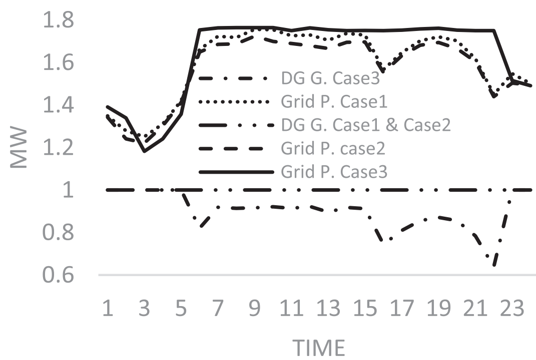

Fig. 7. Expected purchased power from the upper networks & power produced by DG units.

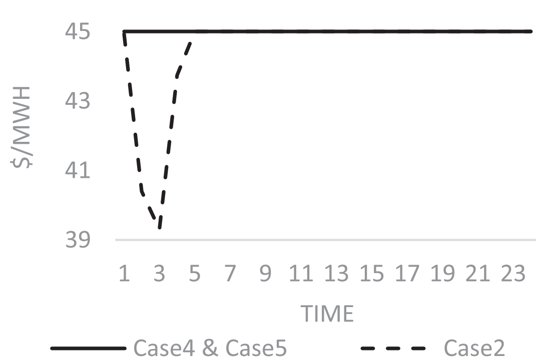

Fig. 8. Expected price of energy at bus 1 in case 2, case 4 and case 5.

Table 8
DisCo's profit and revenue for different demand elasticity.

Elasticity	Expected Profit ($)	Expected Revenue ($)
−0.1	2279.97	4651.54
−0.15	1550.48	3903.50
−0.2	1125.34	3526.72
−0.25	847.44	3315.16
−0.3	847.08	3322.97

case 2, the effect of the limits of the sub-transmission lines on the upper network is considered. In this regard, the thermal capacity of lines 1–2 and 1–3 is assumed to be 2 MW. According to Fig. 3, in this case, the changes in demand are more than the previous cases. In the initial demand curve, the peak demand period was time interval between hour 9 and hour 19. However, in this case, it has changed to hours between hours 7 and 15. For example, at 19:00, demand has decreased by 10.36% with respect to its initial value, because the operating conditions here are different from the previous cases. With the limits on the upper network, the DisCo has no access to other suppliers for purchasing the electricity. Thus, the DisCo uses its potential to prevent increases in price by reducing power consumption during peak hours. According to Fig. 4, the energy price at bus 1 in the first 5 h has decreased by 22.22%. Also, this reduction for the last two hours is 22.22%. The DisCo, in this case, has a monopoly position due to the network limits, so has a greater impact on the market price. Also, the drop in the wholesale market price have directly affected retail prices. As shown in Fig. 5, the retail price has fallen proportionally with the market price. Fig. 6 shows that system loss, has significantly increased compared to case 1 and case 2. The biggest change occurs at hour 1, where the system loss increases by 22.77% compared to case 1. It should be noted that system losses at 7, 12, 14, and 15 reached its maximum level (i.e. 0.115 MW). The reason for the increase in loss is that the demands on bus 2 and bus 3 have limited access to bus 1 and most of the required energy is acquired by the DisCo from bus 1. According to Fig. 7, the energy purchased from the upper network, is higher than case 1 and case 2, and due to purchasing from the upper network, the production of DG units in some hours has decreased. Table 7 states that the energy production of DG units is reduced by 9.95% compared to case 2, which has reduced the operation cost by 9.93%. Although the energy purchased has increased by 4.95%, the cost of purchasing power has been reduced by 0.25%, because the electricity prices are low during the initial and end of the planning period. Due to a significant decline in demand, especially during peak hours, the revenue of the DisCo has dropped by 3.85% compared to case 2. The profit of DisCo has also dropped by 5.31%.

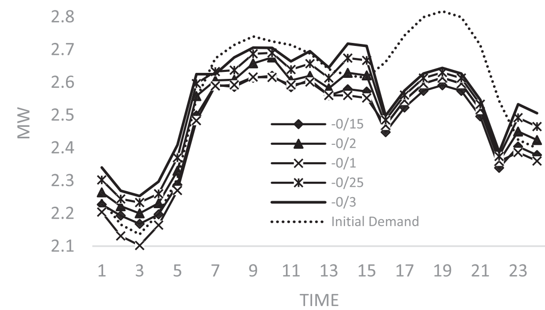

Fig. 9. Expected demand with different self-elasticity coefficients.

4.1.2.5. Case 4. Case 4 is similar to Case 2. However, in this case, the fuel cost for DG units has increased. In fact, impacts of an increase in the cost of fuel for DG units on the DisCo's performance are investigated. Increasing the cost of DG units has reduced the flexibility of the DisCo in the wholesale market. Unlike Case 2, the DisCo has not so much control over the market price as well as the purchases from the wholesale market. Table 7 shows that the production of DG units has dropped, since purchasing energy from the wholesale market is more economical for the DisCo. In this situation, the DisCo is completely dependent on the wholesale market. Fig. 8 shows that the lack of energy production by DG units reduces the ability of the DisCo to control price. It is worth noting that the electricity prices have risen in the early hours of the planning period, compared to case 2. According to Table 7, comparing to case 2, the purchase price and related cost have increased by 41.16% and 41.64%, respectively. The profit of the DisCo has been decreased by 40.38%. Despite the changing conditions of operation and pricing at the retail level, revenue has increased to some extent.

4.1.2.6. Case 5. In this case, the effect of reducing the consumer's elasticity on the performance of the DisCo is examined. To do this, the simulation was performed, assuming that the elasticity coefficient is equal to −0.1. Other features of this case are similar to case 4.

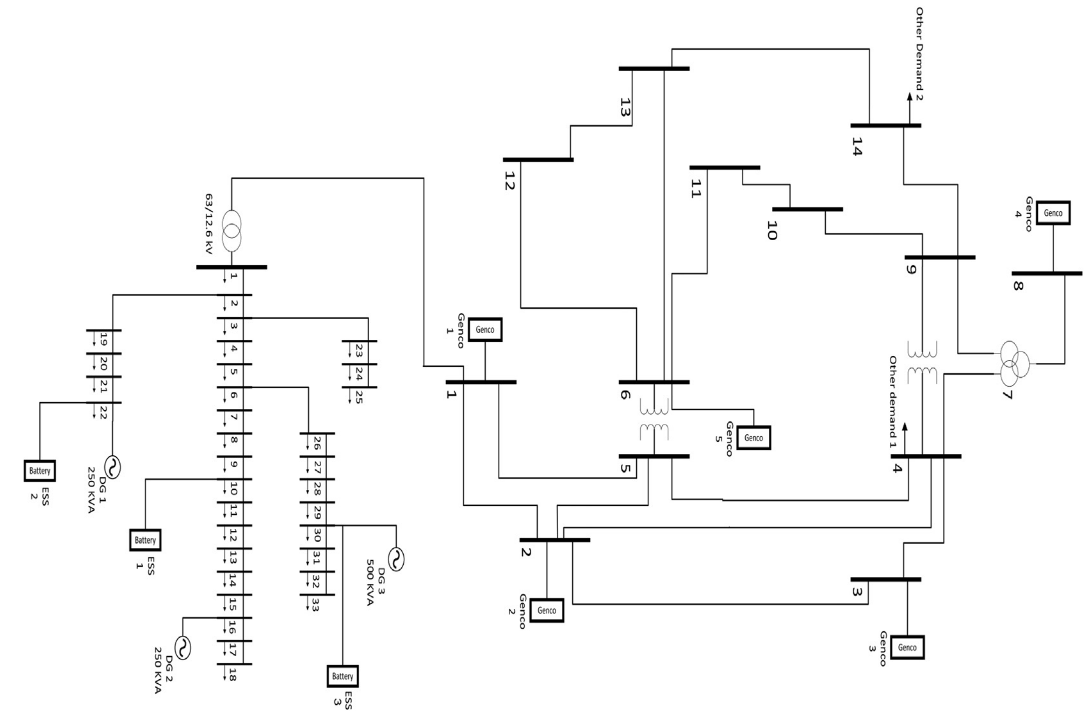

Fig. 10. IEEE 33-BUS distribution network and the IEEE 14-BUS sub-transmission network.

Table 9
Characteristics of DG units.

DG No.	Bus No.	P^{max}(MW)	P^{min}(MW)	Q^{max}(MVar)	Q^{min}(MVad)	Ramp up (MW/min)	Ramp down (MW/min)
1	16	7	0	5	−5	5	5
2	22	7	0	5	−5	5	5
3	30	5	0	5	−5	3	3
DG No	SU Ramp (MW/min)	SD Ramp (MW/min)		Case	α	β	γ
1	5	5		**#2**	0	30	0.1
2	5	5		**#5**	10	50	1
3	3	3					

Table 10
Characteristics of ESSs.

ESS No.	Bus No.	P^{chmax}(KW)	P^{dismax}(KW)	η^{ch}	η^{dis}	Initial level (KW)	Capacity (KW)
1	10	300	300	0.9	0.9	500	1500
2	22	300	300	0.9	0.9	500	1500
3	30	300	300	0.9	0.9	500	1500

Decreasing the self-elasticity coefficient has led to fewer peak shaving and demand shifting. This has two consequences. Firstly, it reduces the power of the DisCo in the wholesale market, to control the market price. As shown in Fig. 8, unlike case 4, the DisCo is no longer able to reduce market prices to less than 45 ($/MWh). Secondly, reducing the demand elasticity, escalates the energy consumption at the retail level, especially at peak hours, which increases the DisCo's profit and revenue as well as customer payments. In this case, the cost of purchasing energy is lower than the revenue of energy soled, because the company still has the ability to apply market power. According to Table 7, the revenue and profit of DisCo have increased by 23.66% and 63.31%, respectively, compared to case 4. It should be noted that the total

Table 12
Data of Gencos in the wholesale electricity market.

GENCO	Bus No.	Cap. B1 (MW)	Cap. B2 (MW)	Cap. B3 (MW)	Offer Price B1 ($/MWh)	Offer Price B2 ($/MWh)	Offer Price B3 ($/MWh)
1	1	4	3	2	10	45	70
2	2	4	3.5	2.5	15	55	65
3	3	4	3	1	35	65	79
4	6	5	2	1	35	65	89
5	8	4	2.5	1	45	65	89

Table 11
Data of demands in the wholesale electricity market.

Demands	Bus No.	Cap. B1 (MW)	Cap. B2 (MW)	Cap. B3 (MW)	Bid Price B1 ($/MWh)	Bid Price B2 ($/MWh)	Bid Price B3 ($/MWh)
1	4	5	3	1	60	45	20
2	14	6	4	2	60	30	10
Strategic DisCo	1	7	5	3			

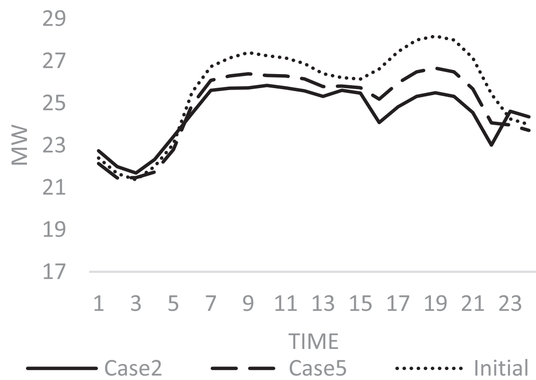

Fig. 11. Expected demand in the distribution system (case study 2).

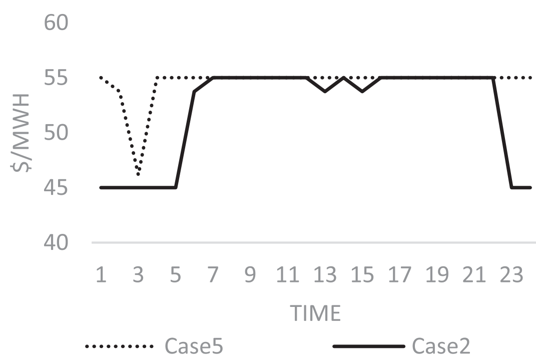

Fig. 12. The expected electricity price on bus1 of sub-transmission network.

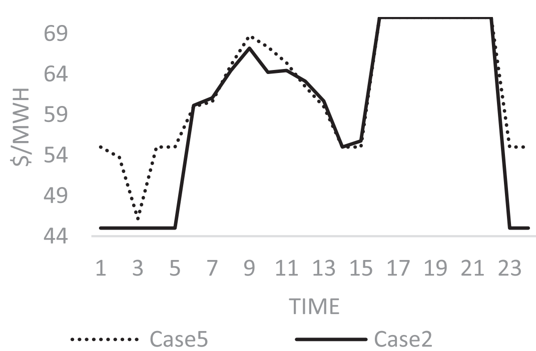

Fig. 13. Expected retail price offered to customers.

purchase of energy and total demand are less than those in case 4, which is caused by the DR program. In fact, the demand curves in both cases are almost the same, however, there are more peaks in this demand curve, compared to the case 4. Also, the demand reduction led to a reduction in the purchasing cost from the wholesale market, (i.e. $ 56.88), which is negligible compared to $ 1157.68 profit.

Additionally, in this case, the impact of elasticity change on the profit of the DisCo and also the demand curve is investigated. Consumers' behavior can be studied using the elasticity matrix. Retail prices and demand response are highly dependent on the elasticity matrix. Simulations were performed for a variety of the self-elasticity coefficient (i.e., −0.1 to −0.3). Table 8 show that reducing the demand

Table 14
The output of CPLEX 12.2.0.1 for each study case.

Case	Model	Best solution	Absolute gap	Computation time (Min.)
Base case	MILP	1197.343080	0%	00:27
#1	MILP	1103.018657	0%	11:00
#2	MILP	1125.340428	0%	9:37
		16563.55658		36:15
#3	MILP	1065.230477	0%	20:27
#4	MILP	670.866123	0%	15:33
#5	MILP	1828.524151	0%	04:50
		21221.51425		25:34

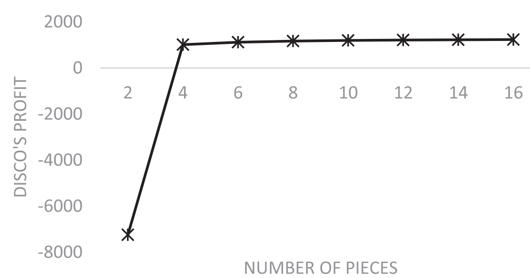

Fig. 14. Evaluation of the DisCo's profit versus the number of pieces.

elasticity leads to an increase in the profit of DisCo and consumer payment. With declining in elasticity from 0.2 to 0.1, the consumers' payments increased by 24.18% and the company's profits increased by 50.64%.

Fig. 9 shows the demand curves for different elasticity coefficients. By increasing elasticity, consumption decreases during peak hours, and increases in at valley hours. This change in the consumption can create new peak and valley hours. As can be seen in Fig. 9, a new peak is created at −0.3 elasticity.

4.2. Case study 2

In the second case study, a power system larger than the one in the first case study has been selected. For this purpose, the IEEE-14-bus test system is used as the upper-network. In the distribution level, the IEEE-33-bus test system is used, except in this case the consumption within the distribution network is increased to10 times greater than what was used in the first case study. It is assumed that the distribution network feeders have sufficient thermal capacity to conduct the required power. Case 2 and case 5 of the first case study 1 are investigated in this case study.

4.2.1. Data

Fig. 10 shows a single-line diagram of the network, in which the distribution network is connected to bus 1 of the sub-transmission network. The data of the DisCo are represented by Tables 9 and 10. Also, data of the market participants' are shown in Tables 11 and 12. The data related to pricing is similar to the first case study. The ESS is connected to buses 10, 22 and 30 of the distribution network.

4.2.2. Results

Fig. 11 shows the expected level of consumption of consumers in the distribution network for case 2 and case 5. The amount of demand has

Table 13
Simulation results for case study 2 (case 2 and case 5).

Case	Expected revenue ($)	Expected profit ($)	Expected grid purchase cost ($)	Expected LC cost ($)	Expected DG generation cost ($)	Expected generated energy (MWh)	Expected purchased energy (MWh)
#2	40262.31	16563.55	13206.84	0	10491.91	343.89	245.46
#5	54312.7	21221.51	19535.29	0	13555.88	240.21	357.53

changed at different times, according to the RTP scheme and the operation conditions and constraints. It can be seen that the DR program reduced the peak consumption. In the case 2, the peak consumption was shifted from hour 19 to hour 10. Also, at the beginning and the end of the planning period, consumption has increased.. Similar results were obtained for the first case study.

Fig. 12 shows the electricity price at bus 1 of the sub-transmission grid, which is similar to Fig. 4. It is observed that as the DisCo uses more of its resources as well as DR program, electricity prices in bus 1 is decreased. In case 5, the operation cost of DGs is high, so the use of these resources has become limited. Consumer elasticity has also been reduced in this case. Due to these limitations, the influence of the DisCo on the wholesale market has actually decreased, and this can be seen in the price of energy on bus 1.

Fig. 13 illustrates the retail price of electricity in the distribution network. Similar to the results of the first case study (i.e. Fig. 5), the retail price of energy is higher during peak hours to reduce consumption. Also, the electricity price in case 5 is higher than the price in case 2.

Table 13 shows the profit, revenue, and the costs of the DisCo. The cost of purchasing energy in case 5 is higher than case 2 because the operation cost DG units has increased. Additionally, the company has earned much more in case 5. Also, when consumers have low elasticity, the consumption would be high, especially during peak hours. So, it will increase the DisCo's profit and revenue, as well as customer payments.

4.3. Verification of results' accuracy

As mentioned before, for solving the MILP model, CPLEX 12.2.0.1 [48] under GAMS [49]. The tests executed by a single processor Intel core-i3-3217U clocking at 1.8 GHz and 4 GB of RAM. Because of the huge number of constrains, binary variables, and linearization the convergence and accuracy of the algorithm and the model should be justified. Table 14 illustrates the "best solution" answer and the "absolute gap ", as well as "computation time", of the results in a relatively sufficient precision for different cases. As can be seen, the results yielded without any gap, and it shows the convergence of the algorithm.

In addition, because there were so many nonlinear terms in the equations and constraints in the model, we had to use the linearization methods for many times to linearize them. The determination of the required pieces is a challenging issue. If a high number of pieces be used in approximation, it can lead to extra computation. However, too few pieces can cause inaccuracy in results. According to Fig. 14, it is observed that the profit of the DisCo starts to remain at the same level around 10 pieces.so, 10 pieces are sufficient enough for accuracy of the result and its justification.

5. Conclusion

In this paper, a new framework is proposed to optimize the participation strategy of distribution companies owning distributed generation units and storage systems in wholesale and retail energy markets under uncertainty associated with consumption and RTP scheme. The DisCo acts as a price-maker in the wholesale electricity market because of its DG units and ESS. Uncertainty of the customers' behavior within the distribution network is modeled by a set of scenarios and the demand response program is modeled by considering the customers' elasticity, as well. The problem is modeled in the form of a bi-level optimization problem, which involves maximizing distribution's profits in its upper-level and the lower-level of problem consist of the market-clearing problem with constraints of DC network and maximizing social welfare as its objective. By formulating the bi-level problem to the MPEC and linearizing it using the dual theory and the KKT conditions, the proposed model is transformed into the MILP. The simulation results in the example system are as follows:

- The DisCo, using its resources and demand response programs, is trying to impose on the wholesale market by lowering market prices and increasing its profits.
- The presence of the ESS reduces energy losses within the distribution network. The profit of the DisCo will also increase. Besides, these resources increase the ability of the DisCo to exercise power in the wholesale market.
- The limits of the upper network put the DisCo in a monopolistic position. These conditions increase the influence of the DisCo, which reduces energy costs. In these circumstances, it is expected that the company will buy all the required power at a cheap price from the wholesale market, but the constraints on the upper network will limit the amount of available energy for the company too, which compels the DG units to be turned on in order to provide some part of the required energy.
- The rise in fuel prices for DG units will make the company lose its market power and more dependent on the wholesale market for its energy acquisition. In this situation, the price of energy in the market is high, which leads to a reduction in the DisCo's profits.
- Reducing the demand side elasticity leads to an increase in the strategic DisCo's profit. Consumers will not be able to benefit from the demand response program due to a reduction in the elasticity. Therefore, retail prices will not be moderated at the peak of consumption, and consumer payments will also increase. This increase in consumer payments can increase the revenue of the DisCo. Also, the strategic nature of the DisCo does not let the cost of energy to soar up. Hence, reducing the elasticity in these situations will increase the company's profits.

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.

Acknowledgements

The authors acknowledge the funding support of Babol Noshirvani University of Technology through Grant program No. BNUT/370117/98.

References

[1] Kirschen Daniel S, Strbac Goran. Fundamentals of power system economics. 1 ed. Wiley; 2004.
[2] Bahramara SalahMaziar, Yazdani-Damavandi Javier, Contreras Miadreza, Shafie-Khah João PS Catalão. Modeling the strategic behavior of a distribution company in wholesale energy and reserve markets. IEEE Trans Smart Grid 2017;9(4):3857–70.
[3] Safdarian Amir, Fotuhi-Firuzabad Mahmud, Lehtonen Matti. Integration of price-based demand response in DisCos' short-term decision model. IEEE Trans Smart Grid 2014;5(5):2235–45.
[4] Marvasti Amin Kargarian, Fu Yong, DorMohammadi Saber, Masoud Rais-Rohani. Optimal operation of active distribution grids: A system of systems framework. IEEE Trans Smart Grid 2014;5(3):1228–37.
[5] Triki Chefi, Violi Antonio. Dynamic pricing of electricity in retail markets. 4OR 2009;7(1):21–36.
[6] Mohiti Maryam, Monsef Hassan, Anvari-Moghaddam Amjad, Guerrero Josep, Lesani Hamid. A decentralized robust model for optimal operation of distribution companies with private microgrids. Int J Electr Power Energy Syst 2019;106:105–23.
[7] Bahramara Salah, Sheikhahmadi Pouria, Mazza Andrea, Chicco Gianfranco, Shafie-khah Miadreza, Catalão João PS. A risk-based decision framework for the distribution company in mutual interaction with the wholesale day-ahead market and microgrids. IEEE Trans Indus Informat 2019.
[8] Zhang Chunyu, Wang Qi, Wang Jianhui, Korpås Magnus, Khodayar Mohammad E. Strategy-making for a proactive distribution company in the real-time market with demand response. Appl Energy 2016;181:540–8.
[9] Khazaei Hossein, Vahidi Behrooz, Hosseinian Seyed Hossein, Rastegar Hasan. Two-level decision-making model for a distribution company in day-ahead market. IET Generat Transmiss Distrib 2015;9(12):1308–15.
[10] Wei Wei, Liu Feng, Mei Shengwei. Energy pricing and dispatch for smart grid

retailers under demand response and market price uncertainty. IEEE Trans Smart Grid 2014;6(3):1364–74.
[11] Zhang Chunyu, Wang Qi, Wang Jianhui, Pinson Pierre, Morales Juan M, Østergaard Jacob. Real-time procurement strategies of a proactive distribution company with aggregator-based demand response. IEEE Trans Smart Grid 2016;9(2):766–76.
[12] Zhang Chunyu, Wang Qi, Wang Jianhui, Korpås Magnus, Pinson Pierre, Østergaard Jacob, et al. Trading strategies for distribution company with stochastic distributed energy resources. Appl Energy 2016;177:625–35.
[13] Sheikhahmadi P, Bahramara S, Moshtagh J, Yazdani Damavandi M. A risk-based approach for modeling the strategic behavior of a distribution company in wholesale energy market. Appl Energy 2018;214:24–38.
[14] Cerbantes Marcel Chuma, Fernández-Blanco Ricardo, Ortega-Vazquez Miguel A, Mantovani José Roberto Sanches. Short-term operation of a distribution company: A pseudo-dynamic tabu search-based optimisation. IET Generat Transmiss Distrib 2018;12(12):2995–3004.
[15] Safdarian Amir, Fotuhi-Firuzabad Mahmud, Lehtonen Matti. A stochastic framework for short-term operation of a distribution company. IEEE Trans Power Syst 2013;28(4):4712–21.
[16] Algarni Ayed AS, Bhattacharya Kankar. A generic operations framework for discos in retail electricity markets. IEEE Trans Power Syst 2009;24(1):356–67.
[17] Cerbantes Marcel Chuma, Mantovani José Roberto Sanches, Fernández-Blanco Ricardo, Ortega-Vazquez Miguel A. Optimal short-term operation of a DisCo including voltage-sensitive loads. 2016 Power Systems Computation Conference (PSCC). Ieee; 2016. p. 1–7.
[18] Safdarian Amir, Fotuhi-Firuzabad Mahmud, Lehtonen Matti. A medium-term decision model for DisCos: Forward contracting and TOU pricing. IEEE Trans Power Syst 2014;30(3):1143–54.
[19] Badri Ali, Kashefi Hazhir. Optimal bidding strategy of retailers in a mixed pool-bilateral market considering demand response programs. 2013 Smart Grid Conference (SGC). IEEE; 2013. p. 131–7.
[20] Lujano-Rojas Juan M, Monteiro Cláudio, Dufo-López Rodolfo, Bernal-Agustín José L. Optimum residential load management strategy for real time pricing (RTP) demand response programs. Energy Policy 2012;45:671–9.
[21] Zugno Marco, Morales Juan Miguel, Pinson Pierre, Madsen Henrik. A bilevel model for electricity retailers' participation in a demand response market environment. Energy Econ 2013;36:182–97.
[22] Gottwalt Sebastian, Ketter Wolfgang, Block Carsten, Collins John, Weinhardt Christof. Demand side management—A simulation of household behavior under variable prices. Energy Policy 2011;39(12):8163–74.
[23] Mazidi Mohammadreza, Monsef Hassan, Siano Pierluigi. Incorporating price-responsive customers in day-ahead scheduling of smart distribution networks. Energy Convers Manage 2016;115:103–16.
[24] Moghaddam M Parsa, Bahramara S, Damavandi MY, Haghifam MR. Distribution company and microgrids behaviour in energy and reserve equilibirum. 2015 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC). IEEE; 2015. p. 1–5.
[25] Bahramara Salah, Moghaddam Mohsen Parsa, Haghifam Mahmoud Reza. Modelling hierarchical decision making framework for operation of active distribution grids. IET Gener Transm Distrib 2015;9(16):2555–64.
[26] Bahramara S, Parsa Moghaddam M, Haghifam MR. A bi-level optimization model for operation of distribution networks with micro-grids. Int J Electr Power Energy Syst 2016;82:169–78.
[27] Toutounchi Amir Naebi, Seyedshenava Seyedjalal, Contreras Javier, Akbarimajd Adel. A stochastic bilevel model to manage active distribution networks with multi-microgrids. IEEE Syst J 2019.
[28] Fateh H, Safari A, Bahramara S. A Bi-level optimization approach for optimal operation of distribution networks with retailers and micro-grids. J Operat Automat Power Eng 2019.
[29] Sadati S Muhammad Bagher, Moshtagh Jamal, Shafie-khah Miadreza, Rastgou Abdollah, Catalão João PS. Operational scheduling of a smart distribution system considering electric vehicles parking lot: A bi-level approach. Int J Electr Power Energy Syst 2019;105:159–78.
[30] Nguyen Duong Tung, Nguyen HieuTrung, Le Long Bao. Dynamic pricing design for demand response integration in power distribution networks. IEEE Trans Power Syst 2016;31(5):3457–72.
[31] Henríquez Rodrigo, Wenzel George, Olivares Daniel E, Negrete-Pincetic Matías. Participation of demand response aggregators in electricity markets: Optimal portfolio management. IEEE Trans Smart Grid 2017;9(5):4861–71.
[32] Manshadi Saeed D, Khodayar Mohammad E. A hierarchical electricity market structure for the smart grid paradigm. IEEE Trans Smart Grid 2015;7(4):1866–75.
[33] López-Lezama Jesús María, Padilha-Feltrin Antonio, Contreras Javier, Muñoz José Ignacio. Optimal contract pricing of distributed generation in distribution networks. IEEE Trans Power Syst 2010;26(1):128–36.
[34] Palma-Behnke Rodrigo, Vargas Luis S, Jofré Alejandro. A distribution company energy acquisition market model with integration of distributed generation and load curtailment options. IEEE Trans Power Syst 2005;20(4):1718–27.
[35] Haghighat Hossein, Kennedy Scott W. A bilevel approach to operational decision making of a distribution company in competitive environments. IEEE Trans Power Syst 2012;27(4):1797–807.
[36] Li Haiying, Li Yuzeng, Li Zuyi. A multiperiod energy acquisition model for a distribution company with distributed generation and interruptible load. IEEE Trans Power Syst 2007;22(2):588–96.
[37] Carrión Miguel, Arroyo José M. A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem. IEEE Trans Power Syst 2006;21(3):1371–8.
[38] Safdarian Amir, Fotuhi-Firuzabad Mahmud, Aminifar Farrokh, Lehtonen Matti. A new formulation for power system reliability assessment with AC constraints. Int J Electr Power Energy Syst 2014;56:298–306.
[39] Safdarian Amir, Firuzabad Mahmud Fotuhi, Aminifar Farrokh. A novel efficient model for the power flow analysis of power systems. Turkish J Electr Eng Comput Sci 2015;23(1):52–66.
[40] Ghasemi Ahmad, Mortazavi Seyed Saeidollah, Mashhour Elaheh. Hourly demand response and battery energy storage for imbalance reduction of smart distribution company embedded with electric vehicles and wind farms. Renewable Energy 2016;85:124–36.
[41] Kirschen Daniel S, Strbac Goran, Cumperayot Pariya, de Paiva Mendes Dilemar. Factoring the elasticity of demand in electricity prices. IEEE Trans Power Syst 2000;15(2):612–7.
[42] Kazempour S Jalal, Conejo Antonio J, Ruiz Carlos. Strategic bidding for a large consumer. IEEE Trans Power Syst 2014;30(2):848–56.
[43] Kazempour S Jalal, Conejo Antonio J, Ruiz Carlos. Strategic generation investment using a complementarity approach. IEEE Trans Power Syst 2010;26(2):940–8.
[44] Fortuny-Amat José, McCarl Bruce. A representation and economic interpretation of a two-level programming problem. J Oper Res Soc 1981;32(9):783–92.
[45] 33-bus test system - available [online]: https://egriddata.org/dataset/33-bus-radial-distribution-system/resource/602529ca-b81d-4c1e-b4eb-071158f42ebc.
[46] Growe-Kuska Nicole, Heitsch Holger, Romisch Werner. Scenario reduction and scenario tree construction for power management problems. 2003 IEEE Bologna Power Tech Conference Proceedings, vol. 3. IEEE; 2003. p. 7.
[47] Yu Nanpeng, Liu Chen-Ching, Tesfatsion Leigh. Modeling of suppliers' learning behaviors in an electricity market environment. 2007 International conference on intelligent systems applications to power systems. IEEE; 2007. p. 1–6.
[48] 'The ILOG CPLEX', http://ilog.com-productsplex.
[49] Available [online]: https://gams.com.