
Deep Reinforcement Learning to Minimize Long-Term Carbon Emissions and Cost

Anonymous Author(s)

Affiliation

Address

email

Abstract

Placeholder: A transition from a high carbon electricity supply to a low-carbon system is central to avoiding catastrophic climate change [1]. Much of the work in decarbonisation relies on a low-carbon electricity supply, such as cooling, heating and automotive, amongst others. Such a transition must be made in a gradual approach to avoid frequent collapse of the electricity supply.

Renewable energy costs, such as solar and wind sources, have dropped in price making them cost competitive with fossil fuels. This is projected to continue into the future [2]. The future cost of generation, demand and fuel prices, however, remain uncertain over the long-term future. These uncertainties are risks which investors must analyse while making long-term investment decisions.

1 Introduction

A transition from a high carbon electricity supply to a low-carbon system is central to avoiding catastrophic climate change [1]. Much of the work in decarbonisation relies on a low-carbon electricity supply, such as cooling, heating and automotive, amongst others. Such a transition must be made in a gradual approach to avoid frequent collapse of the electricity supply.

Renewable energy costs, such as solar and wind sources, have dropped in price making them cost competitive with fossil fuels. This is projected to continue into the future [2]. The future cost of generation, demand and fuel prices, however, remain uncertain over the long-term future. These uncertainties are risks which investors must analyse while making long-term investment decisions.

In this paper, we use the deep deterministic policy gradient (DDPG) reinforcement learning algorithm to simulate the behaviour of investors over a 50-year horizon [3]. The environment used was a modified version of the FTT:Power model [4]. FTT:Power is a global power systems model that uses logistic differential equations to simulate technology switching.

We modified the FTT:Power model to use the DDPG algorithm in place of the logistic differential equations to simulate investment decisions, as well as only simulating two countries: the United Kingdom and Ireland. The DDPG algorithm allowed us to simulate the decisions made by investors made under imperfect information over a 50-year period. That is, they did not have knowledge of future demand, fuel and generation costs.

The reinforcement learning algorithm enabled us to model the behaviour of an investor without perfect knowledge of the future, with a view to reduce carbon emissions and overall cost of the system. The reinforcement learning agent is a single actor that invests in both the UK and Ireland. This work enabled us to see whether a low-carbon mix is possible over the next 50-years to avert climate change.

34 Through this work, it is possible to assess whether a low cost, low-carbon electricity mix is viable
 35 over the long-term using a deep reinforcement learning investment algorithm. Our approach is in
 36 contrast to a mixed-integer linear programming problem, where full knowledge of the time-horizon is
 37 required.

38 Oliveira *et al.* also use reinforcement learning for the capacity expansion problem [5]. They, however,
 39 focus on a 20-year time horizon. Kazempour *et al.* use a mixed integer linear programming approach
 40 to solve the generation investment problem [6].

41 **2 Model and methodology**

42 In this paper, we used the Future Technology Transformations for the power sector model
 43 (FTT:Power). This model represents global power systems based on market competition, induced
 44 technological change and natural resource use and depletion. Induced technological change occurs
 45 through technological learning produced by cumulative investment and leads to nonlinear path depen-
 46 dent technological transitions [4]. The model uses a dynamic set of logistic differential equations for
 47 competition between technology options.

48 For this work, however, we modified the FTT:Power model to use the deep reinforcement learning
 49 investment algorithm, DDPG. That is, the size of the investment made in each technology was made
 50 by the DDPG algorithm. In addition, we reduced the model to only consider the countries of Ireland
 51 and the UK. This enabled us to iterate through enough episodes for the reinforcement learning to
 52 converge to an optimal reward.

53 **Reinforcement Learning**

54 The investment decision making process can be formulated as a Markov Decision Process (MDP) [7].
 55 In an MDP environment, an agent receives an observation about the state of their environment s_t ,
 56 chooses an action a_t and receives a reward r_t based upon their environment and action. Solving an
 57 MDP consists of maximising the cumulative reward over the lifetime of the agent.

58 For our simulation environment, the agent makes a continuous investment decision for each energy
 59 technology, in each region and each year, starting from 2017 until 2060. Technology switching is
 60 modelled using a pairwise comparison of flows of market shares of different electricity generation
 61 capacity. That is, how much capacity flows from one technology to another.

62 The agent’s observation space is a matrix consisting of the electricity produced by each technology,
 63 total capacity, total CO₂ emissions over the simulation, levelized cost of electricity (LCOE) both
 64 including taxes and not including taxes, cumulative investment in each technology, investment in new
 65 capacity, carrier prices by commodity, fuel costs and carbon costs.

66 The reward r is defined as:

$$r = - \left(1000 \times \text{CO}_{2e} + \frac{LCOE}{1000} \right), \quad (1)$$

67 where CO_{2e} is equal to total CO₂ emissions over the simulation, and $LCOE$ is equal to LCOE,
 68 excluding taxes. The scaling factors were used to place the $LCOE$ and CO₂ on the same scale. The
 69 reward was multiplied by -1 due to the RL algorithm maximising reward and our requirement to
 70 reduce both LCOE and CO₂ emissions.

71 RL approaches have been used to solve MDP [8]. In recent times, however, RL has been extended to
 72 incorporate Deep Reinforcement Learning (DRL). DRL relies on deep neural networks to overcome
 73 the problems of memory complexity and computation complexity [9].

74 We applied the deep deterministic policy gradient (DDPG) DRL algorithm [3] from the Ray RLlib
 75 package [10]. The DDPG algorithm is made up of an actor and critic network. We designed both of
 76 these to have two hidden layers, made up of 400 and 300 units per layer. The training batch size was
 77 set to 40,000. We chose these numbers due to their implementation in Ray RLlib and our inability to
 78 tune the hyper-parameters, due to a two week running time.

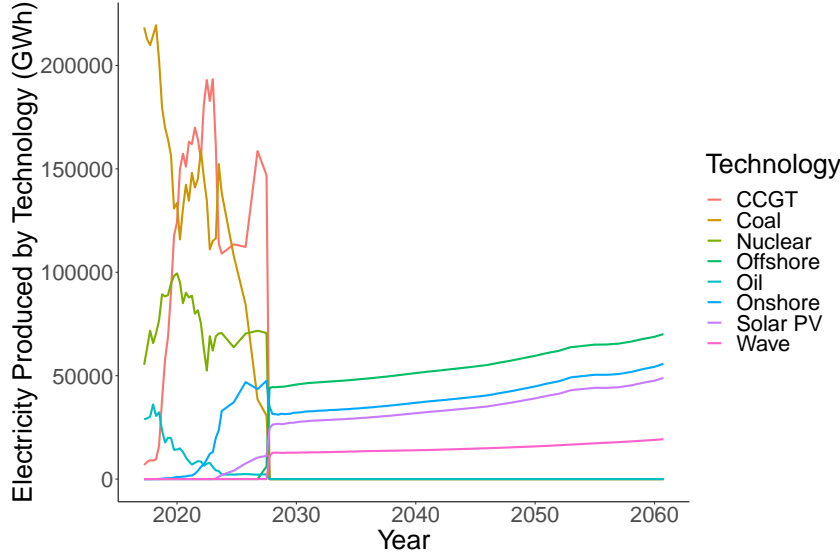


Figure 1: Mean outputs of various technologies vs. mean absolute error from 2018 to 2035 in ElecSim.

3 Results

- Display electricity mix over time-horizon
- Investment in solar, offshore, onshore and wave
- Visualise carbon emissions and electricity price over time
- Discuss these results and detail

References

- [1] A. J. M. Kell, M. Forshaw, and A. S. McGough, "Long-Term Electricity Market Agent Based Model Validation using Genetic Algorithm based Optimization," *The Eleventh ACM International Conference on Future Energy Systems (e-Energy'20)*, 2020.
- [2] IEA, "Projected Costs of Generating Electricity," p. 215, 2015.
- [3] J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous learning control with deep reinforcement," *International Conference on Learning Representations (ICLR)*, 2016.
- [4] J. F. Mercure, "FTT:Power A global model of the power sector with induced technological change and natural resource depletion," *Energy Policy*, vol. 48, pp. 799–811, 2012.
- [5] F. S. Oliveira and M. L. Costa, "Capacity expansion under uncertainty in an oligopoly using indirect reinforcement-learning," *European Journal of Operational Research*, vol. 267, no. 3, pp. 1039–1050, 2018.
- [6] S. J. Kazempour, A. J. Conejo, and C. Ruiz, "Strategic generation investment using a complementarity approach," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 940–948, 2011.
- [7]
- [8] R. S. Sutton and A. G. Barto, "An introduction to reinforcement learning," *The MIT Press*, 2015.
- [9] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "A Brief Survey of Deep Reinforcement Learning," *IEEE Signal Processing Magazine*, pp. 1–16, 2017.
- [10] E. Liang, R. Liaw, P. Moritz, R. Nishihara, R. Fox, K. Goldberg, J. E. Gonzalez, M. I. Jordan, and I. Stoica, "RLlib : Abstractions for Distributed Reinforcement Learning," 2014.