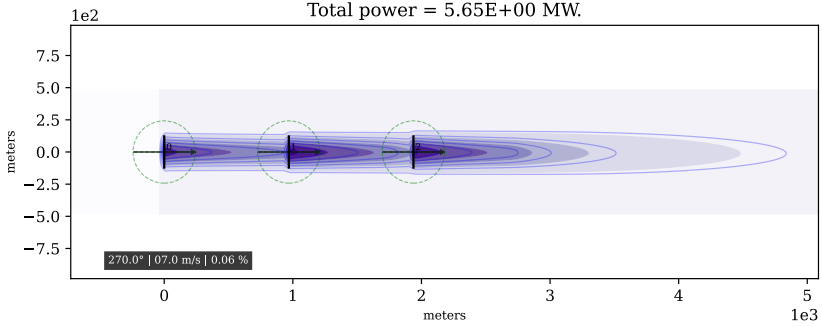
Helping mitigate climate change through efficient reinforcement learning-based wind farm flow control

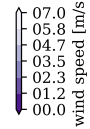
Elie Kadoche[1,2] **Pascal Bianchi**[1] **Florence Carton**[2] **Philippe Ciblat**[1] **Damien Ernst**[1,3]
[1]Polytechnic Institute of Paris, 19 Place Marguerite Perey, 91120 Palaiseau, France
[2]TotalEnergies OneTech, 2 Place Jean Millier, 92400 Courbevoie, France
[3]Montefiore Institute, University of Liège, 4000 Liège, Belgium
elie.kadoche@totalenergies.com

Wind farm flow control

NEURAL INFORMATION PROCESSING SYSTEMS

WIND TRACKING

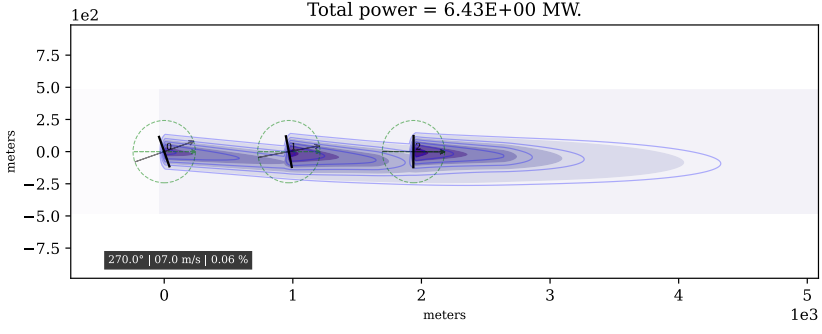


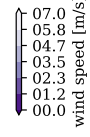


Local velocities	Power outputs
$v^0 = 6.97$ m/s	$P^0 = 4.19$ MW
$v^1 = 4.12$ m/s	$P^1 = 0.70$ MW
$v^2 = 4.26$ m/s	$P^2 = 0.76$ MW

Total power = 5.65 MW.

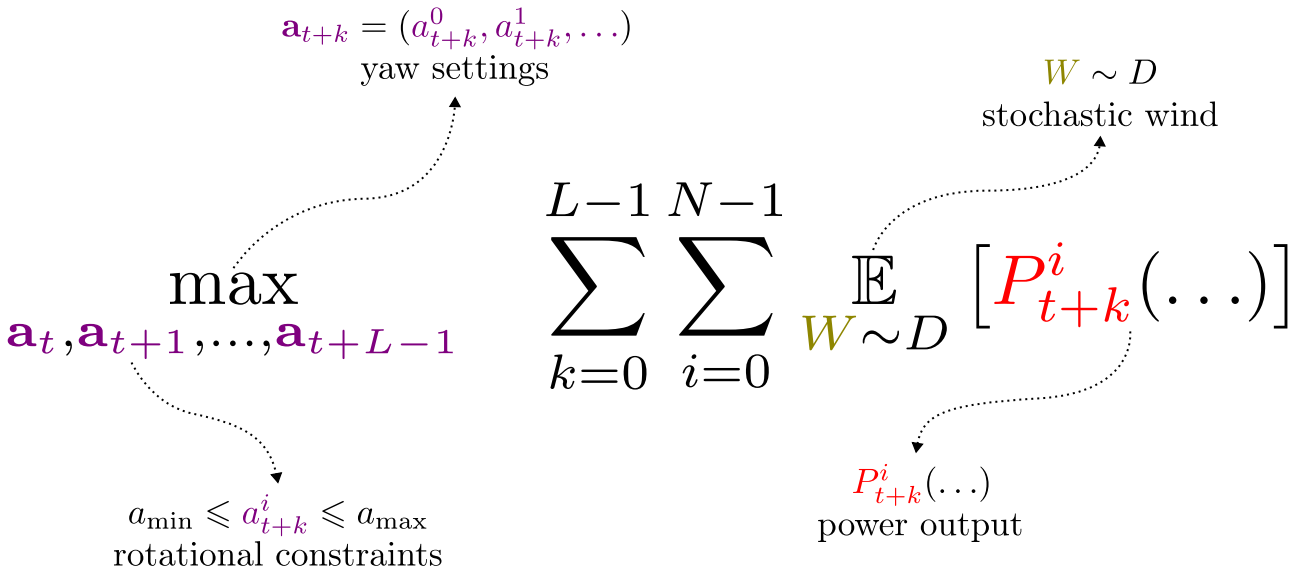
WAKE STEERING





Local velocities	Power outputs
$v^0 = 6.97$ m/s	$P^0 = 3.71$ MW
$v^1 = 4.92$ m/s	$P^1 = 1.27$ MW
$v^2 = 5.07$ m/s	$P^2 = 1.47$ MW

Total power = 6.45 MW.


$\mathbf{a}_{t+k} = (a^0_{t+k}, a^1_{t+k}, \ldots)$
yaw settings
$W \sim D$
stochastic wind
$\max_{\mathbf{a}_t, \mathbf{a}_{t+1}, \ldots, \mathbf{a}_{t+L-1}}$
$\sum_{k=0}^{L-1} \sum_{i=0}^{N-1} \mathbb{E}_{W \sim D} \left[P^i_{t+k}(\ldots) \right]$
$a_{\min} \leqslant a^i_{t+k} \leqslant a_{\max}$
rotational constraints
$P^i_{t+k}(\ldots)$
power output

Existing methods

$$\max_{a_t, a_{t+1}, \dots, a_{t+L-1}} \sum_{k=0}^{L-1} \sum_{i=0}^{N-1} \mathbb{E}_{W \sim D} \left[P^i_{t+k}(\dots) \right]$$

Myopic control

$$\max_{a_t} \sum_{i=0}^{N-1} \mathbb{E} \left[P^i_t \right]$$

Simplistic approach.

Predictive control

$$\max_{a_t, a_{t+1}, \dots, a_{t+L-1}} \sum_{k=0}^{L-1} \sum_{i=0}^{N-1} \mathbb{E} \left[P^i_{t+k} \right]$$

Model dependent.

Reinforcement learning

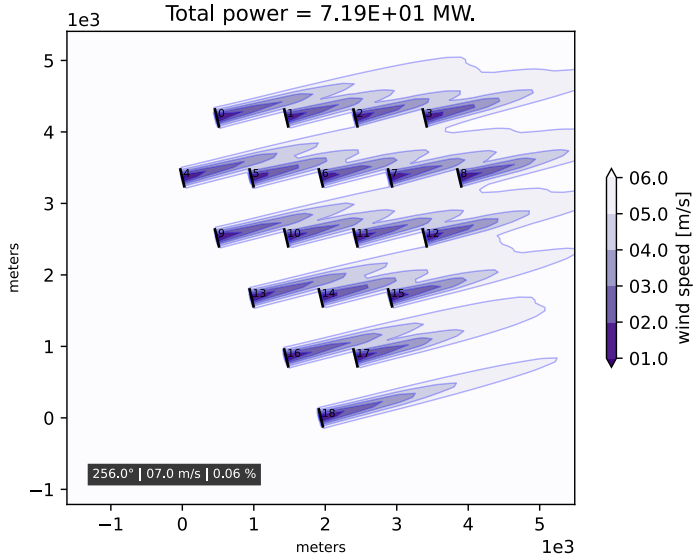
$$\max_{\theta} \mathbb{E}_{\pi_\theta} \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \right]$$

$$\pi_\theta(s_t) = (\bar{\mu}_t, \bar{\kappa}_t, v_t)$$

$$a^i_t \sim \text{von Mises}(\mu^i_t, \kappa^i_t)$$

Data-driven, model-free.

Reinforcement learning limitations



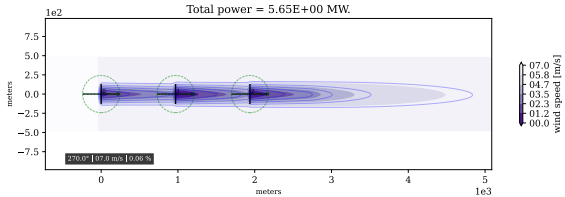
Generalization
Current RL methods only focus on limited sets of wind conditions.

Sample efficiency
Training a model requires a large number of simulations.

Generalization issue

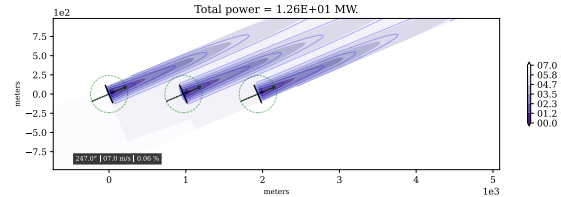
Bad reward shaping

$$r_{t+1} = \frac{1}{N} \sum_{i=0}^{N-1} \frac{P_t^i}{\Phi(V_t)}$$



Baseline wake losses $\bar{\mathcal{L}}_t = 55.03$ %.
Baseline power output $\bar{P}_t = 5.65$ MW.
Theoretical maximum $\Phi(V_t) = 12.56$ MW.

Optimal reward: close to 0.5.
Wake steering is optimal.



Baseline wake losses $\bar{\mathcal{L}}_t = 00.03$ %.
Baseline power output $\bar{P}_t = 12.55$ MW.
Theoretical maximum $\Phi(V_t) = 12.56$ MW.

Optimal reward: close to 1.
Wind tracking is optimal.

The policy learns a wind tracking solution, and it is enough to get high rewards.

Generalization improvement

Better reward shaping

$$\Delta_{P_t} = (P_t - \bar{P}_t)/\bar{P}_t$$

$$r_{t+1} = \Delta_{P_t}\mathbf{1}_{\Delta_{P_t}<0} + \exp(-p\bar{\mathcal{L}}_t)\Delta_{P_t}\mathbf{1}_{\Delta_{P_t}\geq 0}$$

penalize policies below baseline

restrict Δ_{P_t} when it is supposed to be high (strong wake conditions)

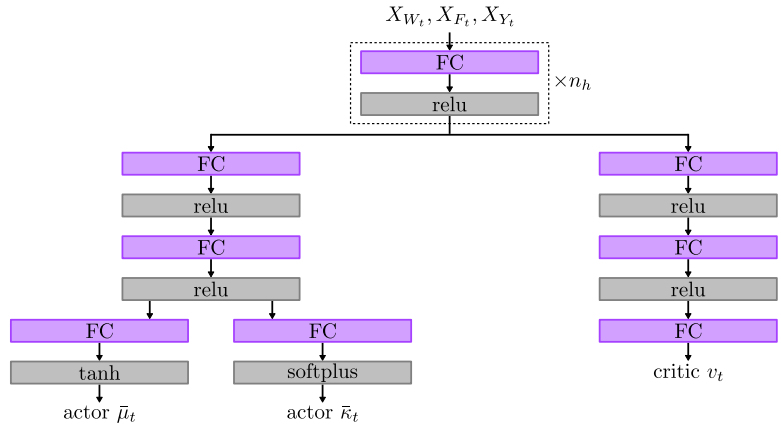
encourage policies above baseline

Sample efficiency issue

Fully-connected neural network architecture

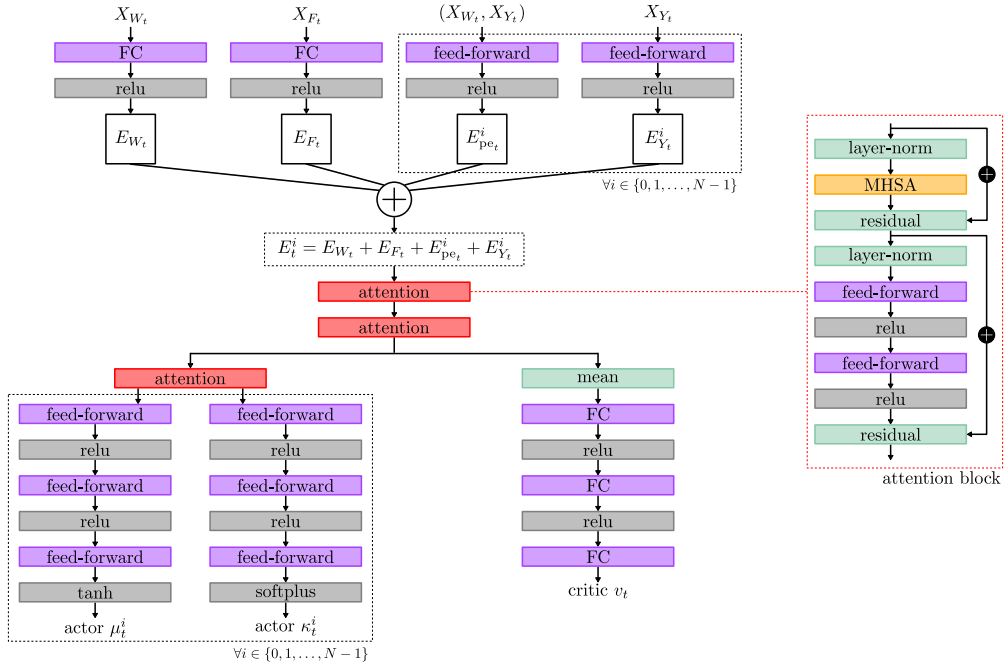
$$s_t = (X_{W_t}, X_{F_t}, X_{Y_t})$$

- X_{W_t} current wind data (noisy).
- X_{F_t} wind forecast (noisy).
- X_{Y_t} yaw angles.



Sample efficiency improvement

Attention-based neural network architecture



Training

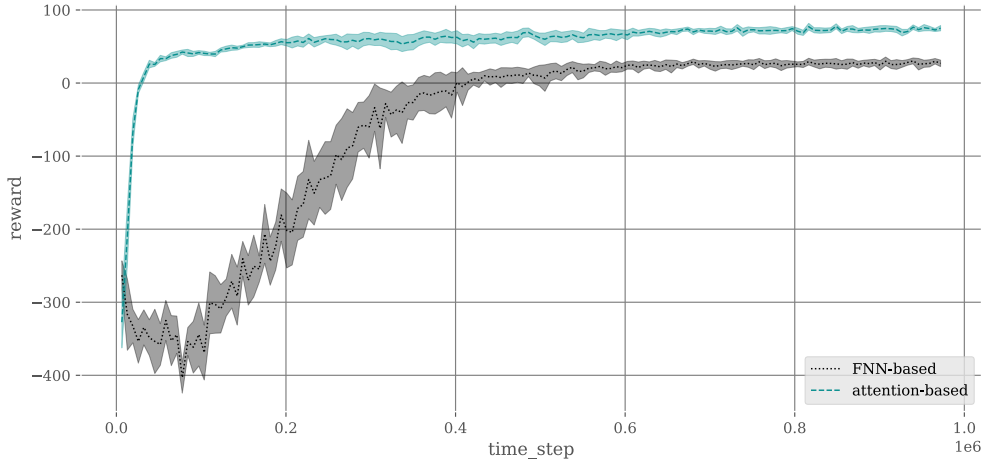


Figure: training curves of each model for 10 different seeds.

* The attention-based model requires ≈ 10 times fewer simulation steps than the FNN-based model and it reaches superior performance.

Testing

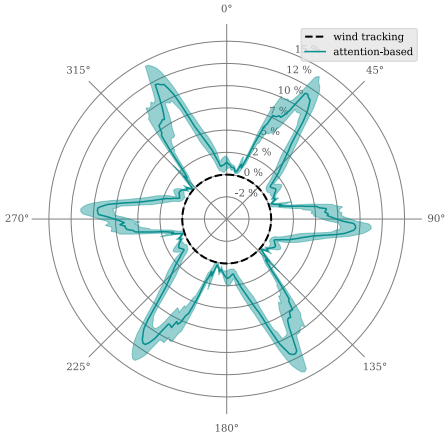


Figure: attention-based model.

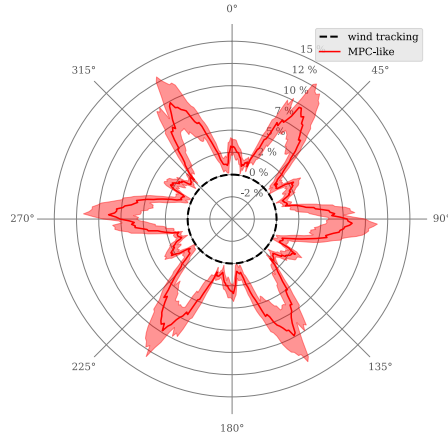


Figure: MPC-like solution.

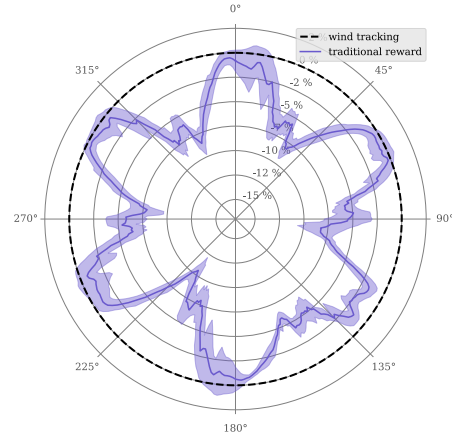


Figure: traditional reward.

* The attention-based model achieves performance comparable to the MPC-like solution, but with lower variance and higher gains.

* When using the traditional reward, the attention-based model completely fails to learn a wake steering strategy.


Thank you
Salomé