Appendix

* 1. *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:



Where,  and  are empirical distribution and real probability distribution, respectively. is a joint distribution of two random variables,  and  with marginal distributions  and . is the support set of random variables.

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



The fuzzy set is a Wasserstein ball with  as the center and  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:







Where, is the confidence level.  and  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:



Where,  and  are the values of samples and deviation, respectively.  denotes the uncertainty loss[1].

* 1. *The Proof of ADMM Algorithm Convergence*

From the convergence theorem of the ADMM algorithm, it follows in the following form of convex optimization problems:



Where A and B are linear constraint matrices, c is a constant vector.

If f(x) and g(z) are closed convex functions (Condition 1), the problem has a feasible solution (Condition 2), and strong duality holds (Condition 3), the sequence  generated by ADMM satisfies:

(1) Original residual convergence: 

(2) Objective function convergence: , where  is the optimal solution

(3) Dual variable convergence: , where  is the dual optimal solution[2].

In the MGSC-SEPS optimization model, the objective function is Eqs. (41)-(42).  contains linear and quadratic terms, both convex;  is a linear function, obviously convex; the constraints are all linear equations/inequalities, convex, and thus satisfy condition 1. In the system design, the MGSC-SEPS system satisfies the base load demand, there exists at least one feasible solution, condition 2 is satisfied, and the system has strong duality, and condition 3 is satisfied.

In summary, the algorithm in this paper converges globally.

* 1. *The Proof of Proposition 1*

Firstly we construct the centralized optimization problem in Eq. (38) to minimize the total system cost.

Secondly, we establish optimization models for microgrid clusters and shared energy storage plants, respectively. For simplicity, we ignore the index , and  is a collection of  for all ; other symbols without  are defined similarly.

The MG’s problem is equivalent to



Based on the variational inequality, if  is the optimal solution, then for all , we have



The SEPS’s problem is equivalent to



Based on the variational inequality, if is the optimal solution and  is the corresponding dual variable, then for all , we have







The optimality conditions (C2), (C4)-(C6), and (37) constitute the condition for a MGSC-SEPS joint system equilibrium.

Similarly, we suppose  is the optimal solution of (38). Then, for all ,we have







If  is a system equilibrium associated with price , according to (C5), we have  for all and .

If we let





then (C2), (C4) and (C6) become (C7), (C8) and (C9), respectively.

Then, the optimality conditions (C2), (C4)-(C6) and the supply-demand balance (37) 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 (38).

## Basic Data



Fig. D1. The WT’s samples curve of MG1.



Fig. D2. The PV’s samples curve of MG2.



Fig. D3. The WT’s samples curve of MG3.



Fig. D4. The electricity load prediction curves.



Fig. D5. The heat load prediction curves.



Fig. D6. The purchasing price sample curves.

****

Fig. D7. The selling price sample curves.

TABLE D1

THE PARAMETERS OF SYSTEM

|  |  |
| --- | --- |
| Parameters | Values |
| // | 3.021/0.992/6.063kW/m3 |
| // | -0.500/1.487/0.493 kW/m3 |
| // | 1/0.85/0.9 |
| / | 0/20% |
| // | 0.218/0.2/0.269 |
| / | 0.04kg/ kW, 35.46 kW/kg |
| / | 1000 /1000kW |
| / | 1000/1000 kW |
| / | 720/80 kWh |
| / | 200/200 kWh |
| / | 0.95/0.95 |
| / | 300/30kg |
| / | 80/80kg |
| // | 0.968/0.968/0.889 |
| // | 0.9/0.1/0.2 |
| // | 400kW/400kW/1200 kWh |
| /// | 0.112/0.067/0.067/0.798 |
| // | 0.2 ¥/kg, 800kg, 0.5 |
| / | 16.705/10.075 ¥ |
| / | 3.2 ¥/m3,0.01 ¥/kW |

## Residual Optimization Results

Fig. E1. The electric power balance of MG2.

Fig. E2. The thermal power balance of MG2.



Fig. E3. The electric power balance of MG3.

Fig. E4. The thermal power balance of MG3.

 (a) MG2 in Mode 1 (b) MG2 in Mode 3

 (c) MG3 in Mode 1 (d) MG3 in Mode 3



Fig. E5. The results of ES for MG2/MG3.

References

1. Mohajerin Esfahani P, Kuhn D. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations[J]. Mathematical Programming, 2018, 171(1): 115-166.
2. Boyd S, Parikh N, Chu E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. Foundations and Trends® in Machine learning, 2011, 3(1): 1-122.