Funded by the European Union

AVIATION TWIN TRANSITION CLUSTER

A EUROPEAN INITIATIVE FOR A SUSTAINABLE FUTURE

AVIATION
TWIN TRANSITION
CLUSTER

Funded by
the European Union

RefMap Clustering Event 2025

Advancing Sustainable Aviation & Urban Air Mobility

Aircraft Trajectory Planning for Climate Impact Mitigation Considering Air Traffic Complexity: A Constrained Multi-Agent Reinforcement Learning Approach

Fateme Baneshi, Universidad Carlos III de Madrid

29–30 July 2025, Athens, Greece

Funded by
the European Union

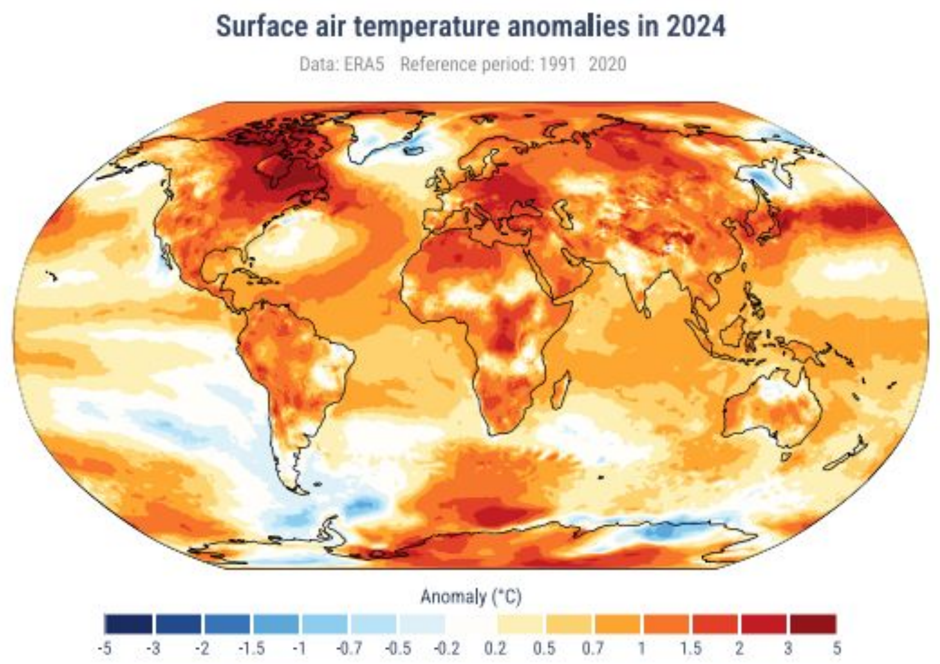
Motivation

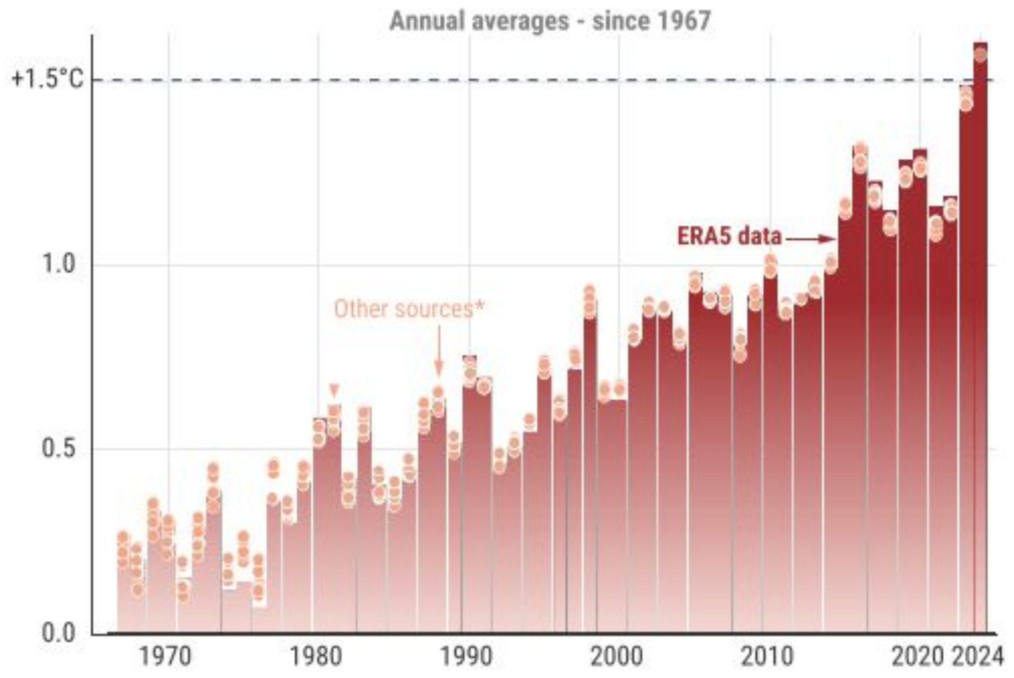
In 2024, the annual average temperature exceeded 1.5°C above pre-industrial levels for the first time.

Aviation is one of the contributors to global warming:

- Carbon dioxide (CO_2)
- Non-CO_2 effects, e.g., Contrails

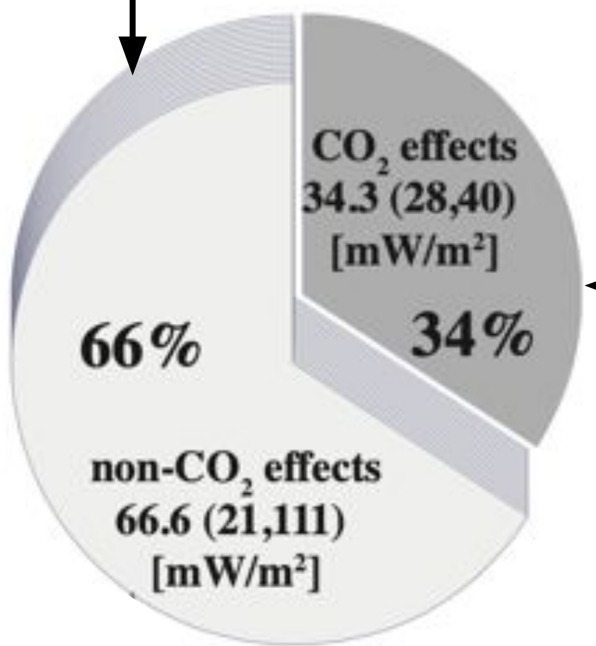
Due to the rapid growth of the aviation sector, it faced increasing pressure to reduce its climate footprint.





Global surface temperature anomalies and trends.

- Operational measures (e.g., Flight Planning)



- Mitigation measures:
  - Alternative fuels
  - Technological advancements
- Can be beneficial for non-CO_2 effects
- **Medium-to long-term measures**

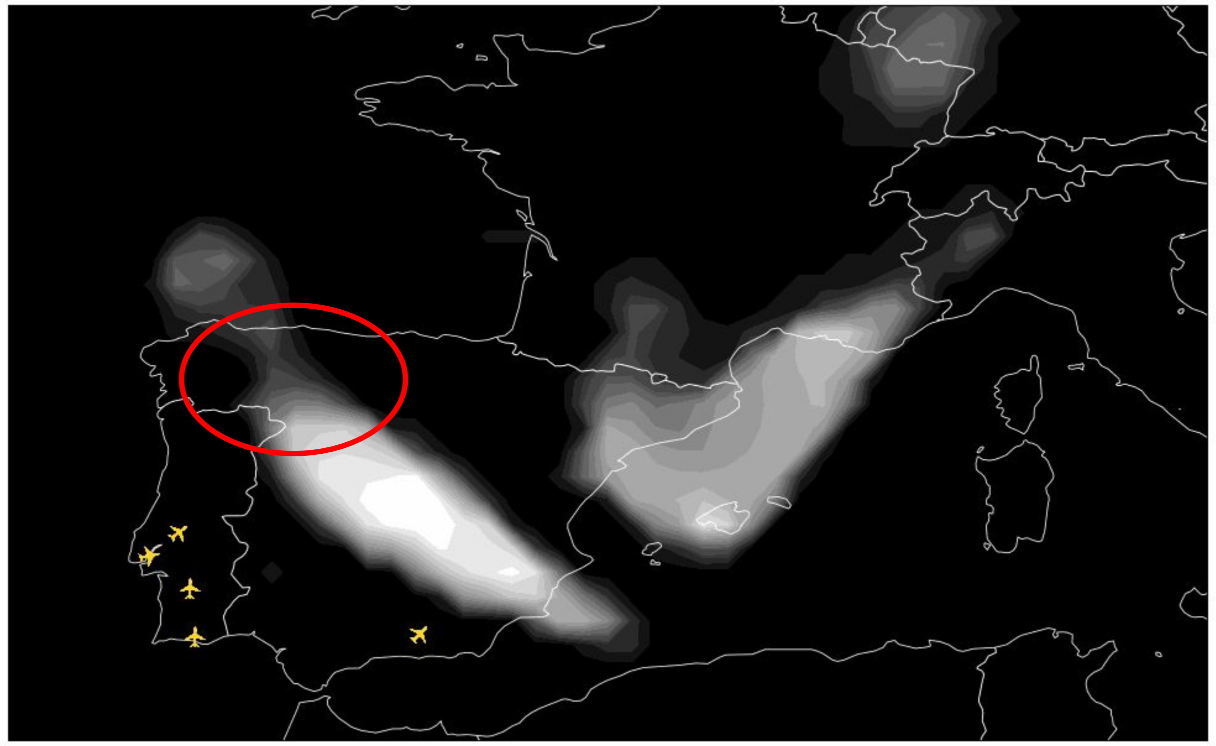
- Feasible with existing infrastructures
- **Relatively immediate solutions**

Lee, D. S., Fahey, D. W., Skowron, A., Allen, M. R., Burkhardt, U., Chen, Q., ... & Wilcox, L. J. (2021). The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmospheric environment*, *244*, 117834.

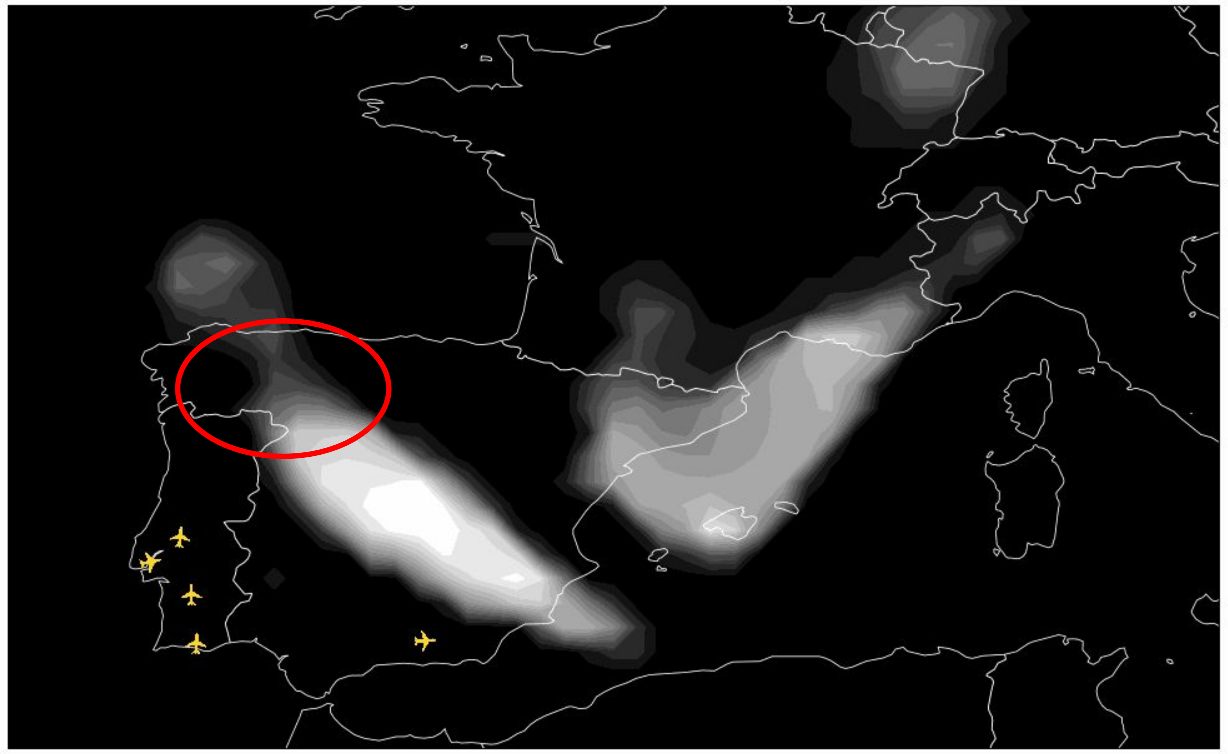
Research question

Is optimizing individual trajectories to reduce climate impact operationally feasible?

Business as usual trajectories



Climate-optimized trajectories



Adopting individually optimized trajectories:

- **Redistributes** traffic flow
- **Increases** the **congestion** in specific areas
- It can **adversely affect** the overall performance of the ATM system

Climate-optimal flight planning needs to be studied at the network-scale

Funded by
the European Union

State-of-the-art

Most recent studies on climate-optimized flight planning:

Study	*Forcing agents*	*Model*	*Routing*	*Uncertainty*	*Opt. scale*
Yamashita et al. (2020) [50]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale
Lührs et al. (2021) [58]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale
Yamashita et al. (2021) [51]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale
Yin et al. (2022) [45]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale
Castino et al. (2024) [62]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale
Sausen et al. (2023) [59]	Contrails	ISSR	Structured	–	Micro-scale
Frias et al. (2024) [20]	Contrails	CoCiP	Structured	–	Micro-scale
Simorgh et al. (2022) [22]	CO_2 and non-CO_2	aCCFs	Structured	MET	Micro-scale
Simorgh et al. (2024) [61]	CO_2 and non-CO_2	aCCFs	FFRA	MET	Micro-scale
Simorgh et al. (2024) [19]	CO_2 and non-CO_2	aCCFs	Structured	MET	Micro-scale

Existing studies have focused on micro-scale flight planning

The reported climate benefits in the literature might not be achievable in practice -> not reliable indicators to incentivize stakeholders

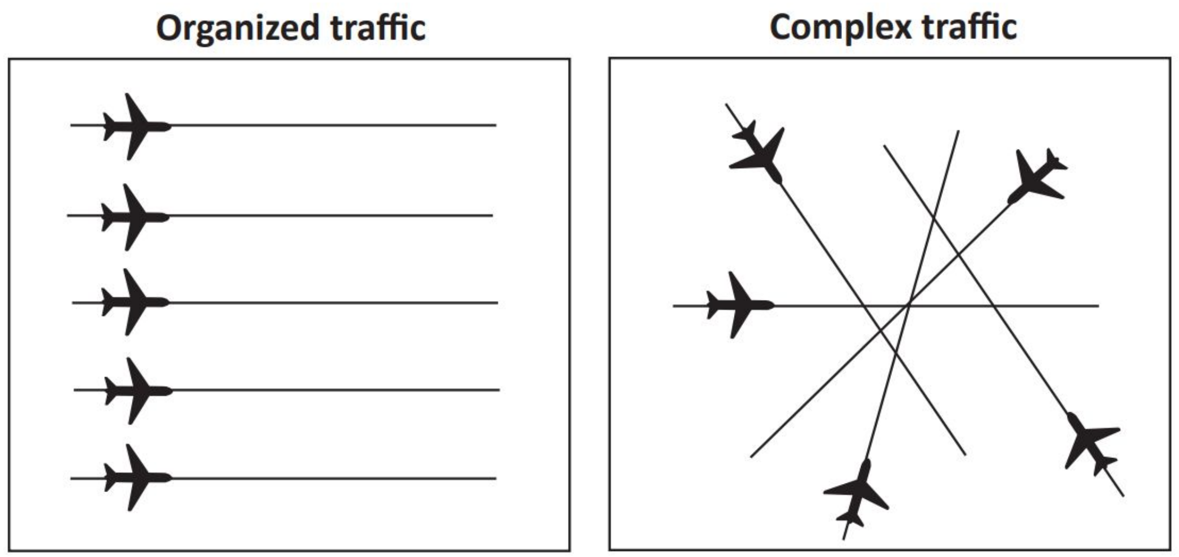
Research Gap: Climate-optimized flight planning at the network level to account for operational manageability of optimized flight plans.

Objectives

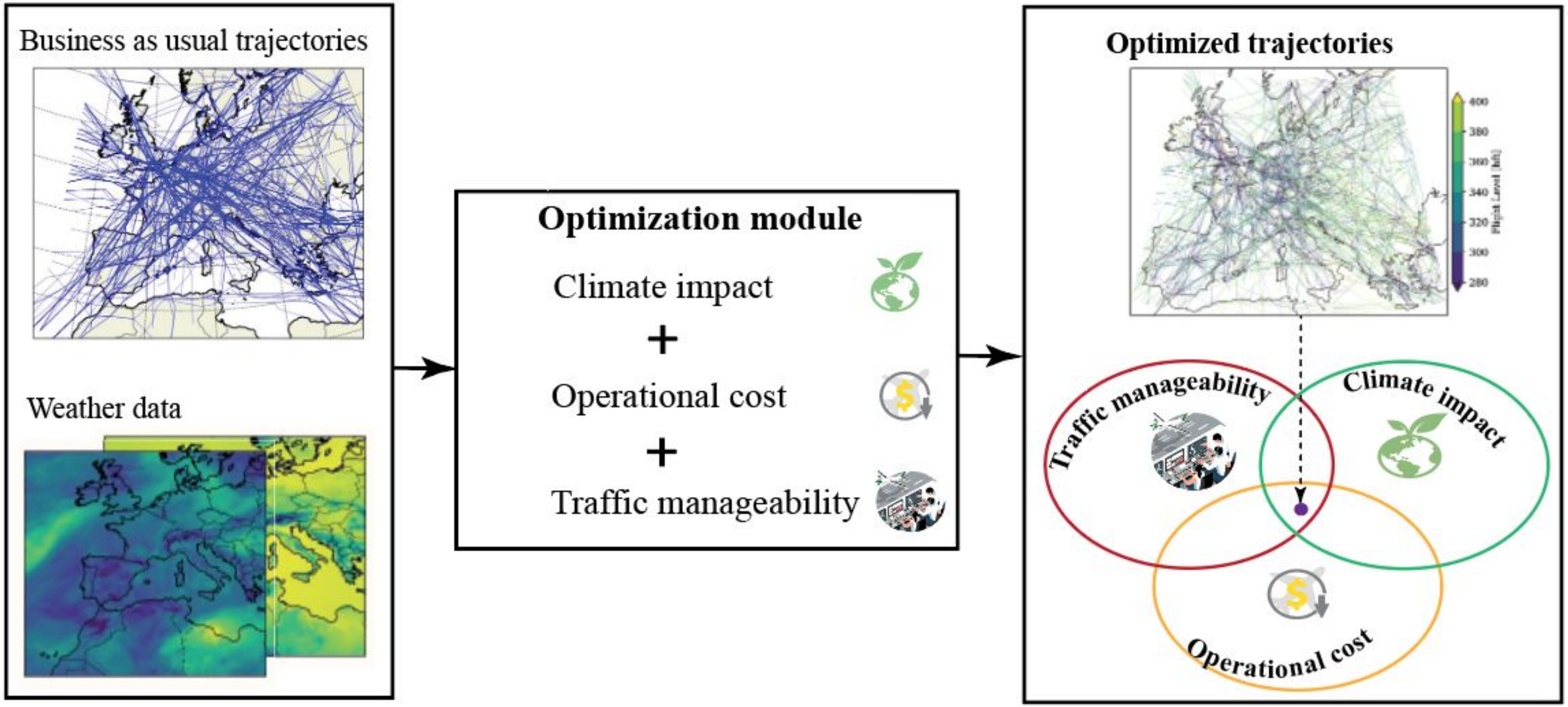
The goal is to plan aircraft trajectories at the network scale to:

- **Mitigate the climate impact**: The aim is to reduce their negative climate footprint.
- **Maintain manageable traffic**: Ensure that the complexity of the new trajectories remains at a level that can be easily managed and implemented.
- **Ensure feasible operational cost**: This serves as a measure of cost-effectiveness.

- **Complexity:** The level of difficulty in managing air traffic safely and efficiently



Framework for climate-optimal flight planning



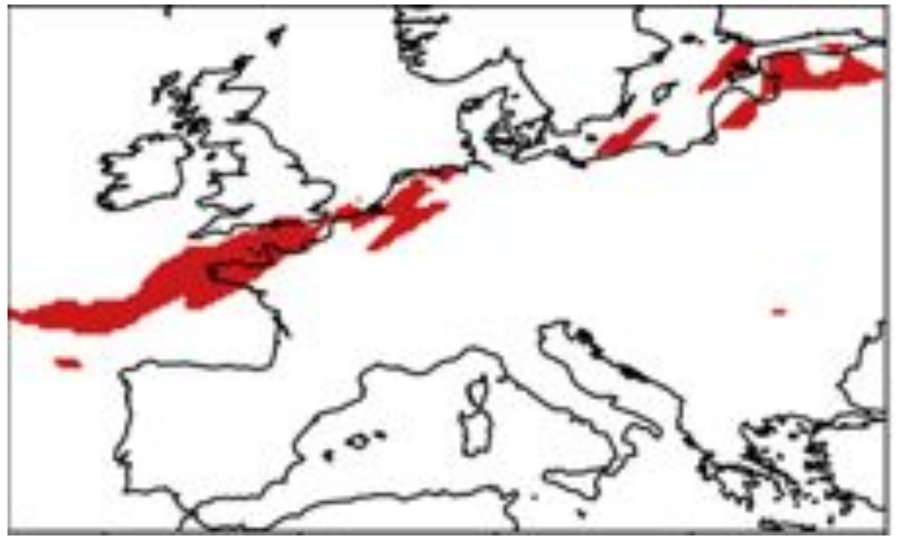
Problem modeling

- **Mitigate climate impact:**

  Identify climate hotspot areas

  Modify flight trajectories to avoid climate hotspot areas

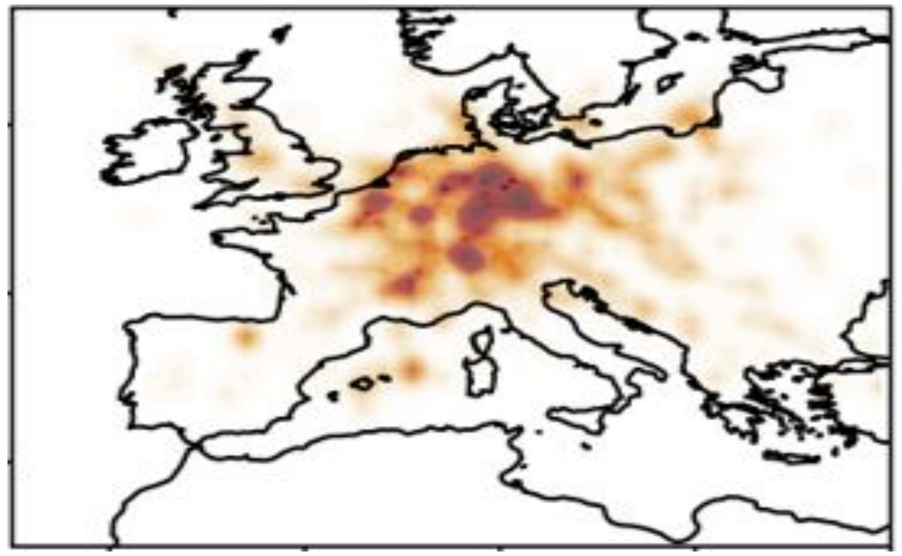
Constraints



- **Maintain manageability:**

  Minimize/maintain the complexity of air traffic

Objective function



This problem can be addressed within the framework of constrained multi-agent reinforcement learning

Constrained multi-agent reinforcement learning framework

Formally, multi-agent reinforcement learning is modeled as a Constrained Markov Decision Process:

(N, A, O, R, C, c)

N: Number of aircraft

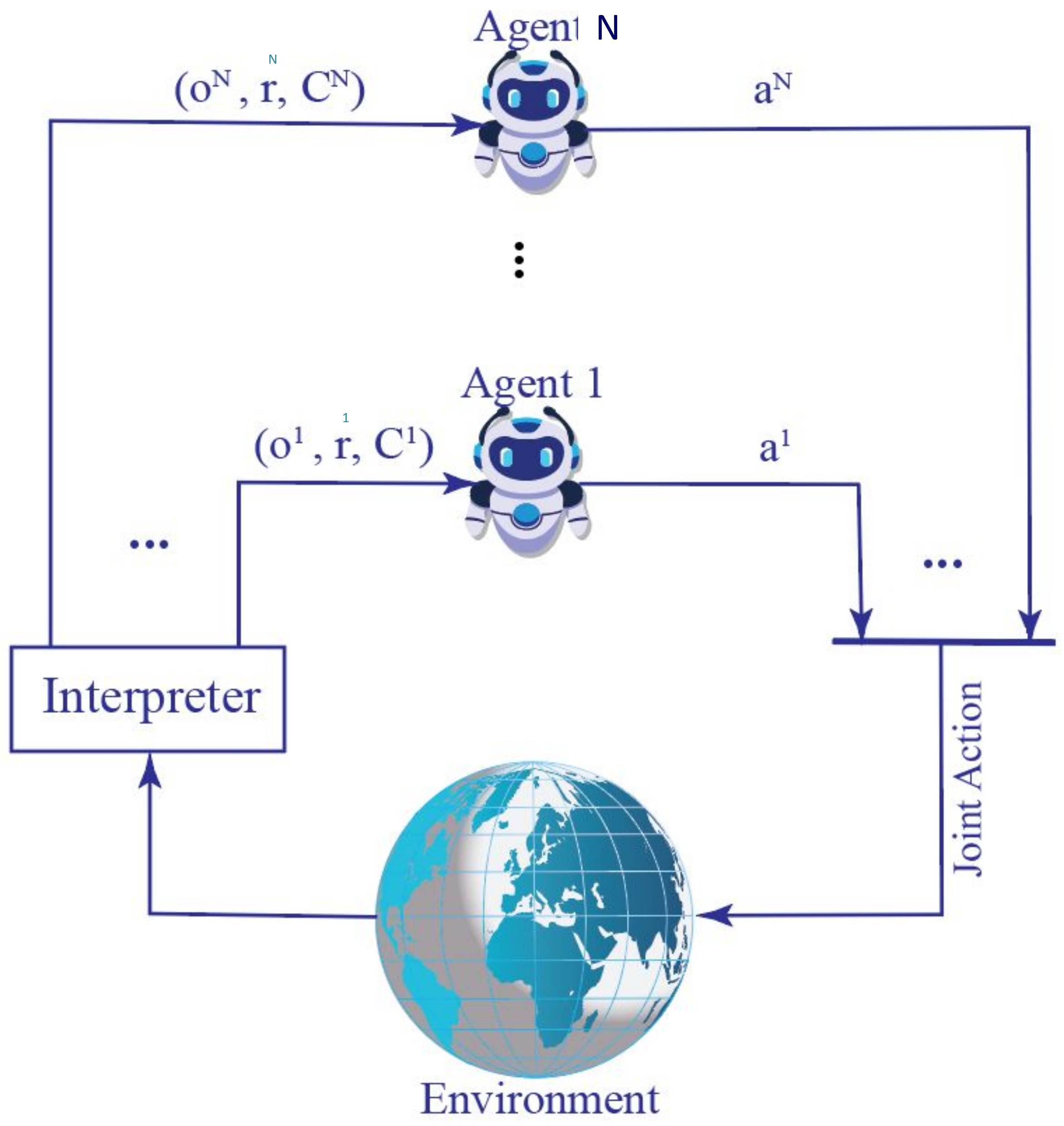
A: Joint action space

O: Observation space

R: Reward function

C: Cost of constraint violation

c: Threshold value



- At each time step t:
- Each agent receives a local observation o_t^i
- Takes action $a_t^i = \pi^i(a_t^i|o_t^i)$ according to its policy
- The joint action $a_t = (a_t^1, \ldots, a_t^N)$ is applied to the environment
- Each agent receive a reward $R(o_t^i, a_t^i)$ and a cost $C(o_t^i, a_t^i)$

The goal is to find the optimal policy that:

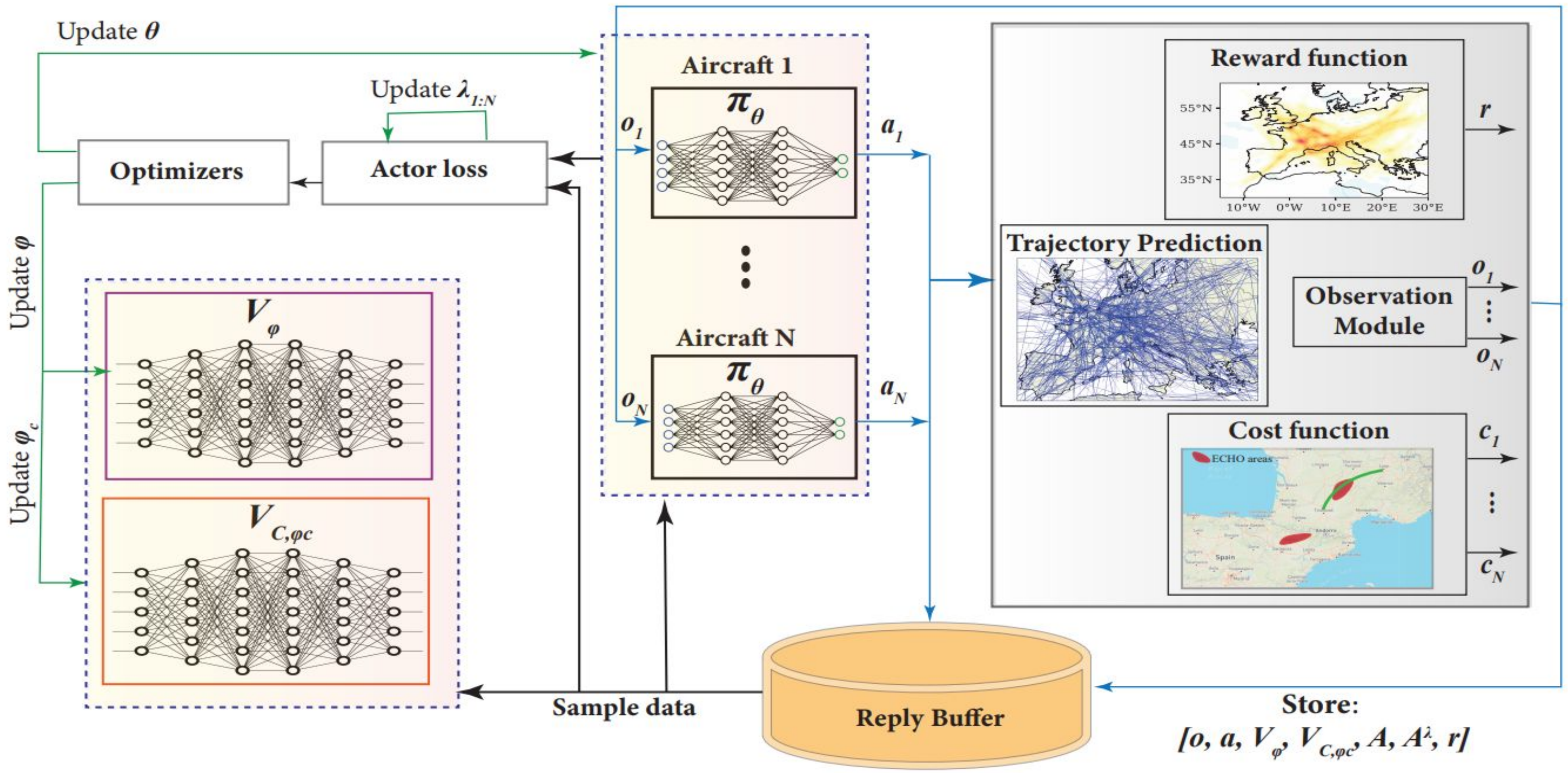
Maximize

$$J(\pi) = \mathbb{E}_{a_t \sim \pi}\left[\sum_{t=0}^{T_f} R(o_t, a_t)\right]$$

s.t.

$$J^i(\pi) = \mathbb{E}_{a_t^i \sim \pi^i}\left[\sum_{t=0}^{T_f} C(o_t^i, a_t^i)\right] < c$$

Casting climate-optimal trajectory planning at network scale as a constrained MARL problem



Observations

$$o_t^i = [\tau_t^i, E, \chi_t^i, p_t^i, v_t^i, T_t^i, I_t]$$

- Trajectory data

$$\tau_t^i = [(\varphi_0, \lambda_0, h_0), \dots, (\varphi_k, \lambda_k, h_k)]$$

- Hotspot areas

$$E = [(\varphi_{e_1}, \lambda_{e_1}, h_{e_1}), \dots, (\varphi_{e_{n_h}}, \lambda_{e_{n_h}}, h_{e_{n_h}})]$$

- Neighboring Aircraft

$$I_t = ((r_v^1, r_\chi^1, r_\gamma^1), \dots, (r_v^m, r_\chi^m, r_\gamma^m))$$

Actions

- Lateral adjustment

  $\varphi,\ \lambda + \Delta l$

  Δl = [0.4, 0.2, 0, -0.2, -0.4]
- Altitude modification

  $FL + \Delta FL$

  ΔFL = [20, 40, 0, -20, -40]
- Speed regulation

  M+ΔM

  ΔM = [-0.3, 0, 0.3]

Reward

$$R_t = \Psi_0 - \sum_{i=1}^{N} \sum_{k=1, k\neq i}^{N} \Psi_t^{i,k}$$

$$\Psi_t^{i,k} = \sum_{g_t}^{g_{t+\Delta t}} (\nu^{i,k} + \varkappa^{i,k} + \upsilon^{i,k})$$

Cost penalty

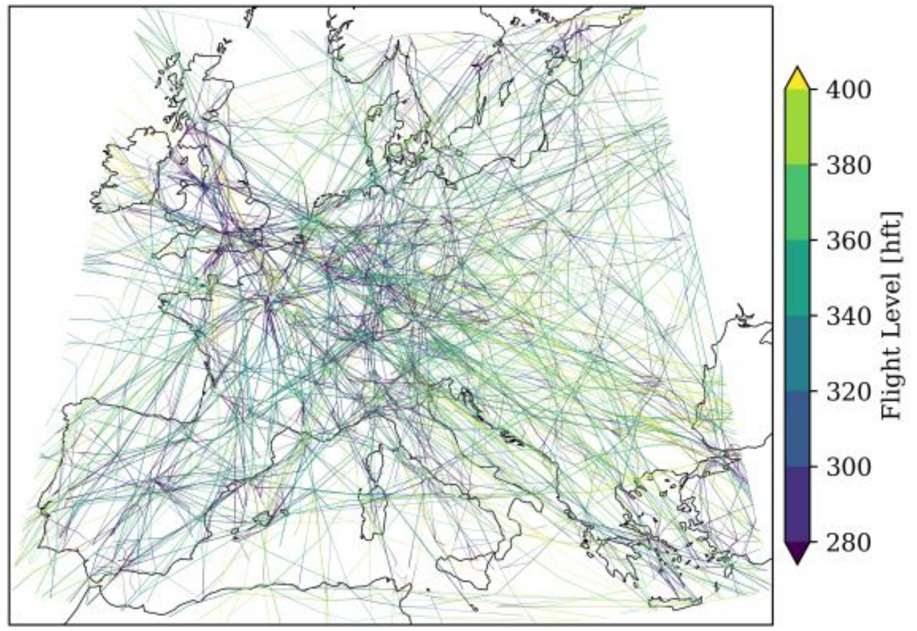
$$C_t^i = \begin{cases} c_h & \text{if } \tau_t^i \in E \\ 0 & \text{otherwise} \end{cases}.$$

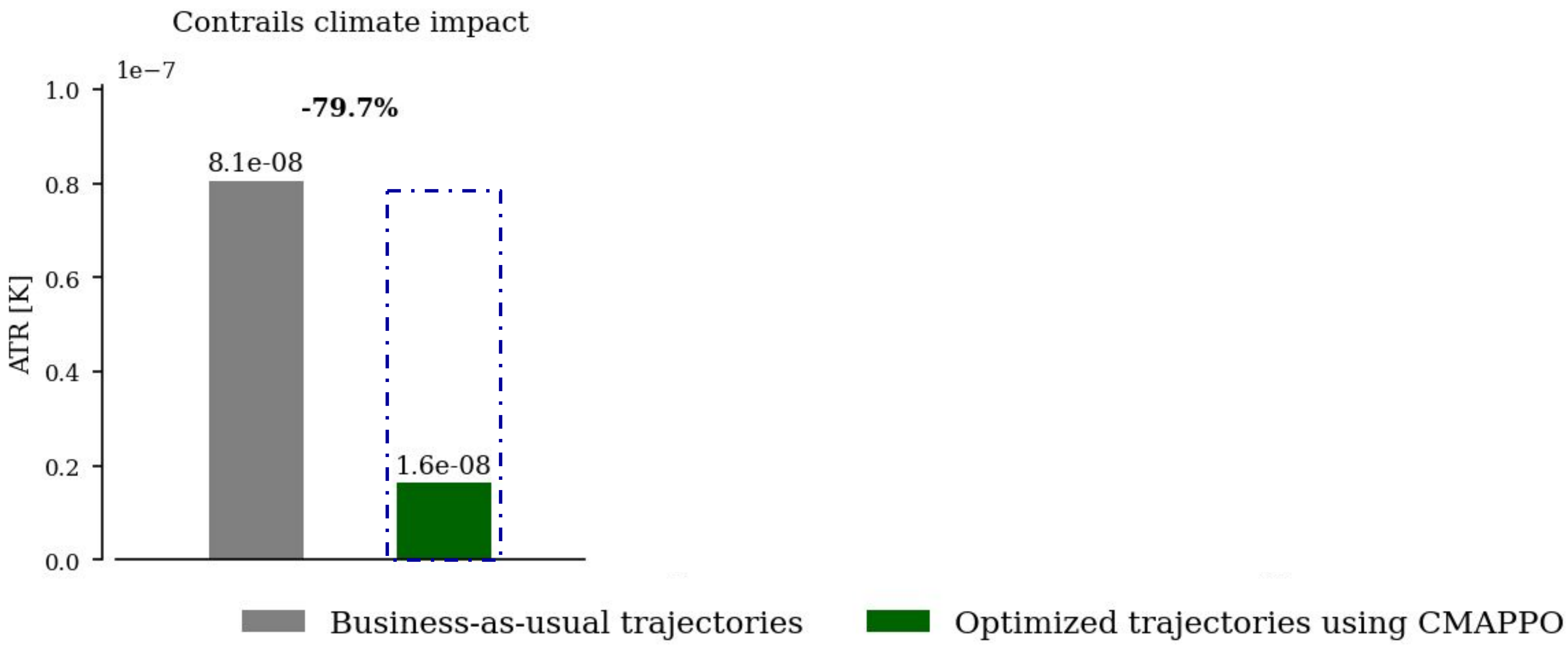
E: Hotspot areas (ECHO areas)

Results

Case study:

- **Date:** December 20th, 2018
- **Time:** 12:00 to 14:00
- **Weather data:** ERA5 reanalysis data
- **Region:** ECAC airspace
- **Data source:** DDR2
- **Routing:** Structured

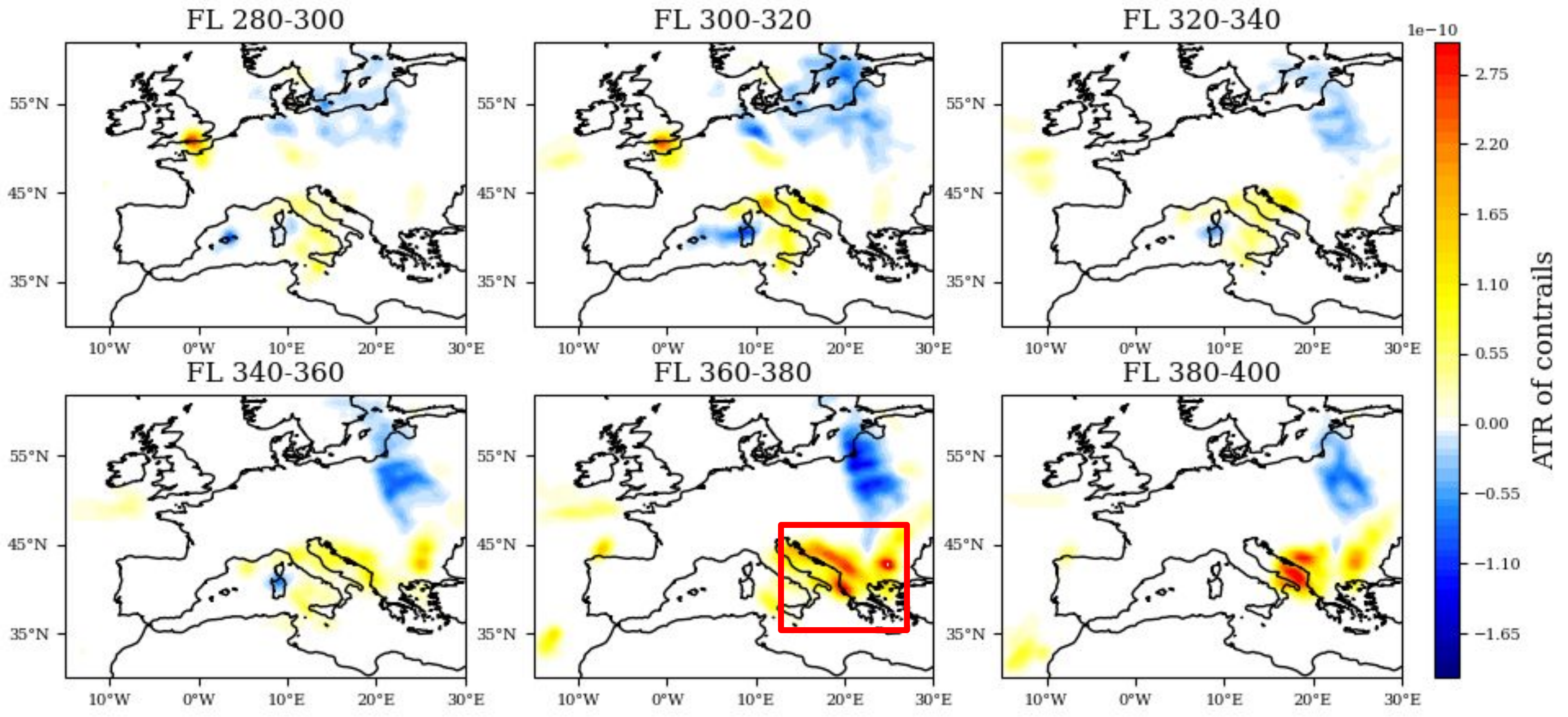


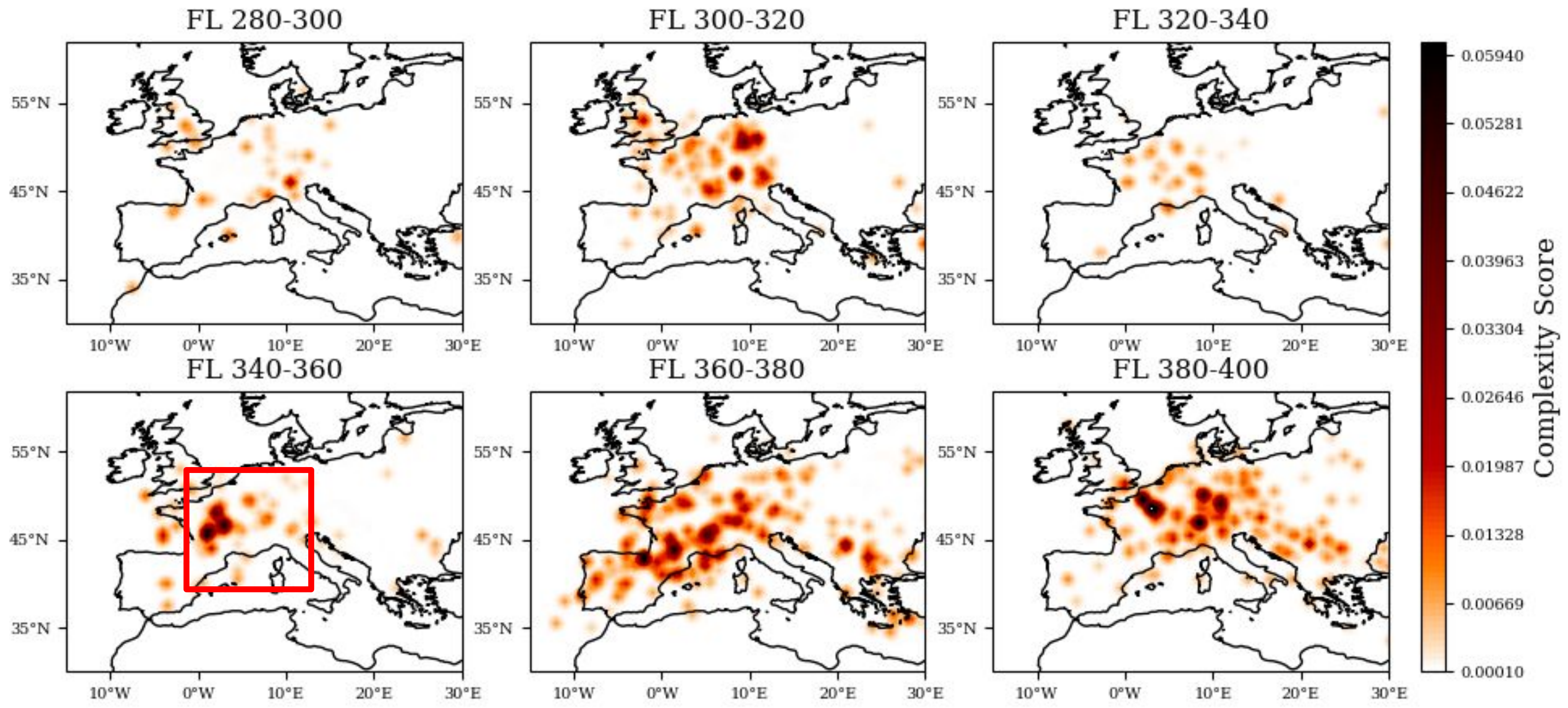


Results

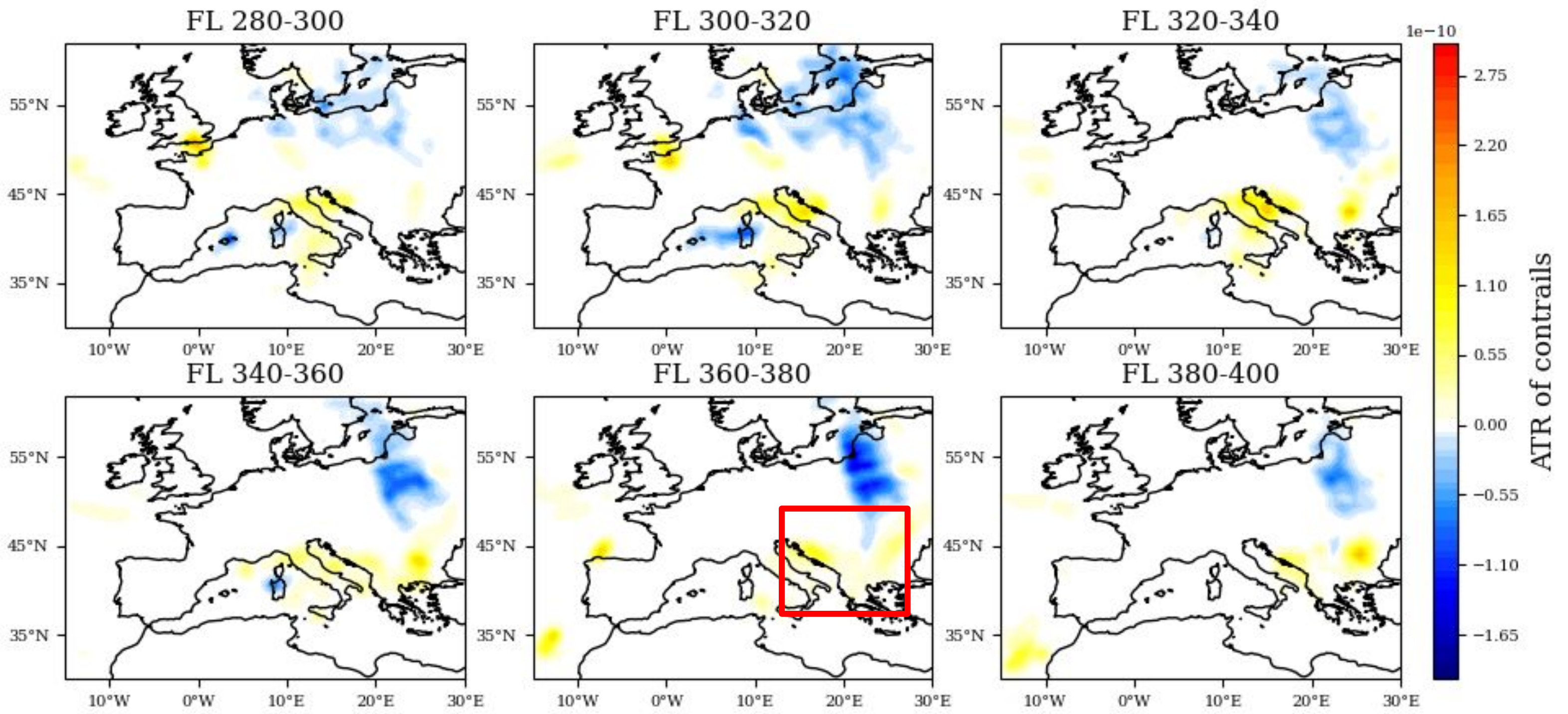
Climate impact | **Air traffic complexity**

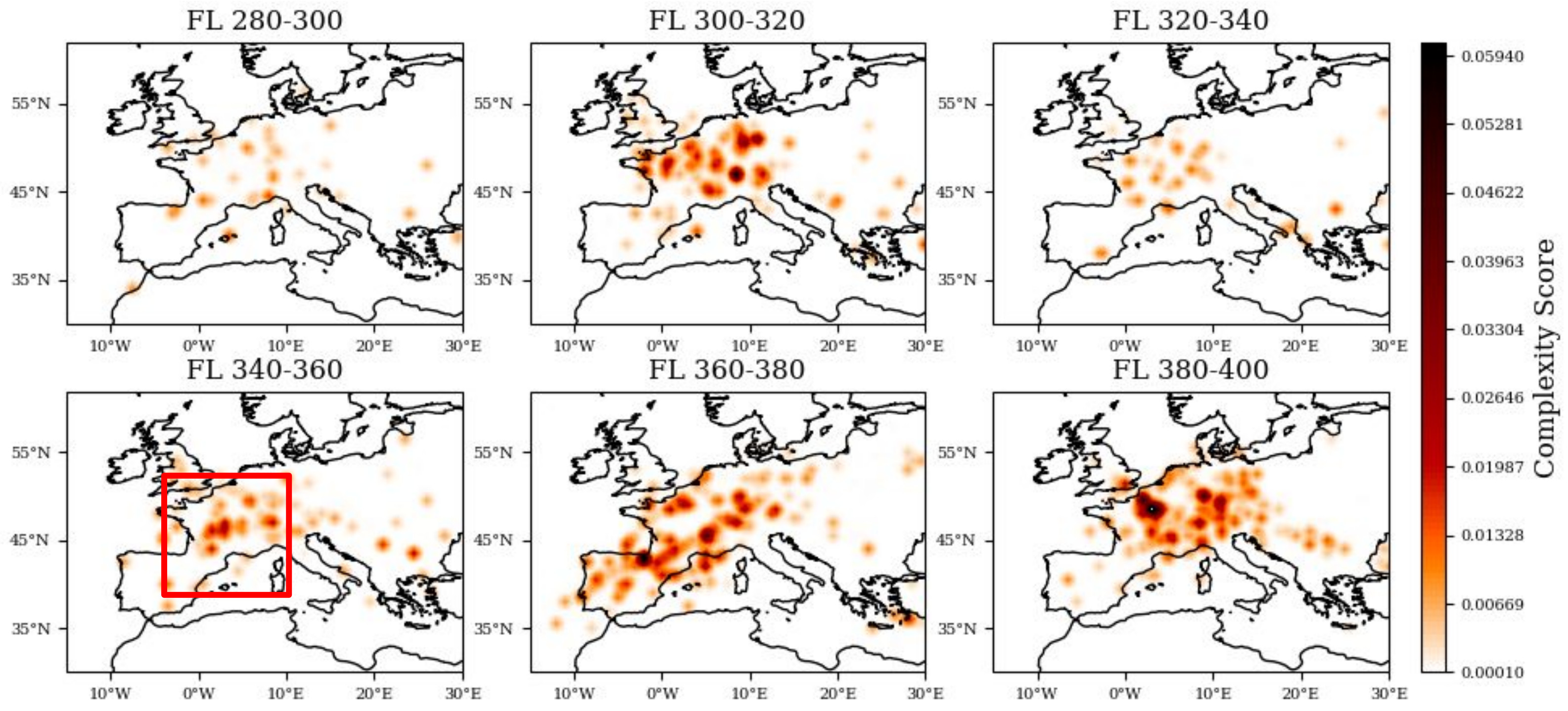
Initial trajectories (Business-as-usual-trajectories)





Optimized trajectories using the proposed Constrained MAPPO





Funded by
the European Union

Conclusions

- Planning aircraft trajectories to avoid climate-sensitive areas poses operational challenges, including increased traffic complexity.

- A framework was introduced to plan climate-optimized trajectories at the air traffic network scale.

- The presented approach employs the MARL algorithm and adapts it to handle constraints related to climate hotspot avoidance.

- The proposed constrained MARL has the potential to plan operationally feasible climate-optimal trajectories, simultaneously considering both climate impact and air traffic complexity.

AVIATION
TWIN TRANSITION
CLUSTER

refmap

Funded by
the European Union

Thank you
for your attention!

Fbaneshi@pa.uc3m.es

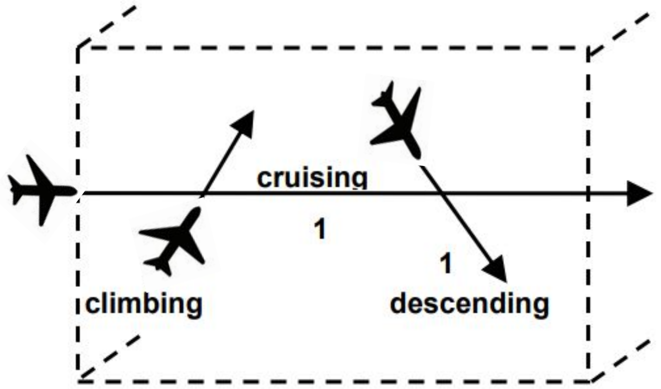
For more information:
www.refmap.eu

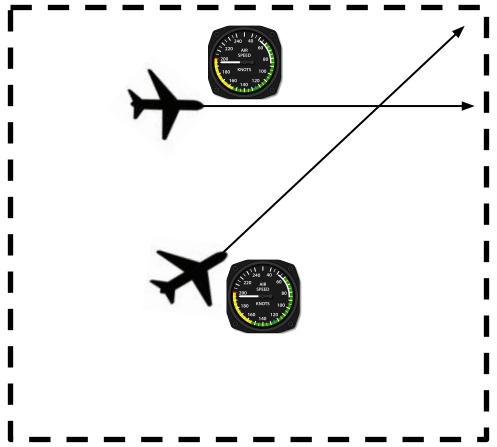
Q&A / Closing

AVIATION TWIN TRANSITION CLUSTER

Air traffic complexity

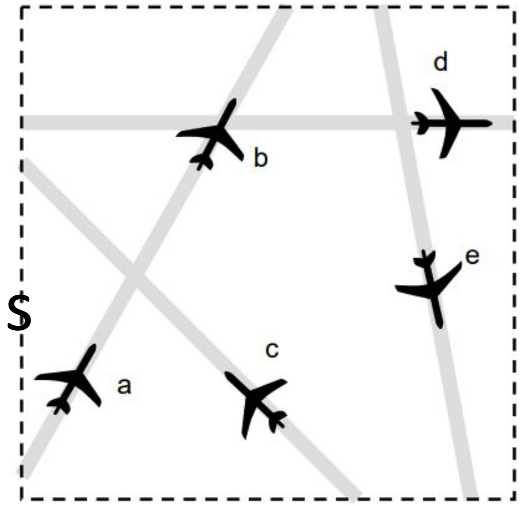
Vertical Interactions

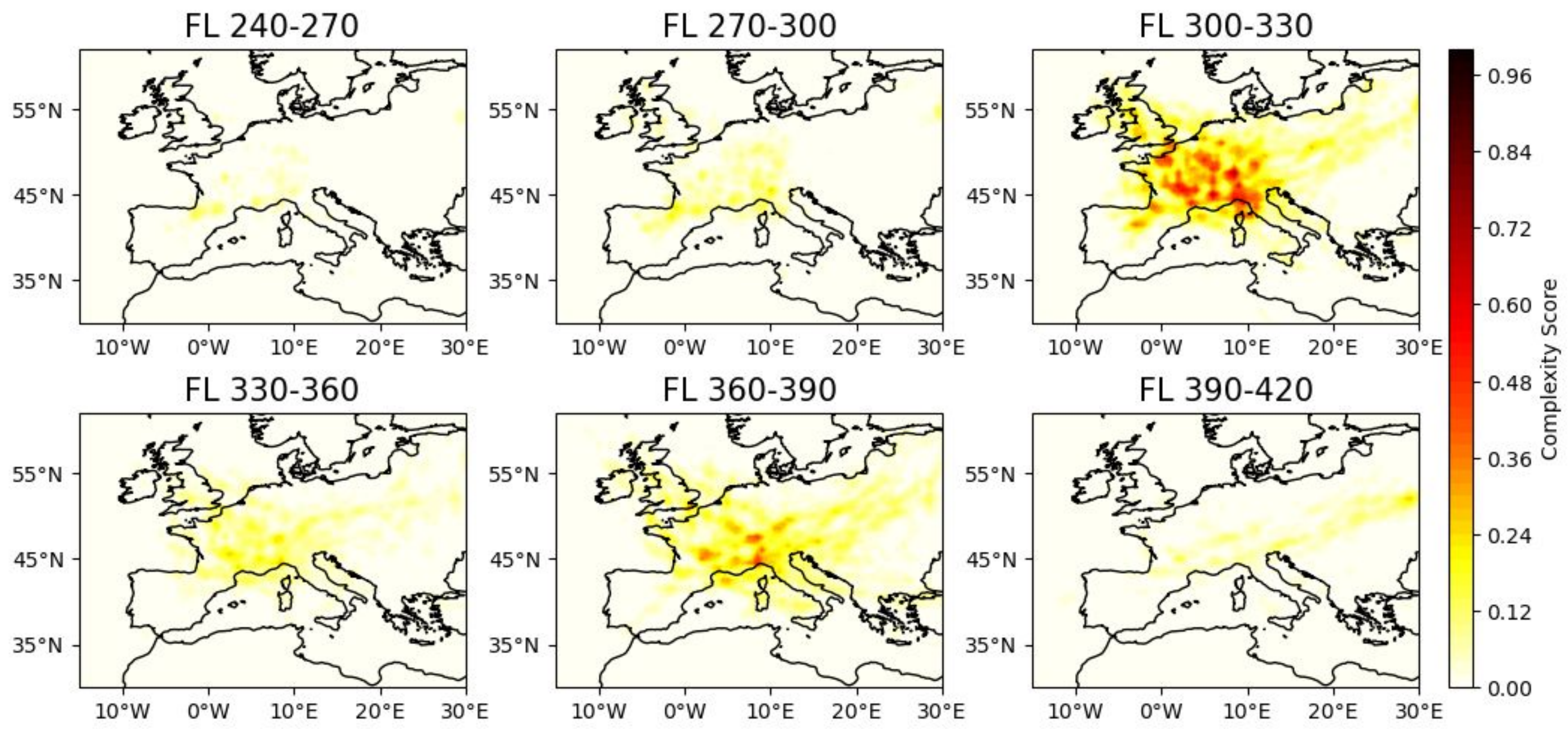




Speed Interactions

Horizontal Interactions





Complexity metric = Horizontal Interactions + Speed Interactions + Vertical Interactions

$$C^i_{clx} = \sum_{k=1,k\neq i}^{N} \sum_{g_0}^{g_f} \nu + \varkappa + \upsilon$$

$$\nu_i = \begin{cases} 2\tau^2/((t_x{}^i - t_e{}^i) + (t_x{}^k - t_e{}^k)) & \text{if} \quad ([t_e{}^i, t_x{}^i] \cap [t_e{}^k, t_x{}^k] \neq \varnothing) \quad and \quad (P^i \neq P^k) \\ 0 & \text{else} \end{cases}$$

$$\varkappa_i = \begin{cases} 2\tau^2/((t_x{}^i - t_e{}^i) + (t_x{}^k - t_e{}^k)) & \text{if} \quad ([t_e{}^i, t_x{}^i] \cap [t_e{}^k, t_x{}^k] \neq \varnothing) \quad and \quad (|\chi^i - \chi^k| > 20°) \\ 0 & \text{else} \end{cases}$$

$$\tau = [t_e{}^i, t_x{}^i] \cap [t_e{}^k, t_x{}^k].$$

$$\upsilon_i = \begin{cases} 2\tau^2/((t_x{}^i - t_e{}^i) + (t_x{}^k - t_e{}^k)) & \text{if} \quad ([t_e{}^i, t_x{}^i] \cap [t_e{}^k, t_x{}^k] \neq \varnothing) \quad and \quad (|v^i - v^k| > 35kts) \\ 0 & \text{else} \end{cases}$$

Constrained Multi-agent Proximal policy Optimization

Objective:

$$\max_{\pi} \mathbb{E}_{\boldsymbol{a}_t \sim \boldsymbol{\pi}} \left[\sum_{t=0}^{T_f} R(o_t, a_t)\right]$$

s.t.

$$\mathbb{E}_{a_t^i \sim \pi^i} \left[\sum_{t=0}^{T_f} \mathrm{C}(o_t^i, a_t^i)\right] < c$$

Assumptions

- **Fully cooperative setting**

  The reward function ***R*** depends on the joint actions of all agents, and all agents receive the same reward.

- **Homogeneous agents**

  The policy is parametrized by $\boldsymbol{\theta}$ and shared between agents.

Definitions:

$$V_R^{\pi}(s) := \mathbb{E}_{\mathbf{a}_t \sim \pi, s_t \sim \mathcal{P}} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, \mathbf{a}_t) \mid s_0 = s\right]$$

$$V_{C^i}^{\pi}(s) := \mathbb{E}_{\mathbf{a}_t \sim \pi, s_t \sim \mathcal{P}} \left[\sum_{t=0}^{\infty} \gamma^t C^i(s_t, a_t^i) \mid s_0 = s\right]$$

Using adaptive Lagrange multipliers:

$$\max_{\theta} \min_{\lambda^i \geq 0,\, i \in \mathcal{N}} V_R^{\pi_\theta}(s_0) - \sum_{i \in \mathcal{N}} \lambda^i (V_{C^i}^{\pi_\theta}(s_0) - c^i)$$

Constrained Multi-agent Proximal policy Optimization

Objective:

$$\max_{\theta} \min_{\lambda^i \geq 0, i \in \mathcal{N}} V_R^{\pi_\theta}(s_0) - \sum_{i \in \mathcal{N}} \lambda^i \left(V_{C^i}^{\pi_\theta}(s_0) - c^i\right)$$

Proposed constrained multi-agent Proximal policy optimization (MAPPO):

- By clipping the probability ratio ($\frac{\pi_\theta}{\pi_{\theta_{\text{old}}}}$) within $(1-\epsilon, 1+\epsilon)$ it ensures that the new policy remains close to the old policy.

$$L\left(\theta, \{\lambda^i\}_{i \in \mathcal{N}}\right) := \mathbb{E}_{\mathbf{a} \sim \pi_\theta, s \sim p}\left[\sum_{i=1}^{N} \min\left(\frac{\pi_\theta(a^i \mid o^i)}{\pi_{\theta_{\text{old}}}(a^i \mid o^i)} A_{\lambda^i}^{\pi_\theta}(s, \mathbf{a}), \boxed{\operatorname{clip}\left(\frac{\pi_\theta(a^i \mid o^i)}{\pi_{\theta_{\text{old}}}(a^i \mid o^i)}, 1-\epsilon, 1+\epsilon\right)} A_{\lambda^i}^{\pi_\theta}(s, \mathbf{a})\right)\right]$$

$$\boxed{A_{\lambda^i}^{\pi_\theta}(s, \mathbf{a}) := \frac{A_R^{\pi_\theta}(s, \mathbf{a})}{N} - \lambda^i \left(A_{C^i}^{\pi_\theta}(s, a^i) - c^i\right)}$$

- $\boldsymbol{A}(\boldsymbol{s}, \boldsymbol{a})$ is the advantage function evaluates the benefit of taking action $\boldsymbol{a}$ in state $\boldsymbol{s}$ relative to the baseline value.

We iteratively apply the following update rules:

$$\lambda^i \leftarrow \lambda^i - \alpha_\lambda \nabla_{\lambda^i} L\left(\theta, \{\lambda^i\}_{i \in \mathcal{N}}\right), \forall i \in \mathcal{N},$$
$$\theta \leftarrow \theta + \alpha_\theta \nabla_\theta L\left(\theta, \{\lambda^i\}_{i \in \mathcal{N}}\right),$$