Journal of Renewable Energy and Smart Systems **Vol.1, No.2, 2024: 13-27**

Renewable Energy and Smart Systems

journal homepage: https://sanad.iau.ir/journal/res

Developing Environmental Based Incentive Mechanism to Increase Wind Penetration Rate in Competitive Smart Power System

Shoorangiz Shams Shamsabad Farahani[1], E. Ghaeini Hasaroueiye[2]

Roohollah Barzamini[3], Mohammad Arabian[3,*]

[1]*Department of Electrical Engineering, Islamshahr branch, Islamic Azad University, Islamshahr, Iran*
shoorangiz_shams@yahoo.com, Shoorangiz@iau.ac.ir,

[2]*Hormozgan Electric Distribution Company (HEDC) Ehsan.Ghaeni@hotmail.com*

[3]*Department of Electrical Engineering, Islamic Azad University Tehran Central Branch, Tehran, Iran;*
r.barzamini.eng@iauctb.ac.ir, m.arabian.eng@iauctb.ac.ir

Article info

Keywords:
Wind Generation
Environmental based incentive
Smart Grid
Generation expansion planning

Article history:
Received: 9 Aug 2024
Accepted: 9 Oct 2024

Abstract

Smart grids play a crucial role in transitioning to a low-carbon energy sector by ensuring the efficient and sustainable utilization of natural resources, such as wind generation. The high penetration rate of intermittent wind generation in competitive smart power systems necessitates the development of more sophisticated support schemes for this resource. Most current policies rely on financial incentives for wind generation. However, determining the cost of these support schemes in a competitive electricity market poses a significant challenge, which is addressed in this paper. In this paper, the Generation Expansion Planning (GEP) model is developed to design an environmentally-based incentive mechanism for wind generation in a smart, competitive power system. The GEP model examines the non-cooperative competition of generators at two layers. At the top layer, the generation investment game is analyzed, and at the bottom layer, the Cournot game at the power network operational level is examined. A solution algorithm based on Q-learning is used for the two-layer model, demonstrating how these layers interact to obtain environmental incentives by solving the multi-year generation expansion problem. This framework is implemented on a test system to demonstrate the effectiveness of the proposed approach. The outputs of the numerical studies include the expansion strategies of generation firms, the total profit of the wind power plant firm, the cost of the incentive mechanism for wind generation, the Wind Penetration Index (WPI)

* Corresponding author.
E-mail address: s.masoud_seyedi@aliabadiau.ac.ir (S.M. Seyyedi).

as an index of wind expansion rate, the average annual price (AAP) of electricity, and the Herfindahl-Hirschman Index (HHI) as an index of market power.

1. Introduction

A Smart Grid (SG) is an advanced electricity network that leverages communication and information technologies within the power system. Smart Grid technology addresses several critical issues in the electric power industry, including the limitation of fossil fuels and air pollution emissions, by incorporating renewable energy resources [A1]. The place of renewable energy in the new area is illustrated in Fig. 1 [A2].

To achieve the critical objectives of smart grids, the ElectriNetSM framework is designed to integrate smart grids with low-carbon central generation, local energy networks, and electric transportation, as illustrated in Fig. 2 [A3].

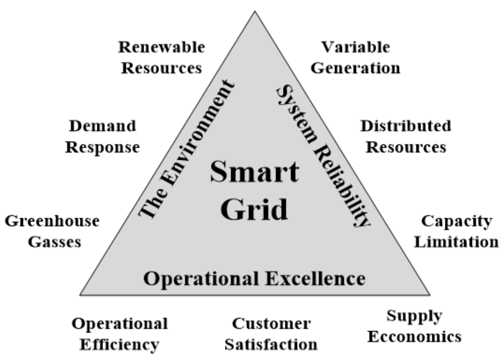

Fig. 1: The role of renewable energy in smart grid.

As shown in this figure, a fundamental component of the ElectriNetSM is low-carbon central generation. The successful implementation of the ElectriNetSM depends on meeting performance and deployment targets for several advanced technologies, which are crucial for estimating CO2 emissions reduction potential. For central generation, this includes expanded use of renewable energy and widespread adoption of CO2 capture and storage post-2020. Improving renewable energy policies, particularly for wind energy, is increasingly important. Governments frequently revise policy designs and implementations to encourage renewable electricity generation [A4]. Unlike other forms of renewable energy, advancements in wind technology have led to wind generators that are comparable to conventional units in both cost and capacity ratings. Wind energy has the potential to significantly reduce fuel costs, enhance system adequacy, and provide security against price volatility and dependence on imported fuel for many industrialized and developing nations [A5-A7]. Previous studies have explored various methodologies for short-term load forecasting in national power systems using advanced techniques like multi-layer perceptron and fuzzy inference systems [A8-A9].

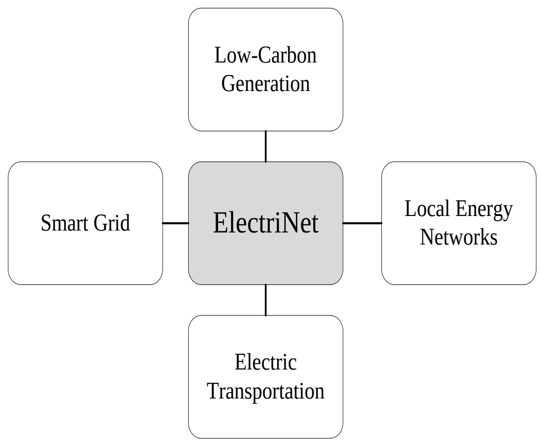

Fig. 2: The components of ElectriNet framework [2].

The production from wind units is intermittent in nature, and these resources have high investment costs. In competitive electricity markets, generation companies (GENCOs) aim to maximize their expected profits during both the operation and planning periods. Within this competitive framework, each firm strives to achieve the highest possible benefit. However, high investment costs, intermittency, and uncertainty of wind power generation are significant obstacles to the high penetration of wind power.

In [A10-A12], the investment dynamic in electricity markets of the United States and electricity crisis in California are investigated. In [A13], the incentive mechanism effect in Colombian power market is studied. Moreover, in [A14-A16], markets of Tradable Green Certificates (TGCs) and effects of environmental rules on investment is studied.

The capacity payment is considered to prevent long-term fluctuations in investment planning for traditional power plants where a fixed payment is paid to a firm which is available for electricity generation to recover a part of its investment in the long-term. Capacity payment is considered as a fixed incentive mechanism in most studies. However, in [A17], variable capacity payment and its effects on long-term power market stability is studied.

In [A2], Feed-in-Tariff support scheme is utilized to promote wind generation firm to participate in power market. In [A2], optimum reliability-based incentive for wind generation is considered. Moreover, wind farm power management is studied by high penetration of electrical vehicles (EVs) as "vehicle-to-grid (V2G). In [A5], a framework is presented to study the impacts of DR programs to increase the flexibility and wind power owners' benefits in power market. Moreover, a modified model is developed to integrate the operational planning of wind generation along with implementing DR programs. In this regard, the impact of implementing DR programs on wind's incentive cost is studied in [A5]. In [A7], a modified model is obtained to study the impact of energy efficiency approaches on support scheme cost of wind generation where energy efficiency programs are presented in two perspectives as the supply side and the demand side. Implementing energy efficiency programs in supply side is accomplished to increase the capacity factor of wind generation. However, on demand side, implementing the energy efficiency programs is presented as strategies to reduce the peak load levels. In [A18], a framework is presented on the basis of a combination of stochastic dynamic programming (SDP) algorithm and game theory. In [A18], regulatory policies including Feed-in-Tariff (FIT) incentive, quota and tradable green certificate are considered. Moreover, a model is gained to study the regulatory impacts on wind generation expansion planning where the probabilistic nature of wind generation is modeled which can calculate the optimal investment strategies including wind power uncertainty. In [19], (V2G) technology is used which increases the benefits of wind power in power market. V2G stabilizes the output power of the wind farm by providing storage capacity for electric energy during high wind speed. Moreover, a FIT incentive mechanism is considered to promote wind power owner for electricity power market participation. The suggested framework is implemented on a test system to illustrate effectiveness of the scheme. In [A20], improved incentive mechanisms for wind power investment is presented where the mechanism is addressed based on the system dynamic model. In this mechanism, wind power generations recover a part of the investment cost through incentive mechanisms. In this regard, different incentive methods are presented which is dependent on market conditions. In [A21], a reliability-based incentive mechanism is presented to promote wind power generation in a competitive electricity market. In this paper, the optimum amount of incentive is gained based on system dynamic approach. Moreover, game theory is used to model strategic uncertainties among market players in spot power market. In this regard, the Cournot game concept is suggested and the Nash equilibrium is gained for each state. Finally, LOEEpu is considered as the reliability index in the generation sector.

Global environmental concerns, driven by greenhouse gas emissions, have led governments to develop policies encouraging energy generation from renewable energy sources (RES). The primary goal of these policies is to reduce carbon emissions. To this end, various laws and policy directives have been established, such as the Renewable Portfolio Standard (RPS), which mandates that a certain percentage of energy supplied in the power system comes from RES [A22]. Mechanisms must be implemented to incentivize investment in these resources to increase the penetration rate of wind generation in smart power systems. The most commonly used mechanism to date is the fixed incentive [A19]. However, this support scheme conflicts with the competitive nature of the power market. The main issue with this type of incentive is its failure to promote competition and efficiency once wind power generation has reached a certain maturity, potentially resulting in excessive costs for society.

Therefore, the provision of support scheme costs is a critical issue investigated in this paper. To address this problem, a modified generation expansion planning model is employed to develop an environmentally-based incentive for wind power generation. In this developed mechanism, the cost of the incentive is covered by the revenue generated from penalties imposed on other polluting power generation sources. Implementing this mechanism could be an effective solution for covering the incentive costs of wind generation in competitive electricity markets. This support scheme could also enhance the wind penetration rate in smart power systems.

This paper proposes a hybrid framework based on game theory and dynamic programming (DP) to design an environmentally-based incentive mechanism grounded in the Generation Expansion Planning (GEP) problem within a competitive smart power system. Generation Expansion Planning in a restructured market involves determining the type, location, and timing of new generation capacities to be installed by competing generators in response to expected demand growth, changes in network conditions, and market design incentives. This paper addresses this challenge by developing a comprehensive game-theoretic model that incorporates market features such as multiple competing generators, a multi-year planning horizon, demand elasticity, and an environmentally-based incentive mechanism as a regulatory intervention to increase the penetration rate of wind generation. The model employs a two-layer matrix game construct. Expansion strategies and environmental incentives for wind power are derived using a reinforcement learning-based value function approximation algorithm to solve matrix games [A23-A24].

The rest of this paper is organized as follows. Section 2 presents the proposed framework for developing an environmentally-based incentive mechanism for wind generation in smart power systems. The mathematical formulation of the two-layer game model in GEP is detailed in Section 3. Section 4 introduces the Reinforcement Learning (RL) based solution algorithm for matrix games. In Section 5, this method is implemented on a test system. Finally, the last section is dedicated to the conclusion.

2. The proposed Framework for Developing Environmental Based Incentive Mechanism

As mentioned earlier, increasing environmental concerns in smart power systems have prompted governments to support large-scale integration of renewable generation by introducing mandatory Renewable Portfolio Standards (RPS) or equivalent policies. Various mechanisms have been employed to promote renewable energy worldwide [A13-A17]. The most commonly used mechanism is the Feed-In-Tariff (FIT) incentive. Under FIT, customers are required to purchase renewable energy at a predefined price or a premium on energy spot prices [A4]. These prices are generally offered in a non-discriminatory manner for every kWh of electricity produced and can be differentiated based on several parameters, such as the type of technology, the size of the installation, the quality of the resource, the location of the project, and other project-specific variables.

Quotas and tradable green certificates represent the second support scheme, where minimum shares of renewable energy are mandated for customers or producers, with penalties for noncompliance. The appeal of quotas is generally less than that of FIT. In some markets, trading of quotas is permitted, leading to the development of a green certificate market [A13].

As shown in Fig. 3, a support scheme package which includes mandatory policy and incentive-based policies was developed in [A13,A17,A19]. Under mandatory policies, power producers or customers must ensure that a portion of their production or consumption is sourced from wind generation. Incentive-based mechanisms can be categorized into market-independent incentives, market-based incentives, and reliability-based incentives. In market-independent incentives, a fixed price is paid to wind investors [A11]. In market-based incentives, a premium price above the regular electricity price is paid to wind firms [A18]. However, the high penetration of wind power resulting from these support schemes significantly impacts power system reliability. Therefore, identifying an incentive mechanism that promotes wind power expansion while maintaining reliability at a predetermined level is crucial. In [A19], to achieve a reliability-based incentive mechanism, the reliability index (LOEE) is considered.

As mentioned before, one of the most important problems of previous incentive mechanisms to increase the penetration rate of wind generation in smart environment is providing their related cost. Furthermore, these incentives have some conflict with competitive electricity market. As shown in Fig. 3, in this paper a support scheme for wind generation is developed based on generation expansion planning problem. The developed environmental based incentive mechanism reduces the problems about previous support schemes. In this incentive mechanism, the wind generation with no carbon production receive some persuasive, however, other power plants which produce emissions, are penalized.

The environmental based incentive could be introduced as an essential method in smart competitive power system; since the incentive is paid to wind generation based on adequacy. Furthermore, by considering some penalty obtained from pollutant resources, some portion or total amount incentive cost of wind generation could be provided.

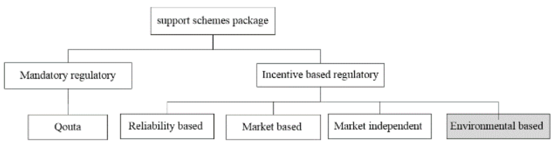

Fig. 3: Proposed support schemes package for wind generation.

The developed framework for the environmentally-based incentive mechanism is illustrated in Fig. 4, structured into fourteen blocks. The elastic demand considered in this model is depicted in block one. The profit of the investor is influenced by fluctuations in the spot price resulting from the elastic demand curve. The demand within each year is divided into four seasons, each with three states: base, medium, and peak demand. It is assumed that a portion of the demand within each season is price-responsive up to a certain price level (block 1).

Regulatory policies are illustrated in blocks 2, 3, and 4. The system regulator is assumed to set a price cap (block 2), which determines the price in the power market. This cap should provide accurate price signals for investment in new power generation. Alternatively, the price cap could be set at lower values to protect customers from significantly high prices. Block 3 illustrates the environmentally-based incentive for wind power. The amount of incentive mechanism in environmental based incentive mechanism is the function of penalty considered for other resources because of environmental pollutions (block 4). This dependency is illustrated in Fig. 4. Data required for solving the optimization problem, such as information on existing and candidate generating technologies selected for expansion planning, are indicated in block 5. The generation expansion planning model in a competitive electricity market, developed in this paper, is designed to create an environmentally-based incentive and consists of two layers. The top layer represents the investment competition among generators. This competitive decision-making scenario is modeled as an investment game (block 6).

On the other hand, the bottom layer represents the competition among generators to supply electricity to the network during the operation state. Modeling the output of wind power generation during operation is required in the GEP problem, as shown in block 7. Since an investor's objective is to maximize profit, the revenue from the power market must be calculated during the operation period. This involves calculating the electricity price at each stage and state of the DP.
The method for evaluating other hand, the bottom layer represents the competition amongst generators to supply electricity into the network. This scenario is also modeled as a
the electricity price is also crucial. For this, equilibrium analysis is applied, using a matrix game to model the strategic behavior of market players (block 8). It is also important to consider the pollution penalties of polluting power plants in the operation problem (block 9). Dynamic programming is used to solve the investment optimization problem, as indicated in block 10.

Some outputs such as expansion strategies of investors, environmental based incentive mechanism cost, wind penetration index and the revenue of emission penalty from pollutant power plant are illustrated in blocks 11, 12, 13 and 14. The advantage of this mechanism is that in such a method, supplying incentive cost is determined and revenue from fines could provide some or total amount of wind incentive cost (block 12 and 14).

3. Mathematical Basis of a two-layer game model in GEP

As mentioned before, the algorithm for designing environmental based incentive is developed based on generation expansion planning problem. The generation expansion planning model proposed in this paper consists of two layers, as shown in Fig. 5. The top layer of the model represents the investment competition amongst generators. This competitive decision-making scenario is modeled as a matrix game and is therefore referred to as investment game. On the other hand, the bottom layer represents the

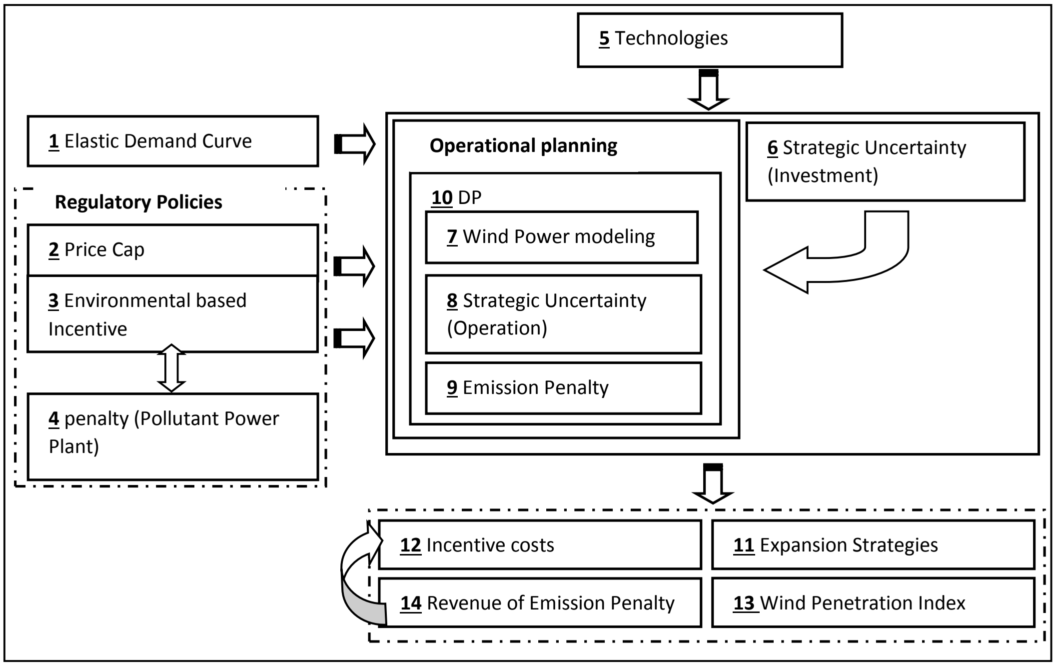

Fig. 4: The developed framework for deriving environmental based incentive Mechanism.

competition amongst generators to supply electricity into the network. This scenario is also modeled as a matrix game and is referred to as Cournot matrix game.

Each strategy that is combination of the investment game represents a possible generation capacity expansion alternative. Hence, for each of the alternatives, there exists a corresponding Cournot game, which when solved permits the examination of the profitability of each expansion alternative.

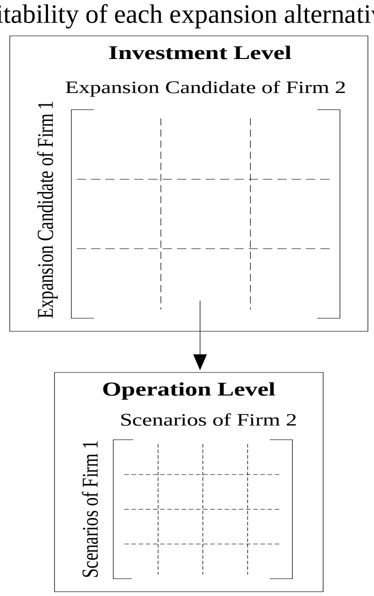

Fig. 5: Generation Expansion planning model.

3.1. Generation expansion planning problem formulation

The investment optimization problem is formulated in Equations (1) to (5). The objective function, represented by Equation (1), indicates the total discounted profits over the planning period. Equation (2) represents the capacity vector for wind power generation at time step t. Equation (3) represents the capacity vector for other power plants at time step t. Constraints (4) and (5) account for the annual variation in demand and fuel prices.

$$\Psi_0 = Max\mathrm{E}[\sum_{t=0}^{T}[(1+r)^{-t}.\Omega_t(G_t, Gw_t, D_t, Gex_t, \pi_t)]] \tag{1}$$

S.T.

$$Gw_{t+1} = Gw_t + Gwex_{t+1} \tag{2}$$

$$G_{t+1} = G_t + Gex_{t+1} \tag{3}$$

$$D_{t+1} = D_t + \Delta D_t \tag{4}$$

$$FC_{t+1} = FC_t + \Delta(FC)_t \tag{5}$$

The total expected profit of the investor for each time step is represented by Equation (6).

$$\Omega_i^* = \mathrm{E}\left[\Omega_{energy,i,t}\right] + \mathrm{E}\left[\Omega_{Reg.,i,t}\right] - C_{inv,i,t} - C_{var,i} - C_{c-tax,i} - C_{c-penalty,i} \quad (6)$$

The first term of Equation (6) represents an investor's revenues from energy sales in the spot market. The second term represents the revenues resulting from the environmentally-based incentive. The investment cost and variable cost (operation cost) are illustrated by the third and fourth terms, respectively. The fifth term implies the carbon tax, and finally, the sixth term represents the emission penalty for polluting resources.

3.1.1. Short term optimization problem

The short-term optimization problem for each firm at any stage and state of the dynamic programming problem is represented by Equations (7) to (13). Equation (7) represents the investor's revenues obtained in the power market. Equations (8) and (9) represent the operation cost and carbon tax, respectively. The revenue obtained from penalties imposed on pollutant power plants is represented in Equation (10). By imposing these penalties, a portion or the total incentive cost for wind generation can be covered by the revenue obtained from these penalties. Equation (11) represents the demand constraint. Constraints (12) and (13) set the bounds on the decision variables. Finally, Equation (14) represents the price cap constraint to prevent an increase in electricity prices, as determined by the regulatory body.

$$\Omega_{energy,i,t} = \sum_{s=1}^{N_s}\sum_{l=1}^{N_l}\sum_{f=1}^{F}\sum_{n=1}^{S_N}\left[d_{tsl}.\left(P_{Ge,tsl,n}\right).Prob_n.\pi_{tsl,n}\right] + \sum_{s=1}^{N_s}\sum_{l=1}^{N_l}\sum_{f=1}^{F}\sum_{n=1}^{S_N}\left[d_{tsl}.\left(P_{Gew,tsl,n}\right).Prob_n.\pi_{tsl,n}\right] \quad (7)$$

$$C_{var,i} = \sum_{s=1}^{N_s}\sum_{l=1}^{N_l}\sum_{f=1}^{F}\sum_{n=1}^{S_N}\left[d_{tsl}.\left(P_{Ge,tsl,n}\right).Prob_n.Var_f\right] \quad (8)$$

$$C_{c-tax,i} = \sum_{s=1}^{N_s}\sum_{l=1}^{N_l}\sum_{f=1}^{F}\sum_{n=1}^{S_N}\left[d_{tsl}.\left(P_{Ge,tsl,n}\right).Prob_n.C_{tax}\right] \quad (9)$$

$$C_{c-penalty,i} = \sum_{s=1}^{N_s}\sum_{l=1}^{N_l}\sum_{f=1}^{F}\sum_{n=1}^{S_N}\left[d_{tsl}.\left(P_{Ge,tsl,n}\right).Prob_n.C_{penalty}\right] \quad (10)$$

$$\sum_{f=1}^{F} P_{Ge,tsl,i} + \sum_{f=1}^{F} P_{Gew,tsl,i} \leq D_{tsl} \quad (11)$$

$$P_{Ge,min} \leq P_{Ge,tsl} \leq P_{Ge,max} \quad (12)$$

$$P_{Gew,min} \leq P_{Gew,tsl} \leq P_{Gew,max} \quad (13)$$

$$\pi_{tsl} \leq PC \quad (14)$$

As mentioned before, the elastic demand curve is considered in this paper. Eq. (15) presents the elastic demand curve. To find constants $A_{s,l}$ and $B_{s,l}$, the base demand $\left(D_{base,s,l}\right)$ and the reference price $\left(\prod_{base,s,l}\right)$ are used as shown in (16) and (17) [A20].

$$D_{tsl} = -A_{s,l}.\pi_{s,l} + B_{s,l} \quad (15)$$

$$A_{s,l} = \varepsilon.\frac{D_{base,s,l}}{\pi_{base,s,l}} \quad (16)$$

$$B_{s,l} = D_{base,s,l}.(1+\varepsilon) \quad (17)$$

3.1.2. Representation of the investment cost

The investment cost is represented in Equation (18). The proportion of the new generation technologies' lifetime remaining in the planning period, the inflation rate, and the fixed annuity for all time steps in the planning period are considered for adjusting the investment cost [A21].

$$C_{inv,i,t} = \sum_{f=1}^{F}\left[inv_{t,f}.Gex_{t,f}\right].\frac{\sum_{m=1}^{T-t}(1+r)^{-m}}{\sum_{n=1}^{lt,f}(1+r)^{-n}} + \sum_{\acute{f}=1}^{F}\left[inv_{t,\acute{f}}.Gwex_{t,\acute{f}}\right].\frac{\sum_{m=1}^{T-t}(1+r)^{-m}}{\sum_{n=1}^{lt,\acute{f}}(1+r)^{-n}} \quad (18)$$

3.1.3. Income obtained from environmental based incentive mechanism

The revenue of wind firms obtained from environmental based incentive mechanism is represented in Eq. (18). As mentioned before, the environmental based incentive mechanism is the function of emission penalty. One of the most important benefits of this mechanism is providing incentive mechanism cost of wind generation by penalty revenue earned from pollutant power plant (Eq. 10).

$$\Omega_{Reg,\acute{f},t} = \sum_{s=1}^{N_s}\sum_{l=1}^{N_l}\sum_{n=1}^{S_N}[d_{tsl}.(P_{Gew,tsl,n}).Prob_n.\pi_{(\pi_{tsl},C_{Penalty})}] \quad (19)$$

3.2. Top layer of matrix game: investment game

The investment matrix game can be defined by the set $\{N, A^1, \dots, A^N, R^1, \dots, R^N\}$ where, N represents the number of generators, A^i represents the set of expansion alternatives available to generator i and R^i is the payoff function for generator i. R^i can be written in the form of *N*-dimensional matrices representing the investment matrix game as illustrated by (20).

$$R^i: A^1 \times \dots \times A^N \to R \quad (20.a)$$

$$R^i = [r^i(a^1, a^2, \dots, a^N)], |A^1|, \dots, |A^N| \quad (20.b)$$

The generators select expansion alternatives from the set of available choices with the goal of maximizing their payoffs, which depend on the selection of all other generators. The concept of Nash equilibrium is used to illustrate a strategy as the most rational behavior by the generators acting to maximize their payoffs.

So, for the investment matrix game, a pure strategy Nash equilibrium is a collection of expansion alternatives $a^* = (a_*^1, \dots, a_*^N)$ for which $r^i(a_*^i, a_*^{-i}) \geq r^i(a_*^i, a_*^{-i}), \forall\, a^i \in A^i, i = 1,2, \dots, N,$ where a^i indicates the selection of a non-Nash equilibrium alternative by generator i and a_*^{-i} indicates the Nash equilibrium choice of all the other generators.

3.3. Bottom layer of matrix game: Cournot game

The Cournot matrix game can be defined by the set $\{N, \tilde{A}^1, \dots, \tilde{A}^n, \tilde{R}^1, \dots, \tilde{R}^n\}$, where $\tilde{A}^i$ represents the set of bid choices available to generator i, $\tilde{R}^i$ is the payoff function for generator i when an element $\tilde{r}^i = (b^1, \dots, b^N)$ of $\tilde{R}^i$ is the profit of generator i when the generators choose bids b^1 through b^N. $\tilde{R}^i$ for all i can be written in the form of *N*-dimensional matrices representing the Cournot matrix game as illustrated in (21).

$$\tilde{R}^i: \tilde{A}^1 \times \dots \times \tilde{A}^N \to R \quad (21.a)$$

$$\tilde{R}^i = [\tilde{r}^i(b^1, \dots, b^N)], \left|\tilde{A}^1\right|, \dots, \left|\tilde{A}^N\right| \quad (21.b)$$

The pure strategy Nash equilibrium for the Cournot game is defined as the bid choice profile b^* for which (22.b) is satisfied.

$$b^* = (b_*^1, \dots, b_*^N) \quad (22.a)$$

$$\tilde{r}^i(b_*^i, b_*^{-i}) \geq \tilde{r}^i(b^i, b_*^{-i}), \forall b^i \in \tilde{A}^i, i = 1,2, \dots, N \quad (22.b)$$

The generator profits $\tilde{r}^i = (b^1, \dots, b^N)$ where, including the Cournot matrix game are calculated by Eqs. (7),(19). The payoffs for each generator are used to populate the *N*-dimensional payoff matrices for the Cournot game. Then, the reinforcement learning algorithm is proposed to get the equilibrium bids, corresponding price and quantity allocations.

3.4. Solution algorithm for the GEP considering wind

To solve the two-layer matrix game model for GEP, the following algorithm is used.

Step 1: In the beginning of every year, potential investors evaluate the future demand growth plans, profits from previous years and the regulatory interventions to develop a set of feasible generation expansion investment alternatives.

Step 2: Let $a^i: i = 1, \dots, N$ indicate the number of investment alternatives accessible to generator i. Then, the investment matrix game A, is an N-dimensional matrix of size $a^1 \times a^2 \times \dots \times a^N$.

Step 3: For each element of matrix game $\boldsymbol{A}$, there is a corresponding matrix game of size $\prod_{i=1}^{N} \boldsymbol{b^i}$,where $\boldsymbol{b^i}$ indicates the number of bids of generator $\boldsymbol{i}$.

Step 4: After solving the corresponding optimization problem, the profit for each element of the matrix games $\left(\tilde{\boldsymbol{r}}^{\boldsymbol{i}} = (\boldsymbol{b^1}, \boldsymbol{b^1}, \dots, \boldsymbol{b^N})\right)$ will be obtained.

Step 5: Once the profits for each element of the games are obtained, a value approximation-based reinforcement learning algorithm is used to find the equilibrium profits (Ω_i^*) for the generators.

Step 6: Ultimately, the reinforcement learning algorithm, which is presented in next section, is used in matrix game A to obtain the equilibrium solution.

This procedure (steps 1-6) is represented once a year, till the generation expansion strategy for the entire planning horizon is obtained for each generator.

4. Reinforcement learning based solution algorithm for matrix games

Matrix games can be viewed as recursive stochastic games with a single state. A stochastic game can be introduced by a set $\{n, S, A^1, \dots, A^n, P, \tilde{R}^1, \dots, \tilde{R}^n\}$, which differs from matrix games through holding the additional elements. In this set, S is a finite set of states (S) of the environment, and P is the set of transition probability matrices.

In this section, we present an approach to obtain Nash equilibrium of n-player matrix games. Let $R^k(a)$ represent the reward matrix of the R^{th} player of which $r^k = (a_*^1, \dots, a_*^N)$ are the matrix elements. The value of an action a^k to player k is defined as shown in (23).

$$Val[R^k(a^k)] = \sum_{\{a^1,\dots,a^n \backslash a^k\}} p(a^{-k}, a^k) r^k.(a^1, \dots, a^k, \dots, a^n) \tag{23}$$

In this equation, $p(a^{-k}, a^k)$ indicates the probability of choice of an action combination a^{-k} by all the players while player k chose action a^k In this paper, for matrix games that have multiple players and a single state, it is assumed that there exist optimal values for all actions of the players which can give pure and mixed NE strategies. However, the probabilities $\left(p(a^{-k}, a^k)\right)$ needed to compute these values are impossible to obtain for real life problems without prior knowledge of players' behavior. Therefore, a learning approach is employed to estimate the values of the actions. By using learning method, the Eq. (23) can be rewritten by (24). In this equation, γ_t is learning parameter [A12-A13]. The algorithm presented below utilizes the value learning scheme (24) to derive pure and mixed Nash Equilibrium (NE) strategies for n-player matrix games.

$$Val[R_{t+1}^k(a^k)] = (1 - \gamma_t)[R_t^k(a^k)] + \gamma_t[r^k(a^1, \dots, a^k, \dots, a^n)] \tag{24}$$

4.1. Algorithm to obtain Nash equilibrium for n-player matrix games

In this study, it is assumed that the game has n-players and each player k has a set of A_k action choices. Thus, n reward matrices of size $|A_1| \times |A_2| \times \dots \times |A_n|$ are available. The steps to obtain the Nash equilibrium of n-player matrix games are as follows:

Step1: Eliminate elements of the matrices associated with non-rational strategies. These are strategies that will never be adopted by a rational player, regardless of the choices of other players.

Step 2: Let iteration count $t = 0$. Initialize the R-values for all player and action combinations $R(k, a)$ to an identical small positive value. Also initialize the learning parameter γ_0, exploration parameter ϕ_0, and parameters γ_t, ϕ_t needed to obtain suitable decay rates of learning and exploration. Let Maxsteps denote the maximum iteration count.

Step 3: If $t \leq Maxsteps$, continue learning of the R-values through the following steps.

(a) Greedy action selection for pure strategy Nash equilibrium: Each player k with probability $(1-\phi_t)$, chooses a greedy action for which $R^k(a) \geq R(k,\tilde{a})$. With probability ϕ_t, the player chooses an exploratory action from the remaining elements of A_k (excluding the greedy action), where each exploratory action is chosen with equal probability. Probabilistic action selection for mixed strategy Nash equilibrium can be computed using the ratio of R-values at iteration t as follows. For each player k, the probability of choosing the action $a \epsilon A_k$ is given by $\frac{R(k,a)}{\sum_{b \epsilon A_k} R(k,b)}$.

(b) R-Value Updating: Update the specific R-values for each player k corresponding to the chosen action a using the learning scheme given below.

$$R_{t+1}(k,a) \leftarrow (1-\gamma_t)R_t(k,a) + \gamma_t\big(r(k,a)\big) \quad (25)$$

(c) $set \leftarrow (t+1)$

(d) Update the learning parameters γ_τ and exploration parameter ϕ_τ, following the decay scheme given in (26) [A12-A13].

$$\theta_t = \left(\frac{\theta_0}{1+\mu}\right) \quad (26.a)$$

$$\mu = \left(\frac{t^2}{\theta_\tau + t}\right) \quad (26.b)$$

In (26), θ_0 denotes the initial value of a learning/exploration rate, and θ_τ is a large value chosen to obtain an appropriate decay rate for the learning/exploration parameters. Exploration rate generally has a large starting value and a quicker decay, whereas learning rate has a small starting value and very slow decay rate. An exact option of these values depends upon the application [A23-A24].

(e) If $t \leq Maxsteps$, go to Step 3(a), else go to Step 4.

Step 4: Equilibrium Strategy Determination: For each player k the pure strategy is action a for which $R(k,a) \geq \ max_{b \epsilon A_k} R(k,b)$. The combined strategies for all players constitute the pure strategy equilibrium.

5. Numerical studies

5.1. Description of the test system

The proposed method is implemented on a test system. Initial data for the test system is collected from [A25-A26]. The test system has a total installed capacity of 15,800 MW and a peak load of 15,000 MW in the initial year of investment. In this study, demand is considered elastic. The planning horizon is assumed to be five years. The probabilistic distribution function (pdf) of wind power generation is evaluated for each season. The electricity market considered here consists of four firms. Data regarding the ownership of the units by generation firms are shown in Table 1.

The discount rate and annual growth of demand are assumed to be 5%. The seasonal factors are 1.0, 1.1, 1.2, and 0.9 for the four seasons, respectively. The load coefficients are assumed to be 2, 1.5, and 1 for peak-load, medium-load, and base-load sub-periods, respectively. In this study, the price cap is considered to be \$80/MWh.

The outputs of the numerical studies include the expansion strategies of generation firms, the total profit of the wind power plant firm, the cost of the incentive mechanism for wind generation, the Wind Penetration Index (WPI) as an index of wind expansion rate, the average annual price (AAP) of electricity, and the Herfindahl-Hirschman Index (HHI) as an index of market power [A27]. As previously mentioned, an environmentally-based incentive mechanism based on the generation expansion planning problem is developed in this study. To compare the results, three case studies are considered in this paper (Table 2).

Tables 1: Data of generation technologies

firm	Generation technology	Variable cost (US$/MWh)	Capacity (MW)	Expansion candidate capacity (MW)	Life Time (yrs.)	F.O.R	CO2 (lbs/MW)
1	Nuclear	6.6	4000	400	40	0.02	0
2	Coal-steam	15	8000	200	40	0.04	1840
3	Gas	39.1	3000	50	20	0.01	889
4	Wind	0	800	50	30	0.03	0

Tables 2: Three case studies considered in this study.

Case 1	No incentive mechanism
Case 2	Market based incentive mechanism
Case 3	Environmental based incentive mechanism

5.2. Simulation results

Case study 1:
In this study, the impacts of the incentive mechanism on expansion planning are not considered. The results, as illustrated in Table 3, show that the wind power investor has no willingness to invest. Since no incentive is considered in this case, the WPI exhibits a descending trend over the planning period. This is in contrast to the policies of regulators and governments. Additionally, the wind firm revenue, HHI, and AAP, along with WPI variations over five years, are illustrated in Table 3.

Case study 2:
In the second case, the required market-based incentive is derived for expanding wind power based on maximizing the wind firm's revenue. The amount of the market-based incentive mechanism is the minimum incentive needed to encourage a wind firm to invest in wind generation expansion.
If the incentive mechanism is lower than this amount, wind investors would not be motivated to invest. In this case, the penetration of wind power accumulates year by year. Results such as the generation expansion strategies, wind firm revenue, incentive cost, AAP, HHI, and WPI variations over five years are illustrated in Table 4. As observed, the WPI increased from 5.06% in the initial year to 6.67% in the fifth year. The calculated HHIs are 3571 and 3622.09 for the beginning and the end of the planning period, respectively. These results indicate that the market power indices are reduced compared to case 1.

The most important problem is that in this case the cost of incentive mechanism must be provided by customers. So, the environmental based incentive mechanism will be presented in case 3.

Case study 3:
As mentioned in previous section, the way that incentive costs are provided is another key problem which is considered in this case study by developing environmental based incentive mechanism. The cost of this incentive mechanism is provided by revenue earned by penalties of other pollutant power generation. The environmental based incentive which obtained based on generation expansion problem is illustrated in Table 5.

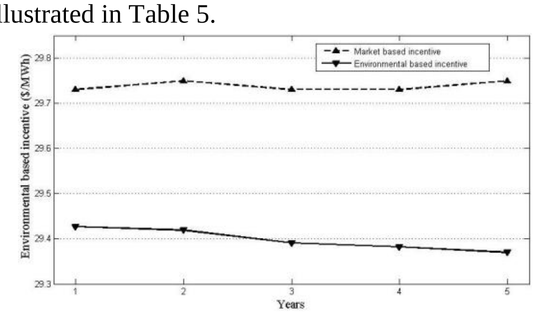

Fig. 6: Comparison of incentive mechanism in case 2 and 3.

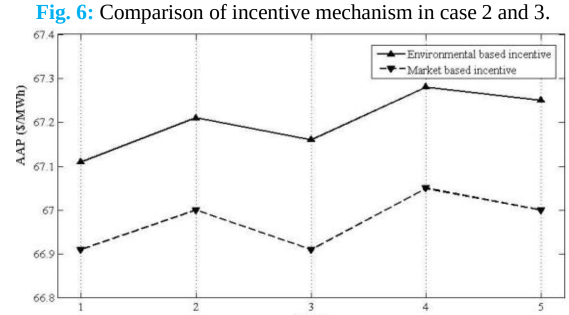

Fig. 7: Comparison of average annual price in case 2 and 3.

Tables 3: The results of case 1: No incentive mechanism

	Year 1	Year 2	Year 3	Year 4	Year 5
Strategies (firm)	400(1), 2*200(2)	2*200(2)	400(1), 2*200(2)	2*200(2)	400(1), 2*200(2)
WPI (%)	4.82	4.71	4.49	4.4	4.21
AAP ($/MWh)	66.7	67.01	66.7	67.08	66.7
Wind Benefits (M$)	93.49	93.93	93.49	94.01	93.49
HHI	3613	3683.05	3702.82	3768.87	3786.15

Tables 4: The results of case 2: (Market based incentive mechanism).

	Year 1	Year 2	Year 3	Year 4	Year 5
Strategies (firm)	400(1), 2*200(2), 2*50(4)	2*200(2), 2*50(4)	400(1), 2*200(2), 2*50(4)	2*200(2), 2*50(4)	400(1), 2*200(2), 2*50(4)
Incentive($/MWh)	29.73	29.75	29.73	29.73	29.75
WPI(%)	5.39	5.81	6.08	6.45	6.67
AAP($/MWh)	66.91	67	66.91	67.05	67
Wind Incentive Cost(M$)	42.97	48.21	53.39	58.6	63.85
Wind Revenue(M$)	135.45	147.45	159.11	171.19	182.97
HHI	3575.96	3610.06	3598.49	3631.63	3622.09

Tables 5: The results of case 3: (Environmental based incentive mechanism).

	Year 1	Year 2	Year 3	Year 4	Year 5
Strategies (firm)	400(1), 2*200(2), 2*50(4)	2*200(2), 2*50(4)	400(1), 2*200(2), 2*50(4)	2*200(2), 2*50(4)	400(1), 2*200(2), 2*50(4)
Incentive($/MWh)	29.427	29.419	29.391	29.382	29.37
WPI(%)	5.39	5.81	6.08	6.45	6.67
AAP($/MWh)	67.11	67.21	67.16	67.28	67.25
Wind Incentive Cost(M$)	42.53	47.68	53.39	57.91	63.03
Wind Revenue(M$)	135.45	147.45	159.11	171.19	182.97
HHI	3575.96	3610.06	3598.49	3631.63	3622.09

Furthermore, electricity price, wind firm profit, and wind penetration index are shown in this Table. As shown in Fig. 6, the amount of incentive is decreased in comparison with case 2 (market-based incentive). So, the incentive cost is reduced, when pollutant generation is penalized. In addition, by considering some penalties for pollutant power plant, marginal cost is increased and then the electricity price will be more than case 2 as shown in Fig. 7.

6. Numerical studies

In this paper, a generation expansion planning (GEP) model is developed to design an environmentally-based incentive mechanism for wind generation in a smart, competitive power system. The GEP model examines the non-cooperative competition of generators at two layers. At the top layer, the generation investment game is analyzed, and at the bottom layer, the Cournot game at the power network operational level is examined. A solution algorithm based on Q-learning is used for the two-layer model, demonstrating how these layers interact to obtain environmental incentives by solving the multi-year generation expansion problem.

In this developed incentive mechanism, power generation that produces pollution must pay penalties, while wind generation receives incentives for producing clean energy. One of the most important benefits of this mechanism is that the cost of the wind generation incentive is covered by the penalty revenue earned from polluting generation. Furthermore, by implementing this support scheme, investors are encouraged to invest in wind generation with lower incentives because the penalties imposed on polluting generation increase the marginal cost of power production for these resources. This penalty system further encourages firms to invest in wind generation with reduced incentives. Thus, in the developed environmentally-based incentive mechanism, the cost of the incentive is decreased and funded by penalties collected from polluting power plants.

References

[A1] P. Subhodeep;D. Tathagata Dey;Pallab Saha;Snehasish Dey;Ruma Sen, “Review on the development scenario of renewable energy in different country,” 2021 Innovations in Energy Management and Renewable Resources (52042), (2021), Kolkata, India.

[A2] E. Alishahi, S. Sayali, “Reliability-Based Short Term Scheduling of Wind Power Considering the Impacts of Electrical Vehicles and FIT Incentive Mechanism,” CIRED workshop, (2012), Lisbon, Portugal.

[A3] C.W. Gellings, “The smart grid: enabling energy efficiency and demand response,” first ed., Fairmont Press, (2009).

[A4] B. Liu;L. Sun;X. Guo;X. Han;G. Chen, “Higher Renewable Generation Utilization via the Energy Trading Mechanism in Electricity Market,” IEEE 7th Conference on Energy Internet and Energy System Integration (EI2), (2023).

[A5] E. Ghaeni, S. Shams Shamsabad Farahani, “The Operational Planning of Wind Generation and DR Programs Considering the Impacts of Incentive Mechanism,” Majlesi Journal of Energy Management, (2020) 9: 16-24.

[A6] E. Ghaeni, S. Shams Shamsabad Farahani, “Determining the location and size of the wind resources in the distribution network from the perspective of the distribution network operator using a hunting search (HUS) algorithm,” Majlesi Journal of Energy Management, December (2020) 9:1-9.

[A7] S. Yousefi sadr, S. Shams Shamsabad Farahani, “Developing Incentive Mechanism for Wind Generation Considering Energy Efficiency Programs,” Majlesi Journal of Energy Management, (2022) 10:17-31.

[A8] R. Barzamini, M.B. Menhaj, A. Khosravi, S.H. Kamalvand, “Short term load forecasting for Iran national power system and i.ts regions using multi layer perceptron and fuzzy inference systems,” Proceedings. (2005) IEEE International Joint Conference on Neural Networks.

[A9] R. Barzamini, F. Hajati, S Gheisari, M.B. Motamadine, “Short term load forecasting using multi-layer perception and fuzzy inference systems for Islamic countries,” Journal of Applied Sciences (2012) 12: 40-47.

[A10] A. Ford, “Cycles in competitive electricity markets: a simulation study of the western United States,” Journal of Energy Policy (1999) 27: 637-658.

[A11] A. Ford, “Waiting for the boom: a simulation study of power plant construction in California,” Journal of Energy Policy (2001) 29: 847-869.

[A12] A. Ford, “Boom and bust in power plant construction: lessons from the California electricity crisis,” Journal of Industry, Competition and Trade (2002) 2: 59-74.

[A13] S. Arango, “Simulation of alternative regulations in the Colombian electricity market,” Journal of Socio-Economic Planning Sciences (2007) 41: 305-319.

[A14] K. Vogstad, “Combining system dynamics and experimental economics to analyze the design of tradable green Certificates,” Proceedings of the 38th Hawaii International Conference on system Sciences (HICSS’05) (2005).

[A15] A. Dimitrovski, M. Gebremicael, K. Tomsovic, A. Ford, K. Vogstad, “Comprehensive long term modeling of the dynamics of investment and growth in electric power systems,” EPNES Workshop, Mayaguez (2004), Puerto Rico.

[A16] A. Ford, K. Vogstad, H. Flynn, “Simulating price patterns for tradable green certificates to promote electricity generation from wind,” Journal Of Energy Policy (2007) 35:91-111.

[A17] M. Assili, M.H. Javidi, R. Ghazi, “An improved mechanism for capacity payment based on system dynamics modeling for investment planning in competitive electricity environment,” Journal of Energy Policy (2008) 36: 3703-37013.

[A18] E. Alishahi, M. Parsa Moghaddam, M. K. Sheikh-El-Eslami, “An investigation on the impacts of regulatory interventions on wind power expansion in generation planning,” Energy Policy(2011) 39: 4614-4623.

[A19] E. Alishahi, “Wind Power Scheduling with Considering the Impacts of V2G Technology and FIT Incentive Mechanism,”26th International Power System Conference, Tehran, (1390).

[A20] E. Alishahi, M. Parsa Moghaddam, M. K. Sheikh-El-Eslami, “A System Dynamics Approach for Investigating Impacts of Incentive mechanisms on Wind Power Investment,” Renewable Energy (2012) 37: 310-317.

[A21] E. Alishahi, M. Parsa Moghaddam, M. K. Sheikh-El-Eslami, “A System Dynamics Approach for Evaluating the Optimum Value of Reliability-Based Incentive Mechanism for Wind Generation in GEP," General Meering (2012), San Diego, USA.

[A22] J.Lin ; Z. Fan ,"Economic analysis of complying Renewable Portfolio Standard in the State of North Carolina," North American Power Symposium (NAPS), (2015) , Charlotte, NC , 1-5.

[A23] V. Nanduri, T.K. Das, "A reinforcement learning approach for obtaining the Nash equilibrium of multiplayer matrix games," IIE Trans. 41 (2009): 158–167.

[A24] V. Nanduri, T.K. Das, P. Rocha, "Generation Capacity Expansion in Energy Markets Using a Two-Level Game-Theoretic Model," IEEE Trans. Power Syst. (2009) 24: 1165-1172.

[A25] L.A. Barroso, H. Rudnick, F. Sensfuss, P. Linares, "Economic and market impacts of renewable energy development in Spain, Germany, and Latin America," IEEE power and energy magazine (2010) 8: 22-35.

[A26] T. Couture, Y. Gagnon, "An analysis of feed-in tariff remuneration models: Implications for renewable energy investment," Energy Policy (2010) 38: 955–965.

[A27] G. Cristian Lazaroiu, M. Roscia, "Key Elements to Create Renewable Energy Communities (REC)," 12th International Conference on Renewable Energy Research and Applications (ICRERA), (2023).

[A28] A. Bouraiou; A. Dekhane;A. Slimani;A. Rouabhia;A. Neçaibia;A. Dahbi;S. Lachtar;N.Boutasseta; E. Sakher, "Windscapes of Algeria: Assessing Potential, Challenges, and Future Prospects in Renewable Energy," 14th International Renewable Energy Congress (IREC), (2023).

[A29] T. Barforoushi, M. Parsa Moghaddam, M.H. Javidi, M.K. Sheikh-El-Eslami, "Evaluation of regulatory impacts on dynamic behavior of investments in electricity markets: A new hybrid DP/GAME framework," IEEE Trans. Power Syst. (2010) 25: 1978-1986.

[A30] A. Botterud, M.D. Ilic, I. Wangensteen, "Optimal investments in power generation under centralized and decentralized decision making," IEEE Trans. Power Syst. (2005) 20: 254–263.

[A31] Y. Huang, K. Fun Rachael Ip, W. Liang, "Game Model for Optimizing Microgird PPP Project," 2019 IEEE 6th International Conference on Industrial Engineering and Applications (ICIEA), (2019).

[A32] T. K. Das, A. Gosavi, S. Mahadevan, N. Marchalleck, "Solving semi-Markov decision problems using average reward reinforcement learning," Management Science, (1999) 45: 560–574.

[A33] A. Gosavi, N. Bandla, T. K. Das, "A reinforcement learning approach to airline seat allocation for multiple fare classes with overbooking," IIE transactions, Special issue on advances on large-scale optimization for logistics, production and manufacturing systems, (2001) 34: 729–742.

[A34] S. Kannan, S. Baskar, J.D. McCalley, P. Murugan, "Application of NSGA-II Algorithm to Generation Expansion Planning," IEEE Trans. Power Syst. (2009) 24: 454-461.

[A35] H. Tekiner, D. Coit, F. Felder, "Multi-period multi-objective electricity generation expansion planning problem with Monte-Carlo simulation," Electr. Power Syst. Research (2010) 80: 1394–1405.

[A36] D. Kirschen, G. Strbac, "Fundamentals of power system economics," Wiley, (2004).

Nomenclature

Indices

l load level

S Season

t time step (year)

Constants

A is an *N*-dimensional matrix of size $a^1 \times a^2 \times \ldots \times a^N$.

A^i set of expansion alternatives available to generator i

a^i number of investment alternatives available to generator i

$\tilde{A}^i$ set of bid choices available to generator i

b^i bid of generator i

D electricity demand (MW)

$D_{base,s,l}$ base demand (MW)

D_{tsl} time duration in tsl (hr)

$\acute{f}$ number of wind generation firms

F number of other generation firms

G_t power generation except wind power (MW)

FC fuel price($/MBtu)

lt life time of new technology (year)

N number of generators

N_i number of load levels

N_s number of seasons

PC price cap ($/MWh)

$P_{Ge,min}$ minimum power of generation (MW)

$P_{Ge,max}$ maximum power of generation (MW)

$P_{Gew,min}$ minimum power of wind generation (MW)

$P_{Gew,max}$ maximum power of wind generation(MW)

$Prob_i$ probability of each scenario in wind generation modeling

r discount rate

Var variable cost ($/MWh)

$\prod_{base,s,l}$ reference price ($/MWh)

θ_0 initial value of a learning/exploration rate

ε elasticity coefficient

Variables

a^i selection of a non-Nash equilibrium alternative by generator i

a_*^{-i} Nash equilibrium choice of all the other generators

E expected benefits of investment planning

Gex_t other generation expansion in time step t (MW)

$Gwex_t$ wind generation expansion in time step t (MW)

P set of transition probability matrices

$P_{Ge,tsl,n}$ power generation in n'th scenario in tsl (MW)
$P_{Gew,tsl,n}$ wind power generation in n'th scenario in tsl (MW)
P_{Ge} total power generation (MW)
S finite set of states (s)
γ_t learning parameter
ϕ_t exploration parameter
$\pi_{(\pi_{tsl}, C_{Penalty})}$ amount of reliability based incentive ($/ MWh)
$\prod$ price of electricity($/ MWh)

Functions

$\Omega_{energy,f,\acute{f},t}$ investor's revenues from energy sales ($)
$\Omega_{Reg,\acute{f},t}$ wind firm revenues from incentive ($)
Ω_t expected net profit in year t ($)
Ω_i^* total expected profit of investor i for each time step
$C_{c-tax,f}$ carbon tax cost ($)
$C_{inv,f,\acute{f},t}$ investment cost in year t ($)
$C_{c-penalty,i}$ Penalty cost of pollutant power plant ($)
$C_{var,f}$ variable cost(operation cost) ($)
$p(a^{-k}, a^k)$ probability of choice of an action combination a^{-k} by all the players while player k chose action a^k.
R^i pay off function for generator i
$R^k(a)$ reward matrix of the k^{th} player
$\tilde{r}^i = (b^1, b^1, \ldots, b^N)$ profit of generator i when the generators choose bids b^1 through b^N.
Ψ_0 total discounted profits over the planning period