Appendix

A. Wasserstein-distance-based Distributionally Robust

In this paper, the distributionally robust method based on Wasserstein distance is used to deal with uncertain variables.

The Wasserstein distance is defined as follows:

$$d_{\mathrm{W}}\left(P_{\mathrm{N}}, P\right)=\min_{\Phi} \int_{\Xi^2}\| \varsigma-\xi \| \Phi(\mathrm{d}\varsigma, \mathrm{d}\xi) \tag{1}$$

where, P_{N} and P are empirical distribution and real probability distribution, respectively. Φ is a joint distribution of two random variables, ς and $\boldsymbol{\xi}$ with marginal distributions P_{N} and P. $\boldsymbol{\Xi}$ is the support set of random variables.

A fuzzy set of uncertain variables is constructed based on Wasserstein distance:

$$B_{\varepsilon}\left(P_{\mathrm{N}}\right)=\{P: P(\xi \in \Xi)=1\} \cap\left\{P: d_{\mathrm{w}}\left(P_{\mathrm{N}}, P\right) \leqslant \varepsilon\right\} \tag{2}$$

The fuzzy set is a Wasserstein ball with P_{N} as the center and ε as the radius of probability distribution space. Random variables are included in this ball at a higher confidence level. The relationship between radius and confidence level is as follows:

$$P\left\{d_{\mathrm{W}}\left(P_{\mathrm{N}}, P\right) \leqslant \varepsilon\right\} \geqslant \rho=1-\exp \left(-N \frac{\varepsilon^{2}}{2 H^{2}}\right) \tag{3}$$

$$\varepsilon=H \sqrt{\frac{2}{N} \ln \left(\frac{1}{1-\rho}\right)} \tag{4}$$

$$H=\min_{\eta \geqslant 0} 2 \sqrt{\frac{1}{2 \eta}\left[1+\ln \left(\frac{1}{N} \sum_{m=1}^{N} e^{\eta\left\|\varsigma_{m}-\mu\right\|}\right)\right]} \tag{5}$$

where, ρ is the confidence level. μ and N are the mean and number of samples, respectively.

The equivalent transformation theorem of Wasserstein fuzzy constraints:

If the optimization problem is convex, the uncertainty loss (the worst case) of the distributionally robust problem with Wasserstein fuzzy sets is consistent with the optimal value of the following convex problem:

$$\begin{cases} \sup\limits_{q_i} \dfrac{1}{N} \sum\limits_{i=1}^{N} \ell\left(\hat{\xi}_{i}-q_{i}\right) \\ \text{s.t. } \dfrac{1}{N} \sum\limits_{i=1}^{N}\left\|q_{i}\right\| \leqslant \varepsilon \\ \hat{\xi}_{i}-q_{i} \in \Xi \quad \forall i \leqslant N \end{cases} \tag{6}$$

where, q_i and $\hat{\xi}_i$ are the values of samples and deviation, respectively. ℓ denotes the uncertainty loss.

B. The Proof of Proposition 1

The MG's problem is equivalent to

$$\begin{aligned} &\min C_{\mathrm{MG},i}^{2}+\sum_{t=1}^{T} \rho_{i,t} P_{i,t}^{\mathrm{ses}} \\ &s.t. \{P_{i,t}^{\mathrm{ses}}, \forall t\} \in \mathcal{D}_{i} \end{aligned} \tag{A1}$$

where, $C_{\mathrm{MG},i}^{2}$ is the cost of MG*i* without considering the transaction cost $C_{\mathrm{ses},i}$. Based on the variational inequality, if $P_{i,t}^{*\mathrm{ses}}$ is the optimal solution, then for all $\{P_{i,t}^{\mathrm{ses}}, \forall t\} \in \mathcal{D}_i$, we have

$$C_{\mathrm{MG},i}^{2}(P_{i}^{\mathrm{ses}})-C_{\mathrm{MG},i}^{2}(P_{i}^{*\mathrm{ses}})+\sum_{t=1}^{T} \rho_{i,t}(P_{i,t}^{\mathrm{ses}}-P_{i,t}^{*\mathrm{ses}}) \geq 0 \tag{A2}$$

For simplicity, we ignore the index t, and P_i^{ses} is a collection of $P_{i,t}^{\mathrm{ses}}$ for all t; other symbols without t are defined similarly.

The SEPS's problem is equivalent to

$$\begin{aligned} &\min C_{\mathrm{SES}}^{1}-\sum_{t=1}^{T} \sum_{i=1}^{M} \rho_{i,t} P_{i,t}^{\mathrm{ses}} \\ &s.t. P_{i,t}^{\mathrm{ses}}-Q_{\mathrm{e},i,t}=0: \phi_{i,t} \\ &\{Q_{\mathrm{e},i,t}, \forall i, \forall t\} \in \mathcal{P} \end{aligned} \tag{A3}$$

Based on the variational inequality, if $(P_i'^{\mathrm{ses}}, \forall i; Q_{\mathrm{e},i}^{*}, \forall i)$ is the optimal solution and $\phi_i^*, \forall i$ is the corresponding dual variable, then for all $\{P_i^{\mathrm{ses}}, \forall i; Q_{\mathrm{e},i}, \phi_i, \forall i\} \in \mathbb{R}^{M\times T} \times \mathbb{R}^{M\times T}$, we have

$$C_{\mathrm{SES}}^{1}(Q_{\mathrm{e},i})-C_{\mathrm{SES}}^{1}(Q_{\mathrm{e},i}^{*})+\sum_{t=1}^{T} \sum_{i=1}^{M} \phi_{i,t}^{*}(Q_{\mathrm{e},i,t}-Q_{\mathrm{e},i,t}^{*}) \geq 0 \tag{A4}$$

$$(-\rho_{i,t}+\phi_{i,t}^{*})(P_{i,t}^{\mathrm{ses}}-P_{i,t}'^{\mathrm{ses}}) \geq 0, \forall t \tag{A5}$$

$$\sum_{t=1}^{T} \sum_{i=1}^{M}(\phi_{i,t}-\phi_{i,t}^{*})(P_{i,t}'^{\mathrm{ses}}-Q_{\mathrm{e},i,t}^{*}) \geq 0 \tag{A6}$$

The optimality conditions (B2), (B4)-(B6), and (39) constitute the condition for a MGSC-SEPS joint system equilibrium.

Similarly, we suppose $(\bar{P}_{i,t}^{\mathrm{ses}}, \forall i, \forall t)$ is the optimal solution of (40). Then, for all $\{Q_{\mathrm{e},i}, \forall i; P_i^{\mathrm{ses}}, \forall i; \mu_i\} \in \mathcal{P} \times \prod_i \mathcal{D}_i \times \mathbb{R}^{M\times T}$, we have

$$C_{\mathrm{SES}}^{1}(Q_{\mathrm{e},i})-C_{\mathrm{SES}}^{1}(\bar{Q}_{\mathrm{e},i})+\sum_{t=1}^{T} \sum_{i=1}^{M} \bar{\mu}_{i,t}(Q_{\mathrm{e},i,t}-\bar{Q}_{\mathrm{e},i,t}) \geq 0 \tag{A7}$$

$$C_{\mathrm{MG},i}^{2}(P_{i}^{\mathrm{ses}})-C_{\mathrm{MG},i}^{2}(\bar{P}_{i}^{\mathrm{ses}})+\sum_{t=1}^{T} \bar{\mu}_{i,t}(P_{i,t}^{\mathrm{ses}}-\bar{P}_{i,t}^{\mathrm{ses}}) \geq 0, \forall i \tag{A8}$$

$$\sum_{t=1}^{T} \sum_{i=1}^{M}(\mu_{i,t}-\bar{\mu}_{i,t})(\bar{P}_{i,t}^{\mathrm{ses}}-\bar{Q}_{\mathrm{e},i,t}) \geq 0 \tag{A9}$$

If $(P_i^{*\mathrm{ses}}, \forall i; P_i'^{\mathrm{ses}}, \forall i)$ is a system equilibrium associated with price $(\rho_i, \forall i)$, according to (B5), we have $\rho_{i,t}=\phi_{i,t}^*$ for all $i \in M$ and $t \in T$. (B5) says that $\forall \{P_{i,t}^{\mathrm{ses}}, \forall i, \forall t\} \in \mathbb{R}^{M\times T}$, $(-\rho_{i,t}+\phi_{i,t}^{*})(P_{i,t}^{\mathrm{ses}}-P_{i,t}^{*\mathrm{ses}}) \geq 0$. Therefore, we have $\rho_{i,t}=\phi_{i,t}^*$.

If we let

$$\bar{P}_{i,t}^{\mathrm{ses}}=P_{i,t}^{*\mathrm{ses}}=P_{i,t}'^{\mathrm{ses}}, \forall i, \forall t \tag{A10}$$

$$\bar{\mu}_{i,t}=\phi_{i,t}^{*}=\rho_{i,t}, \forall i, \forall t \tag{A11}$$

then (B4), (B2) and (B6) become (B7), (B8) and (B9), respectively. Similarly, if $(\bar{P}_i^{\mathrm{ses}}, \forall i)$ is the optimal solution of (40) and $(\bar{\mu}_i, \forall i)$ is the corresponding dual variable. If we let

$$P_{i,t}^{*\mathrm{ses}}=P_{i,t}'^{\mathrm{ses}}=\bar{P}_{i,t}^{\mathrm{ses}}, \forall i, \forall t \tag{A12}$$

$$\rho_{i,t}=\phi_{i,t}^{*}=\bar{\mu}_{i,t}, \forall i, \forall t \tag{A13}$$

then, the optimality conditions (B2), (B4)-(B6) and the supply-demand balance (39) are all satisfied. Therefore, the MGSC-SEPS joint system equilibrium can be constructed when a price is consistent with the value of the dual variable in the model (40).

C. Basic Data

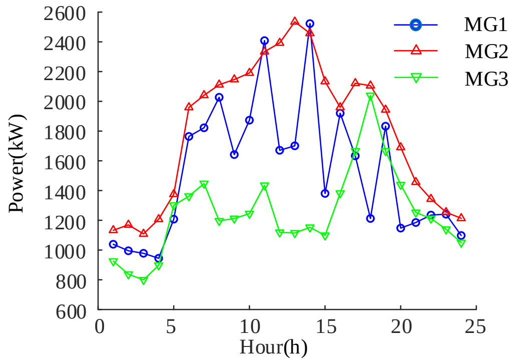


Fig. C1. The electricity load prediction curves.

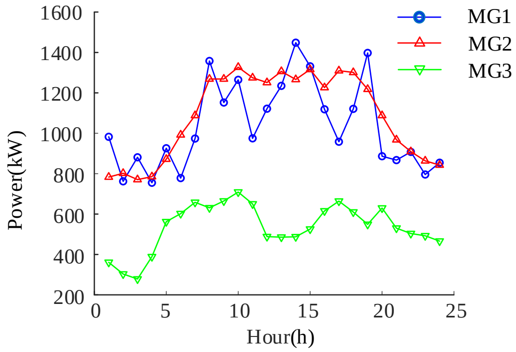


Fig. C2. The heat load prediction curves.

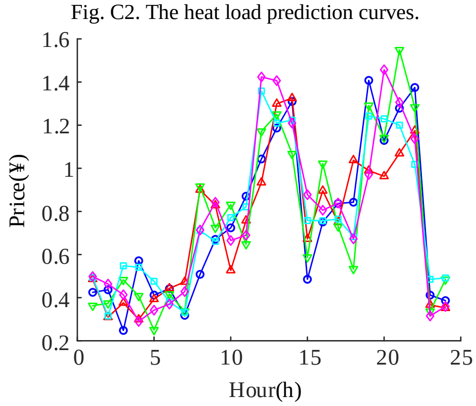


Fig. C3. The purchasing price sample curves.

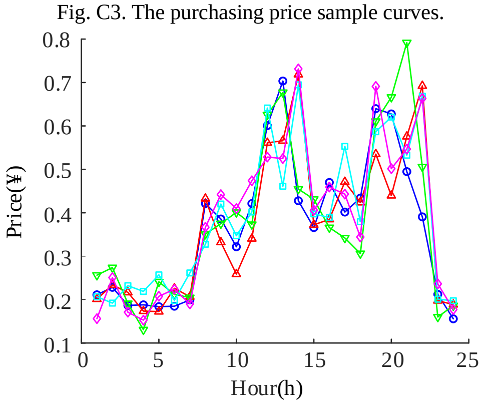
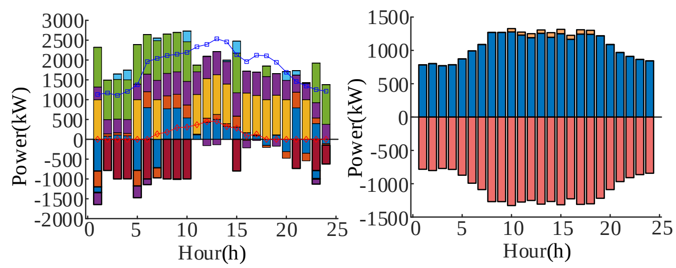


Fig. C4. The selling price sample curves.

TABLE C1
THE PARAMETERS OF SYSTEM

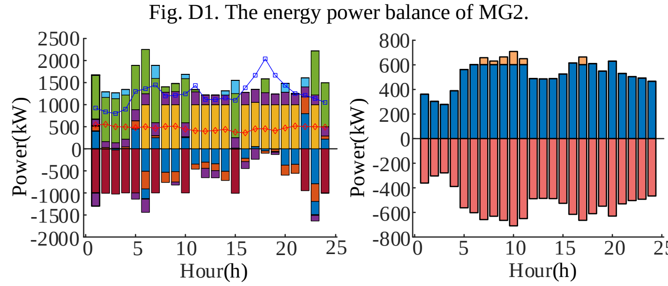
Parameters	Values
$a_{e}/b_{e}/a_{water}$	3.021/0.992/6.063kW/m^3
$b_{water}/a_{smoke}/b_{smoke}$	-0.500/1.487/0.493 kW/m^3
$\eta_{GT,water}/\eta_{GT,smoke}/\eta_{GB}$	1/0.85/0.9
$\gamma_{GT}^{min}/\gamma_{GT}^{max}$	0/20%
$a_{P2G}/a_{co_2}/a_{CCS}$	0.218/0.2/0.269
a_{P2H}/a_{H2P}	0.04kg/ kW, 35.46 kW/kg
$P_{i,t}^{P2H,max}/P_{i,t}^{H2P,max}$	1000 /1000kW
$P_{i,t}^{buy,max}/P_{i,t}^{sell,max}$	1000/1000 kW
$s_{max}^{ES}/s_{min}^{ES}$	720/80 kWh
$s_{i,initail}^{ES}/s_{i,expect}^{ES}$	200/200 kWh
η_{ch}/η_{dis}	0.95/0.95
$s_{max}^{H_2}/s_{min}^{H_2}$	300/30kg
$s_{i,initail}^{H_2}/s_{i,expect}^{H_2}$	80/80kg
$\chi_{GT}/\chi_{GB}/\chi_{buy}$	0.968/0.968/0.889
$s_{max}^{SES}/s_{min}^{SES}/s_{0}^{SES}$	0.9/0.1/0.2
$P_{ses,max}^{ch}/P_{ses,max}^{dis}/E_{max}$	400kW/400kW/1200 kWh
$\chi_{GT}^{qua}/\chi_{GT,h}^{qua}/\chi_{GB}^{qua}/\chi_{buy}^{qua}$	0.112/0.067/0.067/0.798
$c/d/v$	0.2 ¥/kg, 800kg, 0.5
c_{H2P}/c_{P2H}	16.705/10.075 ¥
c_{gas}/c_{es}	3.2 ¥/m^3,0.01 ¥/kW

D. Residual Optimization Results



(a) The electric power balance (b) The thermal power balance

Fig. D1. The energy power balance of MG2.



(a) The electric power balance (b) The thermal power balance

Fig. D2. The energy power balance of MG3.