Sailing the Wind: Evaluating the Impact of COMA on Multi-Agent Active Wake Control in Wind Farms

Mihai Filimon (M.R.Filimon-1@student.tudelft.nl) | Supervised by Grigory Neustroev and Mathijs de Weerdt

TU Delft Delft University of Technology

1. Introduction

The total power output of a wind farm can be decreased because of wake effects - areas of high turbulence and low wind speed caused by the extraction of wind by the first turbine [1].

This issue can be tackled through Active Wake Control (AWC), the process of turning the turbine to **redirect the wake away from the downstream turbines**. One way to optimise this is through Multi-Agent Reinforcement Learning (MARL). Unlike single-agent approaches that are time-consuming and rely on individual observations, MARL enables the agents to learn from each other's experiences to accelerate the process and improve efficiency.

2. Research Question

What is the effect of COMA on the problem of AWC compared to single-agent RL algorithms?

- What is the difference in performance between COMA and TD3?
- What are the limitations of COMA?

3. COMA

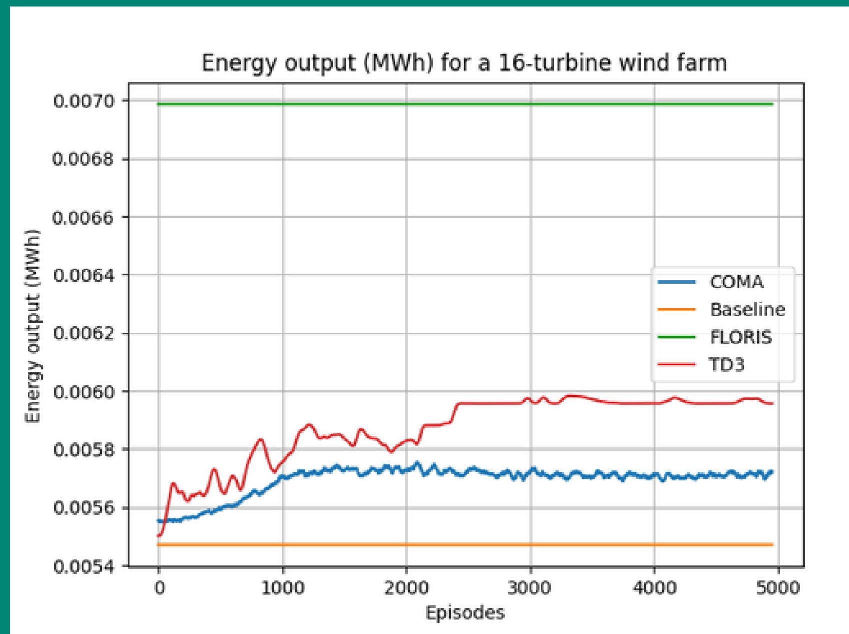
Counterfactual Multi-Agent Policy Gradients (COMA) uses a **centralised critic** to estimate the Q-function and decentralised actors to optimise the agents' policies [2]. The key aspect is that it uses a **counterfactual baseline** - for each agent the global reward is compared to a scenario where that agent's action changes and all the rest stay the same [2].

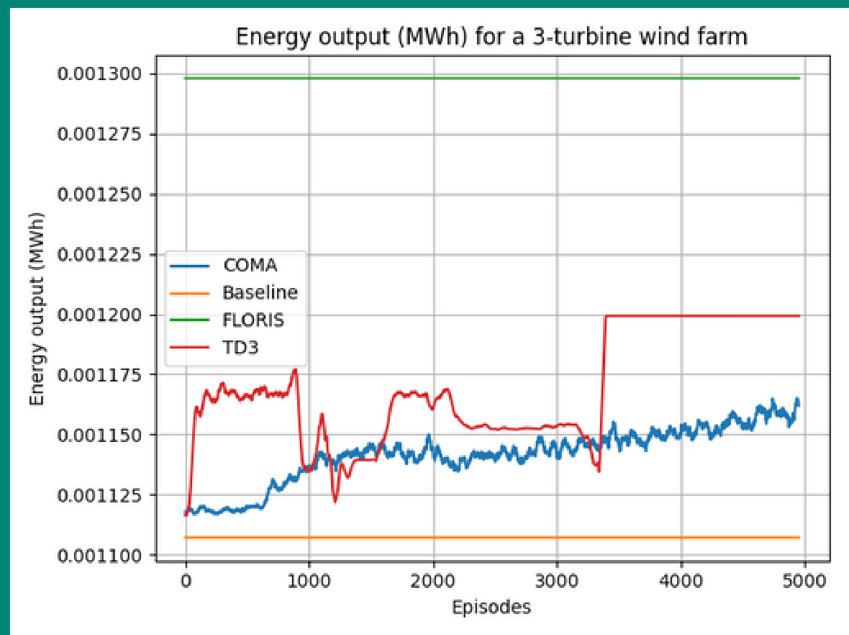
The centralised critic encourages communication and coordination and allows the agents to understand the global consequences of their actions.
As a cooperative MARL algorithm, it is fit for a problem with interdependent objectives like Active Wake Control.

4. Methodology

COMA was applied on a "wind tunnel" of 3 turbines and a 4-by-4 grid of 16 turbines through the wind farm environment created by Grigory Neustroev. It was trained using 5000 episodes with a maximum of 100 steps each and a discount factor of 0.99.

5. Results





6. Analysis

- Despite COMA's centralised-critic approach combined with counterfactual comparisons, it is not able to learn a more efficient policy on any setting.
- **Initial hypothesis** that it performs better than TD3 for 3 turbines **was rejected**.
- Besides the unsatisfactory learning curve, concerns that it **cannot scale well with size** were confirmed.

- COMA was not tested on an already existing wind farm.
- It is fully verifiable and reproducible as all the code parts are publicy available.
- Experiments indicated limitations (~8 hours of computation for 3 turbines & ~32 hours for 16 turbines)

7. Conclusion

- COMA **did not provide an improvement** compared to single-agent Reinforcement Learning.
- It faced significant computation constraints that make it unfit for large wind farms.

- While the experiments provided relevant information, it is **not possible to draw a complete conclusion** yet regarding COMA's potential for Active Wake Control.

8. Future Work

- COMA **must be applied on a real-world wind farm** like "Princess Amalia" or "Gemini".
- Test multiple wind farm layouts and turbine placements.
- Test environments where the wind is coming from different directions .
- Transition to a continuous action space.

REFERENCES

[1] G NEUSTROEV, M. M. DE WEERDT ET AL. **"DEEP REINFORCEMENT LEARNING FOR ACTIVE WAKE CONTROL"**. *INTERNATIONAL CONFERENCE ON AUTONOMOUS AGENTS AND MULTIAGENT SYSTEMS*, AA-MAS 2022, 322(10):944–953, 2022
[2] J FOERSTER, G. FARQUHAR ET AL. **"COUNTERFACTUAL MULTI-AGENT POLICY GRADIENTS"**. *AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE*, AAAI-18, VOL. 32. NO. 1. 2018.