Revealing the Oil Majors' Adaptive Capacity to the Energy Transition with Deep Multi-Agent Reinforcement Learning

Dylan Radovic* Imperial College London

Lucas Kruitwagen University of Oxford

Christian Schroeder de Witt University of Oxford

Abstract

A low-carbon energy transition is transpiring to combat climate change, posing an existential threat to oil and gas companies, particularly the Majors[1]. Though Majors yield the resources and expertise to adapt to low-carbon business models, meaningful climate-aligned strategies have yet to be enacted. A 2-degrees pathways (2DP) wargame was developed to assess climate-compatible pathways for the oil Majors. Recent advances in deep multi-agent reinforcement learning (MARL) have achieved superhuman-level performance in solving high-dimensional continuous control problems. Modeling within a Markovian framework, we present the novel 2DP-MARL model which applies deep MARL methods to solve the 2DP wargame across a multitude of transition scenarios. Designed to best mimic Majors in real-life competition, the model reveals all Majors quickly adapt to low-carbon business models to remain robust amidst energy transition uncertainty. The purpose of this work is provide tangible metrics to support the call for oil Majors to diversify into low-carbon business models and, thus, accelerate the energy transition.

1 Introduction

An energy transition is manifesting in efforts to combat climate change and meet the goals set out by the Paris Agreement. The Integrated Assessment Modeling Consortium (IAMC) and Institute for Applied Systems Analysis (IIASA) [8] exhibited a multitude of potential climate change mitigation pathways in line with climate targets. Though pathways vary, all scenarios are predicated on low-carbon energy sources supplying the majority of global energy consumption. Deep decarbonization, however, has yet to be realized as fossil fuel growth persists, thus prolonging an energy transition. Carney [3] urges the need for massive reallocation of capital toward low-carbon solutions and McCollum [9] quantifies the investment gaps needed to be filled to achieve Paris. As individual countries act to answer these calls and deliver a collective transition to a low-carbon economy, risks will arise and companies may need to rapidly adapt. The Task Force on Climate-Related Financial Disclosures [15] highlighted such transition risks and their potential financial impacts. Evidently, the risks inherent to a low-carbon energy transition pose an existential threat to the fossil fuel industry.

Since the 2014 oil price collapse, international oil companies (IOCs) have significantly underperformed in contrast to economic growth realized in the same period, one of the many signs suggesting a faltering business model [14, 17]. A crucial financial recovery for IOCs becomes increasingly unlikely as the impending energy transition unfolds. Though posing an existential threat to the industry, the energy transition gives rise to significant opportunities for such actors, particularly the oil and gas

*Corresponding Author: Dylan Radovic, dylan.radovic19@imperial.ac.uk

[1]The oil and gas Majors comprise of the seven largest publicly-traded international oil companies by market capitalization: ExxonMobil, Chevron, ConocoPhillips, Shell, BP, Total, and Eni

Tackling Climate Change with Machine Learning workshop at NeurIPS 2020.

Majors. The Majors are the most susceptible to energy transition risks due to their dependence on high-cost reserves vulnerable to asset stranding and shareholder capital at risk of divestment [4, 5] yet offer the balance-sheets and project management expertise necessary to finance and scale low-carbon technologies that could largely fill the Paris-aligned investment gaps [7]. Central to this argument is the question of when will the Majors begin to transition towards low-carbon business models. Despite clear signals, the Majors have yet to make any significant moves into low-carbon business models which would prove resilient to carbon demand shocks and, thus, produce stable returns, drawing concerns from investors and governments alike [6].

Literature regarding oil and gas companies in the energy transition echo the same sentiments: Majors can and should adapt to a low-carbon world. However, besides outlining potential low-carbon strategies, literature fails to provide tangible first low-carbon mover risks and rewards. Quantifying such outcomes may suggest an end to the Majors' capital reallocation waiting game by reducing diversification uncertainties. Founded on the shareholder primacy model of capitalism, the Majors operate to maximize growth and returns to shareholders. This work aims to provide tangible insights into the Majors' adaptive capacity to the energy transition by identifying first low-carbon mover potential outcomes in relation to such metrics and, thus, strategy pathways robust to transition uncertainty. In efforts to achieve this, this work presents a novel deep MARL model utilized to solve a continuous microeconomics wargame.

2 Solving a Wargame

The Oxford Sustainable Finance Programme and E3G developed the 2DP wargaming tool to "help inform company, investor, government and civil society thinking around the pathways the oil and gas majors can take to become 1.5°C/2°C-compatible" by simulating oil Majors in competition within varying transition scenarios [2]. 2DP wargame participants role play as IOCs tasked to responsibly allocate capital across oil, gas, and 'green' markets in efforts to minimize asset stranding while maximizing long-term market capitalization. Though the action space and objectives are clear, the requirement of human players reveals more about human bias than the Majors' robust strategy pathways and first mover outcomes. Aware of this limitation, the authors proposed a future 2DP development to involve "agent-based simulation to discover both optimal and scenario-robust strategies." This proposal prompts the purpose of this work going forward: to solve the 2DP wargame.

Game theoretic approaches solve strategic interactions between a set of agents. Applying such methods to the 2DP wargame requires framing the long-term player relationships as an infinitely-repeated game. Conventional game theoretic algorithms, however, fail to effectively solve games with no end, such as the 2DP wargame. Moreover, modeling the 2DP wargame's high dimensional state space and non-linear payoffs with game theory proves analytically and computationally intractable.

To overcome these barriers, developments in reinforcement learning are explored. The 2DP wargame is a continuous control problem framed as a Partially Observable Stochastic Game (POSG)—a Markovian framework resemblant of real-life market competition. Recent advances in deep MARL [16, 1] have led to superhuman-level achievements in solving high-dimensional, continuous control problems with multiple competing agents. In efforts to achieve this work's aim, deep MARL methods are applied to solve the 2DP wargame.

3 2DP-MARL

We present the 2DP-MARL model which utilizes deep MARL methods, including the application of the Advantage Actor-Critic [11] and Proximal Policy Optimization [12] algorithms, to solve the 2DP wargame under different 1.5°C scenarios and, thus, provide insight from the Majors' emergent robust strategy pathways in response to energy transition uncertainty. Appendix A details adjustments and additions made to the 2DP wargame to enable the implementation of deep MARL and best mimic real-life market competition. To the best of our knowledge, there is not yet a deep MARL model that solves a microeconomics simulation relevant to the oil and gas energy transition discussion.

Initialized with real financial and industry data, agents compete as Majors within the 2DP wargame environment under a given 1.5°C scenario extracted from the IAMC/IIASA scenario ensemble. Following the POSG framework, agents are granted incomplete observations of their environment and opponents. The agents' available action set draws directly from the original 2DP wargame. Actions

Table 1: Overview of the 2DP-MARL baseline, dividend-focused reward function.

Condition	Value	Comments
Negative Return on Equity	-5	Encourages positive net income, mitigates debt engulfment
Negative Total Enterprise Value	-5	Mitigates debt engulfment
Insufficient, Low Dividend Payouts	-5	Majors maintain high dividends in efforts to keep investors satisfied despite experiencing volatility
If above conditions are not met, positive rewards are granted		
Yearly Dividend Payouts	$\frac{Dividend\ Payouts}{1e5}$	Scaled to the magnitude of negative rewards

include capital allocation across oil, gas, and 'green' markets as well as cash borrowing and debt payoff, dividend payouts and player-to-player trading. For our baseline results, the reward function is designed in parallel with the goals of a Major and the desires of its investors: to maximize shareholder value. The only tangible method for 2DP-MARL agents to achieve this in-game, therefore, is to maximize total dividend payouts. Focusing the reward function on dividend payouts requires a company to maximize its capital efficiency in response to its competitors' choices and the scenario at play. Negative rewards are introduced to encourage robust strategies and realistic agent behavior akin to the Majors they represent. Table 1 tabulates an overview of the applied baseline reward function.

2DP-MARL agents are trained across 408 IAMC/IIASA energy transition scenarios. Upon reaching the endgame, at year 2040, the environment resets to generate a new climate-aligned scenario at random. Unique to each IAMC/IIASA scenario are the projections with respect to oil and gas demand, dictating the respective carbon asset's production and pricing values. Appendix B displays all unique IAMC/IIASA oil and gas demand projections. Other global scenario metrics that shape the environment's dynamics, such as debt interest, 'green' return on investment, and available 'green' assets for acquisition, are drawn directly from the original 2DP wargame.

4 Baseline Results & Analysis

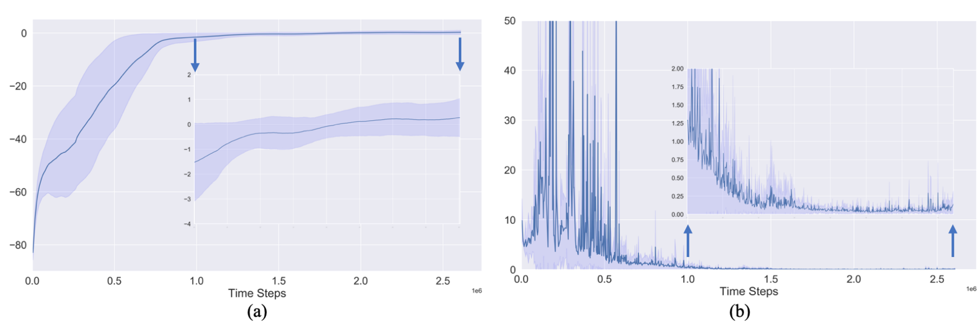


Figure 1: Convergence of (a) average rewards and (b) loss function captured while training the 2DP-MARL model across several random seeds.

Upon convergence, after over 130,000 epochs, Figure 1, trained 2DP-MARL agents were evaluated across all IAMC/IIASA scenarios to reveal robust strategy pathways unique to each climate-aligned scenario. Individual Majors' yearly average dividend payouts and market valuation as well as oil and gas, 'green' and debt asset holdings across all IAMC/IIASA scenarios are displayed in Figure 2.

While merely a baseline test, meaningful insights can immediately be drawn from the results. Amidst energy transition uncertainty, Majors successfully realize robust strategies as their market valuations, Figure 2(a), and dividend payouts, Figure 2(b), are maximized by the endgame. To further understand the emerged robust business models, average oil and gas, 'green' and debt asset holdings are examined.

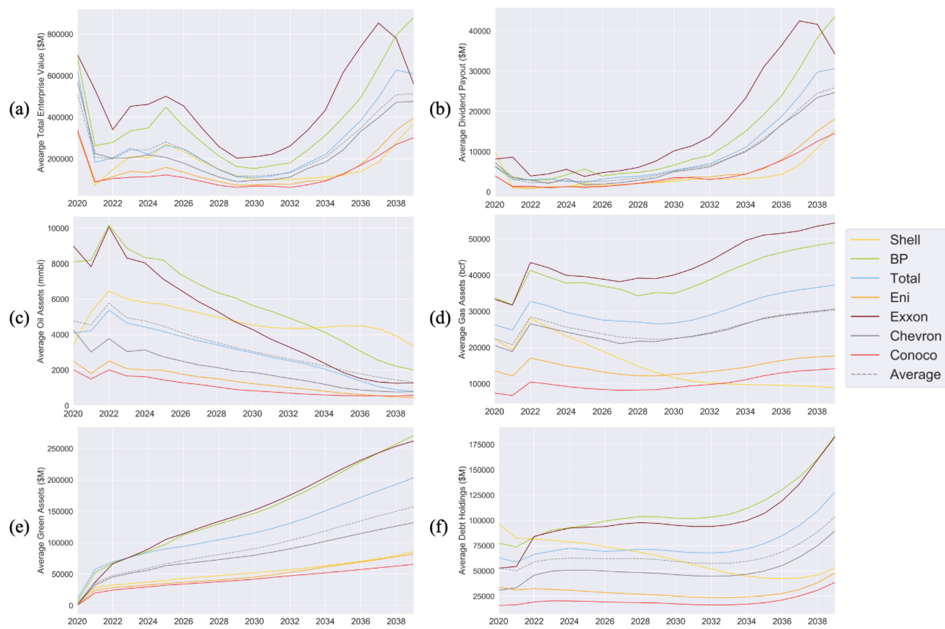


Figure 2: Majors' average yearly (a) market valuation and (b) dividend payouts as well as (c) oil and (d) gas, (e) 'green' and (f) debt asset holdings across all IAMC/IIASA scenarios.

Oil. At the game's start, Majors largely divest from their oil asset portfolios, Figure 2(c), suggesting both the desire to leave the market quickly as well as exploit oil's high, short-term returns before effects of oil demand shocks take place.

Gas. Within gas markets, Figure 2(d), Majors pursue gradual investment strategies highlighting both the carbon asset's influence on market valuation and potential in realizing stable, long-term returns.

'Green'. Across all IAMC/IIASA scenarios, Majors find solace in 'green' markets, Figure 2(e), a result of Majors learning the robust returns inherent to the game's asset. All Majors exhibit first low-carbon mover behavior and continue large-scale, low-carbon investment throughout the game's entirety, supporting the case for full diversification. All Majors become prominent low-carbon leaders.

Debt. Observing debt, Figure 2(f), it is evident that leveraged financial plays prove vital for the Majors to adapt to new business models. Early peaks in debt accumulation coincide with first low-carbon mover behaviors suggesting Majors exploit low-cost cost capital early such to realize stable, long-term 'green' returns and, thus, allow for manageable debt burdens.

Evidently, the results of 2DP-MARL indicate that quickly adapting to low-carbon business models proves a robust strategy amidst energy transition uncertainty. Majors achieve these by exploiting oil's short-term returns and enacting high leverage plays to acquire significant levels of 'green' assets—a strategy deemed as a leveraged transition. While proving ideal in the long-term, such early 'green' diversification strategies risk considerable short-term market valuation decline and dividend cuts. Understanding these upside and downside risks inherent to first low-carbon movement is critical for Majors and their shareholders to enact a smooth transition and mitigate loss in worst-case scenarios.

5 Conclusion

The 2DP-MARL model supports the call for oil Majors to diversify into low-carbon business models by identifying first low-carbon mover potential outcomes and, thus, the robust strategies Majors should take in response to energy transition uncertainty. Though the model serves as a generalization as to what types of capital allocation strategies would prove most beneficial for the company and its shareholders, meaningful insights from the emergent robust strategies can be drawn with respect to such actors. Future developments of 2DP-MARL seek to explore 'green' assets at a more granular level and answer corporate governance questions central to the industry's future in a climate-aligned world, greatly increasing this work's impact. While largely responsible for climate degradation, the Majors have the opportunity to serve as energy transition vanguards. The 2DP-MARL model has the potential to help guide the Majors' decision-makers in realizing carbon-neutral business pathways to accelerate the the global low-carbon energy transition.

Appendix

A 2DP-MARL Adjustments & Additions

While a majority of the game logic seen in the original 2DP wargame has be replicated for this work's model, this appendix details essential adjustments and additions made to enable the implementation of deep MARL methods and best mimic real-life market competition. A high-level overview of the presented 2DP-MARL model is displayed in Figure 3.

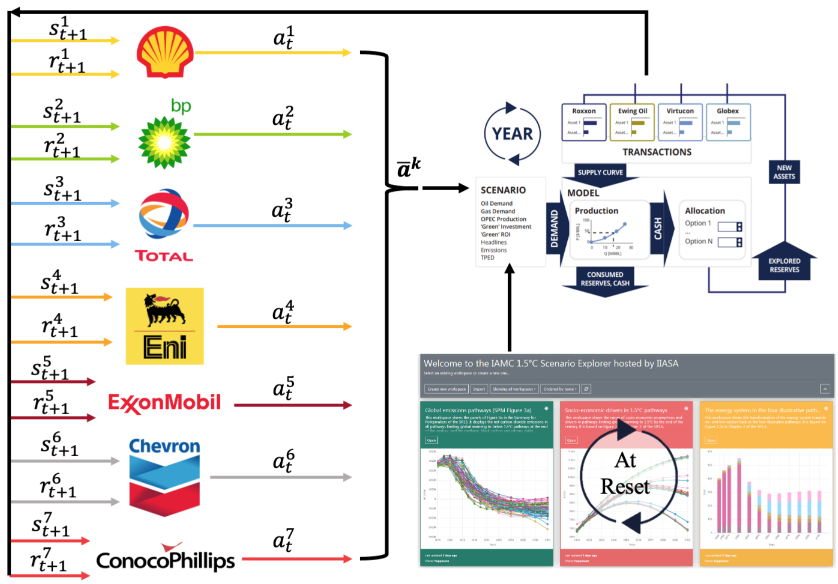


Figure 3: High-level view of the 2DP-MARL model where Majors act as agents within the 2DP wargame environment under a given IAMC/IIASA scenario. Each agent k takes an action within the environment at each timestep t, representing a single year within the game. Subsequently, agents receive a new observation, s_{t+1}^k, and reward, r_{t+1}^k, both of which predicated on the agent's actions as well as those of others, $\bar{a}^k$. Upon reaching the endgame at 2040, the game resets agents to their initial state at 2020 and a new IAMC/IIASA scenario is generated at random.

A.1 Asset Initialization

The original asset categories utilized in the 2DP wargame remain the same for the developed model. However, initial assets are drawn from each Majors' most recent financial statements as well as other publicly available data. Table 2 tabulates each asset along with the respective source and justification.

A.2 Decision-Making Metrics

Most decision-making metrics were drawn directly from the 2DP wargame, however, critical changes were made to the Return on Investment (ROI), now Return on Assets (ROA), and Total Enterprise Value (TEV) equations in order to achieve an accurate representation of real-world values. Table 3 tabulates the decision-making metrics, their equations, and respective source of justification.

Table 2: 2DP-MARL asset initialization.

Asset	Source	Justification
Cash	Cash and Equivalents from Balance Sheet	Represents cash on hand readily available to allocate towards capital expenditures, debt payoff, dividend payoff, etc.
Debt	Total Debt from Balance Sheet	Total debt includes short-term obligations and long-term liabilities
Green	Extrapolated from [6] and 2020 Annual Reports	Total, disclosed low-carbon investments
Undeveloped Low (Oil)	1/6 of Proven Undeveloped Oil	Proven undeveloped oil assets reduced by a factor of 6 in parallel with 2DP's explore ratios
Undeveloped Medium (Oil)	1/3 of Proven Undeveloped Oil	Proven undeveloped oil assets reduced by a factor of 3 in parallel with 2DP's explore ratios
Undeveloped High (Oil)	1/2 of Proven Undeveloped Oil	Proven undeveloped oil assets reduced by a factor of 2 in parallel with 2DP's explore ratios
Developed Low (Oil)	1/6 of Proven Developed Oil	Proven developed oil assets reduced by a factor of 6 in parallel with 2DP's explore ratios
Developed Medium (Oil)	1/3 of Proven Developed Oil	Proven developed oil assets reduced by a factor of 3 in parallel with 2DP's explore ratios
Developed High (Oil)	1/2 of Proven Developed Oil	Proven developed oil assets reduced by a factor of 2 in parallel with 2DP's explore ratios
Undeveloped Gas	Proven Undeveloped Gas	-
Developed Gas	Proven Developed Gas	-

Table 3: 2DP-MARL decision-making metrics.

Decision-Making Metric	Equation	Source
Reserves-to-Production Ratio (R/P Ratio)	$\frac{Reserves_{oil}}{Production_{oil}}$	[2]
Gas-to-Oil Reserves Ratio (R/G Ratio)	$\frac{Reserves_{oil}}{Reserves_{gas}}$	[2]
Total Enterprise Value (TEV)	$\frac{Cash * Dividend\ Policy}{Cost\ of\ Capital} - Debt + Carbon\ Asset\ Costs$	[2], [17]
Return on Assets (ROA)	$\frac{Net\ Income}{Assets\ on\ Hand}$	[13]
Return on Equity (ROE)	$\frac{Net\ Income}{TEV}$	[2]
Debt-to-Equity Ratio (D/E Ratio)	$\frac{Debt}{TEV + Debt}$	[2]
Cost of Capital	$\begin{cases} 20\%, \text{ if D/E Ratio = 2.0} \\ lin \\ 4\%, \text{ if D/E Ratio = 0.0} \end{cases}$	[2]

Table 4: 2DP-MARL global scenario metrics.

Global Scenario Metric	Value Range	Source
Oil Demand (mbbl/d)	Dependent on IAMC/IIASA scenario	[8]
Gas Demand (bcf/yr)	Dependent on IAMC/IIASA scenario	[8]
OPEC & Others' Production Share (%)	88 – 98	[10, 13]
Available 'Green' Assets ($M)	260,000 – 420,000	[2]
'Green' Return on Investment (%)	8 – 10	[2]
Debt Interest (%)	4	[2]

A.3 Global Scenario Metrics

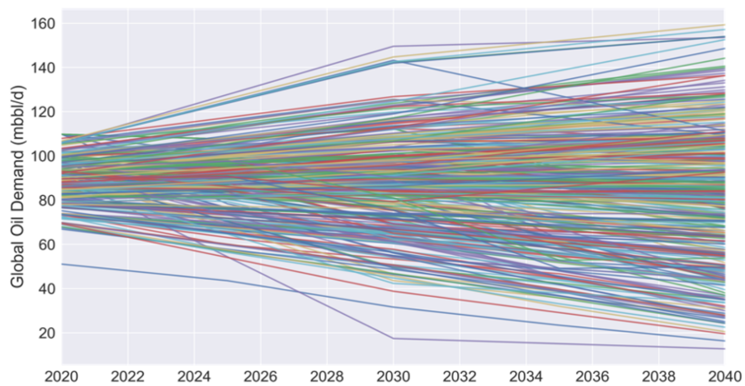
Global scenario metrics determine the dynamics of the 2DP-MARL environment. As stated in Section 3, oil and gas demand values are contingent on the IAMC/IIASA scenario being played. Organization of Petroleum Exporting Countries (OPEC) & others' production share, available 'green' assets for acquisition, and 'green' return on investment value ranges were kept the same across all scenarios. Table 4 tabulates these metrics as well as their respective value range and source.

A.4 Agent Observations

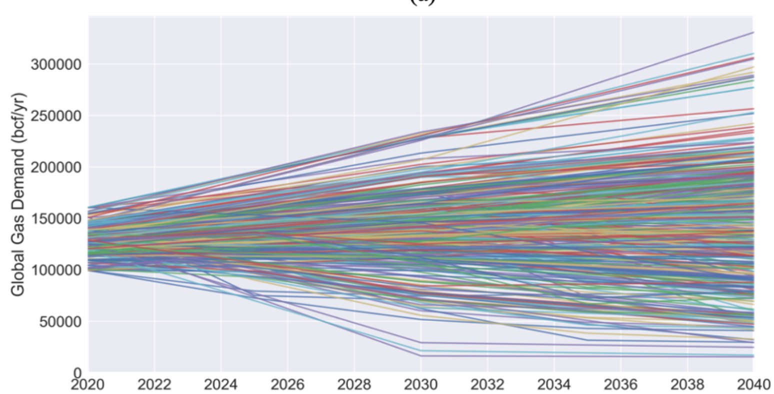
Agent observations describe the elements in which each agent has access to with respect to its environment and opponents. The 2DP-MARL model follows the POSG framework, restricting agent observations to an incomplete view of its surroundings, thus mimicking company observations in real-life market competition. An agent's observations within the 2DP-MARL model include all its own assets and decision-making metrics as well as complete information of the scenario at hand. However, an agent maintains an incomplete observation for each competing agent, only observing their on-hand assets.

B All IAMC/IIASA Scenario Demand Projections

As stated in Section 3, unique to each IAMC/IIASA scenario at play are the projections with respect to oil and gas demand. Figure 4 charts these metrics for each IAMC/IIASA scenario.



(a)



(b)

Figure 4: All IAMC/IIASA scenarios' global (a) oil demand and (b) gas demand projections.

References

[1] Christopher Berner, Greg Brockman, Brooke Chan, Vicki Cheung, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, Rafal Józefowicz, Scott Gray, Catherine Olsson, Jakub Pachocki, Michael Petrov, Henrique Pondé de Oliveira Pinto, Jonathan Raiman, Tim Salimans, Jeremy Schlatter, Jonas Schneider, Szymon Sidor, Ilya Sutskever, Jie Tang, Filip Wolski, and Susan Zhang. Dota 2 with Large Scale Deep Reinforcement Learning. 2019.

[2] Ben Caldecott, Ingrid Holmes, Lucas Kruitwagen, Dileimy Orozco, and Shane Tomlinson. Crude Awakening: Making Oil Major Business Models Climate-Compatible. Technical report, 2018.

[3] Mark Carney. Breaking the tragedy of the horizon – climate change and financial stability. 2015.

[4] CTI. Mind the Gap: The $1.6 trillion energy transition risk. 2018.

[5] CTI. Balancing the Budget: Why deflating the carbon bubble requires oil and gas companies to shrink. Technical report, Carbon Tracker Initiative, 2019.

[6] Luke Fletcher. Beyond the cycle: what's on the horizon for oil and gas majors?, 5 2019.

[7] IEA. The Oil and Gas Industry in Energy Transitions. Technical report, 2020.

[8] IIASA and IAMC. IAMC 1.5°C Scenario Explorer hosted by IIASA, 2019.

[9] David L McCollum, Wenji Zhou, Christoph Bertram, Harmen-Sytze De Boer, Valentina Bosetti, Sebastian Busch, Jacques Després, Laurent Drouet, Johannes Emmerling, Marianne Fay, Oliver Fricko, Shinichiro Fujimori, Matthew Gidden, Mathijs Harmsen, Daniel Huppmann, Volker Krey, Elmar Kriegler, Claire Nicolas, Shonali Pachauri, Simon Parkinson, Miguel Poblete-Cazenave, Peter Rafaj, Narasimha Rao, Julie Rozenberg, Andreas Schmitz, Wolfgang Schoepp, Detlef Van Vuuren, and Keywan Riahi. Energy investment needs for fulfilling the Paris Agreement and achieving the Sustainable Development Goals. *Nature Energy*, 2018.

[10] OPEC. OPEC : Historical Production Data, 2020.

[11] OpenAI. OpenAI Baselines: ACKTR & A2C, 2017.

[12] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms. 7 2017.

[13] S&P CapitalIQ. CapitalIQ, 2020.

[14] Paul Stevens. International Oil Companies The Death of the Old Business Model. Technical report, Chatham House, 2016.

[15] TCFD. Recommendations of the Task Force on Climate-related Financial Disclosures. 2017.

[16] Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, Junhyuk Oh, Dan Horgan, Manuel Kroiss, Ivo Danihelka, Aja Huang, Laurent Sifre, Trevor Cai, John P. Agapiou, Max Jaderberg, Alexander S. Vezhnevets, Rémi Leblond, Tobias Pohlen, Valentin Dalibard, David Budden, Yury Sulsky, James Molloy, Tom L. Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Yuhuai Wu, Roman Ring, Dani Yogatama, Dario Wünsch, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps, and David Silver. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, 11 2019.

[17] Jens Zimmerman. The Oil and Gas Industry Guide. Technical report, CFA, 2016.