Designing Incentive-Compatible Cooperation for the Global Commons:
A Reinforcement Learning-Integrated Mechanism Design Framework for Arctic Conservation

I. MOTIVATION

Resolving the commons dilemma in shared global resources: the Arctic as a strategic case study

Key Issue:
1. Arctic ice melt rate: 12.5% per decade.
2. Arctic ice shrinkage: 77,000 km^2/year.
3. Expected first ice free Arctic summer: September 2030-2050 (CMIP6).

Impact:
1. Accounts for 29% of the baseline decadal global temperature rise of 0.36°F.
2. Sea level rising at 4.5 mm/year.
3. Economic impact estimated at $98.2-$221.8 billion.

TRAGEDY OF THE COMMONS[1]

Arctic conservation requires sustained international participation, yet individual nations face short-term incentives to defect and free-ride on collective efforts. When defection reaches a critical mass, stability of conservation is threatened.

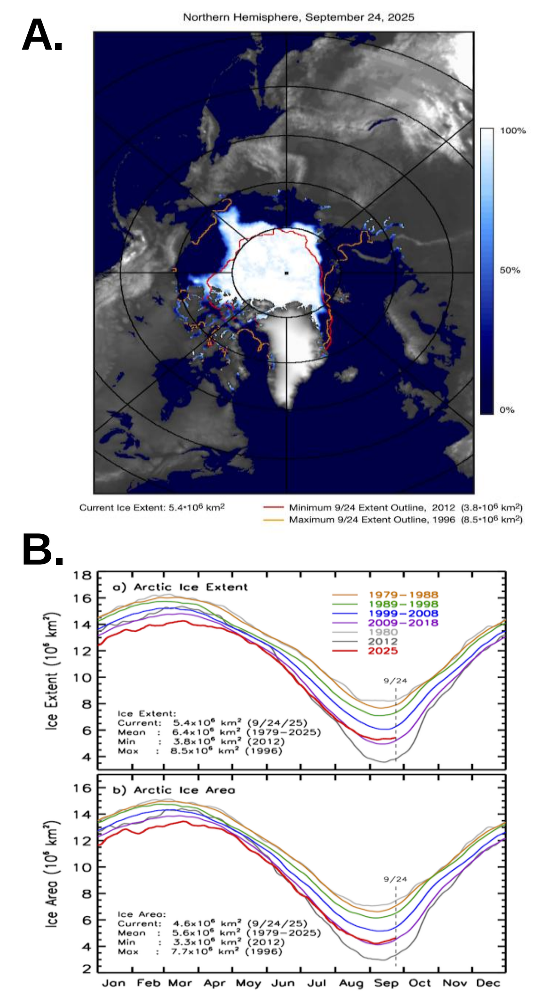

Figure 1. Ice loss (A) & evolution (B)
Source: Comiso, J. C. (2023), "Bootstrap Sea Ice Concentration," NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 09/2025
https://earth.gsfc.nasa.gov/cryo/data/current-state-sea-ice-cover

II. RESEARCH GOAL

Reframe the Arctic as a **coupled socio-ecological system** of interacting geophysical regions and national actors, and design mechanisms that render conservation **incentive-compatible**.

Current Research Gaps	Solution Proposed
1. Gen 1 Physics-based Ising models: focus on geophysical evolution.[2]	1. **Graph theory**: dual graph links Arctic geophysics and human actions with an iterative feedback loop.
2. Gen 2 Deep-learning based IceNet model: short-term melt prediction (3-6 months).[3]	2. **Game theory**: forward-looking nations optimize strategies based on discounted future payoffs.
3. Human behavior is exogenous in both models, limiting applicability for policy design.	3. **Reinforcement learning**: endogenous Q-updating as an alignment mechanism.

III. MODEL & IMPLEMENTATION

Linked dual graph with game theory and reinforcement learning

I. Dual Graphical Setup

A. Physical Graph

a. $G_r = (N_r, E_r)$: geophysical nodes
b. $\{N_r\}_{r=1}^{K}$: discrete Arctic regions
c. $E_r \subseteq N_r x N_r$: edges
d. W_r: $E_r \rightarrow [0,1]$: edge weights (temp gradient, ocean current)
e. $\{I(r,t), \Phi(r,t)\}$: physical state

B. Human Graph

a. $G_p = (N_p, E_p)$: stakeholder nodes
b. $\{N_p\}_{p=1}^{M}$: nations/players
c. $E_p \subseteq N_p x N_p$: edges
d. W_p: $E_p \rightarrow [0,1]$: edge weights (economic & political connections)
e. $\{S(p,r,t), \Pi(p,s,t)\}$: human action/payoff

II. Game Theory Integration

C. Strategy Space

a. $s \in S(p,r,t) = 0$: defection
b. $s \in S(p,r,t) = 1$: conservation

D. Payoff Matrix

a. $\Pi(p,s,t) = \{\pi^c(p,t), \pi^d(p,t)\}$
b. $\pi^c(p,t) = \frac{1}{|N_r|}\sum_{r\in N_r}\kappa_{pr}\left(-P(r,t) + \Theta(I(r,t))\right)$: conservation payoff
c. $\pi^d(p,t) = \frac{1}{|N_r|}\sum_{r\in N_r}\kappa_{pr}\left(R(r,t) - \mathfrak{h}(I(r,t))\right)$: defection payoff

E. Impact on Stressors

$H(r,t) = \sum_{p\in N_p}\kappa_{pr}\mathfrak{h}(I(r,t))\left(1 - S(p,r,t)\right)$: anthropogenic stress

III. Strategy Updating with RL

ε-greedy algorithm

$$s = \begin{cases} random\,(0,1), & with\,probability\,\varepsilon \\ argmax_s Q_p(s_p,r,t), & with\,probability\,1-\varepsilon \end{cases}$$

Q-updating

$$Q_p(s_p,r,t) \leftarrow (1-\gamma)Q_p\,(s_p,r,t) + \gamma\left(\pi_p(s_p,s_{-p},r,t) + (1-\delta)\,max_{s'_p}Q_p(s'_p,r,t+1)\right)$$

```
Input: Number of episodes, states, actions, reward function π_p, transition
function, learning rate γ, discount factor δ, exploration rate ε
For each episode in range (number_of_episodes):
        Initialize state = Random choice from states
        While state != terminal
                If Random number in [0, 1) < ε:
                        action = Random choice from actions
                Else:
                        action = Action that maximizes the Q-
                                 value based on the current
                                 state
                Reward = π(state, action)
                Next state = Transition to new state based on
                        state & action
                Q(state, action): Based on Q-updating
                Current state = Next state
        End While
 End For
 Policy = {state: action that maximizes Q(state, action) for all actions}
Output: Optimal Policy
```

Figure 2. RL implementation pseudocode
Source: Pseudocode based on code written by finalist in Python, 02/2025

IV. RESULTS

Result 1. Predictive Power

Graph-game model **improves long-term predictive power**

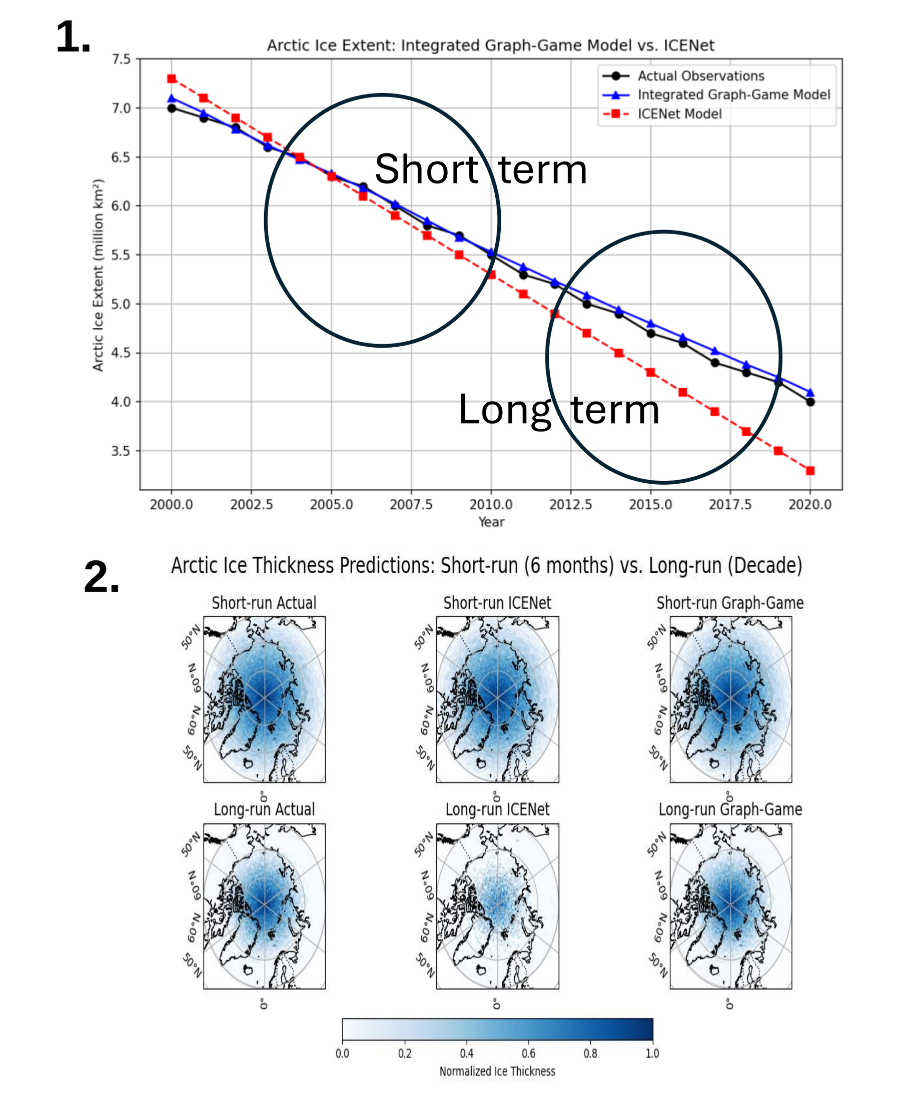

Figure 3. Ice extent (1) & thickness (2): data & model simulations
Source: Figure created by finalist using Matplotlib and Cartopy, 03/2025.

1. **Short-term predictive power:** forward looking model **preserves** short-term predictive power (comparable to IceNet).
2. **Long-term predictive power:** forward looking model **improves** long-term predictive accuracy by 11.25%-23.10%.
3. **Regional heterogeneity:** explicitly captured in the stressors, which allows strategy variations through differential payoffs.

Result 2. Conservation Potential

Forward looking strategizing **reduces the decadal ice melt rate**, stabilizing the Arctic.

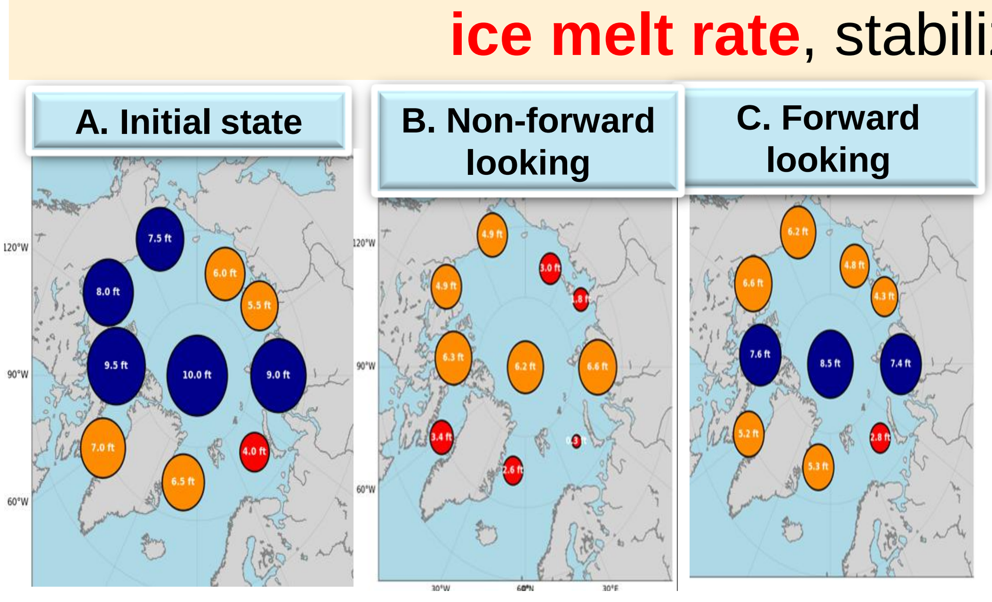

Figure 4. Arctic conservation: data & model simulations
Source: Figure created by finalist using model simulations, Matplotlib, and Python Basemap toolkit, 03/2025.

	Current Data	Graph-Game
Mean	12.51%	4.07%
Std. Dev	3.95%	0.90%
Max	18.50%	8.03%
Min	5.32%	2.30%

Table 1. Arctic conservation: data & model simulations
Source: Table generated by finalist using model simulations in Python, 03/2025.

Result 3. Policy Implications

Incentive compatible strategizing promotes voluntary conservation, **reducing reliance on Pigouvian taxation**.

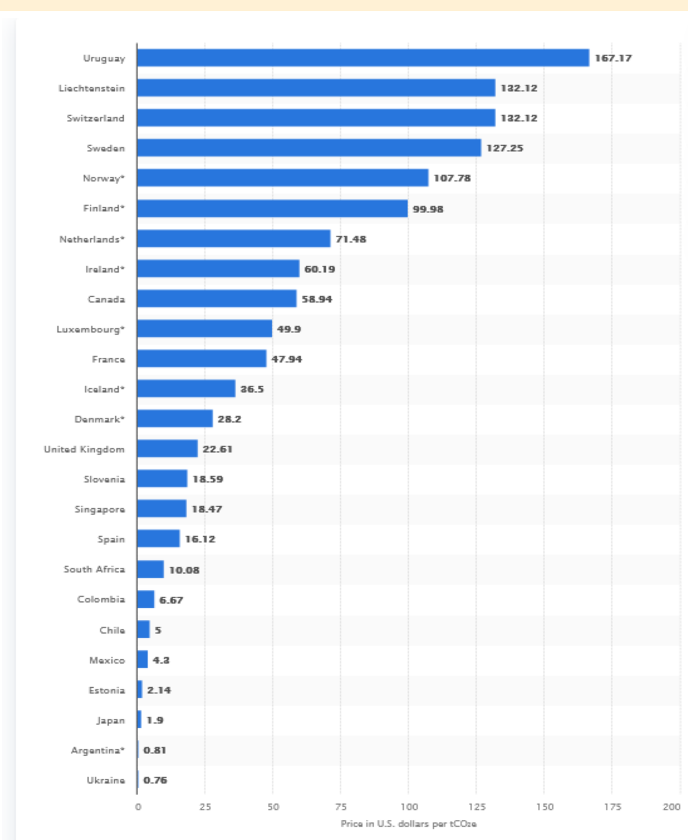

Figure 5. Carbon taxes across nations
Source: "Carbon taxes worldwide by select country 2025," Statista, accessed: 04/2025
https://www.statista.com/statistics/483590/prices-of-implemented-carbon-pricing-instruments-worldwide-by-select-country/

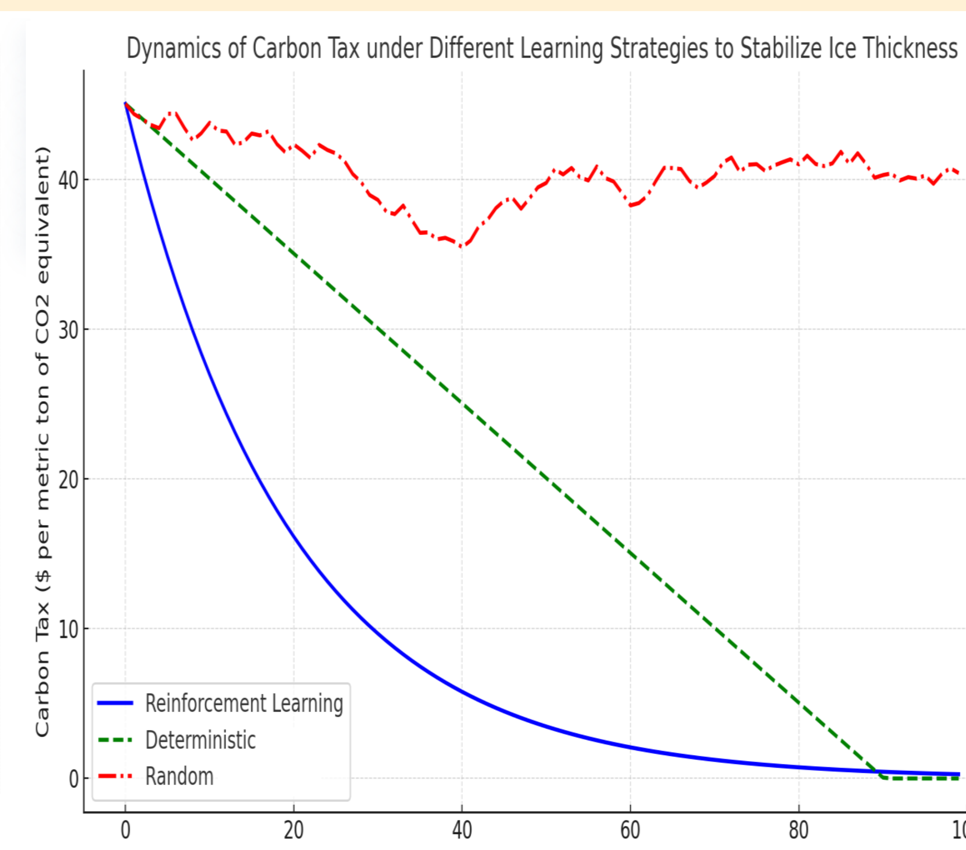

Figure 6. Counterfactual experiment: evolution of carbon taxes under random, deterministic, and RL-informed strategizing with a goal to stabilize Arctic ice thickness.
Source: Figure created by finalist using model simulations in Python, 05/2025

1. **Policy conundrum**: low compliance with carbon quotas or taxation due to short term incentives (IPCC).
2. **Potential solution**: incentive-compatible mechanisms where voluntary conservation is an optimal Nash outcome.

V. DISCUSSION & CONCLUSION

Contribution:
- Captures interplay of ice evolution and national strategies to design incentive-compatible conservation mechanisms.
- Flexible framework, adaptable to IceNet and extendable to other global commons (atmosphere, high seas, etc.).

Limitations:
- Computational complexity.
- Restrictive assumptions.

VI. REFERENCES

[1] Hardin, G. The Tragedy of the Commons. *Science*, New Series 162:3859 (1968)
[2] Ma, Y.P. et al. Ising model for melt ponds on Arctic sea ice. *New Journal of Physics*, 21:063029 (2019)
[3] Andersson, T.R. et al. Seasonal Arctic sea ice forecasting with probabilistic deep learning. *Nature Communications*, 12:5124 (2021)

Chopra.R