Overview
How to mitigate the Greenhouse Gas (GHG) emissions from power sector is critical to a nation’s overall environmental achievement. In order to gain insight into the changes of electricity market due to certain environmental policies and to assess the effectiveness of the policies this work employs a multi-agent model based on reinforcement learning to simulate the liberalized electricity market under environmental regulations. In this paper we discuss the Carbon Tax (CT) policy and Emission Trading (ET) policy. By conducting computer simulations we evaluate the impacts of these policies on the wholesale electricity market of and assess their contribution on CO₂ emissions reductions. We also do some discussions on the CO₂ free electricity trading policy which will be introduced to the Japan Electric Power Exchange (JEPX) from April 2009.

The structure of the paper is as follows: After the introduction section we give a brief overview about Japan’s electricity market in section two. The third section describes the multi-agent model for the wholesale electricity market. The next is the principles of reinforcement learning method that we use in our model. In section five we show the results of simulations. The last is the conclusion part.

Methods
We use multi-agent model based on reinforcement learning in this work. In our model there are a set of supplier agents \( G = \{ A_i : i = 1, \ldots, n \} \) and one demander agent \( D = \{ A_0 : j = 1 \} \). The bidding function of a supplier agent is based on its marginal cost as Equation (1) shows. The demander agent bids for the market following its marginal utility function expressed by Equation (2). In Equation (2) \( \eta \) is the electricity price elasticity.

\[
P_{gi}(q) = MC_{gi}(q) + \alpha_{gi} = 2a_{gi}q + b_{gi} + \alpha_{gi}
\]

\[
p = (P_{gi}/Q_{gi}) \times q^{1/\eta}
\]

In equation (1), \( p_{gi} \) is the bidding price of supplier agent \( A_i \), \( MC_{gi} \) is its marginal cost, and \( \alpha_{gi} \) is the bias value which indicates the bidding strategy and as we assume constant marginal cost \( a_{gi} \) equals to zero. When the market is exposed to no environmental regulations \( b_{gi} \) equals to the unit fuel cost. We add environmental cost to \( b_{gi} \). Equation (3) and (4) present the calculation of \( b_{gi} \) under the CT policy and the ET policy respectively.

\[
b_{gi} = P_{gi} + e_i \times P_{ct}
\]

\[
b_{gi} = P_{gi} + (e_i - e_{cap}) \times P_{ct}
\]

Where, \( P_{gi} \) is the fuel cost, \( e_i \) is the emission rate, \( P_{ct} \) is carbon tax rate, \( e_{cap} \) is the cap for emission rate and \( P_{ct} \) is the price for CDM credits. We assume that supplier agents with emission rate below the cap can get credits. The basic unit of the reinforcement learning is called an episode. Within one episode \( k \) the supplier agent gets feedback from the last episode and selects the optimal bias value \( \alpha_{gi} \) that results the maximum action value \( Q_k(s, \alpha) \), which is equal to maximize reward \( R_k \). We use Boltzmann distribution function (Equation (5)) to determine the optimal \( \alpha_{gi} \).

\[
\pi(s, \alpha) = \exp\left[\frac{Q(s, \alpha)}{T}\right] / \sum_{\alpha} \exp\left[\frac{Q(s, \alpha)}{T}\right]
\]

Where, \( \pi(s, \alpha) \) is the probability to choose \( \alpha_{gi} \), \( N \) is the number of options we have and \( T \) is the Boltzmann temperature. A larger \( Q(s, \alpha) \) results a higher \( \pi(s, \alpha) \). Thus, the bias value \( \alpha_{gi} \) which makes \( Q(s, \alpha) \) the largest has the highest probability to be chosen.

Results
Fig. 1~ Fig.15 show how the market share, market price and CO₂ emissions (and also the required CDM credits in the emission trading case) change with the increasing environmental cost (the carbon tax from 1,000JPY/t-CO₂ to 8,000JPY/t-CO₂ and the price of CDM credit from 1,000JPY/t-CO₂ to 8,000JPY/t-CO₂) (JPY: Japanese Yen, 1JPY≈0.01 US dollar).
As environmental cost increases there is a switching from coal to LNG (Fig. 1–Fig. 5). The fuel switching causes the PX market price to rise, but the RT market is almost unaffected (Fig. 6–Fig. 10). In the ET case, the PX market price is lower under a higher $e_{cap}$. Benefit of the fuel switching is that it results in an large progress on CO₂ emissions reduction, but after the switching further CO₂ emissions reductions are difficult (Fig. 11–Fig. 15).

In the CO₂ free electricity market all the electricity is carbon free, which means that if the electricity is generated from fossil fuel power plants the electricity has to be traded with the CDM credits that offset the CO₂ emissions. In our model, we devise the PX market into a normal market and a CO₂ free market. Supplier agents in one market can not bid for the other market. There is one demander agent who bids for both markets, and then adjusts its bidding quantities for the two markets according to the unit cost of each market. The amount of CO₂ emissions is capped, and the demander agent has to pay penalty for its excessive emissions.

Fig. 16–Fig. 18 show the traded electricity quantity and the power source structure in each market in the BAU case. Fig. 19 and Fig. 20 show how the traded electricity quantity in each market changed under different environmental regulations.

In the BAU case the demander agent only bid for electricities generated from hydro and nuclear in the CO₂ free market because they are cheaper compared with other types of plants. When the environmental regulation goes stricter, the CO₂ free market becomes more appealing to the demander agent.

**Conclusions**

In the short term, the increasing environmental cost provides enough economic incentive to cause a switching from coal to LNG. This switching results effective CO₂ emissions reductions. However, the burden of environmental cost will be passed to customers by higher PX market price. Under the CO₂ free electricity trading policy, how the demander agent change its allocation strategy is discussed in the paper.

**References**


