Day-ahead Risk Averse Market Clearing Considering Demand Response with Data-driven Load Uncertainty Representation: A Singapore Electricity Market Study

Yuanzheng Li[a], Jingjing Huang[b], Yun Liu[c], Tianyang Zhao[d], Yue Zhou[e], Yong Zhao[a], Chau Yuen[f]

[a]*School of Artificial Intelligence and Automation, Key Laboratory on Image Information Processing and Intelligent Control of Ministry of Education, Huazhong University of Science and Technology, Wuhan 430074, China*
[b]*China-EU Institute for Clean and Renewable Energy, Huazhong University of Science and Technology, Wuhan 430074, China*
[c]*School of Electric Power Engineering, South China University of Technology, Guangzhou 510641, China.*
[d]*School of Electrical Engineering, Jinan University, Zhuhai, China.*
[e]*School of Engineering, Cardiff University, Cardiff, U.K.*
[f]*Engineering Product Development, Singapore University of Technology and Design, Singapore.*

Abstract

Demand response program is being implemented in the National Electricity Market of Singapore, which boosts the flexibility of demand side to actively participate in the real-time electricity market. Meanwhile, it is also significant to implement such a program in the day-ahead market, since generation companies could arrange their generating plans and load providers are able to adjust their hourly purchasing schedules. However, uncertain factors should be considered in the demand response program of the day-ahead market, such as the uncertain electricity load. Regarding the issue, this paper proposes a day-ahead bidding and clearing framework considering demand response with uncertain and correlated nature of electricity loads. To this end, a data-driven Dirichlet process mixture model is introduced to represent the load uncertainty, which might bring about the economic risk. To further reduce such a risk, a worst-case conditional value at risk is integrated into our proposed framework, and a WCVaR based two-step risk averse market clearing model is proposed. Finally, we conduct numerical studies based on the Singapore electricity market. Numerical studies demonstrate the outperformance of Dirichlet process mixture model for the load uncertain representation, and also verify that the worst-case conditional value at risk based market clearing model could effectively reduce the economic risk while maximizing the social welfare.

Keywords: Market clearing, demand response, data-driven, uncertainty representation, economic risk.

1. Introduction

1.1. Research Background

National Electricity Market of Singapore (NEMS), as the first liberalized electricity market in Asia, is well-known for its distinguished efficiency and maturity. Recently, this well-structured market is undergoing an important reform to implement the demand response (DR) program in the real-time market (RTM) [1]. DR is defined as the end-users' intentional modifications of energy consumption stimulated by electricity price fluctuations or incentive payment mechanisms [2]. It could bring about multiple benefits to power system operations, such as lowering electricity prices, mitigating peak-valley load gaps, enhancing operational security of power grids, etc. [3, 4].

Market clearing (MC) plays a crucial role for electricity market operations, as the electricity supply and demand should be cleared and priced. With the incorporation of DR program in NEMS, generation companies and load providers are both required to submit their bidding offers to the power system operator (PSO) during each dispatch period in RTM. On the generation side, each unit should bid on energy, reserve, and regulation resources [5]. On the demand side, load providers in the wholesale market are capable to submit bidding offers as well, which include the information of total load demand and load curtailment. The load curtailment information is in the form of price-quantity tranches, which is a salient feature of Singapore's DR program. These tranches specify the load providers' willingness to voluntarily reduce their electricity load at a certain price level.

1.2. Literature Review

Because of the significant role in electricity markets, numerous studies have investigated DR participation in the MC [6–9]. For instance, a price-based self-scheduling scheme for the multi-energy market is proposed in [6], which aims to maximize the total profit while considering characteristics of various DR consumers. Similarly, [7] presents an optimization framework, with the adoption of a newly developed heuristic algorithm, to obtain the optimal bidding strategy in the day-ahead market. A competition framework in a retail energy market is formulated in the presence of DR [8], which verifies that it could reduce prosumers cost and increased retailers profit with DR. [9] proposed an improved incentive-based DR model, in which the incentive value is not a fix value and relates to the peak intensity of each hour. Nevertheless, the above works adopt a deterministic optimization approach to study the MC problem with DR, which may be inadequate considering various factors, such as the load uncertainties.

Moreover, some studies have adopted stochastic optimization methods to deal with the MC problem considering load uncertainties [10–13]. For instance, [10] studies an economic dispatch framework based on the scenario analysis, which explores the uncertainties brought by DR providers in the day-

☆This work is supported by National Natural Science Foundation of China (62073148) and Tencent Rhinoceros Foundation of China (RAGR20210102)
Email address: liuyun1988@gmail.com (Yun Liu)

ahead MC. Also, [11] presents a multi-stage stochastic model in order to analyse behaviours of electricity market participants. Considering the load uncertainties, [12] presents a bi-level optimization problem to clear the market after strategic scheduling of a multi-energy system. To cope with the uncertainties such as load under the multi-level electricity market, the bi-level stochastic programming approach is exploited to maximize the profit of both retailers and DR aggregators [13].

However, two aspects shall be taken into account while considering uncertain loads, i.e., the uncertainty characterization and its impact on the MC. On the one hand, some studies have focused on probabilistic representations of load uncertainties. For instance, [13] and[14] employs the Normal probability distribution to model the uncertainty of electrical load. To take full advantages of load data, an ensemble method is proposed by clustering time-series profiles into multiple groups to obtain the aggregated load information [15]. Moreover, in [16], the Gaussian mixture model (GMM) is used to represent electricity load uncertainty, and the covariance matrix denotes correlations among different time periods. However, the accuracy of GMM significantly relies on the number of Gaussian components, which is usually set by the prior knowledge.

On the other hand, the risk induced by load uncertainties should not be ignored in the MC, and the conditional value at risk (CVaR) has been widely used to deal with this issue [17–20]. For instance, [17] adopts a bidding strategy based on a bi-level optimization approach, in which the MC is solved and CVaR is used as a risk assessment metric against load uncertainties. Similarly, [18] models a strategic behaviour in the power market while considering load uncertainties, which is solved via a risk-based stochastic optimization approach with the CVaR. A bidding strategy model is proposed for microgrid to participate in the day-ahead energy markets considering the uncertainties of load[19]. Moreover,the risk of participation in the day-ahead and real-time energy market is assessed using CVaR criteria and sensitivity analysis. In addition, in order to maximize profits of combined heat and power units when they are participating in the multi-energy market, [20] proposes a bidding strategy, in which the CVaR indicator measures the economic risk brought by load uncertainties.

It should be mentioned that the usage of CVaR is usually based on the standard probabilistic distribution of random variables, as shown in the above studies. However, it is difficult to obtain the exact probabilistic distribution of electricity load, which is usually data-dependent. Indeed, a mixture probability distribution may be obtained from historical data [21], which refers to a set of probabilistic components and corresponding weights. Therefore, whether CVaR is capable for controlling the economic risk in such a case is unknown.

1.3. Motivation and Contribution of This Paper

According to the above literature review, we identify the research gaps listed as follows:

1. The NEMS has implemented a load curtailment based DR mechanism in the wholesale electricity market. It is verified to be effective in the practice, yet based on the single period merely [5]. Therefore, we aim to extend such DR mechanism to multi-period horizon in DAM. However, due to the day-ahead forecasting errors of electricity load, its uncertainty might bring about the economic risk that impacts the social welfare.
2. When dealing with the characterization of uncertain load, standard probabilistic distributions are usually adopted to generate scenarios, whose accuracies are limited. Moreover, GMM is used to characterize the electricity loads in recent researches. However, the accuracy of GMM is dependent on the number of mixture components, which is usually manually set.
3. The conventional risk indicators, such as CVaR, are usually adopted based on the standard probabilistic distribution of random variables. However, they might not be suitable to address the economic risk in data-dependent mixture distribution.

In order to deal with these research gaps, we conduct our research in the following aspects, which are the contributions of our work.

1. This paper proposes a day-ahead bidding and clearing framework based on the DR mechanism of Singapore's wholesale electricity market, in which load providers could bid on load curtailments. This framework includes the day-ahead bidding mechanism with the DR program, the load uncertainty representation and a risk averse market clearing model (MCM).
2. A data-driven Dirichlet process mixture model (DPMM) is adopted to capture the uncertainties and time-series correlation of electricity load for the first time, to the best of authors' knowledge. Compared with traditional probabilistic model, the number of mixture probability distribution components could be automatically and optimally obtained without using prior knowledge.
3. We introduce a worst-case conditional value at risk (WCVaR) of the social welfare to measure the economic risk induced by load uncertainties. Then, a WCVaR based two-step risk averse MCM is proposed to maximize the expected social welfare while reducing the economic risk. Note that WCVaR is a risk indicator that represents the worst possible CVaR value among a set of probability distribution components [22]. In addition, the impact of the WCVaR index on the MC process is further studied.

The rest of this paper is organized as follows. Section 2 presents an overview of the DR mechanism in NEMS and introduces the day-ahead bidding and clearing framework considering DR. Then, the data-driven DPMM for representing time-series load uncertainty is introduced in Section 3. Next, Section 4 shows the proposed WCVaR based two-step risk averse MCM. Numerical simulation results are presented and analyzed in Section 5. Finally, the conclusion is drawn in Section 6.

2. Problem Description

In this section, we firstly introduce the mechanism of DR program in the NEMS's wholesale market in 2.1. Participating such a program, load providers are capable to bid on the load curtailments for each dispatch period. On this basis, we propose a bidding and clearing framework considering DR for the day-ahead wholesale market, which is presented in 2.2.

2.1. DR Mechanism in the NEMS

When participating in the DR program of the NEMS, licensed load providers are qualified to submit bidding offers to

the PSO. The most important information in the bidding offers includes the total load and load curtailment, in the form of price-quantity tranches. The total load is the overall electricity demand estimated by the load provider during the dispatch period. Furthermore, these tranches specify the quantities of electricity demand that the load provider is willing to curtail at certain price levels. For instance, as shown in Figure 1 (a), when electricity price is equal to or higher than $LP_{n,1}$, i.e., the higher price of the two tranches offered by the load provider n, it tends to curtail the total quantity of both tranches as $LQ_{n,1} + LQ_{n,2}$. It is notable that all the quantities for such tranches of a bidding offer could be deemed as curtailable loads for a load provider, which is submitted to PSO.

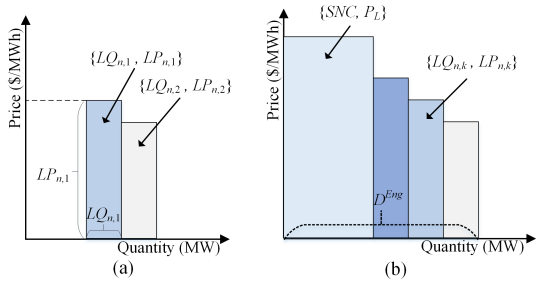

Figure 1: (a) Price-quantity tranches of bidding offer of load provider n; (b) All bidding offers on demand side in descending order.

Then, in order to determine the inelastic electricity demand, i.e., the *system non-curtailable (SNC)* load, the electricity load forecast information from PSO D^{Eng} is used. Accordingly, *SNC* can be calculated through subtracting total quantities of curtailable (elastic) bidding offers from the forecast electricity load. It is formulated as follows.

$$SNC = D^{Eng} - \sum_{n=1}^{N} \sum_{k=1}^{K_n} LQ_{n,k} \tag{1}$$

where n is the index of bidding offers ($n = 1, 2, \cdots, N$), note that each load provider submit one bidding offer and N is the total number of load providers. k is the index of price-quantity tranches for bidding offer $n(k = 1, 2, \ldots, K_n)$ and K_n is the number of tranches for bidding offer n. $LQ_{n,k}$ are the quantity of the curtailable load of tranche k in bidding offer n.

Since *SNC* is the total inelastic electricity load of whole market, which is not subject to any curtailments, the corresponding price that load providers are willing to pay is set as a fixed large value P_L. Afterwards, bidding offers on the demand side would be arranged in descending orders, as shown in Figure 1(b), consisting of the price-quantity tranches that represent all potential curtailments and the *SNC*. Finally, PSO clears the market according to the bidding offers of both the generation and demand sides, and determine the energy prices.

2.2. Day-ahead Bidding and Market Clearing Considering DR

In NEMS, the DR program is verified to be effective on the single-period basis in RTM [5]. Nevertheless, implementing the DR program on multi-period basis in DAM is also significant, since load providers could flexibly adjust their hourly purchasing schedules according to day-ahead price signals [23]. Therefore, we aim to expand the aforementioned DR mechanism, which is featured by the bidding of load curtailment, into DAM considering multiple periods. However, load uncertainties usually exist in DAM, which brings the economic risk to

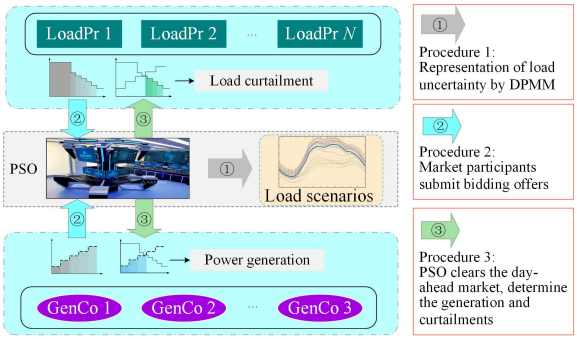

Figure 2: The day-ahead bidding and clearing framework.

the social welfare. *Note that the economic risk refers to the lower welfare than the expected one.*

To reduce such kind of risk, we propose a day-ahead bidding and clearing framework considering DR in the wholesale electricity market, in which a WCVaR based two-step risk averse MCM is further proposed. In this framework, market participants include generation companies and load providers (licensed electricity retailers or large consumers). As shown in Figure 2, the day-ahead bidding and clearing framework is conducted as following three procedures.

Procedure 1: First of all, the day-ahead electricity load forecasting information is published to the market participants by the PSO. However, due to the existence of load forecast errors, the uncertainty of correlated time-series load should be better characterized. To this end, we adopt a data-driven DPMM method to capture the characteristics of load information based on historical data in this paper. Then, the probabilistic scenarios are generated by the scenario-reduction based Monte-Carlo simulation, in order to conduct the stochastic optimization. This part will be further presented in Section 3.

Procedure 2: The market participants for both generation and demand sides will submit bidding offers for the next day to PSO. On the generation side, companies will submit day-ahead bidding offers for individual units with price-quantity tranches that are in the ascending order [24]. Each tranche specifies the expected selling price and corresponding supplying quantity. It should be notable that there are three kinds of transacting products in the wholesale market, including the energy, reserve and regulation. To be specific, the energy refer to the electricity that could be consumed by users. The reserve and regulation are both ancillary services purchased by PSO, in which reserve is needed for dealing with unexpected generation outages, and regulation represents the capacity to compensate load variations. Thus, generation companies should bid on such products for each unit as well.

On the demand side, load providers bid on total load and curtailable load of every periods for the next day, in the form of price-quantity tranches as stated in 2.1. Bidding offers on the demand side are based on the historical price information, i.e., relative high load curtailment bidding occurs within the periods when the historical price is high. In other words, load providers are willing to consume less electricity with a higher price.

Procedure 3: Considering the bidding information of both sides, PSO clears the day-ahead market by maximizing the social welfare with the adoption of MCM. However, the load uncertainty might bring about the economic risk. Therefore, a WCVaR based two-step risk averse MCM is proposed to

clear the market, by dispatching on power quantities of generation units, as well as load curtailments of providers. It will be specifically formulated in Section 4.

3. Data-driven DPMM Based Load Uncertainty Model

Note that the load uncertainty usually exists as it is difficult to forecast the load with the perfect accuracy. Usually, such uncertainties are presented by assuming that the load forecast error follows certain probabilistic distributions. However, this is an ideal assumption, which may not reflect actual situations [25, 26]. Also, the uncertain load with different time scales are correlated [16]. Therefore, it should be better to characterize the load uncertainty, while considering the correlations among different time periods. To this end, we adopt a data-driven DPMM by leveraging historic load to better represent the load uncertainty in our work. Details of the data-driven DPMM is presented as follows.

3.1. Dirichlet Process with Stick-Breaking Representation

DPMM is based on Dirichlet Process (DP), which is described via a collection of base probabilistic distributions and their corresponding weights [27]. DP could be expressed explicitly by the well-known stick-breaking representation via discrete distributions [28]. A discrete distribution G is defined to be DP distributed if

$$G\,|\{\varphi, H\} \sim \mathrm{DP}(\varphi, H) \tag{2}$$

where φ represents a concentration parameter, and H stands for the base distribution parameterized by matrix Λ. We consider two collections of independently generated stochastic variables β_m and Ω_m with respect to φ and H.

$$\beta_m \sim \mathrm{Beta}(\beta_m|1, \varphi) \tag{3}$$

$$\Omega_m \sim H\,(\Omega_m|\Lambda) \tag{4}$$

where $m = \{1, \cdots, \infty\}$. It should be mentioned that variable set $\{\beta_m\}$ are parameters for (8), and $\mathrm{Beta}(\beta_m|1, \varphi)$ represents Beta distribution, which is expressed as follows.

$$\mathrm{Beta}(\beta_m|1, \varphi) = \frac{\Gamma(1+\varphi)}{\Gamma(1)\Gamma(\varphi)}\beta_m^0(1-\beta_m)^{\varphi-1} \tag{5}$$

where, $\Gamma(\cdot)$ is the Gamma function. Then, we introduce an impulse function δ_{Ω_m} as the indicator, to distinguish whether the distribution parametrized by Ω_m is chosen by G.

$$\delta_{\Omega_m}(\omega) = \begin{cases} 1, & \omega{=}\Omega_m \\ 0, & \text{elsewhere} \end{cases} \tag{6}$$

where ω is the continuous variable of discrete distribution G. Then, the stick-breaking representation of G is defined as follows.

$$G = \sum_{m=1}^{\infty} \pi_m \delta_{\Omega_m} \tag{7}$$

$$\pi_m = \beta_m \prod_{s=1}^{m-1} (1-\beta_s) \tag{8}$$

where $\sum_{m=1}^{\infty} \pi_m = 1$. In the process of breaking a unit length stick into an infinite number of pieces, π_m could be considered as the length of m-th piece regarding the breaking proportion β_m. In summary, G is a distribution indicating whether $H\,(\Omega_m|\Lambda)$ weighted by π_m is selected.

3.2. DPMM for Multi-Period Load

Indeed, according to the historical multi-period load data, DPMM is able to update its mixture components determined by $\{\pi_m\}$ and $\{\Omega_m\}$ without using the prior knowledge. By adopting a T-dimensional vector $\boldsymbol{\varepsilon} = \{\varepsilon_t\}$ to represent random load variables, the probability density function (PDF) of $\boldsymbol{\varepsilon}$ is formulated as follows.

$$p(\boldsymbol{\varepsilon}|\boldsymbol{\pi}, \boldsymbol{\Omega}) = \sum_{m=1}^{\infty} \pi_m p_m(\boldsymbol{\varepsilon}|\Omega_m) \tag{9}$$

where $\boldsymbol{\pi} = \{\pi_m\}_{m=1}^{\infty}$ and $\boldsymbol{\Omega} = \{\Omega_m\}_{m=1}^{\infty}$, $p_m(\boldsymbol{\varepsilon}|\Omega_m)$ stands for the PDF of the mixture component m, and π_m indicates the corresponding weight.

The most crucial purpose of introducing DPMM is to characterize load uncertainties and their correlations among multiple time periods. To this end, we set the prior distribution of H as a T-dimensional Gaussian distribution ($T = 24$), since it could indicate the correlations of multi-period loads through covariance matrices [16]. Hence, the DPMM can be expressed as an infinite set of 24-dimensional Gaussian mixture distribution.

$$p(\boldsymbol{\varepsilon}|\boldsymbol{\pi}, \boldsymbol{\Omega}) = \sum_{m=1}^{\infty} \pi_m \mathbb{N}_m(\boldsymbol{\varepsilon}|\mu_m, \Sigma_m) \tag{10}$$

in which $\boldsymbol{\Omega} = \{\mu_m, \Sigma_m\}_{m=1}^{\infty}$, where μ_m is the mean vector, Σ_m is the covariance matrix, and $\mathbb{N}_m(\boldsymbol{\varepsilon}|\mu_m, \Sigma_m)$ represents the 24-dimensional Gaussian component m.

Then, we consider a set of observed load dataset $\mathscr{E} = \{\boldsymbol{\varepsilon}^{[r]}\}_{r=1}^{R}$, $\boldsymbol{\varepsilon}^{[r]}$ is the independent r-th sample, $r \in \{1, \cdots, R\}$, and R is the number of collected samples. Therefore, the DPMM on the observed data can be expressed as follows.

$$p(\mathscr{E}|\boldsymbol{\pi}, \boldsymbol{\Omega}) = \prod_{r=1}^{R} \left\{ \sum_{m=1}^{\infty} \pi_m \mathbb{N}_m(\boldsymbol{\varepsilon}^{[r]}|\mu_m, \Sigma_m) \right\} \tag{11}$$

However, it is unknown that which component the individual sample $\boldsymbol{\varepsilon}^{[r]}$ belongs to. Therefore, the binary variable vector $\boldsymbol{z}^{[r]} = [z_1^{[r]}, \cdots, z_m^{[r]}, \cdots]$ is introduced to indicate which mixture component that the samples are associated with. For instance, $z_m^{[r]} = 1$ if load sample $\boldsymbol{\varepsilon}^{[r]}$ belongs to the m-th component. Then, given the indicating variable set $\boldsymbol{Z} = \{z_m^{[r]}\}_{r=1,m=1}^{r=R,m=\infty}$ for all samples and the parameter set $\boldsymbol{\Omega}$ for all components, the conditional PDF of $\mathscr{E}$ is formulated as follows.

$$p(\mathscr{E}|\boldsymbol{Z}, \boldsymbol{\Omega}) = \prod_{r=1}^{R} \prod_{m=1}^{\infty} \mathbb{N}_m(\boldsymbol{\varepsilon}^{[r]}|\mu_m, \Sigma_m)^{z_m^{[r]}} \tag{12}$$

Conventionally, the parameter of prior distribution could be estimated according to statistical characteristics of the observed data. However, it is difficult to directly determine the hyperparameters, i.e., the latent variable set $\mathscr{W} = \{\boldsymbol{Z}, \boldsymbol{\Omega}, \boldsymbol{\beta}, \boldsymbol{\Psi}\}$, where $\boldsymbol{\beta} = \{\beta_m\}_{m=1}^{\infty}$ and $\boldsymbol{\Psi} = \{\varphi_m\}_{m=1}^{\infty}$. In order to address such a dilemma, the variational inference (VI) is adopted to estimate the posterior probability of latent variables $\mathscr{W}$. The VI provides a deterministic methodology to approximate the posterior PDF $p(\mathscr{W}|\mathscr{E})$ of latent variable set $\mathscr{W}$ [29], by converting the posterior PDF computation to a numerical optimization of Kullback-Leibler (KL) divergence, which is usually deemed as a measure of difference between two probability distributions [30]. Detailed deductions and formulations of VI can be referred to Appendix A.

In brief, VI assumes a group of variational posteriors over its latent variables, and it finds the closest variational posterior $\hat{p}(\mathscr{W})$ to the true posterior $p(\mathscr{W}|\mathscr{E})$. The KL divergence is used to measure the distance from the true posterior $p(\mathscr{W}|\mathscr{E})$ to $\hat{p}(\mathscr{W})$ i.e.,

$$\begin{aligned} & D_{KL}(\hat{p}(\mathscr{W})\|p(\mathscr{W}|\mathscr{E})) \\ =&\mathbb{E}_{\hat{p}(\mathscr{W})}[\ln \hat{p}(\mathscr{W}) - \ln p(\mathscr{W}|\mathscr{E})] \\ =&\mathbb{E}_{\hat{p}(\mathscr{W})}[\ln \hat{p}(\mathscr{W}) - \ln p(\mathscr{W},\mathscr{E})] + \ln p(\mathscr{E}) \end{aligned} \tag{13}$$

Then, we can obtain the closest distribution $\hat{p}(\mathscr{W})$ to $p(\mathscr{W}|\mathscr{E})$ via $\arg\min_{\hat{p}(\mathscr{W})} D_{KL}(\hat{p}(\mathscr{W})\|p(\mathscr{W}|\mathscr{E}))$. In this way, the components and corresponding weights involved in DPMM are determined according to the posterior of latent variable set $\mathscr{W}$.

The dataset in this paper is obtained based on the scaling load data in the Energy Market Company, which is the system operator of Singapore's wholesale electricity market. The better performance of DPMM, compared with traditional GMM, will be presented in Section 5.

3.3. Sampling from DPMM

Indeed, the VI algorithm for the DPMM parameter estimation could determine a truncated distribution with the optimal number of components. Therefore, a fixed maximum number of components $\hat{M}$, i.e., the number of stick-breaking pieces, should be determined in advance. Then, the actual number of component M ($M \le \hat{M}$) is obtained through the inferring process based on real data, accordingly.

In this way, the multi-period load information could be represented by a multivariate Gaussian mixture distribution. However, it is difficult to generate samples from a mixture distribution directly [31]. Therefore, we aim to individually sample from each Gaussian component of the DPMM, and aggregate the samples altogether.

With this consideration, S probabilistic scenarios of load are obtained from the final DPMM $\sum_{m=1}^{M} \pi_m \mathbb{N}_m(\cdot)$, by the scenario-reduction based Monte-Carlo simulation [32]. Specifically, we first assign each Gaussian component $\mathbb{N}_m$ with the number of scenarios $S_m = \pi_m S$, where $m = 1, 2, \cdots, M$. Afterwards, S_m load samples are generated from the $\mathbb{N}_m$, which are denoted by a load scenario set $\boldsymbol{\xi}_m = \{\xi_1, \cdots, \xi_{s_m}, \cdots, \xi_{S_m}\}$. s_m is the scenario index corresponding to $\boldsymbol{\xi}_m$. Finally, all sample sets are collected together, forming a total load scenario set $\boldsymbol{\xi}_{\text{DPMM}} = \{\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_m, \cdots, \boldsymbol{\xi}_M\}$.

4. WCVAR Based Two-step Risk Averse Stochastic Optimization for MCM

As stated in Section 2, after submitting bidding offers for both generation and demand sides, PSO conduct the day-ahead MC in the wholesale market. It is difficult to determine the optimal day-ahead schedule on power generation of units and load curtailments of load providers accurately, due to the existence of load uncertainty. Therefore, we use a two-step stochastic optimization approach to solve this problem. The first step aims to obtain the deterministic dispatch and the unit commitment states based on the forecast information. The second step is to adjust the decisions after realization of probabilistic scenario sets by maximizing the expectation for social welfare. It should be mentioned that lower welfares in certain scenarios are deemed as the economic risk, and thus the risk aversion is taken into consideration in our work.

It is notable that the load uncertainty is represented by the data-driven DPMM, which is realized by weighted multiple scenario sets regarding to the Gaussian mixture components, as stated in Section 3. Therefore, the probabilistic characteristics with respect to the scenario sets are different, and economic risks regarding such scenario sets should be well managed. To this end, we introduce the worst-case conditional value at risk (WCVaR) as the risk indicator. It represent the CVaR against the worst-case scenario set, which is more appropriate to solve the risk averse MCM optimization problem regarding to the DPMM. Therefore, we propose a WCVaR based risk averse MCM using the two-step stochastic optimization, which is another contribution of our work. In the following parts, we will present the details of WCVaR.

4.1. Introduction of WCVaR

CVaR is widely used when the uncertainties are associated with standard probabilistic distributions. Given a confident level α, CVaR is usually defined as the expected value of the profit that is lower than the $(1-\alpha)$-quantile of profit distribution. The specific formulation of CVaR can be referred to Appendix B of this paper.

With the help of discretization based on probability scenarios, an auxiliary function $\mathscr{F}_\alpha$ is introduced [33] to facilitate the solution of CVaR, which is expressed as follows.

$$\mathscr{F}_\alpha(x,\eta) = \eta - \frac{1}{1-\alpha}\sum_s p_s \cdot [\eta - f(x,\xi_s)]^+ \tag{14}$$

where $[\eta - f(x,\xi_s)]^+ = \max\{\eta - f(x,\xi_s), 0\}$, $f(x,\xi_s)$ is the profit function of variable x under scenario ξ_s, and its occurrence probability is p_s. η is an auxiliary variable indicating the threshold that the probability of profit is less than or equal to $1-\alpha$. Then, CVaR can be calculated via the following formula.

$$\Phi_\alpha(x) \equiv \text{CVaR}_\alpha(x) = \max_{\eta\in\mathbb{R}} \mathscr{F}_\alpha(x,\eta) \tag{15}$$

In our work, the multivariate Gaussian mixture distribution P_M with M components is adopted to represent the PDF of uncertain loads. It is obtained from DPMM in Section 3 and expressed as follows.

$$P_M = \left\{ \sum_m^M \pi_m p_m(\cdot) \middle| \sum_m^1 \pi_m = 1,\ \pi_m \ge 0, m = \{1,2,...,M\} \right\} \tag{16}$$

where $p_m(\cdot)$ represents the PDF of multivariate Gaussian mixture component m, and π_m is the corresponding weight. WCVaR is defined as the worst-case CVaR value among all mixture components [34]:

$$\text{WCVaR}_\alpha(x) = \min_{p_m(\cdot)\in P_M} \Phi_\alpha(x) = \min_{p_m(\cdot)\in P_M} \max_{\eta\in\mathbb{R}} \mathscr{F}_\alpha(x,\eta) \tag{17}$$

4.2. Two-step Risk Averse MCM

With the adoption of the WCVaR, our proposed two-step risk averse MCM is formulated as follows.

$$\max_{\boldsymbol{x},\Delta\boldsymbol{x}} \left\{ \overbrace{F(\boldsymbol{x})}^{\text{First step}} + \overbrace{\min_{p_m(\cdot)\in P_M} \Big[(1-\rho)\cdot F_m(\Delta\boldsymbol{x}|\boldsymbol{\xi}_m) + \rho \cdot \Phi_{\alpha,m}(\Delta\boldsymbol{x}|\boldsymbol{\xi}_m) \Big]}^{\text{Second step}} \right\} \tag{18}$$

where $F(\boldsymbol{x})$ is the social welfare of the first step optimization with respect to decision $\boldsymbol{x}$. $\Delta\boldsymbol{x}$ represents the decision adjustment regarding the scenario set $\boldsymbol{\xi}_m$. $F_m(\Delta\boldsymbol{x}|\boldsymbol{\xi}_m)$ is the expected social welfare regarding the Gaussian mixture component m. $\Phi_{\alpha,m}(\Delta\boldsymbol{x}|\boldsymbol{\xi}_m)$ could be deemed as the expected value that is lower than the $(1-\alpha)$-quantile of social welfare distribution with respect to $\boldsymbol{\xi}_m$. ρ is the risk aversion factor in the range of $[0, 1)$. The second step optimization is a min-max problem, seeking to maximize the expected social welfare while minimizing the risk under the worst case. It can be formulated into a mixed-integer linear programming problem, which can be efficiently solved by the solver of Gurobi [35].

4.3. First Step Optimization of MCM

In order to better describe the optimization of MCM, Figure 3 illustrates a double auction MC process regarding the first-step optimization in the time period t. To be specific, there are two stepwise curves. One in the ascending order is the sorted bidding offer curve on the generation side, and the other in the descending order is the sorted bidding offer curve on the demand side. Market clearing price is determined as the price on the generation side when the supply and demand are balanced in the electricity market [36].

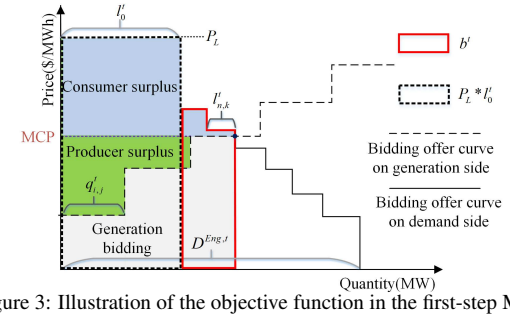

Figure 3: Illustration of the objective function in the first-step MCM

The shaded area denotes the social welfare of the market, which is normally defined as the summation of the producer and consumer surplus. The first-step problem is to maximize the social welfare based on the forecast information of load, which is formulated as follows.

$$\max_{\boldsymbol{x}} F(\boldsymbol{x}) = \sum_{t=1}^{T}\left(P_L \times l_0^t + b^t - g^t\right) \quad \text{s.t.} \quad (22) \sim (34) \tag{19}$$

in which

$$b^t = \sum_{n=1}^{N}\sum_{k=1}^{K_n}\left(LP_{n,k} \times l_{n,k}^t\right) \tag{20}$$

$$g^t = \sum_{i=1}^{I}\Bigg[\sum_{j=1}^{J_i^{Eng}}\left(P_{i,j}^{Eng} \times q_{i,j}^{Eng,t}\right) + \sum_{j=1}^{J_i^{Res}}\left(P_{i,j}^{Res} \times q_{i,j}^{Res,t}\right) + \sum_{j=1}^{J_i^{Reg}}\left(P_{i,j}^{Reg} \times q_{i,j}^{Reg,t}\right)\Bigg] \tag{21}$$

In (19), $\boldsymbol{x}$ is the decision variables set $\{q_{i,j}^{Eng,t}, q_{i,j}^{Res,t}, q_{i,j}^{Reg,t}, \tau_i^{Reg,t}, l_0^t, l_{n,k}^t\}$. Specifically, i is the index of bidding offers and j is the index for price-quantity tranches on the generation side. I is the number of bidding offers (units). J_i^{Eng}, J_i^{Res} and J_i^{Reg} are numbers of tranches of bidding offer i for the energy, reserve, and regulation, respectively. $q_{i,j}^{Eng,t}$, $q_{i,j}^{Res,t}$ and $q_{i,j}^{Reg,t}$ denote the actual quantity of tranche j of bidding offer i for the three products accepted by PSO in the time period t. l_0^t stands for the dispatched system non-curtailable load, and $l_{n,k}^t$ is the accepted quantity of the price-quantity tranche k of bidding offer n.

In addition, T is the number of dispatch time periods, and P_L is a relatively high price corresponding to the non-curtailable load, which would not be influenced by DR. b^t represents the actual energy cost in the time period t for the dispatched curtailable load, as expressed in (20), where $LP_{n,k}$ is the bidding price for tranche k of offer n. g^t is the generation cost calculated by (21), where $P_{i,j}^{Eng}$, $P_{i,j}^{Res}$ and $P_{i,j}^{Eng}$ are the bidding energy, reserve and regulation price for tranche j of offer i, respectively. Meanwhile, the following constraints should be satisfied.

1) Constraints for Generator Capacity and Ramping Limits

The total amount of energy, reserve, and regulation to be dispatched by each generation unit should not exceed its generation capacity, as shown in (22). Meanwhile, the ramping constraint is presented in (23).

$$\sum_{j=1}^{J_i^{Eng}} q_{i,j}^{Eng,t} + \sum_{j=1}^{J_i^{Res}} q_{i,j}^{Res,t} + \sum_{j=1}^{J_i^{Reg}} q_{i,j}^{Reg,t} \le O_i \quad \forall i,t \tag{22}$$

$$-\mathrm{RD}_i \le \sum_{j=1}^{J_i^{Eng}}\left(q_{i,j}^{Eng,t} - q_{i,j}^{Eng,t-1}\right) \le \mathrm{RU}_i \quad \forall i,t \tag{23}$$

where O_i is the capacity for the generation unit regarding the bidding offer i. RU_i and RD_i are the corresponding ramping up and down rates, respectively.

2) Constraints for Energy

The dispatched energy should match the demand as expressed in (24). Constraints (25)~(26) limit the dispatched energy and curtailment quantity of each tranche. (27) denotes the boundaries for the dispatched non-curtailable load.

$$\sum_{i=1}^{I}\sum_{j=1}^{J_i^{Eng}} q_{i,j}^{Eng,t} = l_0^t + \sum_{n=1}^{N}\sum_{k=1}^{K_n} l_{n,k}^t \quad \forall n,k,t \tag{24}$$

$$0 \le q_{i,j}^{Eng,t} \le Q_{i,j}^{Eng} \quad \forall i,j,t \tag{25}$$

$$0 \le l_{n,k}^t \le LQ_{n,k}^t \quad \forall n,k,t \tag{26}$$

$$0 \le l_0^t \le SNC^t \quad \forall n,k,t \tag{27}$$

where $Q_{i,j}^{Eng}$ is the energy bidding quantity for tranche j of offer i. $LQ_{n,k}^t$ denotes the curtailable load bidding quantity for tranche k of offer n, in the time period t. It is notable that SNC^t is calculated via (1) based on the forecast load. Also, the quantity of load curtailment is relatively large when the cleared wholesale electricity price (WEP) is high [37].

3) Constraints for Regulation

Similar to the energy balance in (24), the regulation is balanced with the demand $D^{Reg,t}$, as expressed in (28). Constraints (29) and (30) impose limits on the unit for regulation: (i) the quantity of dispatched energy should be no less than its dispatched regulation by a pre-defined minimum output Reg_i^{Min}; (ii) the total quantity of dispatched energy and regulation should be no more than a pre-defined maximum output

Reg_i^{Max}. Note that the two constraints are valid only when the unit is qualified and dispatched for regulations, as indicated by the binary variable $\tau_i^{Reg,t}$. Similarly, constraint (31) expresses that regulation should not exceed the bidding quantity $Q_{i,j}^{Reg}$ for each tranche.

$$\sum_{i=1}^{I}\sum_{j=1}^{J_i^{Reg}} q_{i,j}^{Reg,t} = D^{Reg,t} \quad \forall t \tag{28}$$

$$\left(\sum_{j=1}^{J_i^{Eng}} q_{i,j}^{Eng,t} - \sum_{j=1}^{J_i^{Reg}} q_{i,j}^{Reg,t}\right) \times \tau_i^{Reg,t} \geq Reg_i^{Min} \quad \forall i,t \tag{29}$$

$$\left(\sum_{j=1}^{J_i^{Eng}} q_{i,j}^{Eng,t} + \sum_{j=1}^{J_i^{Reg}} q_{i,j}^{Reg,t}\right) \times \tau_i^{Reg,t} \leq Reg_i^{Max} \quad \forall i,t \tag{30}$$

$$0 \leq q_{i,j}^{Reg,t} \leq Q_{i,j}^{Reg} \quad \forall i,j,t \tag{31}$$

4) Constraints for Reserve

Different from the energy and regulation, the reserve capacity should be sufficient to cover the power losses due to failure of some generators. It may be further scaled by the risk adjustment factor λ_1 (set by the PSO) to ensure a more secure power supply, as enforced in (32). On the contrary, constraint (33) prevents the situation that the excessive reserve is provided by a unit. That is, it is required that such reserve should not exceed a certain percentage, i.e., λ_2 of the dispatched energy for economic concerns. In addition, the reserve to be dispatched of each unit is bounded by the bidding quantity $Q_{i,j}^{Res}$, as shown in (34).

$$\sum_{i=1}^{I}\sum_{j=1}^{J_i^{Res}} q_{i,j}^{Res,t} \geq \lambda_1 \times \max_{i=1,\ldots,I}\left\{\sum_{j=1}^{J_i^{Eng}} q_{i,j}^{Eng,t} + \sum_{j=1}^{J_i^{Res}} q_{i,j}^{Res,t}\right\} \forall t \tag{32}$$

$$\sum_{j=1}^{J_i^{Res}} q_{i,j}^{Res,t} \leq \lambda_2 \times \sum_{j=1}^{J_i^{Eng}} q_{i,j}^{Eng,t} \quad \forall i,t \tag{33}$$

$$0 \leq q_{i,j}^{Res,t} \leq Q_{i,j}^{Res} \quad \forall i,j,t \tag{34}$$

4.4. Second Step Optimization With WCVaR

After the realization of load uncertainty with probabilistic scenario sets, the adjustments based on dispatch decisions of the first step are conducted considering maximizing the expected social welfare under the worst scenario set $\boldsymbol{\xi}_m$. By introducing an auxiliary variable ζ_{s_m} for each scenario ξ_{s_m}, the second-step problem could be transform into the following solvable form.

$$\begin{aligned}\min_{p_m(\cdot)\in P_M}\Bigg\{\max_{\boldsymbol{x},\Delta\boldsymbol{x},\eta}\Bigg[&(1-\rho)\cdot\sum_{t=1}^{T}\left(P_L\times\Delta l_{s_m}^t+\Delta b_{s_m}^t-\Delta g_{s_m}^t\right)\\&+\rho\cdot\left(\eta-\frac{1}{1-\alpha}\sum_{s_m}p_{s_m}\cdot\zeta_{s_m}\right)\Bigg]\Bigg\}\\ \text{s.t.}\quad&\zeta_{s_m}\geq\eta-\sum_{t=1}^{T}\left(P_L\times\Delta l_{s_m}^t+\Delta b_{s_m}^t-\Delta g_{s_m}^t\right)\\&\zeta_{s_m}\geq 0\\&(39)\sim(51)\end{aligned} \tag{35}$$

where $\Delta l_{s_m}^t$ is the dispatched non-curtailable load with respect to scenario ξ_{s_m}. $\Delta b_{s_m}^t$ is defined as the corresponding deviation of b^t in scenario ξ_{s_m}. $\Delta g_{s_m}^t$ represents the generation cost deviation associated with the second step. They are expressed as (36) and (37), respectively.

$$\Delta b_{s_m}^t = \sum_{n=1}^{N}\sum_{k=1}^{K_n}\left(LP_{n,k}\times\Delta l_{n,k,s_m}^t\right) \tag{36}$$

$$\begin{aligned}\Delta g_{s_m}^t = \sum_{i=1}^{I}\Bigg[&\sum_{j=1}^{J_i^{Eng}}\left(P_{i,j}^{Eng}\times\Delta q_{i,j,s_m}^{Eng,t}\right)+\sum_{j=1}^{J_i^{Res}}\left(P_{i,j}^{Res}\times\right.\\&\left.\Delta q_{i,j,s_m}^{Res,t}\right)+\sum_{j=1}^{J_i^{Reg}}\left(P_{i,j}^{Reg}\times\Delta q_{i,j,s_m}^{Reg,t}\right)\Bigg]\end{aligned} \tag{37}$$

where the variable set $\{\Delta q_{i,j,s_m}^{Eng,t}, \Delta q_{i,j,s_m}^{Res,t}, \Delta q_{i,j,s_m}^{Reg,t}, \Delta l_{n,k,s_m}^t\}$ are adjustments for each scenario ξ_{s_m} in the second step regarding $\{q_{i,j}^{Eng,t}, q_{i,j}^{Res,t}, q_{i,j}^{Reg,t}, l_{n,k}^t\}$. To facilitate the mathematical description, the $SNC_{s_m}^t$ for each scenario ξ_{s_m} in the time period t is introduced as follows.

$$SNC_{s_m}^t = D_{s_m}^{Eng,t} - \sum_{n=1}^{N}\sum_{k=1}^{K} LQ_{n,k}^t \quad \forall s_m,t \tag{38}$$

It should be notable that $D_{s_m}^{Eng,t}$ is the energy demand for scenario ξ_{s_m}. Constraints (39)-(51) are counterparts of (22)-(34) in the second-step stochastic optimization, presented as follows.

$$\begin{aligned}&\sum_{j=1}^{J_i^{Eng}}\left(q_{i,j}^{Eng,t}+\Delta q_{i,j,s_m}^{Eng,t}\right)+\sum_{j=1}^{J_i^{Res}}\left(q_{i,j}^{Res,t}+\Delta q_{i,j,s_m}^{Res,t}\right)\\&\quad+\sum_{j=1}^{J_i^{Reg}}\left(q_{i,j}^{Reg,t}+\Delta q_{i,j,s_m}^{Reg,t}\right)\leq O_i \quad \forall i,s_m,t\end{aligned} \tag{39}$$

$$\begin{aligned}-\mathrm{RD}_{i,j}\leq\sum_{j=1}^{J_i^{Eng}}\Big[&(q_{i,j}^{Eng,t}+\Delta q_{i,j,s_m}^{Eng,t})-(q_{i,j}^{Eng,t-1}+\\&\Delta q_{i,j,s_m}^{Eng,t-1})\Big]\leq\mathrm{RU}_{i,j}\quad\forall i,s_m,t\end{aligned} \tag{40}$$

$$\begin{aligned}\sum_{i=1}^{I}\sum_{j=1}^{J_i^{Eng}}\left(q_{i,j}^{Eng,t}+\Delta q_{i,j,s_m}^{Eng,t}\right)=(l_0^t+l_{s_m}^t)+\sum_{n=1}^{N}\sum_{k=1}^{K_n}\\\left(l_{n,k}^t+\Delta l_{n,k,s_m}^t\right)\qquad\forall s_m,t\end{aligned} \tag{41}$$

$$0\leq\left(q_{i,j}^{Eng,t}+\Delta q_{i,j,s_m}^{Eng,t}\right)\leq Q_{i,j}^{Eng}\qquad\forall i,j,s_m,t \tag{42}$$

$$0\leq l_{n,k}^t+\Delta l_{n,k,s_m}^t\leq LQ_{n,j}^t\qquad\forall n,k,s_m,t \tag{43}$$

$$0\leq l_0^t+l_{s_m}^t\leq SNC_{s_m}^t\quad\forall s_m,t \tag{44}$$

$$\sum_{i=1}^{I}\sum_{j=1}^{J_i^{Reg}}\left(q_{i,j}^{Reg,t}+\Delta q_{i,j,s_m}^{Reg,t}\right)=D^{Reg,t}\qquad\forall s_m,t \tag{45}$$

$$\begin{aligned}\left\{\sum_{j=1}^{J_i^{Eng}}\left(q_{i,j}^{Eng,t}+\Delta q_{i,j,s_m}^{Eng,t}\right)-\sum_{j=1}^{J_i^{Reg}}\left(q_{i,j}^{Reg,t}+\Delta q_{i,j,s_m}^{Reg,t}\right)\right\}\times\tau_i^{Reg,t}\\\geq Reg_i^{Min}\quad\forall i,s_m,t\end{aligned} \tag{46}$$

$$\begin{aligned}\left\{\sum_{j=1}^{J_i^{Eng}}\left(q_{i,j}^{Eng,t}+\Delta q_{i,j,s_m}^{Eng,t}\right)+\sum_{j=1}^{J_i^{Reg}}\left(q_{i,j}^{Reg,t}+\Delta q_{i,j,s_m}^{Reg,t}\right)\right\}\times\tau_i^{Reg,t}\\\leq Reg_i^{Max}\quad\forall i,s_m,t\end{aligned} \tag{47}$$

$$0\leq\left(q_{i,j}^{Reg,t}+\Delta q_{i,j,s_m}^{Reg,t}\right)\leq Q_{i,j}^{Reg}\qquad\forall i,j,s_m,t \tag{48}$$

$$\sum_{i=1}^{I}\sum_{j=1}^{J_i^{Res}}\left(q_{i,j}^{Res,t}+\Delta q_{i,j,s_m}^{Res,t}\right)\geq\lambda_1\times\max_{i=1,\ldots,I}\Bigg\{\sum_{j=1}^{J_i^{Eng}}\Big(q_{i,j}^{Eng,t}+\Delta q_{i,j,s_m}^{Eng,t}\Big)+\sum_{j=1}^{J_i^{Res}}\left(q_{i,j}^{Res,t}+\Delta q_{i,j,s_m}^{Res,t}\right)\Bigg\}\ \forall s_m,t \tag{49}$$

$$\sum_{j=1}^{J_i^{Res}}\left(q_{i,j}^{Res,t}+\Delta q_{i,j,s_m}^{Res,t}\right)\leq\lambda_2\times\sum_{j=1}^{J_i^{Eng}}\left(q_{i,j}^{Eng,t}+\Delta q_{i,j,s_m}^{Eng,t}\right)\quad\forall i,s_m,t \tag{50}$$

$$0\leq\left(q_{i,j}^{Res,t}+\Delta q_{i,j,s_m}^{Res,t}\right)\leq Q_{i,j}^{Res}\qquad\forall i,j,s_m,t \tag{51}$$

Specifically, generation limits and ramping constraints of units are enforced by (39) and (40), respectively. Constraint (41) represents the energy balance, and (42)~(43) describe bounds on generation and demand side variables in the second step. Constraint (44) is the boundary limit of dispatched non-curtailable load. Constraint (45) represents the regulation balance after the adjustment of second step. The limitations on regulation of generating units are shown in (46) and (47). The overall reserve requirement and proportion constraint are presented in (49) and (50), respectively. In addition, (48) and (51) are bounds of regulation and reserve quantities, respectively.

5. Simulation and Results

In this section, the simulation case is set firstly in Subsection 5.1. Performance of load uncertainty represented by DPMM is then presented in Subsection 5.2. Finally, the results of WCVaR based risk averse MCM is shown in Subsection 5.3, and the impact of risk aversion factor is analyzed in Subsection 5.4.

5.1. Simulation Case Setup

In this study, the real electricity load data are collected from Energy Market Company [37] in Singapore electricity market ranging from 1st Aug, 2016 to 31st Dec, 2019. Without loss of generality, the load forecast information on 17th, Mar. 2020 is used to conduct the first step optimization.

Table 1: Bidding Offers on Generation Side

i	O_i (MW)	Reg_i^{Min} (MW)	Reg_i^{Max} (MW)	$Q_{i,1}^{Eng}$ (MW)	$Q_{i,2}^{Eng}$ (MW)	$Q_{i,1}^{Res}$ (MW)	$Q_{i,2}^{Res}$ (MW)
1	330	50	290	150	150	40	20
2	430	45	350	180	200	40	30
3	350	80	300	180	140	60	40
4	340	95	320	170	160	80	50
5	290	80	290	150	120	70	40
6	240	45	220	130	110	20	30
7	260	40	180	90	100	50	25
8	400	90	250	135	130	60	60

$Q_{i,1}^{Reg}$ (MW)	$Q_{i,2}^{Reg}$ (MW)	$P_{i,1}^{Eng}$ ($/MWh)	$P_{i,2}^{Eng}$ ($/MWh)	$P_{i,1}^{Res}$ ($/MWh)	$P_{i,2}^{Res}$ ($/MWh)	$P_{i,1}^{Reg}$ ($/MWh)	$P_{i,2}^{Reg}$ ($/MWh)
10	15	50	105	28.5	33	19	40
10	10	65	110	19.7	32.9	18	35
22	17	70	115	13	15	19	26
18	18	80	120	14	17	17	31
23	14	85	130	15	16	13	25
10	5	90	140	17	20	13	47
12	8	95	150	25	30	25	70
10	10	100	160	10	28	20	30

Bidding offers on the generation side are listed in Table 1, with 8 generation units (i.e., $I = 8$). Each bidding offer includes two price-quantity tranches for the energy, reserve, and regulation. The ramping rates $\mathrm{RD}_{i,j}$ and $\mathrm{RU}_{i,j}$ are set as 35% of the capacity of each unit per hour [38]. Also, the quantities of load curtailment tranches are listed in Table 2. Moreover, the bidding prices for each tranche on demand side $LP_{n,k}$ are listed in Table 3.

Table 2: Bidding Offers on Demand Side

Period	1	2	3	4	5	6	7	8	9	10	11	12
WEP ($/MWh)	94	89	85	80	79	88	95	97	107	113	117	114
$LQ_{1,1}^t$(MW)	0	0	0	0	0	0	0	0	3	7	12	9
$LQ_{1,2}^t$(MW)	0	0	0	0	0	0	0	0	2	6	9	7
$LQ_{2,1}^t$(MW)	0	0	0	0	0	0	0	0	2	5	8	6
$LQ_{2,2}^t$(MW)	0	0	0	0	0	0	0	0	1	3	5	3
$LQ_{3,1}^t$(MW)	0	0	0	0	0	0	0	0	2	6	11	8
$LQ_{3,2}^t$(MW)	0	0	0	0	0	0	0	0	2	6	9	7
Period	13	14	15	16	17	18	19	20	21	22	23	24
WEP ($/MWh)	111	115	118	129	132	113	108	113	109	105	100	96
$LQ_{1,1}^t$(MW)	6	10	14	33	40	7	4	8	5	2	0	0
$LQ_{1,2}^t$(MW)	5	8	10	25	30	6	3	6	3	2	0	0
$LQ_{2,1}^t$(MW)	4	6	8	21	25	5	2	5	3	1	0	0
$LQ_{2,2}^t$(MW)	2	4	5	12	15	3	1	3	2	1	0	0
$LQ_{3,1}^t$(MW)	5	9	12	29	35	6	3	7	4	2	0	0
$LQ_{3,2}^t$(MW)	5	8	10	25	30	6	3	6	3	2	0	0

Table 3: Demand Side Load Curtailment Price Bidding

Load provider	n=1		n=2		n=3	
Bidding tranche	k=1	k=2	k=1	k=2	k=1	k=2
Price ($/MWh)	150	125	145	140	120	115

The other simulation setting are shown as follows. The regulation demand $D^{Reg,t}$ is set to be 130 MW, and the price of inelastic demand P_L is a relatively high value as 200 $/MWh. The confidential level α is 0.9, and risk aversion factor ρ is set as 0.1 originally. The other parameters are set as: $\lambda_1 = 1.5$ and $\lambda_2 = 0.6$. In addition, the total number of load scenarios S is set as 200.

5.2. Representation of Uncertain Load

In this subsection, to verify the effectiveness of the data-driven DPMM, we compare its performance with that of traditional GMM. To this end, the DPMM as well as the compared traditional GMM are adopted to fit the scaled load data. Then, the PDFs of several hourly loads are shown in Figure 4. Specifically, GMM_1, GMM_2, and GMM_3 refer to the Gaussian mixture models with mixture component numbers of 1, 2, and 3, respectively. The VI based DPMM and GMM with different component numbers are compared visually, with the reference of histograms based on the original data set. It should be noted that the probabilistic models regarding the load are established based on 24-dimensional Gaussian distribution. Therefore, the whole models are difficult to be visualized due to the high-dimensional features. Based on this condition, the indication of probability density curves in Figure 4 could be deemed as one-dimensional mapping of multi-dimensional DPMM. It could be observed that DPMM demonstrates better performances in capturing the probabilistic characteristics of load information during several time periods. For instance, the multi-peak features of the load PDF profiles in the 9-th, 12-th, 15-th, 18-th and 21-st periods are well reflected.

However, the PDF curves of the 3-rd, 6-th and 24-th periods obtained by DPMM are with slight differences from those of GMM_1, GMM_2 and GMM_3 from visual observations. Therefore, the fitting performances should be conducted with quantitative analysis. Specifically, we introduce the log-likelihood

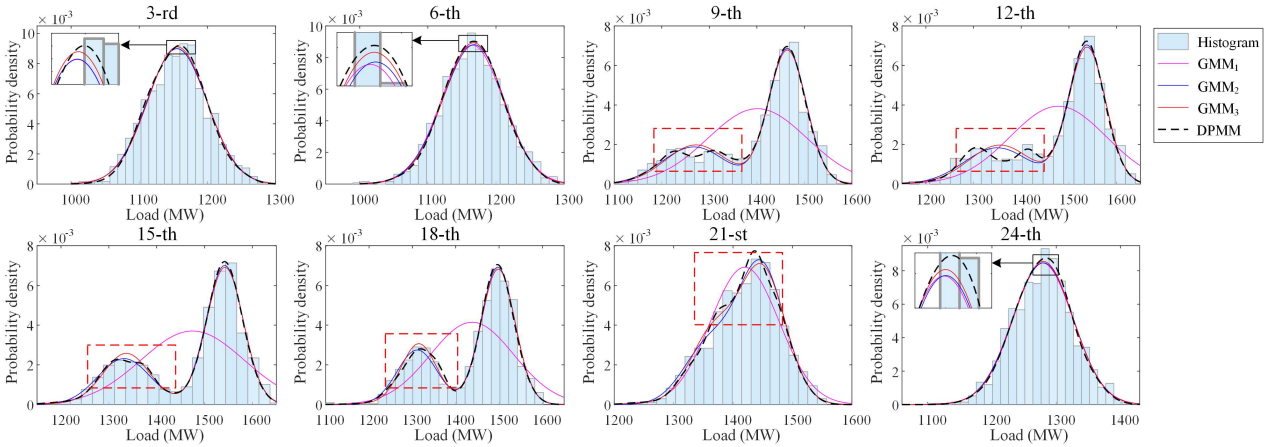

Figure 4: Probability density curves of different probabilistic models in different time periods

Table 4: Numerical Fitting on Load Dataset

	Log-L(10^3)	Chi-square GoF	Log-L(10^3)	Chi-square GoF	Log-L(10^3)	Chi-square GoF	Log-L(10^3)	Chi-square GoF	Log-L(10^3)	Chi-square GoF	Log-L(10^3)	Chi-square GoF	Log-L(10^3)	Chi-square GoF	Log-L(10^3)	Chi-square GoF
Period	3-rd		6-th		9-th		12-th		15-th		18-th		21-st		24-th	
GMM_1	-5.55	57.58	-5.57	40.66	-6.46	569.78	-6.42	513.03	-6.49	641.25	-6.37	524.8	-5.82	73.01	-5.56	50.31
GMM_2	-5.55	57.62	-5.57	42.39	-6.21	80.67	-6.18	54.7	-6.17	44.5	-6.11	66.01	-5.81	46.04	-5.61	50.93
GMM_3	-5.55	58.73	-5.57	42.99	-6.21	81.07	-6.19	56.02	-6.16	45.15	-6.11	66.23	-5.8	47.13	-5.61	52.28
DPMM	**-5.54**	**55.87**	**-5.57**	40.81	**-6.2**	**64.93**	**-6.17**	**29.89**	**-6.16**	**42.81**	**-6.11**	**65.02**	**-5.8**	**42.26**	**-5.61**	**46.97**

(Log-L) and the Chi-square goodness of fit (GoF) to quantify the differences. The larger Log-L value indicates a higher accuracy. For Chi-square GoF, the smaller value means a better fit. The Log-L and GoF values for GMM with different component numbers and DPMM are shown in Table 4. It is observed that Log-L values of DPMM for all periods are lager than such values of GMMs. Also, the Chi-square GoF values of DPMM is obviously smaller than those of GMMs except for the 40.81 in the 6-th period, which is slightly bigger than Chi-square GoF of GMM_1.

Moreover, in order to further evaluate the outperformance of incorporated algorithm, we compare DPMM with GMM, as well as K-means and the only clustering neural network based self-organizing map (SOM) algorithm. Nevertheless, implementing K-means and SOM could not obtain analytic functions. Therefore, it is difficult to visualize and evaluate them by statistical index such as Log-L and the Chi-square GoF. Therefore, we introduce four measurement indices, i.e., Silhouette Coefficient [39], Root-mean-square standard deviation (RMSSTD) [40], R-Square and Hubert's Γ statistic [41], to indicate the priority of DPMM. The specific formulations and explanations of above indices are in Appendix C. The comparison of performance measurement among the different algorithms are listed in Table 5.

Table 5: The Performance Measurement of Different Algorithm

Algorithm	Silhouette Coefficient	RMSSTD	R-Square	Hubert's Γ statistic
SOM	0.524	-	-	-
K-means	0.629	0.436	0.941	250718.249
GMM_1	-	0.687	0.712	-
GMM_2	0.791	0.518	0.907	216245.411
GMM_3	0.338	0.515	0.908	218940.656
DPMM	**0.792**	**0.413**	**0.961**	**263012.216**

In summary, it is easily observed that DPMM is evaluated to be best for the four measurements. Also, it should be notable that only Silhouette Coefficient is measured in SOM algorithm, since the cluster center is not analytical solvable. Moreover, the K-means performs well in measurements of RMSSTD, R-Square and Hubert's Γ statistic, yet worse than DPMM. When it comes to GMMs, Silhouette Coefficient of GMM_2 is relatively good, while GMM_3 performs better in other indices. In summary, DPMM shows a better ability to represent electricity load data. As a matter of fact, DPMM still outperforms through the quantitative analysis.

Then, the 24-period load probabilistic scenarios are generated from DPMM by the scenario-reduction based Monte-Carlo simulation. In this way, these scenarios are adopted in the two-step stochastic optimization, and the related results are in Subsection 5.3.

5.3. Simulation Results of the WCVaR Based Risk Averse MCM

In this subsection, in order to show the outperformance of the proposed WCVaR based MCM, we conduct simulations and obtain the results, compared with that of the CVaR based MCM. Note that the latter MCM is the two-step MCM adopting the risk indicator of CVaR based on all the probabilistic scenarios. Its objective function is shown in (D.1) of Appendix D.

The risk averse term $\Phi_{\alpha,m}(\Delta \boldsymbol{x}|\boldsymbol{\xi}_m)$ in (18) is adopted to reduce the economic risk regarding the social welfare. In other words, it seeks to optimize the expectation of lower social welfare with respect to specific scenarios under the worst case. To verify the advantages of adopting WCVaR as the risk indicator, we compare the performance of WCVaR based MCM with that of CVaR. The comparative results are presented in Figure 5, Figure 6 and Table 6, respectively.

One of the purposes of the day-ahead MC is to determine the electricity prices. Figure 5 shows the clearing price comparisons for the energy, reserve and regulation in each time period after the two-step MC. It is easily observed that the energy prices with WCVaR based MCM are generally lower than or equal to those of the CVaR. It indicates that WCVaR based MCM could achieve an overall lower day-ahead electricity price. In terms of the reserve, its corresponding clear-

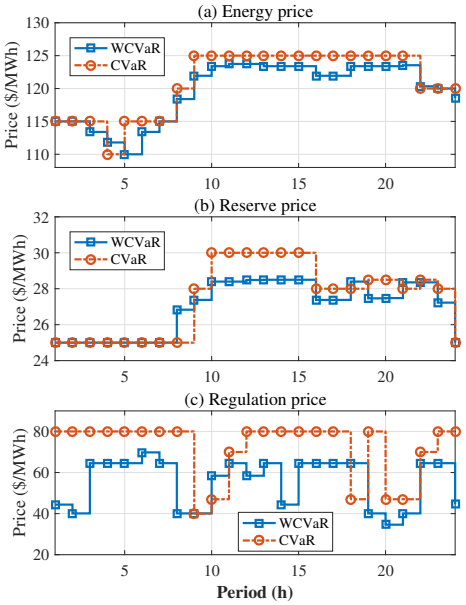

Figure 5: Day-ahead market clearing prices for (a) energy, (b) reserve and (c) regulation by comparing WCVaR and CVaR based MCMs

ing prices are no higher than except for the 9-th, 18-th and 21-st time periods. It reveals that the lower cost of reserve resource is required for the power system by adopting the WCVaR. Furthermore, the regulation prices with respect to CVaR based MCM are as high as 80 $/MWh for most periods. There are only two periods, i.e., the 10-th and 18-th, during which the WCVaR based results lead to higher prices. Therefore, the observations in Figure 6 reveal that lower clearing prices are obtained by the WCVaR based MCM.

Table 6: Dispatched Results under Different Tests

	Generation cost ($)	First-step welfare ($)	Two step expected welfare ($)
WCVaR based model	2.8326×10^6	1.7041×10^6	2.3161×10^6
CVaR based model	2.8352×10^6	2.1563×10^6	2.3154×10^6

On the basis of the clearing prices for energy, reserve and regulation, simulation results for generation cost and social welfare are obtained, which are listed in Table 6. On the one hand, the expected generation cost are lower in the WCVaR based MCM, which is 2.8326×10^6 $ compared with 2.8352×10^6 $ regarding that of CVaR. Which is because that the lower cleared prices are obtained by the WCVaR based MCM. On the other hand, it is notable that the corresponding social welfare of the first-step $F(\boldsymbol{x})$ is 1.7041×10^6 $, much lower than 2.1563×10^6 $ in the CVaR based MCM. It is because that the more robust dispatching decision is conducted in the first step, and more dispatching adjustments are used in the second-step with WCVaR based MCM, to overcome the economic risk brought by load uncertainties. Moreover, its social welfare of the two-step expectation is 2.3161 $\times10^6$ $, higher than 2.3154×10^6 $ regarding the CVaR based results. Therefore, it could be concluded from the simulation results of this part that WCVaR serves as a better risk indicator, due to the higher social welfare and lower economic risk it could obtain.

Furthermore, our proposed MCM is incorporated with a DR program, thus it is also necessary to analyze the dispatched load curtailment for load providers. As shown in Figure 6, the bars with bidding group represent the total quantities of all

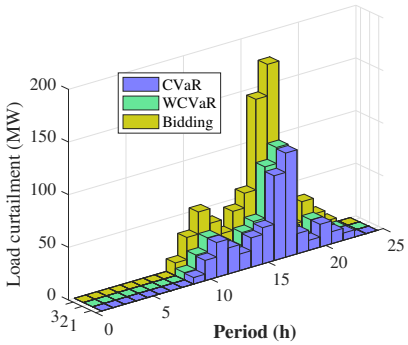

Figure 6: Accepted load curtailment comparison of WCVaR based and CVaR based results

tranches for bidding curtailable load in each time period. The other two groups are dispatched load curtailments regarding CVaR based and WCVaR based results, respectively. It is observed that the quantities of dispatched load curtailments for the WCVaR based MCM are more than CVaR based model in the 9-th~22-nd periods. This phenomenon indicates that the WCVaR based MCM tends to achieve more load curtailments.

5.4. Analysis of Risk Aversion Factor

In this subsection, the impacts of risk averse level on the dispatch results, including the expected social welfare, total generation cost, energy generation cost and curtailment percent, are elaborately analyzed, by adjusting the risk aversion factor at 0.1, 0.3, 0.5, 0.7, 0.9, respectively. The simulation results with different risk averse levels are presented in Table 7, and the relationship between the obtained results and the risk aversion level is illustrated in Figure 7.

Table 7: Results for Different Risk Aversion Factor

Risk averse level	Social welfare ($)	Generation cost ($)	Energy cost ($)	Curtailment percent (%)
$\rho = 0.1$	2.3161×10^6	2.8326×10^6	2.6381×10^6	60.21
$\rho = 0.3$	2.3161×10^6	2.8322×10^6	2.6379×10^6	60.47
$\rho = 0.5$	2.3159×10^6	2.8315×10^6	2.6371×10^6	61.20
$\rho = 0.7$	2.3157×10^6	2.8300×10^6	2.6356×10^6	62.87
$\rho = 0.9$	2.3137×10^6	2.8269×10^6	2.6318×10^6	64.30

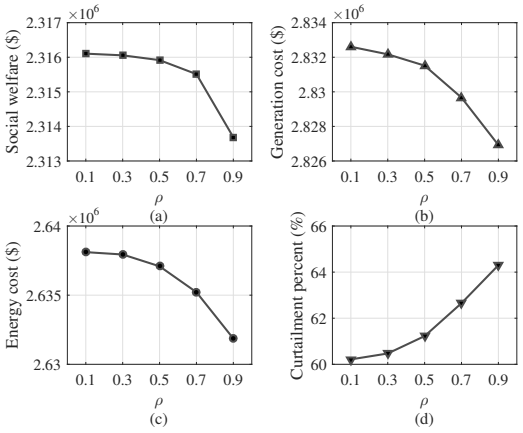

Figure 7: Simulation results under different risk averse levels (a) Social welfare (b) Generation cost (c) Energy cost (d) Curtailment percent.

First, it is easily observed from Figure 7(a) that the social welfare slightly decreases with the increase of risk aversion

factor ρ, from 2.3161×10^6 \$ to 2.3137×10^6 \$. It reveals that there exists a trade-off relationship between the expected social welfare and economic risk. In other words, the expected social welfare will be reduced if we pay more attention to controlling the risk induced by load uncertainty. Furthermore, the generating cost for the energy in Figure 7(c) are decreased when the risk averse level rises, which might be because that more loads are expected to be curtailed, as shown in Figure 7(d). On this basis, the total generation cost are reduced along with the rising of ρ, as illustrated in Figure 7(b). In summary, improving the risk averse level could, to some extent, reduce the expected social welfare, yet it also lowers the generation cost and dispatches more load curtailments. Therefore, decision makers should make reasonable dispatches by adjusting the risk aversion factor, which is also our future research direction.

6. Conclusion

This paper develops a day-ahead bidding and clearing framework considering DR with load uncertainty in the wholesale electricity market. A series of scenario sets are generated based on a data-driven DPMM to represent load uncertainties, considering the correlation of different time periods. In order to control the economic risk induced by load uncertainties, a WCVaR based two-step risk averse MCM is proposed to maximize the social welfare. The first-step optimization is conducted based on system load forecast. Then, the dispatching adjustments are carried out in the second step after the realization of load uncertainty.

Numerical analysis demonstrates the outstanding performance of our proposed framework. First, the load curtailment based DR mechanism is verified to be effective in the day-ahead market. When dealing with the uncertainty in day-ahead load forecasting, the performance of DPMM in representing time-series load uncertainty is better than GMM. In addition, the WCVaR based risk averse MCM could well optimize the social welfare while controlling the economic risk when the load uncertainty is represented by the mixture model. The risk averse level of the MCM could be adjusted by the corresponding factor. In summary, this study could serve as a technical reference for the electricity market operations.

Appendix A. Variational Inference for DPMM

Given observed data set $\mathscr{E}$, the logarithm of posterior probability distribution $p(\mathscr{E})$ could be expressed as follows when we introduce a variational posterior $\hat{p}(\mathscr{W})$.

$$\begin{aligned}\ln p(\mathscr{E}) &= \ln p(\mathscr{E},\mathscr{W}) - \ln p(\mathscr{W}|\mathscr{E}) \\ &= \ln \frac{p(\mathscr{E},\mathscr{W})}{\hat{p}(\mathscr{W})} - \ln \frac{p(\mathscr{W}|\mathscr{E})}{\hat{p}(\mathscr{W})} \\ &= \ln p(\mathscr{E},\mathscr{W}) - \ln \hat{p}(\mathscr{W}) - \ln \frac{p(\mathscr{W}|\mathscr{E})}{\hat{p}(\mathscr{W})} \\ &= \ln p(\mathscr{E},\mathscr{W}) - \ln \hat{p}(\mathscr{W}) + \ln \frac{\hat{p}(\mathscr{W})}{p(\mathscr{W}|\mathscr{E})}\end{aligned} \tag{A.1}$$

The expectation of regarding $\hat{p}(\mathscr{W})$ for both sides in (A.1) could be formulated as follows.

$$\begin{aligned}\int \hat{p}(\mathscr{W}) \ln p(\mathscr{E}) \mathrm{d}\mathscr{W} = \int \hat{p}(\mathscr{W}) \ln p(\mathscr{E},\mathscr{W}) \mathrm{d}\mathscr{W} - \\ \int \hat{p}(\mathscr{W}) \ln \hat{p}(\mathscr{W}) \mathrm{d}\mathscr{W} + \int \hat{p}(\mathscr{W}) \ln \frac{\hat{p}(\mathscr{W})}{p(\mathscr{W}|\mathscr{E})} \mathrm{d}\mathscr{W}\end{aligned} \tag{A.2}$$

Then, the above equation could be reformulated as follows.

$$\begin{aligned}\ln p(\mathscr{E}) = &\overbrace{\mathbb{E}_{\hat{p}(\mathscr{W})}[\ln p(\mathscr{E},\mathscr{W})] - \mathbb{E}_{\hat{p}(\mathscr{W})}[\ln \hat{p}(\mathscr{W})]}^{\text{ELOB}} \\ &+ \underbrace{\mathbb{E}_{\hat{p}(\mathscr{W})} \ln \frac{\hat{p}(\mathscr{W})}{p(\mathscr{W}|\mathscr{E})}}_{D_{KL}(\hat{p}(\mathscr{W}))\|p(\mathscr{W}|\mathscr{E})}\end{aligned} \tag{A.3}$$

Here, $p(\mathscr{E})$ is a constant since it is the real probability density function to be approached. $\mathbb{E}_{\hat{p}(\mathscr{W})} \ln p(\mathscr{E},\mathscr{W}) - \mathbb{E}_{\hat{p}(\mathscr{W})} \ln \hat{p}(\mathscr{W})$ is the evidence lower bound (ELOB), denoted as $\mathscr{L}(\hat{p}(\mathscr{W}))$. $\mathbb{E}_{\hat{p}(\mathscr{W})} \ln \frac{\hat{p}(\mathscr{W})}{p(\mathscr{W}|\mathscr{E})}$ is the Kullback-Leibler (KL) divergence. Therefore, the minimization of KL divergence could be alternatively transformed into the maximization of $\mathscr{L}(\hat{p}(\mathscr{W}))$.

In order to solve the maximization of ELOB regarding $\hat{p}(\mathscr{W})$, latent variables $\boldsymbol{Z}, \boldsymbol{\Omega}, \boldsymbol{\beta}$ and $\boldsymbol{\Psi}$ are assumed to be mutually independent. This assumption has been widely adopted in the variational inference field [29]. Consequently, the variational posterior $\hat{p}(\mathscr{W})$ is factorized into multiple partitions with respect to $\boldsymbol{Z}, \boldsymbol{\Omega}, \boldsymbol{\beta}, \boldsymbol{\Psi}$, expressed as follows.

$$\hat{p}(\mathscr{W}) = \prod_{m=1}^{\infty} \Big\{ \prod_{r=1}^{R} \hat{p}(z_m^{[r]})] \hat{p}(\Omega_m) \hat{p}(\beta_m) \hat{p}(\varphi_m) \Big\} \tag{A.4}$$

The above factorization of $\hat{p}(\mathscr{W})$ regarding the latent variables is feasible. Then, the VI could be implemented by iteratively maximizing $\mathscr{L}(\hat{p}_m(\mathscr{W}_m))$ at one time, which is formulated as

$$\begin{aligned}&\mathscr{L}(\hat{p}_m(\mathscr{W}_m)) \\ =&\mathbb{E}_{\hat{p}(\mathscr{W}_m)}[\ln p(\mathscr{E},\mathscr{W})] - \mathbb{E}_{\hat{p}(\mathscr{W}_m)}[\ln \hat{p}(\mathscr{W}_m)] \\ =&\mathbb{E}_{\hat{p}(\mathscr{W}_m)}\{\mathbb{E}_{\hat{p}(\mathscr{W})_{n\neq m}}[\ln p(\mathscr{E},\mathscr{W})] - \ln \hat{p}(\mathscr{W}_m)\} + \mathrm{C}\end{aligned} \tag{A.5}$$

where $\mathbb{E}_{\hat{p}(\mathscr{W})_{n\neq m}}$ represents the expectation of all variational factors $\hat{p}(\mathscr{W}_n)$ except for $\hat{p}(\mathscr{W}_m)$, n and m are indexes of mixture components, C is the aggregated terms independent on $\hat{p}(\mathscr{W}_m)$. The maximization exists if and only if $\ln \hat{p}(\mathscr{W}_m) = \mathbb{E}_{\hat{p}(\mathscr{W})_{n\neq m}}[\ln p(\mathscr{E},\mathscr{W})]$.

On this basis, the parameter estimation by VI is conducted as follows. First, all the variational factors $\hat{p}(\mathscr{W}_m)$ are appropriately initialized. Then, we compute $\hat{p}(\mathscr{W}_m)$ according to $\ln \hat{p}(\mathscr{W}_m) = \mathbb{E}_{\hat{p}(\mathscr{W})_{n\neq m}}[\ln p(\mathscr{E},\mathscr{W})]$ for every m until $\mathscr{L}(\hat{p}(\mathscr{W}))$ is converged. Eventually, this iterative process could find the optimal posterior distributions. Once the posterior distributions regarding latent variables are determined, the real conditional distribution $p(\mathscr{E}|\mathscr{W})$ is obtained.

Appendix B. Formulation of CVaR

Originally, the conditional value at risk (CVaR) has been developed as the indicator for the higher losses under the uncertain environment [42] in the portfolio optimization field. Reducing such indicator could well manage the risk for decision making. On the contrary, if we seek to improve the lower profit under the uncertain environment, the CVaR with respect to the profit function could also be useful.

To be specific, assuming $f(x, y)$ denote the profit function regarding to the decision vector $x \in X$ and a random vector $y \in Y$, the probability of $f(x, y)$ no less than a threshold η is represented as

$$\psi(x,\eta) = \int_{f(x,y)\geq\eta} p(y)\mathrm{d}y \tag{B.1}$$

where $p(y)$ is the underling probability density of random vector y. On this basis, the VaR for the profit function $f(x, y)$ regarding confidential level α is defined as follows.

$$\mathrm{VaR}_\alpha(x) = \max\{\eta \in R : \psi(x, \eta) \geq \alpha\} \tag{B.2}$$

Therefore, the CVaR regarding $f(x, y)$ is formulated by the expectation of the profits that are lower than VaR with confidential level α, shown as follows.

$$\mathrm{CVaR}_\alpha = \frac{1}{1-\alpha} \int_{f(x,y)\leq \mathrm{VaR}_\alpha(x)} f(x, y)p(y)\mathrm{d}y \tag{B.3}$$

It should be mentioned that the direct computation of CVaR is difficult because of the existence of convoluted and implicit terms. Rockafellar *et. al* have made a significant contribution by introducing an auxiliary function [33] as follows.

$$\mathscr{F}_\alpha(x, \eta) = \eta - \frac{1}{1-\alpha} \int_y [\eta - f(x, y)]^+ p(y)\mathrm{d}y \tag{B.4}$$

where $[\eta - f(x, y)]^+ = \max\{\eta - f(x, y), 0\}$. Then , the CVaR can be obtained as follows (see Theorem in [33]).

$$\mathrm{CVaR}_\alpha(x) = \max_{\eta\in\mathbb{R}} \mathscr{F}_\alpha(x, \eta) \tag{B.5}$$

In practical calculations, the probability density $p(y)$ is usually characterized by the sample-based method. Therefore, with the introduction of probabilistic sample ξ_s, $\mathscr{F}_\alpha(x, \eta)$ could be approximated by

$$\mathscr{F}_\alpha(x, \eta) = \eta - \frac{1}{1-\alpha} \sum_s p_s \cdot [\eta - f(x, \xi_s)]^+ \tag{B.6}$$

Appendix C. Formulation of evaluation measurements

The evaluation measurement used in Table 5 are listed as follows.

Silhouette Coefficient [39]: Silhouette Coefficient compare the average distance to data points in the same cluster to average distance to those in other clusters. Its value range is $[-1, 1]$, the higher the better. Given a sample point x, the definition of Silhouette Coefficient is formulated as:

$$s(x) = \frac{b(x) - a(x)}{\max\{a(x), b(x)\}} \tag{C.1}$$

in which, $a(x)$ is the average distance between x to other points in the same cluster, $b(x)$ is the minimum average distance between x to all points in the different clusters. The Silhouette Coefficient is the average value of all $s(x)$.

Root-mean-square standard deviation (RMSSTD) [40]: Used to measure the homogeneity of clustering results, i.e. compactness, which is defined in (C.2). The smaller the value is, the better clustering performance.

$$RMSSTD = \left\{ \frac{\sum_i \sum_{x\in C_i} \|x - c_i\|^2}{P \sum_i (n_i - 1)} \right\}^{\frac{1}{2}} \tag{C.2}$$

where C_i represent the i-th cluster, c_i is the center of C_i ,

$$x \in C_i$$

means x is a sample point in the cluster C_i , n_i is the number of sample points, P is the number of dimensions of the sample.

R-Square [40]: R-Square represents the improvement of square error after clustering compared with which before clustering. The greater the value, the better. It is formulated as:

$$RS = \frac{\sum_{x\in D} \|x - c\|^2 - \sum_i \sum_{x\in C_i} \|x - c_i\|^2}{\sum_{x\in D} \|x - c\|^2} \tag{C.3}$$

where D represent the whole dataset, c means the centres of all sample points. The rest definitions are the same as (D.2).

Hubert's Γ statistic [41]: The Hubert's Γ statistic has been shown to be effective in assessing fit between data and a priori structures. The larger the Γ value, the more consistent the clustering result is with the original distance of the sample, that is, the higher the clustering quality. It is expressed as follow.

$$\Gamma = \frac{2}{n(n-1)} \sum_{x\in D} \sum_{y\in D} d(x, y) d_{x\in C_i, y\in C_j}(c_i, c_j) \tag{C.4}$$

in which $d(x, y)$ represent the distance between sample point x and y, $d_{x\in C_i, y\in C_j}(c_i, c_j)$ is the distance between the clustering centre c_i which x belonging to and clustering centre c_j which y belonging to. $\frac{n(n-1)}{2}$ is the number of (x, y) sample pairs.

Appendix D. Formulation of CVaR based MCM

In this way, after formulating the CVaR with the sample-based method, the CVaR based MCM is formulated as follows.

$$\max_{\boldsymbol{x},\Delta\boldsymbol{x}} \Big\{ \underbrace{F(\boldsymbol{x})}_{\text{First step}} + \underbrace{(1-\rho)\cdot F(\Delta\boldsymbol{x}|\boldsymbol{\xi}) + \rho \cdot \mathrm{CVaR}_\alpha(\Delta\boldsymbol{x}|\boldsymbol{\xi})}_{\text{Second step}} \Big\} \tag{D.1}$$

References

[1] E. M. Authority, Implementing demand response in the national electricity market of singapore., https://www.ema.gov.sg/cmsmedia/Electricity/Demand_Response/Final_Determination_Demand_Response_28_Oct_2013_Final.pdf.

[2] J. J. Chen, B. X. Qi, Z. K. Rong, K. Peng, Y. L. Zhao, X. Zhang, Multi-energy coordinated microgrid scheduling with integrated demand response for flexibility improvement, Energy 217 (2021) 119387. doi:https://doi.org/10.1016/j.energy.2020.119387.
URL https://www.sciencedirect.com/science/article/pii/S0360544220324944

[3] H. L. Chen, M. B. Liu, Y. Q. Liu, S. J. Lin, Z. B. Yang, Partial surrogate cuts method for network-constrained optimal scheduling of multi-carrier energy systems with demand response, Energy 196 (2020) 117119. doi:https://doi.org/10.1016/j.energy.2020.117119.
URL https://www.sciencedirect.com/science/article/pii/S0360544220302267

[4] Y. L. Wang, Y. Z. Ma, F. H. Song, Y. Ma, C. Y. Qi, F. F. Huang, J. T. Xing, F. W. Zhang, Economic and efficient multi-objective operation optimization of integrated energy system considering electro-thermal demand response, Energy 205 (2020) 118022. doi:https://doi.org/10.1016/j.energy.2020.118022.
URL https://www.sciencedirect.com/science/article/pii/S0360544220311294

[5] S. F. Zhou, Z. Shu, Y. Gao, H. B. Gooi, S. X. Chen, K. Tan, Demand response program in Singapore's wholesale electricity market, Electric Power Systems Research 142 (2017) 279–289. doi:https://doi.org/10.1016/j.epsr.2016.09.022.
URL http://www.sciencedirect.com/science/article/pii/S0378779616303741

[6] P. Y. Liu, T. Ding, Z. X. Zou, Y. H. Yang, Integrated demand response for a load serving entity in multi-energy market considering network constraints, Applied Energy 250 (2019) 512–529. doi:https://doi.org/10.1016/j.apenergy.2019.05.003.
URL http://www.sciencedirect.com/science/article/pii/S0306261919308578

[7] X. Lu, Y. Yang, P. Wang, Y. Fan, F. Yu, N. Zafetti, A new converged Emperor Penguin Optimizer for biding strategy in a day-ahead deregulated market clearing price: A case study in China, Energy 227 (2021) 120386. doi:10.1016/j.energy.2021.120386.

[8] H. Aghamohammadloo, V. Talaeizadeh, K. Shahanaghi, J. Aghaei, H. Shayanfar, M. Shafie-khah, J. P. S. Catalão, Integrated Demand Response programs and energy hubs retail energy market modelling, Energy 234 (2021) 121239. doi:10.1016/j.energy.2021.121239.

[9] E. Shahryari, H. Shayeghi, B. Mohammadi-ivatloo, M. Moradzadeh, An improved incentive-based demand response program in day-ahead and intra-day electricity markets, Energy 155 (2018) 205–214. doi:10.1016/j.energy.2018.04.170.

[10] H. Ming, L. Xie, M. C. Campi, S. Garatti, P. R. Kumar, Scenario-based economic dispatch with uncertain demand response, IEEE Transactions on Smart Grid 10 (2) (2019) 1858–1868. doi:10.1109/TSG.2017.2778688.

[11] M. Shafie-khah, J. P. S. Catalão, A stochastic multi-layer agent-based model to study electricity market participants behavior, IEEE Transactions on Power Systems 30 (2) (2015) 867–881. doi:10.1109/TPWRS.2014.2335992.

[12] N. Nasiri, S. Zeynali, S. N. Ravadanegh, M. Marzband, A hybrid robust-stochastic approach for strategic scheduling of a multi-energy system as a price-maker player in day-ahead wholesale market, Energy 235 (2021) 121398. doi:10.1016/j.energy.2021.121398.

[13] M. Dadashi, S. Haghifam, K. Zare, M.-R. Haghifam, M. Abapour, Short-term scheduling of electricity retailers in the presence of Demand Response Aggregators: A two-stage stochastic Bi-Level programming approach, Energy 205 (2020) 117926. doi:10.1016/j.energy.2020.117926.

[14] N. Rezaei, Y. Pezhmani, A. Khazali, Economic-environmental risk-averse optimal heat and power energy management of a grid-connected multi microgrid system considering demand response and bidding strategy, Energy 240 (2022) 122844. doi:10.1016/j.energy.2021.122844.

[15] Y. Wang, Q. Chen, M. Sun, C. Kang, Q. Xia, An ensemble forecasting method for the aggregated load with subprofiles, IEEE Transactions on Smart Grid 9 (4) (2018) 3906–3908. doi:10.1109/TSG.2018.2807985.

[16] R. Singh, B. C. Pal, R. A. Jabr, Statistical representation of distribution system loads using gaussian mixture model, IEEE Transactions on Power Systems 25 (1) (2010) 29–37. doi:10.1109/TPWRS.2009.2030271.

[17] R. Panda, P. K. Tiwari, Economic risk-based bidding strategy for profit maximisation of wind-integrated day-ahead and real-time double-auctioned competitive power markets, IET Generation, Transmission Distribution 13 (2) (2019) 209–218.

[18] P. Sheikhahmadi, S. Bahramara, J. Moshtagh, M. Y. Damavandi, A risk-based approach for modeling the strategic behavior of a distribution company in wholesale energy market, Applied Energy 214 (2018) 24–38. doi:https://doi.org/10.1016/j.apenergy.2018.01.051.
URL http://www.sciencedirect.com/science/article/pii/S0306261918300631

[19] S. Das, M. Basu, Day-ahead optimal bidding strategy of microgrid with demand response program considering uncertainties and outages of renewable energy resources, Energy 190 (2020) 116441. doi:10.1016/j.energy.2019.116441.

[20] L. Tian, Y. L. Xie, B. Hu, X. P. Liu, T. Y. Deng, H. H. Luo, F. Q. Li, A deep peak regulation auxiliary service bidding strategy for CHP units based on a risk-averse model and district heating network energy storage, Energies 12 (17) (2019) 3314. doi:10.3390/en12173314.
URL https://www.mdpi.com/1996-1073/12/17/3314

[21] W. Labeeuw, G. Deconinck, Residential electrical load model based on mixture model clustering and Markov models, IEEE Transactions on Industrial Informatics 9 (3) (2013) 1561–1569. doi:10.1109/TII.2013.2240309.

[22] S. S. Zhu, M. Fukushima, Worst-case conditional value-at-risk with application to robust portfolio management, Operations Research 57 (5) (2009) 1155–1168. doi:10.1287/opre.1080.0684.
URL https://doi.org/10.1287/opre.1080.0684

[23] J. A. Schachter, P. Mancarella, Demand response contracts as real options: A probabilistic evaluation framework under short-term and long-term uncertainties, IEEE Transactions on Smart Grid 7 (2) (2016) 868–878. doi:10.1109/TSG.2015.2405673.

[24] D. Shah, S. Chatterjee, Optimal placement of time flexible supplier's block bid in a day-ahead electric market using genetic algorithm, in: 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP), 2019.

[25] Y. Li, G. Hao, Y. Liu, Y. Yu, Z. Ni, Y. Zhao, Many-objective distribution network reconfiguration via deep reinforcement learning assisted optimization algorithm, IEEE Transactions on Power Delivery (2021) 1–1doi:10.1109/TPWRD.2021.3107534.

[26] Z. Li, Y. Li, Y. Liu, P. Wang, R. Lu, H. B. Gooi, Deep learning based densely connected network for load forecasting, IEEE Transactions on Power Systems 36 (4) (2021) 2829–2840. doi:10.1109/TPWRS.2020.3048359.

[27] Z. Y. Ma, P. Rana, J. Taghia, M. Flierl, A. Leijon, Bayesian estimation of dirichlet mixture model with variational inference, Pattern Recognition 47 (2014) 3143–3157. doi:10.1016/j.patcog.2014.04.002.

[28] D. M. Blei, M. I. Jordan, Variational inference for dirichlet process mixtures, Journal of Bayesian Analysis 1 (1) (2006) 121–143.

[29] D. Blei, A. Kucukelbir, J. McAuliffe, Variational Inference: A Review for Statisticians, Journal of the American Statistical Association 112 (518) (2017) 859–877. doi:10.1080/01621459.2017.1285773.

[30] S. Kim, T. Kang, Texture classification and segmentation using wavelet packet frame and Gaussian mixture model, Pattern Recognition 40 (2007) 1207–1221. doi:10.1016/j.patcog.2006.09.012.

[31] Z. W. Wang, C. Shen, F. Liu, A conditional model of wind power forecast errors and its application in scenario generation, Applied Energy 212 (2017) 771–785.

[32] Z. Chen, L. Wu, Z. Li, Electric demand response management for distributed large-scale internet data centers, IEEE Transactions on Smart Grid 5 (2) (2014) 651–661. doi:10.1109/TSG.2013.2267397.

[33] R. T. Rockafellar, S. Uryasev, Optimization of conditional value-at-risk, The Journal of Risk 2 (3) (2000) 21–41. doi:10.21314/JOR.2000.038.

[34] X. J. Tong, F. Wu, L. J. Qi, Worst-case CVaR based portfolio optimization models with applications to scenario planning, Optimization Methods and Software 24 (2009) 933–958. doi:10.1080/10556780902865942.

[35] H. Sekhavatmanesh, R. Cherkaoui, Distribution network restoration in a multiagent framework using a convex OPF model, IEEE Transactions on Smart Grid 10 (3) (2019) 2618–2628. doi:10.1109/TSG.2018.2805922.

[36] Z. Zhao, L. Wu, G. Song, Convergence of volatile power markets with price-based demand response, IEEE Transactions on Power Systems 29 (5) (2014) 2107–2118. doi:10.1109/TPWRS.2014.2307872.

[37] E. M. Company, Singapore energy price and demand forecast, https://www.emcsg.com/marketdata/priceinformation.

[38] Q. Wang, J. Watson, Y. Guan, Two-stage robust optimization for N-k contingency-constrained unit commitment, IEEE Transactions on Power Systems 28 (3) (2013) 2366–2375. doi:10.1109/TPWRS.2013.2244619.

[39] P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, Journal of Computational and Applied Mathematics 20 (1987) 53–65. doi:10.1016/0377-0427(87)90125-7.

[40] C. Liu, J. Cen, Training Subset Selection for Support Vector Regression, in: 2019 Federated Conference on Computer Science and Information Systems (FedCSIS), 2019, pp. 11–14. doi:10.15439/2019F363.

[41] C.-Y. Tang, Y.-L. Wu, Y.-C. Lee, Cluster and Clustering Algorithm Validity in Image Retrieval, in: 2006 IEEE International Conference on Systems, Man and Cybernetics, Vol. 4, 2006, pp. 3318–3323. doi:10.1109/ICSMC.2006.384630.

[42] S. Alexander, T. F. Coleman, Y. Li, Minimizing CVaR and VaR for a portfolio of derivatives, Journal of Banking & Finance 30 (2) (2006) 583–605.