




2017/1/18 第45回強化学習勉強会

Application of fuzzy Q-learning for
electricity market modeling by considering
renewable power penetration

Mohammad Reza Salehizadeh, Salman Soltaniyan

担当：慶應義塾大学大学院1年 池田春之介

- 再生可能エネルギーの増加によって、それらの影響が考慮された電力市場のモデリング手法の開発が必要
- 
- 再エネの影響を考慮した1時間前電力市場にfuzzy Q-learningを適用し、test systemで検証
- 
- 提案手法(fuzzy Q-learning)の有用性を2つのcase(normal, stressful)で既存手法(Q-learning)と比較
- 
- 提案手法は、再エネを考慮した状況において、既存手法よりもiterationの数を減らしてモデル化することに成功

再生可能エネルギー(Renewable Energy: RE)

自然界の力によって定常的に補充され枯渇することのないエネルギー
－ 太陽光、水力、風力、地熱、波力、バイオマス etc.

1. Introduction

■ 研究背景

- 現在、電力の分野では送電料金の決定、電力損失の配分、効率的な電力の運用、再エネの統合、電力市場のモデリングなど様々な課題が存在

- 近年、再エネの導入が急速に進んでいる

- ➡ 再エネは、天候による発電量の不確実性があるため、電力システムの構造を複雑にする

- 再エネ導入に伴い、より複雑になった電力市場を解析するための有用なモデルの開発が急がれている



■ 研究目的

- 再エネの影響を考慮した電力市場のモデリングにおいて、有用な手法の提案

1. Introduction

■ 本論文のcontribution

- 再エネの影響を考慮した効果的な電力価格の入札戦略をマルチエージェントのfuzzy Q-learningによって表現

➡ 次元の呪いを解決

➡ 有用性を既存手法と比較して検証

- 提案モデルは、離散点だけではなく連続空間において、ナッシュ均衡点を見つけることができる

➡ どのプレイヤーも自分の戦略を変更することによってより高い利益を得ることができない戦略の組み合わせ

2. Agent-based electricity market modeling

■ 電力系統

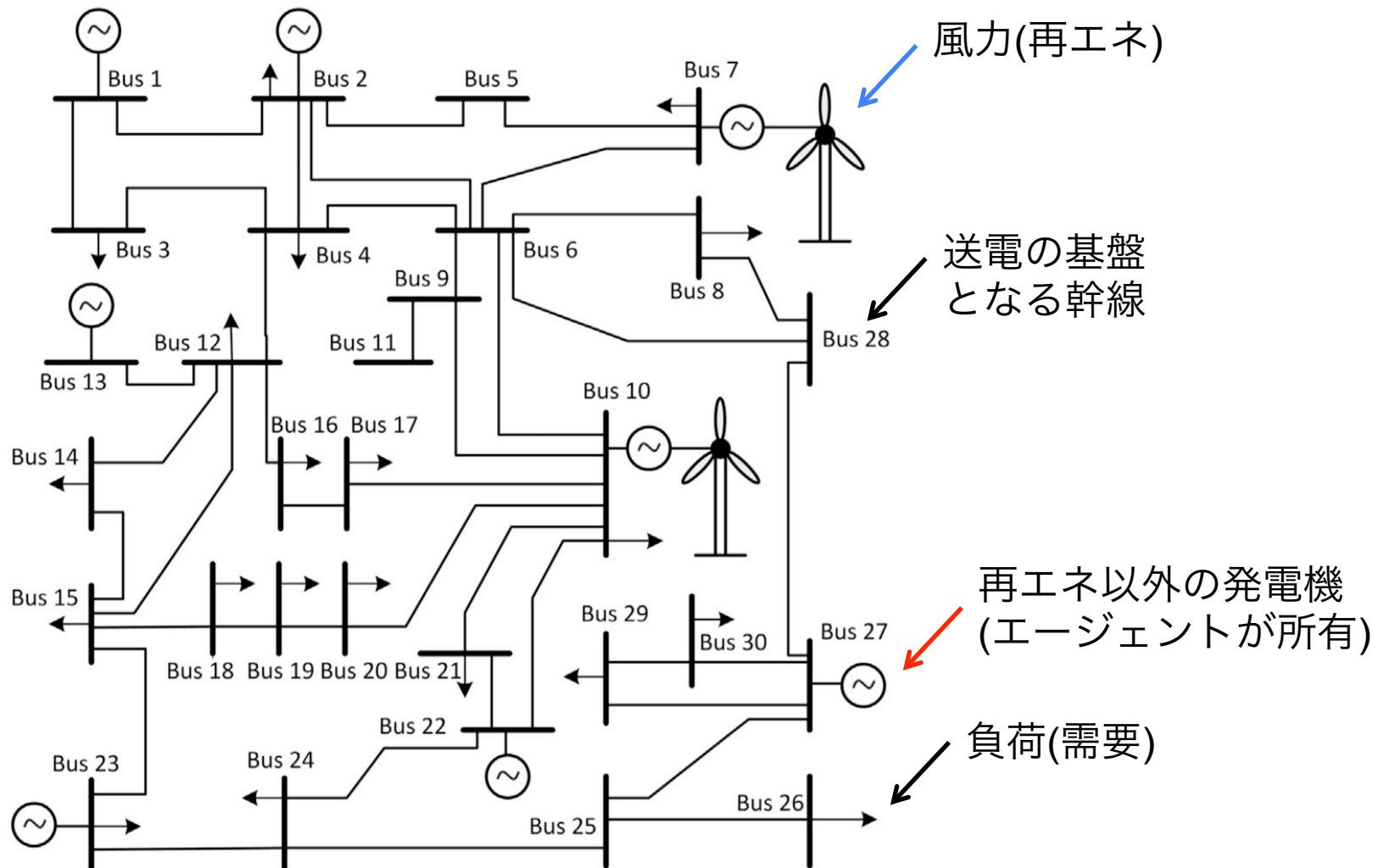
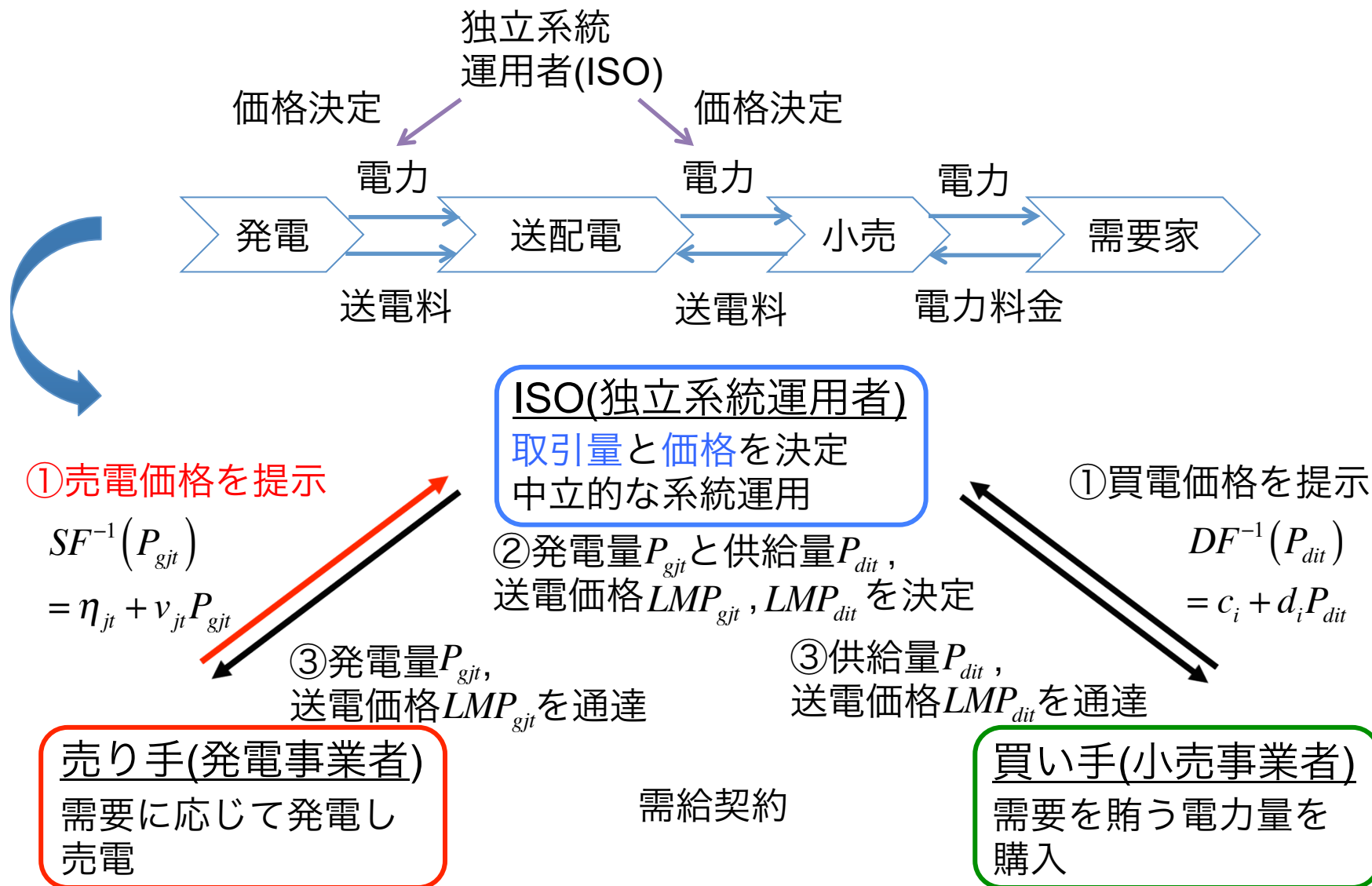


Fig. 1. Diagram of the test system.

2. Agent-based electricity market modeling

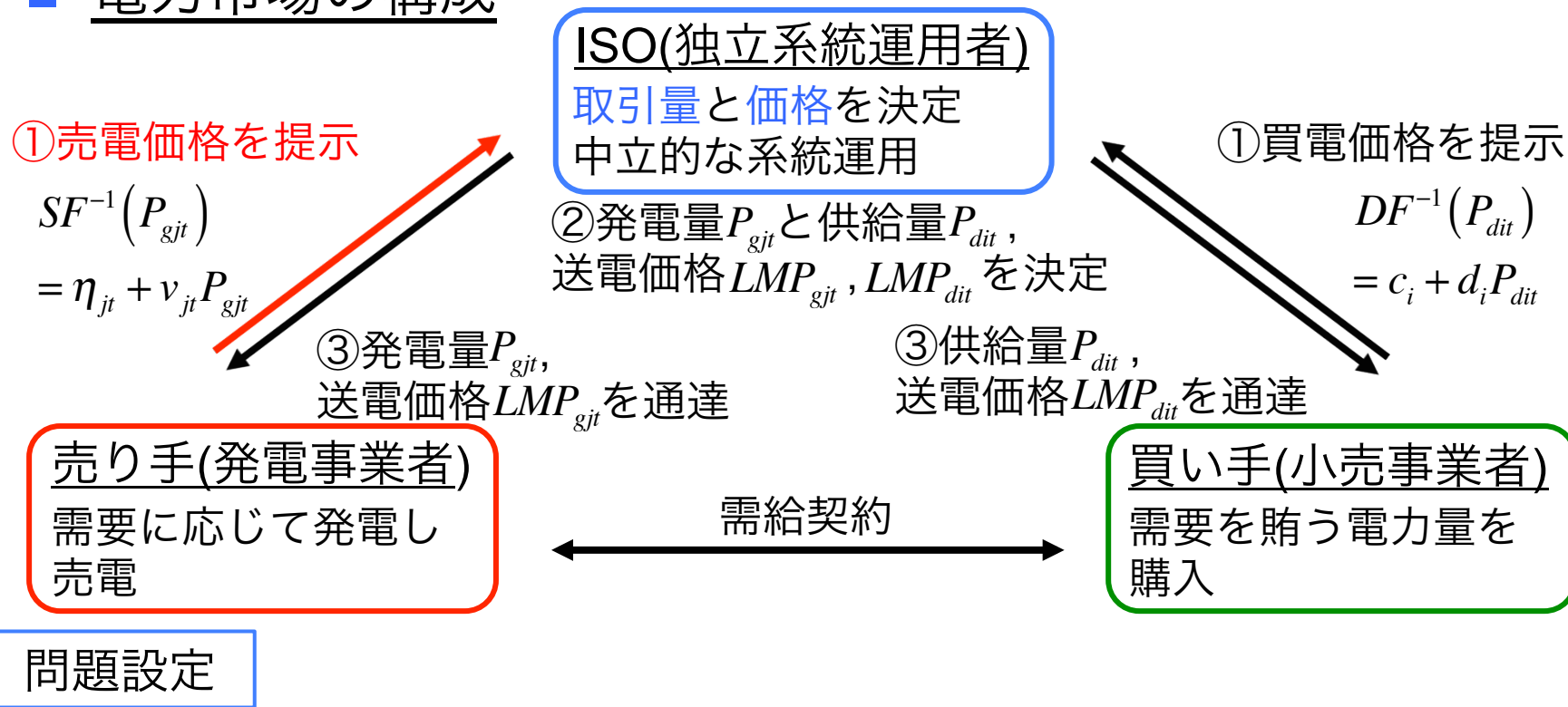
■ 電力市場の構成

□ ISO方式(主にアメリカなどで採用)



2. Agent-based electricity market modeling

■ 電力市場の構成

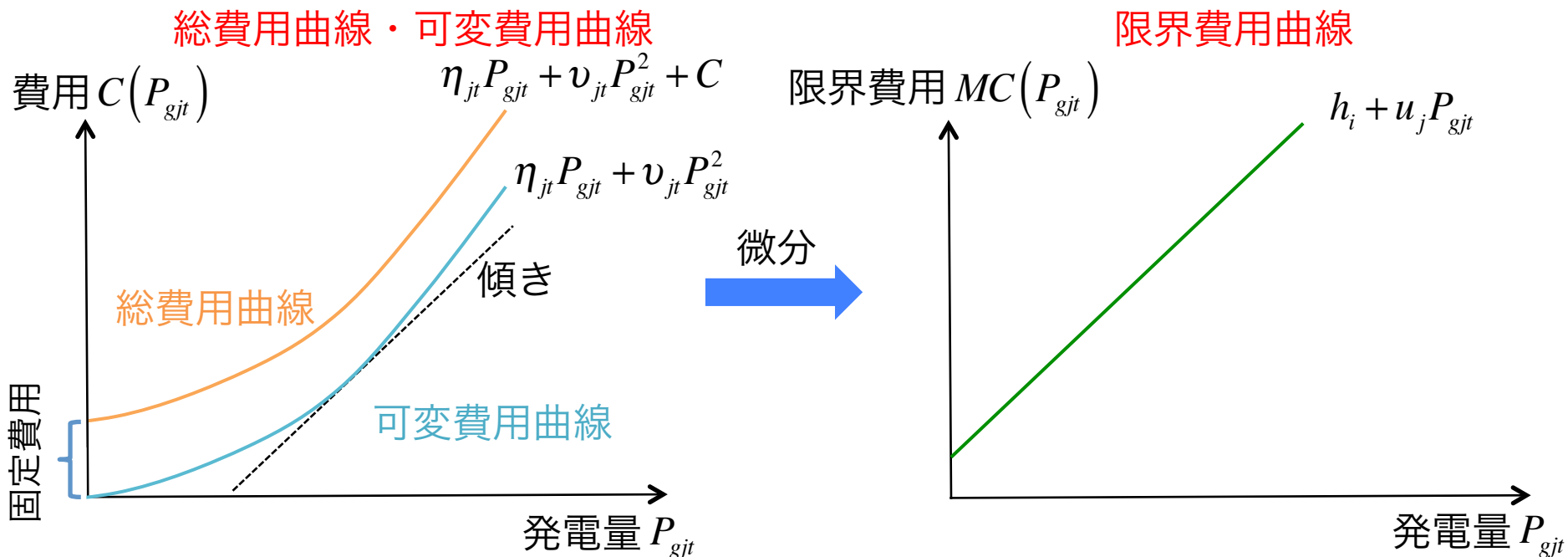


- エージェント：売り手, ISO, 買い手の3者
- 1時間前市場：1時間ごとに入札, 落札が行われる市場 ➡ 1日前市場にも拡張可能
- 意思決定者：売り手(発電事業者)のみ ➡ 買い手も意思決定する場合にも拡張可能
- 1つの発電事業者は1つの発電機を所有 ➡ 複数所有の場合にも拡張可能

※売買価格, 買電価格をとともに関数の形で提示

2. Agent-based electricity market modeling

2.1. The players' bidding problem(売り手の視点)



総費用曲線は、発電量に依らない
固定費用と可変費用の和

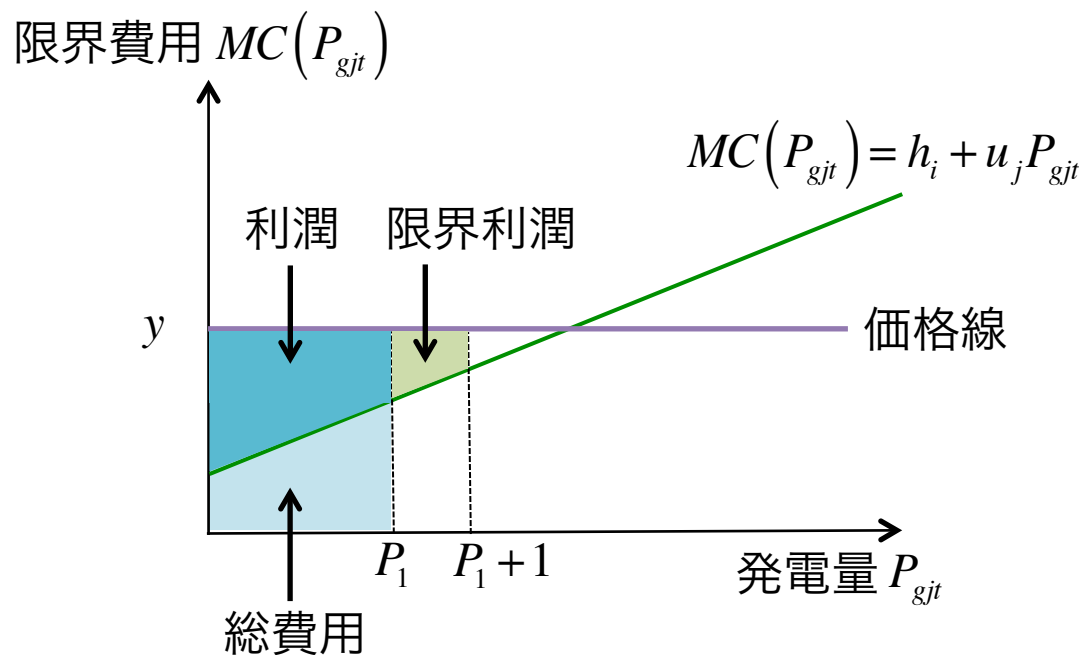
限界費用は可変費用を発電量で
微分したもので、発電量を1単位
増やすのに必要な費用

2. Agent-based electricity market modeling

■ 限界原理

- 利潤を最大化するためには **価格 = 限界費用** となる

限界費用曲線



価格が市場によって、ある価格 y に決まってしまう場合

収入 = 発電量 P_1 × 価格 y

利潤 = 収入 - 総費用

限界利潤 = 価格 - 限界費用

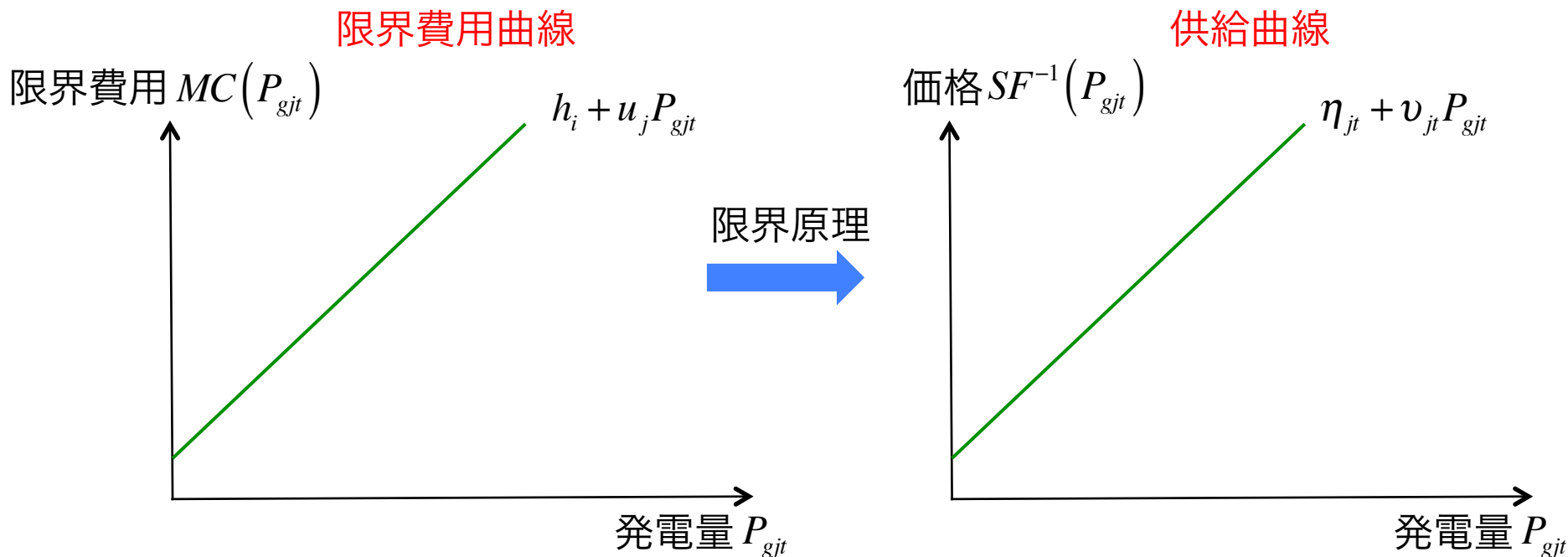
発電量を1単位増やしたときの利潤の増加分

企業は利益を最大化するとき、限界利潤が0になるときである

➡ **価格 = 限界費用**

2. Agent-based electricity market modeling

限界原理(市場によって価格が決められる場合, 価格 = 限界費用)より
限界費用曲線 $MC(P_{gjt}) =$ 供給曲線 $SF^{-1}(P_{gjt})$



発電事業者は, 供給曲線の切片 η_{jt} と傾き v_{jt} を毎時間変化させて,
ISOに供給関数 $SF^{-1}(P_{gjt})$ を受け渡す

$$\eta_{jt} = h_j + mup_{1jt} \times h_j$$

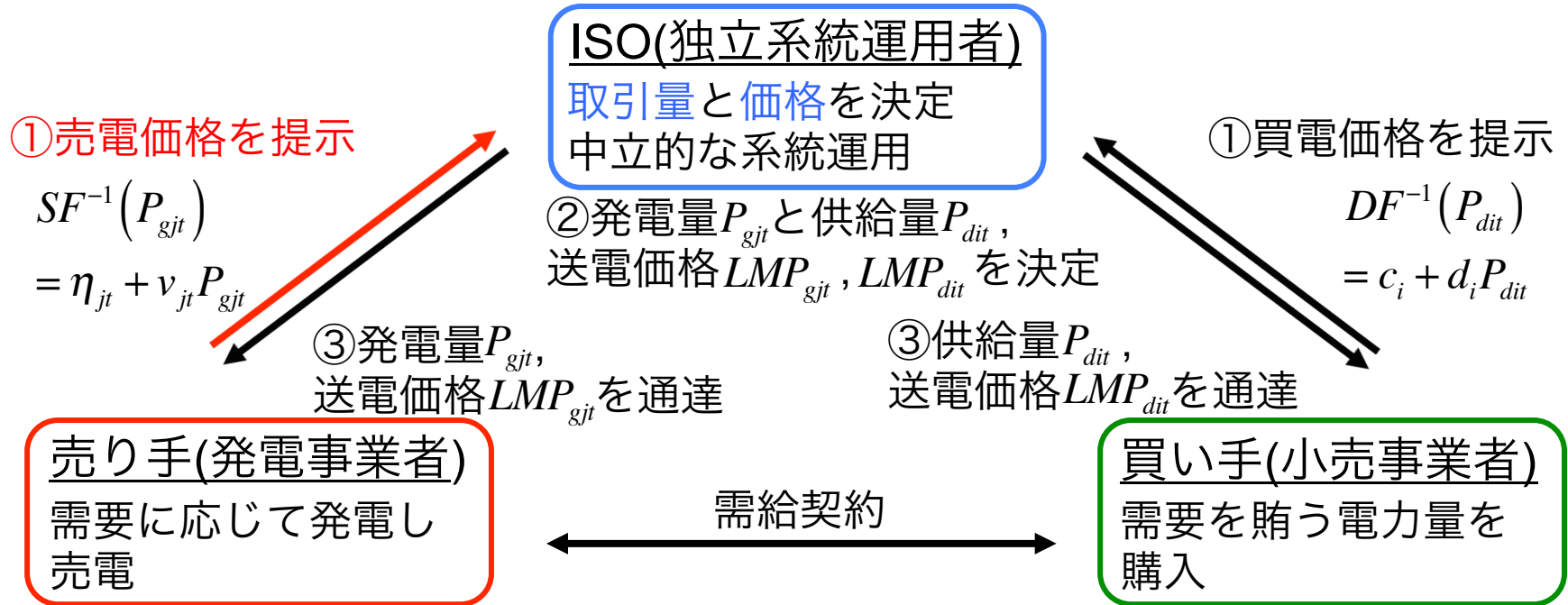
$$v_{jt} = u_j + mup_{2j} \times u_j$$

← mup_{1jt} がエージェントの行動となる

※供給曲線が逆関数であるのは, 供給曲線を価格に対する発電量と定義しているため

2. Agent-based electricity market modeling

2.2. Hourly congestion management model (ISOの視点)



ISOの役割

売り手から供給関数, 買い手から需要関数を受け取り, 最適化問題を双方の利益が最大となるように最適化問題を解き, 発電量 P_{gjt} と供給量 P_{dit} , 送電価格を決定

Maximize $P_{gjt}, \forall j; P_{dit}, \forall i$

$$\underbrace{\sum_{i \in \mathbf{D}} \left(c_i P_{dit} + \frac{1}{2} d_i P_{dit}^2 \right)}_{\text{買い手の総利益}} - \underbrace{\sum_{j \in \mathbf{G}} \left(\eta_{jt} P_{gjt} + \frac{1}{2} v_{jt} P_{gjt}^2 \right)}_{\text{売り手の総費用}}$$

送電容量制限による機会損失分を
売り手と買い手が分担

2. Agent-based electricity market modeling

■ 最適化問題

□ 社会的便益最大化問題

Maximize $P_{gjt}, \forall j; P_{dit}, \forall i$

$$\sum_{i \in \mathbf{D}} \left(c_i P_{dit} + \frac{1}{2} d_i P_{dit}^2 \right) - \sum_{j \in \mathbf{G}} \left(\eta_{jt} P_{gjt} + \frac{1}{2} v_{jt} P_{gjt}^2 \right) \quad (5)$$

買い手の総利益 売り手の総費用

subject to

$$P_{Gzt} + P_{RRzt} - P_{Dzt} = \sum_{m \in \mathbf{Z}} \left(\frac{\theta_{zt} - \theta_{mt}}{x_{mz}} \right) \quad \forall z \in \mathbf{Z} \quad (6)$$

$$P_{Gzt} = \sum_{j \in \mathbf{G}_z} P_{gjt} \quad (7)$$

需要供給バランス

総発電量

$$P_{Dzt} = \sum_{i \in \mathbf{D}_z} P_{dit} \quad (8)$$

総需要量

and

$$\left| \frac{\theta_{mt} - \theta_{zt}}{x_{mz}} \right| \leq P_{mz}^{\max} \quad \forall m, z \in \mathbf{Z} \quad (9)$$

送電線容量制約

$$P_{gj}^{\min} \leq P_{gjt} \leq P_{gj}^{\max} \quad \forall j \in \mathbf{G} \quad (10)$$

発電量の範囲

$$P_{di}^{\min} \leq P_{dit} \leq P_{di}^{\max} \quad \forall i \in \mathbf{D} \quad (11)$$

需要量の範囲

$$P_{gjt}, P_{dit} \geq 0 \quad \forall j \in \mathbf{G}, \forall i \in \mathbf{D} \quad (12)$$

(6)式をもとに送電料金(LMP)を決定

3. Methodology

3.1. Introduction to fuzzy Q-learning

Q-learning

エージェントがQ matrixに基づいて行動を選択し、環境との相互作用によって学習を行う最もよく使われる強化学習のアルゴリズム

問題点

- 状態の集合が離散かつ有限でないと取り扱うのが難しい
- 状態空間が大きくなると探索空間が指数関数的に大きくなり、学習速度が落ちてアルゴリズムの有用性が減ってしまう ➡ 次元の呪い

上記の問題点を解決するために

fuzzy Q-learning

- 関数近似の方法として、Takagi-Sugeno fuzzy systemを採用
- Q matrixを参照する代わりにfuzzy systemを用いて状態行動価値を表現

効率的に連続空間において学習

再エネ導入を考慮した電力市場のモデリングを実現

3. Methodology

3.2. Assumption and definitions

1) 電力市場は1時間ごとにやり取りが行われる

2) 状態 s_{yt} : 再エネの発電量

$$s_{yt} = P_{Ryt}, y = 1, \dots, n$$

(再エネの場所: y , iteration: t)

3) 行動 a_{jt} : 供給曲線の調整 mup_{1jt}

$$A_j = \{a_{j1}, a_{j2}, \dots, a_{jv}, \dots, a_{jN_j}\} \quad (\text{発電機の場所: } j, \text{ iteration: } t)$$

$$\eta_{jt} = h_j + \textcircled{mup_{1jt}} \times h_j$$

$$v_{jt} = u_j + mup_{2j} \times u_j \quad \rightarrow \quad mup_{2jt} = 0.5 \times mup_{1jt}$$

4) 報酬 r_{jt} : 発電事業者の利益

$$r_{jt} = \frac{\text{収入 } LMP_{gjt} P_{gjt}}{\text{総費用 } C(P_{gjt})}$$

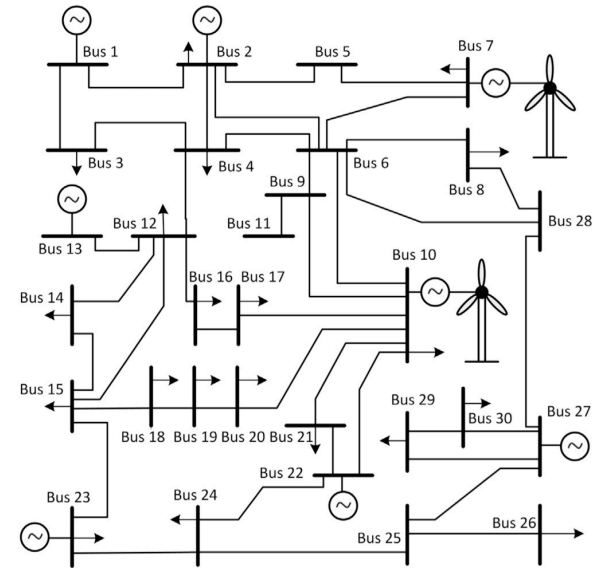
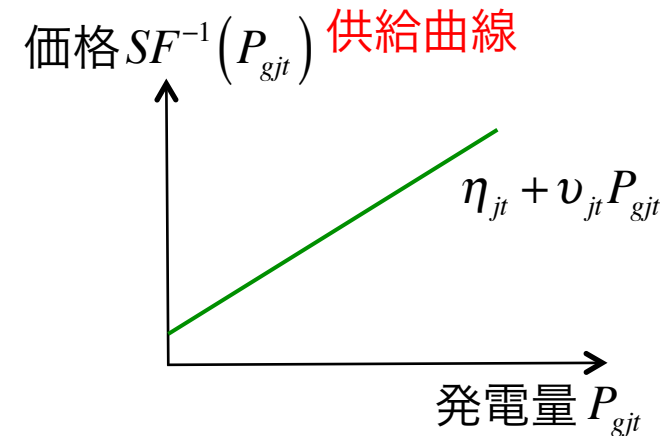


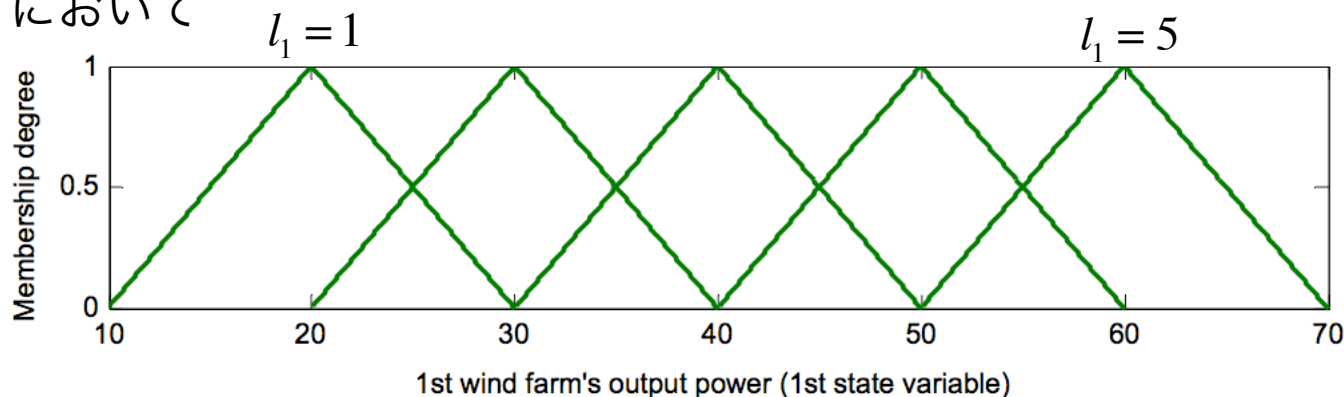
Fig. 1. Diagram of the test system.



3. Methodology

5) Fuzzy 集合 $L_{ly,yj}$:

$y=1$ において



$P_{Ry}^{\min} \leq s_{yt} \leq P_{Ry}^{\max}, y = 1, \dots, n$ の下で

$l_y = 1, \dots, k_{jy}$: Fuzzy集合 $L_{ly,yj}$ のインデックス

6) Fuzzy ルール :

IF s_{1t} is $L_{l_1 1j}$ **and** s_{2t} is $L_{l_2 2j}$ **and** $\dots s_{nt}$ is $L_{l_n nj}$

THEN $\mathbf{O}_{l_1 l_2 \dots l_n jt} = \underbrace{\left[q_{l_1 l_2 \dots l_n jt1} \quad q_{l_1 l_2 \dots l_n jt2} \quad \dots \quad q_{l_1 l_2 \dots l_n jtv} \quad \dots \quad q_{l_1 l_2 \dots l_n jtN_j} \right]}_{\text{行動の数}}$

$\mathbf{O}_{l_1 l_2 \dots l_n jt}$ はFuzzy ルールの最終的な出力(行動価値)

3. Methodology

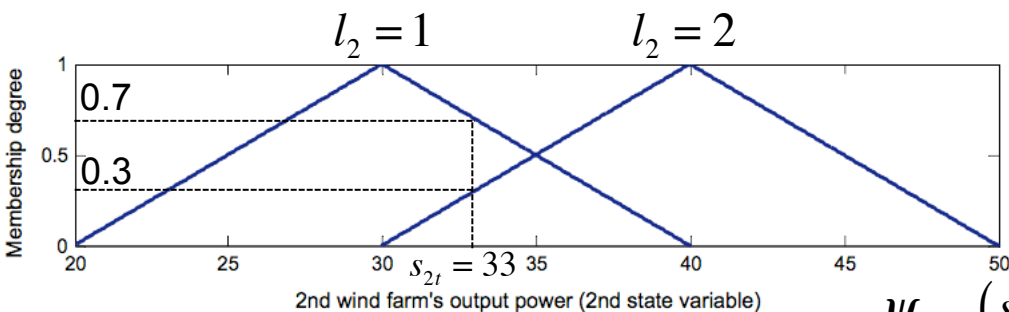
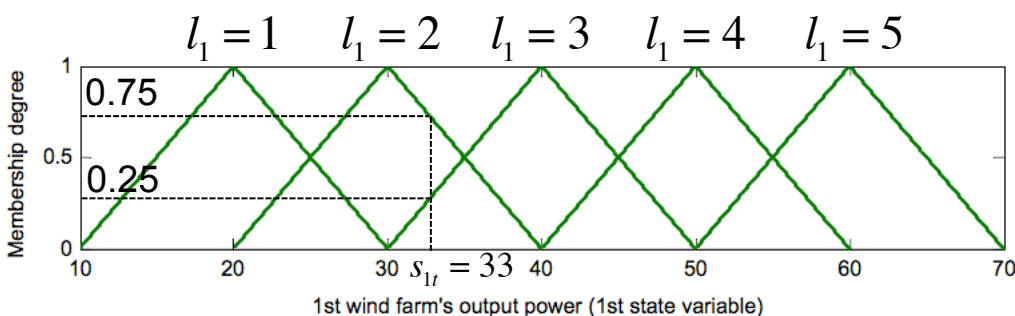
7) 正規化活性度 $\psi_{R_{l_1 l_2 \dots l_{nj}}}(\mathbf{s}_t)$:

Fuzzy ルール $R_{l_1 l_2 \dots l_{nj}}$, メンバシップ関数 $\mu_{L_{lyy_j}, l_y = 1, \dots, k_{jy}, y = 1, \dots, nj \in \mathbf{G}}$ に対して,

活性度は $\phi_{R_{l_1 l_2 \dots l_{nj}}}(\mathbf{s}_t) = \prod_{y=1}^n \mu_{L_{lyy_j}}(s_{yt})$

正規化すると,
$$\psi_{R_{l_1 l_2 \dots l_{nj}}}(\mathbf{s}_t) = \frac{\phi_{R_{l_1 l_2 \dots l_{nj}}}(\mathbf{s}_t)}{\sum_{l_n=1}^{k_{jn}} \sum_{l_{n-1}=1}^{k_{j(n-1)}} \dots \sum_{l_1=1}^{k_{j1}} \phi_{R_{l_1 l_2 \dots l_{nj}}}(\mathbf{s}_t)}$$

Ex. R_{211} の場合



$s_t = (33, 33)$

$$\phi_{R_{211}}(s_t) = \mu_{L_{211}}(s_{1t}) \cdot \mu_{L_{121}}(s_{2t}) = 0.75 \cdot 0.7$$

$$\phi_{R_{221}}(s_t) = \mu_{L_{211}}(s_{1t}) \cdot \mu_{L_{221}}(s_{2t}) = 0.75 \cdot 0.3$$

$$\phi_{R_{311}}(s_t) = \mu_{L_{311}}(s_{1t}) \cdot \mu_{L_{121}}(s_{2t}) = 0.25 \cdot 0.7$$

$$\phi_{R_{321}}(s_t) = \mu_{L_{311}}(s_{1t}) \cdot \mu_{L_{221}}(s_{2t}) = 0.25 \cdot 0.3$$

$$\psi_{R_{211}}(s_t) = \frac{0.75 \cdot 0.7}{0.75 \cdot 0.7 + 0.75 \cdot 0.3 + 0.25 \cdot 0.7 + 0.25 \cdot 0.3}$$

3. Methodology

3.3. The proposed step-by-step procedure

- 1) $t=1$
- 2) $q_{l_1 l_2 \dots l_n j t v}, j \in \mathbf{G}, v = 1, \dots, N_j, l_y = 1, \dots, k_{jy}$ に0を代入
- 3) ランダムな状態生成：iteration t でランダムな状態 \mathbf{s}_t を選ぶ
- 4) 正規活性度 $\psi_{R_{l_1 l_2 \dots l_n j}}(\mathbf{s}_t)$ を各プレイヤー(発電事業者)が計算
- 5) 状態行動価値を計算：

$$Q_{jt}(\mathbf{s}_t, a_{jv}) = \sum_{l_n=1}^{k_{jn}} \sum_{l_{n-1}=1}^{k_{j(n-1)}} \dots \sum_{l_1=1}^{k_{j1}} \psi_{R_{l_1 l_2 \dots l_n j}}(\mathbf{s}_t) q_{l_1 l_2 \dots l_n j t v}$$

- 6) 行動選択：各プレイヤー j は ϵ -greedyによって行動を選択

$$a_{jt}^{**} = \left(\arg \max_{a_{jv} \in A_j} Q_{jt}(\mathbf{s}_t, a_{jv}) \right)$$

その後、需要曲線をISOに提示

- 7) 取引調整：ISOは送電価格LMPと発電量 P_{gjt} ，供給量 P_{dit} を決定
- 8) 報酬を計算： $r_{jt} = LMP_{gjt} P_{gjt} - C(P_{gjt})$
- 9) 学習： $q_{l_1 l_2 \dots l_n j (t+1) v^*} = q_{l_1 l_2 \dots l_n j t v^*} + \alpha_j \delta_{jt} \psi_{R_{l_1 l_2 \dots l_n j}}(\mathbf{s}_t) \quad \delta_{jt} = r_{jt} - Q_{jt}(\mathbf{s}_t, a_{jt}^*)$
- 10) $t = t+1$
- 11) $t \leq T$ なら3)へ

4. Simulation and results

4.1. Data and assumptions

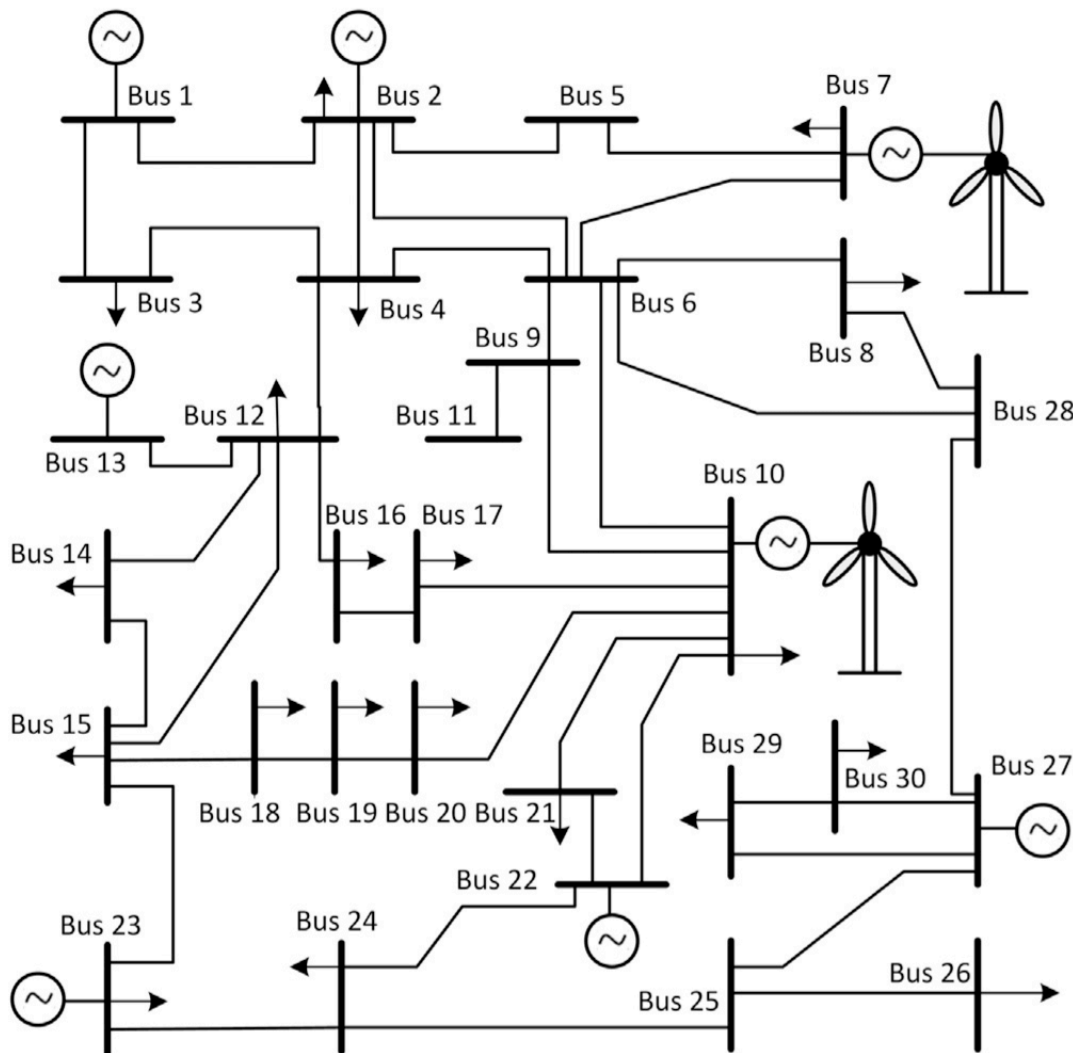


Fig. 1. Diagram of the test system.

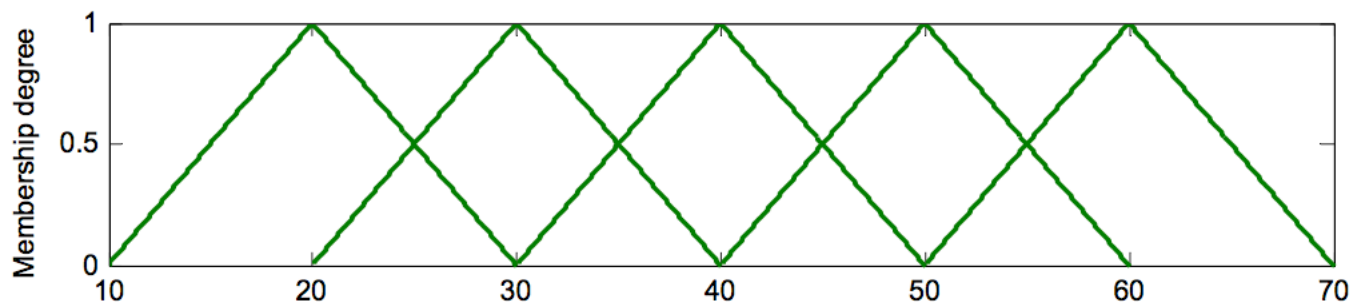
- 系統モデル
 - IEEE 30 bus
- 風力(再エネ)
 - 2箇所(bus7, 10)
- 状態(風力の発電量)
 - $S_1 = [20 \ 60] \text{MW}$ (bus7)
 - $S_2 = [30 \ 40] \text{MW}$ (bus10)
- 需要
 - 2パターン(変動, 固定)
- 行動
 - $\mathbf{A}_j = \{0, 10, 20\} \%, j \in \mathbf{G}$.
- 学習率
 - 全エージェント共通
 - $\alpha = \alpha_j, \varepsilon = \varepsilon_j, \forall j \in \mathbf{G}$.
 - 0.2 0.057

4. Simulation and results

4.2. Implementing fuzzy Q-learning

全発電事業者(6エージェント)はfuzzy Q-learningによって学習

風力1(bus7) : $S_1 = [20 \ 60]$ MW



風力2(bus10) : $S_2 = [30 \ 40]$ MW

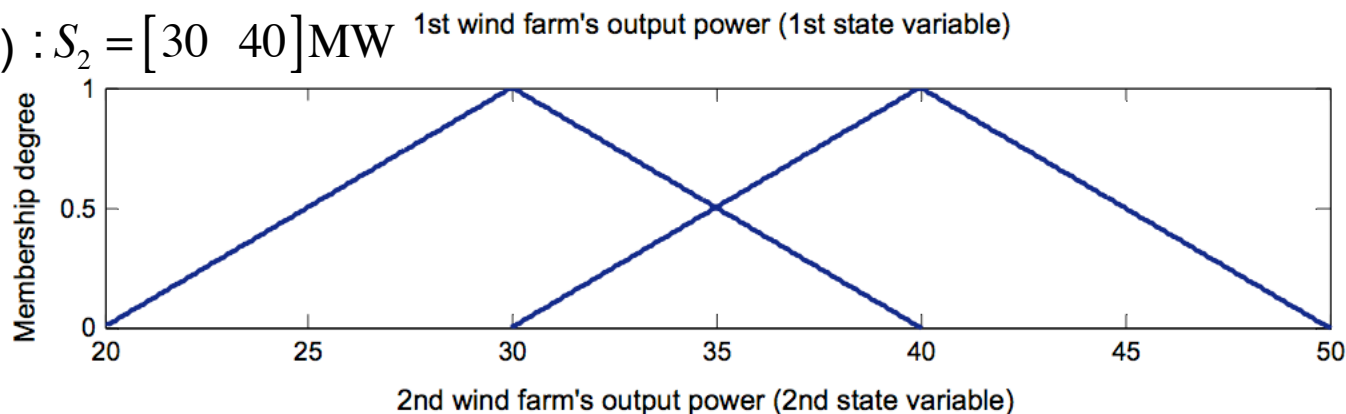


Fig. 2. Fuzzy sets of all generation units defined on the range of renewable resources connected to bus 7 and bus 10.

➡ 5×2 のfuzzyルール

4. Simulation and results

$T = 30000$ iteration後の風力2の行動 mup_{12}

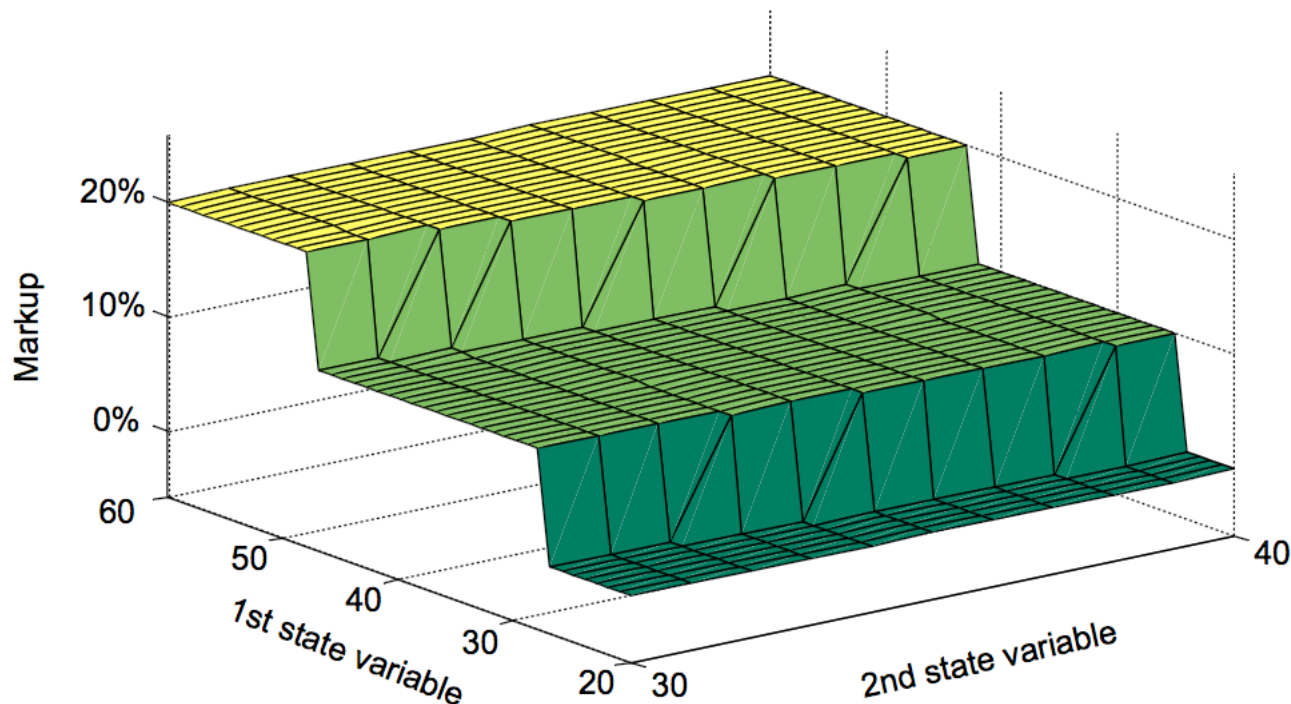


Fig. 3. Intercept markup level of 2nd generation unit's bid function after 30,000 iterations in the state space.

➡ 一般に風力の発電量が大きくなると供給関数のmarkup(行動)も大きくなる
値上げ：市場支配力が弱くなる

but

風力の発電量の増加のために送電線容量制限に引っかかる場合は、利益を減少させる場合もある

4. Simulation and results

□ 社会的便益(売り手と買い手の利益)

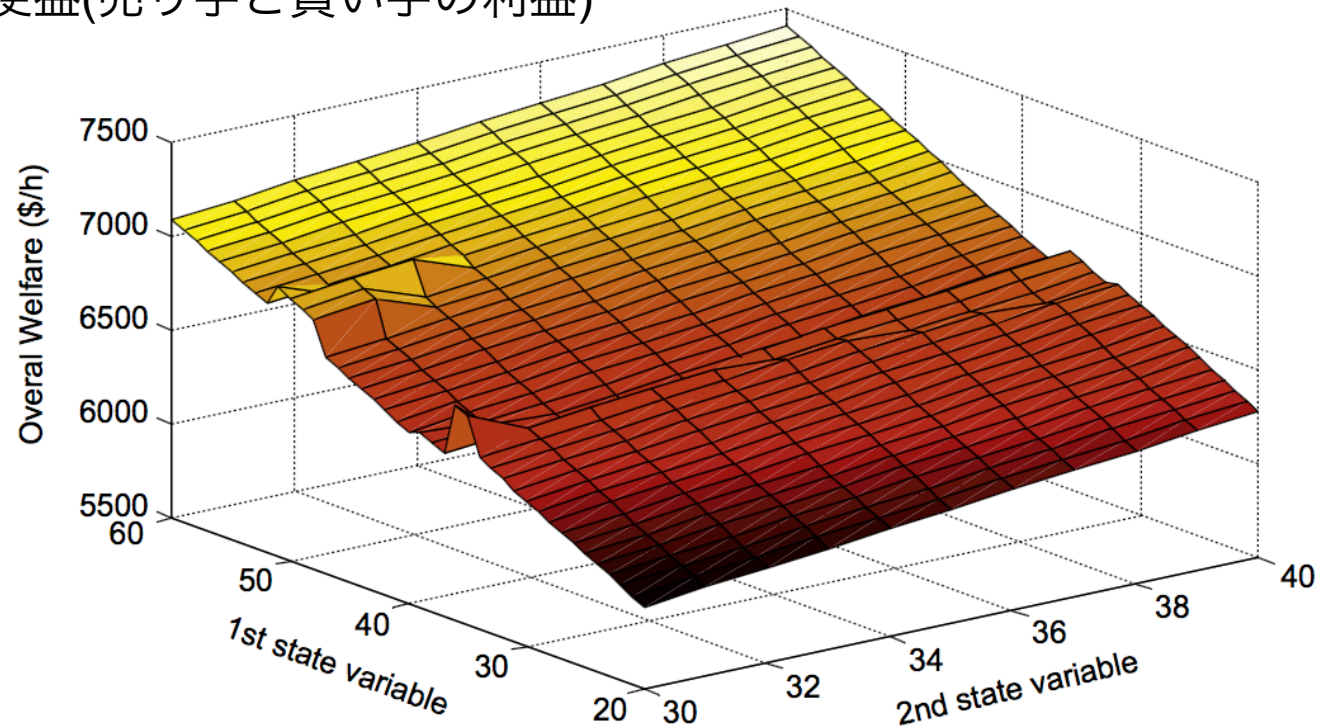


Fig. 4. Overall welfare in state space when all generation units bid based on their final bidding strategy obtained from the proposed approach after 30,000 iterations.

➡ 風力の発電量が大きくなると社会的便益も上昇

4. Simulation and results

4.3. Comparative study

Q-learningとの比較

- 反復回数 T, TQ
Q-learning: $TQ = 2000$, fuzzy Q-learning: $T = 40000$
- 離散化 N_{state}
Q-learning: 状態1,2の領域をランダムに $N_{state} = 50$ 分割
- 総反復回数 NI

• 学習率
 $\mathbf{B} = \{(\alpha, \varepsilon) | \alpha, \varepsilon \in \{0.1, 0.2, 0.3, 0.4\}\}$

組み合わせの数

Q-learning: $NI = T \times |\mathbf{B}| \times N_{run}$

40000 16 10

➡ 6.4×10^6

fuzzy Q-learning: $NI = TQ \times |\mathbf{B}| \times N_{run} \times N_{state}$

2000 16 10

➡ 1.6×10^6

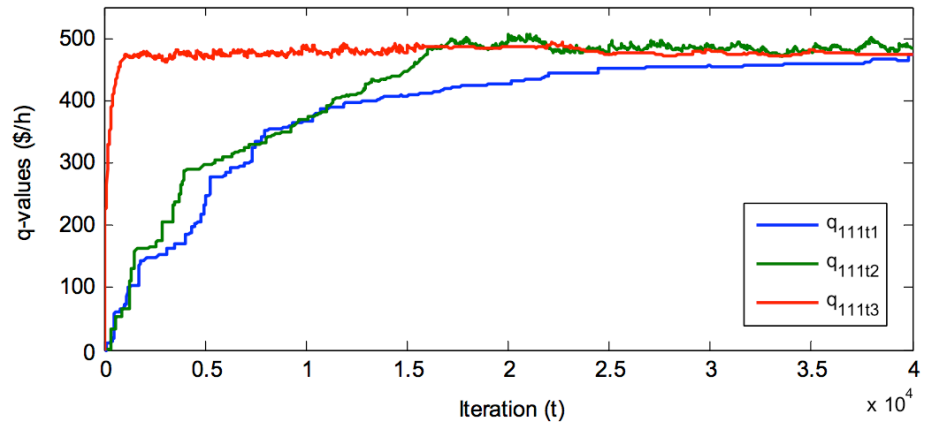
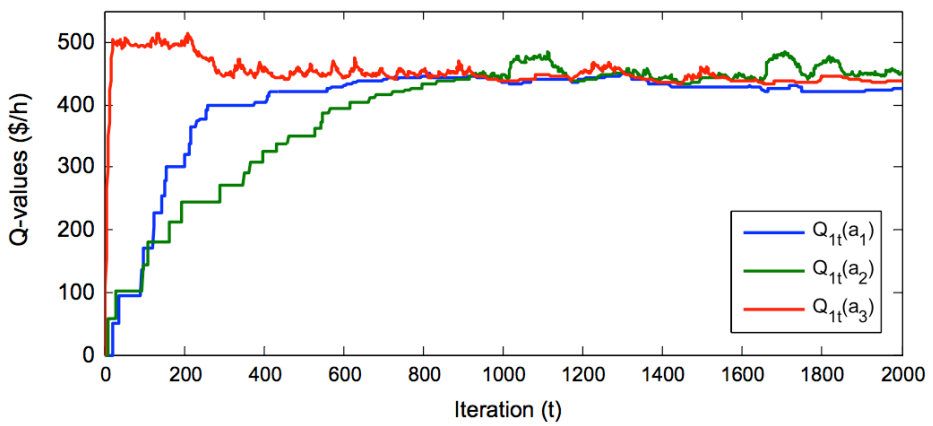


Fig. 6. Development of Q-learning approach's Q-values of generation unit 1 when the state vector is [32.24, 35.21] MW, $\varepsilon = 0.1$ and $\alpha = 0.1$. **Fig. 7.** Development of q -values pertaining to rule R_{111} corresponding to generation unit 1 when applying the proposed approach with $\varepsilon = 0.1$ and $\alpha = 0.1$.

4. Simulation and results

① ナッシュ指標

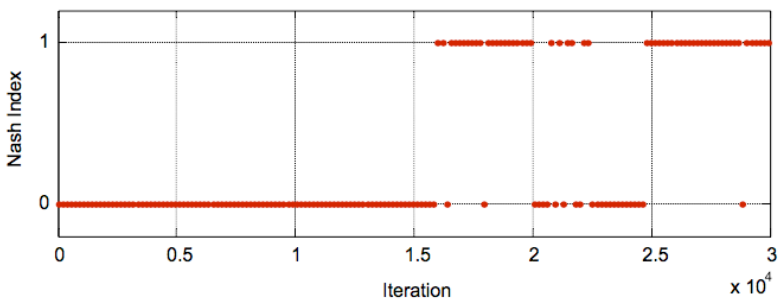


Fig. 5. Nash index during the learning iterations of the proposed method when state vector is [60, 40] MW.

ナッシュ均衡であれば1
そうでなければ0としたときの平均値

② 社会的便益

買い手の利益

$$SW = \sum_{i \in \mathbf{D}} \left(\left(c_i P_{dit} + \frac{1}{2} d_i P_{dit}^2 \right) - LMP_{dit} P_{dit} \right)$$

売り手の利益

$$+ \sum_{j \in \mathbf{G}} \left(LMP_{gjt} P_{gjt} - \left(\eta_{jt} P_{gjt} + \frac{1}{2} v_{jt} P_{gjt}^2 \right) \right)$$

③ 総送電費用

$$CC = \sum_{m \in \mathbf{Z}} \sum_{z \in \mathbf{Z}} P_{mzt} (LMP_{zt} - LMP_{mt})$$

- ④ 総需要量
- ⑤ 送電価格
- ⑥ 送電価格の標準偏差

Table 2 The comparison results. 需要変動 需要最大で固定 送電線容量90%

	Elastic demands and normal condition		Fixed demands and stressful condition	
	Q-learning	Fuzzy Q-learning	Q-learning	Fuzzy Q-learning
Number of total iterations(NI)	16×10^6	6.4×10^6	16×10^6	6.4×10^6
① Avg. Nash index	0.23	0.35	0.72	0.90
② Avg. Social welfare [\$ /h]	4171.62	4181.78	1270.29	1272.71
③ Avg. congestion cost [\$ /h]	6.26	4.42	10,009.23	9815.68
④ Avg. total demand [MWh]	273.56	273.75	289.20	289.20
⑤ Avg. LMP [\$ /MWh]	33.56	33.52	52.99	52.78
⑥ Avg. Std. LMP [\$ /MWh]	0.07	0.05	47.56	46.66

- ナッシュ指標はfuzzyの方が大きくなるが他の指標はそこまで変化しない
- 総反復数はQ-learningの40%で済む
- stressfulの場合には、送電線容量制限のため需要量が多いbusで送電費用が増大する

4. Simulation and results

各busの送電価格LMP

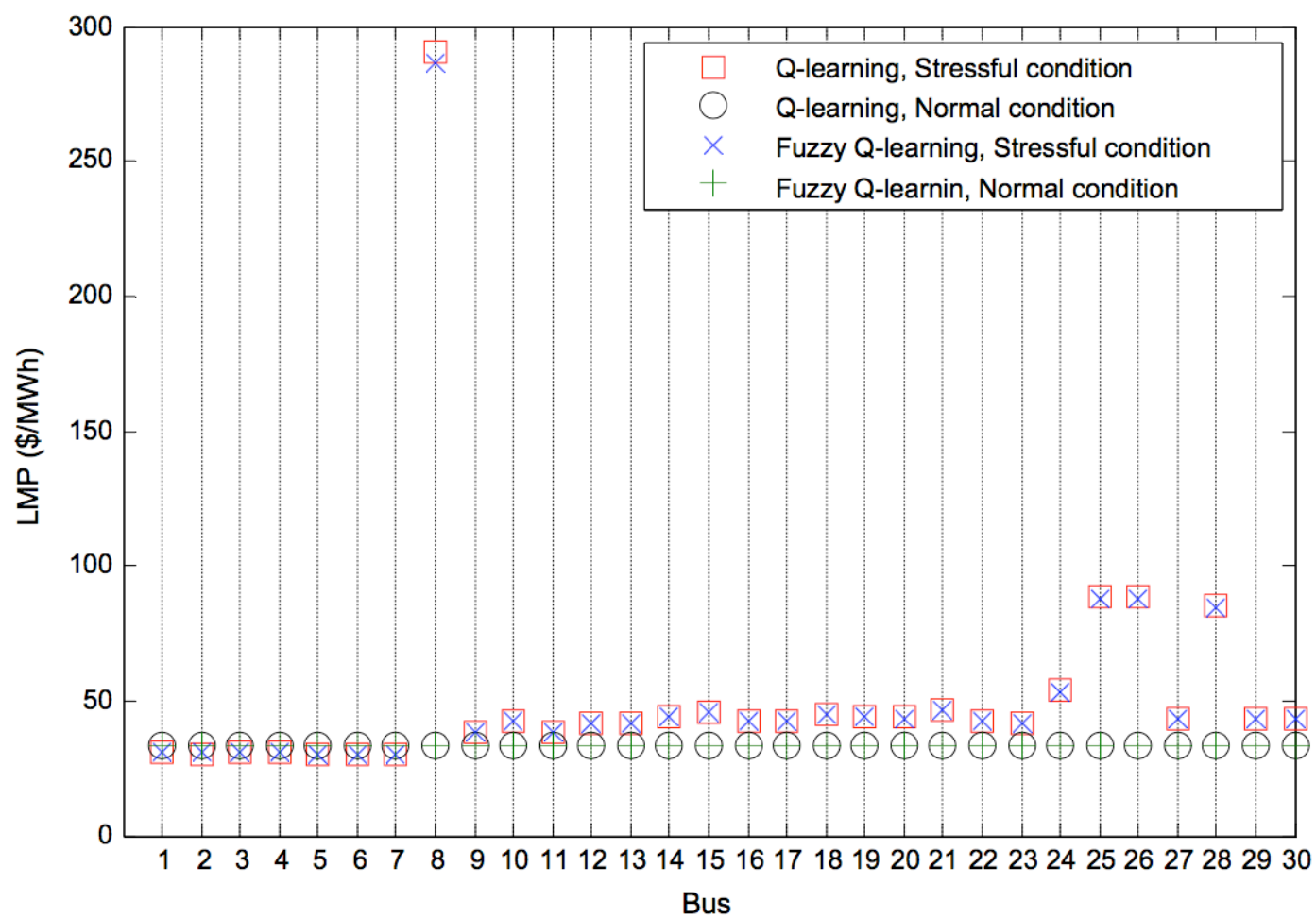


Fig. 8. LMP of each bus in different simulation conditions.

5. Conclusion

- Fuzzy Q-learningを用いることで再エネの発電量を連続変数で考慮できる電力市場モデルを構築
- 提案手法IEEE 30-bus系統でnormal, stressfulの2つのケースでQ-learningと比較して検証
- Fuzzy Q-learningの方がナッシュ指標が向上し, 総反復回数も減少させることができる