Network-Level Aircraft Trajectory Optimization for Climate Impact Mitigation using Multi-Agent Reinforcement Learning

by

Fateme BANESHI

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Aerospace Engineering

Universidad Carlos III de Madrid

Advisors

Manuel SOLER ARNEDO

María CEREZO-MAGAÑA

Tutor

Manuel SOLER ARNEDO

October 2025

Acknowledgements

The research presented in this thesis has been financially supported by the EU project RefMAP. RefMAP received funding from the Horizon Europe program under grant agreement No. 101096698.

This thesis marks the end of a long and challenging journey, one that I could not have completed without the support of others. I'm deeply grateful to all those who supported, guided, and believed in me along the way.

I would like to express my most sincere gratitude to my supervisors, Professors Manuel Soler and María Cerezo. None of this would have been possible without their hard work, dedication, and patience. I will always be indebted to them for their guidance, support, and kindness through all these years. I am especially thankful to Professor Manuel Soler. Working under his guidance has been one of the most rewarding aspects of this PhD. He always welcomed new ideas with openness and encouraged me to pursue them confidently. He is not only an outstanding researcher and mentor but also a deeply admirable human being. I am truly grateful for everything I have learned from him, both professionally and personally.

I am deeply grateful to my beloved husband, Abolfazl, whose unwavering patience, understanding, and support have been invaluable throughout this journey. His continuous encouragement and belief in my potential have been my greatest source of motivation. I feel truly fortunate to have shared not only life but also work with someone so brilliant. His sharp thinking, insightful ideas, constant feedback, and long discussions have had a direct impact on this thesis. He has helped me in more ways than I can count, and his belief in me gave me the strength to continue, even during the most difficult times. There are no words sufficient to express how deeply thankful I am for everything he has done.

I would also like to extend my appreciation to Professor Maryam Kamgarpour for kindly hosting me during my research stay at EPFL. Her expert guidance and constructive feedback significantly enriched my research. I am also thankful to my colleague, Tingting Ni, for her collaboration and thoughtful discussions during that period.

Many thanks to my colleagues at the Aerospace Engineering Department of Universidad Carlos III de Madrid, who made the last four years much more enjoyable. I will not attempt to name all of them, knowing that I would definitely miss someone, but I am truly grateful for all the shared moments, office discussions, lunches, coffee breaks, and PhD gatherings. I am especially thankful to Professor Javier García-Heras for his consistent support and encouragement, and to Dr. Daniel González Arribas, whose help during the early stages of my PhD was truly instrumental.

Last but not least, I am deeply thankful to my family for their endless love, strength, and encouragement. I also wish to honor the memory of my father, whose belief in me and constant support meant more than words can express. Though he is no longer with us, his presence has accompanied me every step of the way.

Published and Submitted Contents

- **Baneshi, F.**, Soler, M., Simorgh, A. (2023). Conflict assessment and resolution of climate-optimal aircraft trajectories at network scale. Transportation Research Part D: Transport and Environment, 115, 103592. Link.
  - The publication is partly included in the Thesis.
  - Material from this publication, related to the development of a centralized conflict resolution strategy based on a heuristic approach within a fully free-routing airspace is included in Chapter 3.
  - Fateme Baneshi conducted the main research under the supervision of Manuel Soler, implemented the code, performed the simulations, and wrote the manuscript.
  - Any material drawn from this source is explicitly cited throughout the thesis.
- **Baneshi, F.**, Cerezo-Magaña, M., & Soler, M. (2024). Integrating Non-CO_2 climate impact considerations in air traffic management: Opportunities and challenges. Transport Policy, Link.
  - The publication is partly included in the Thesis.
  - Material from this publication, related to the evaluation of the operational manageability of climate-optimized trajectories within the structured airspace, is included in Chapter 5.
  - Fateme Baneshi conducted the main research under the supervision of Manuel Soler and María Cerezo Magaña, implemented the code, performed the simulations, and wrote the manuscript.
  - Any material drawn from this source is explicitly cited throughout the thesis.
- **Baneshi, F.**, Cerezo-Magaña, M., & Soler, M. (2025). Network-level aircraft trajectory planning via multi-agent deep reinforcement learning: Balancing climate considerations and operational manageability. Expert Systems with Applications, 126604. Link.
  - The publication is partly included in the Thesis.
  - Material from this publication, related to the development of a distributed framework for conflict mitigation based on a multi-agent deep reinforcement learning approach within a fully free-routing airspace, is included in Chapter 4.
  - Fateme Baneshi conducted the main research under the supervision of Manuel Soler and María Cerezo Magaña, implemented the code, performed the simulations, and wrote the manuscript.
  - Any material drawn from this source is explicitly cited throughout the thesis.
- **Baneshi, F.**, Cerezo-Magaña, M., & Soler, M. (2025). A Deep Multi-Agent Reinforcement Learning Framework for Climate-Aware Aircraft Trajectory Planning Considering Air Traffic Complexity. Knowledge-Based Systems, 114826. Link

- The publication is partly included in the Thesis.
- Material from this publication, related to the development of a multi-agent reinforcement learning framework for complexity management within the structured airspace, is included in Chapter 5.
- Fateme Baneshi conducted the main research under the supervision of Manuel Soler and María Cerezo Magaña, implemented the code, performed the simulations, and wrote the manuscript.
- Any material drawn from this source is explicitly cited throughout the thesis.

Other Research Merits

- **Awards**
  - **Luis Azcárraga Aeronautical Innovation Award**, Fundación EnAire, 2023 (€10,000) to the paper, "Conflict assessment and resolution of climate-optimal aircraft trajectories at network scale." **Baneshi et al. (2023).** Link.
  - **Best Paper Award** at the International Conference on Research in Air Transportation (ICRAT 2022) to "Climate Optimal Trajectory Planning at Network-Scale: Complexity Assessment Based on Probabilistic Conflicts". **Baneshi et al. (2022)**.

- **Journal papers**
  1. González-Arribas, D., **Baneshi, F.**, Andrés, E., Soler, M., Jardines, A., & García-Heras, J. (2023). Fast 4D flight planning under uncertainty through parallel stochastic path simulation. Transportation Research Part C: Emerging Technologies, 148, 104018.

- **Conferences papers**
  1. **Baneshi, F.**, Soler, M., Simorgh, A., Gonzalez-Arribas, D. (2022). Climate Optimal Trajectory Planning at Network-Scale: Complexity Assessment Based on Probabilistic Conflicts. International Conference on Research in Air Transportation (ICRAT 2022). Link.
  2. **Baneshi, F.**, Soler, M., Simorgh, A.,& Martinez, I. (2022). Demand Assessment for Climate Optimal Aircraft Trajectories at Network Scale. In Proceedings of International Workshop on ATM/CNS 2022 International Workshop on ATM/CNS (pp. 200-207). Electronic Navigation Research Institute. Link.
  3. **Baneshi, F.**, Soler, M., Simorgh, A. (2022). Complexity Assessment of Adopting Climate Optimal Aircraft Trajectories at Network Scale. International Conference on Innovation in Aviation & Space towards sustainability today and tomorrow (EASN 2022).
  4. **Baneshi, F.**, & Soler, M. (2023). Network Assessment of the Aviation Climate Impact Considering the European Structured Airspace (No. EGU23-17517). Copernicus Meetings. Link.
  5. **Baneshi, F.**, & Soler, M. (2023). A Reinforcement Learning Approach to Conflict Resolution of Climate Optimal Trajectories". The European Conference for Aeronautics and Space Sciences (EUCASS 2023). Link.

6. **Baneshi, F.**, & Soler, M. (2023). Scalable Reinforcement Learning Algorithm for Strategic Deconfliction of Climate-Optimal Trajectories. 4th European Conference for Aviation and the Environment (ECATS 2023). Link.
7. **Baneshi, F.**, Cerezo Magaña, M., Soler, M., Ni, T., & Kamgarpour, M. (2024). Aircraft Trajectory Planning for Climate Hotspot Avoidance Considering Air Traffic Complexity: A Constrained Multi-Agent Reinforcement Learning Approach. SESAR Innovation Days 2024. Link.
   - The publication is partly included in the Thesis.
   - Material from this publication, related to the development of an integrated framework for climate optimal flight planning at the network scale, is included in Chapter 6.
   - Fateme Baneshi conducted the main research under the supervision of Manuel Soler, María Cerezo Magaña, and Maryam Kamgarpour (supervisor during Fateme Baneshi's research stay at EPFL), implemented the code, performed the simulations, and wrote the manuscript. The other authors contributed with comments and revisions.
   - Any material drawn from this source is explicitly cited throughout the thesis.

- **Research Stay**
  - **Institution:** École Polytechnique Fédérale de Lausanne (EPFL), Systems Control and Multi-Agent Optimization (Sycamore) Laboratory
  - **Period:** April 2024 – July 2024
  - **Topic:** Constrained Multi-Agent Reinforcement Learning for Network-Scale Climate-Optimal Flight Planning
  - **Host Supervisor:** Prof. Maryam Kamgarpour

Abstract

The aviation sector contributes to climate change directly through the emission of carbon dioxide (CO_2), water vapor, sulfur dioxide, and soot, and indirectly through the formation of contrail cirrus and chemically induced changes in ozone, methane, and stratospheric water vapor resulting from nitrogen oxide emissions. Collectively, these effects are estimated to account for approximately 3–5% of total anthropogenic radiative forcing, placing increasing pressure on the aviation industry to mitigate its environmental footprint. Notably, non-CO_2 effects, responsible for nearly two-thirds of aviation's net effective radiative forcing, albeit with considerable uncertainty, are highly sensitive to the time and location of the emissions. Consequently, operational measures such as climate-aware flight planning, aimed at avoiding regions with strong warming effects or intentionally flying through areas associated with cooling impacts, offer a practical and infrastructure-compatible solution to mitigate aviation's climate impact in the short term.

In recent years, there has been a growing interest in exploring the potential of flight planning to mitigate the climate impact of aviation. While these studies demonstrate promising reductions in climate effects, they have limited their focus to individual trajectory optimization, overlooking the interactions between flights and the resulting implications on the air traffic management (ATM) system. Indeed, optimizing individual flight trajectories to account for climate-sensitive areas results in a redistribution of traffic, diverting flows away from regions associated with warming effects and increasing congestion in areas characterized by cooling effects. Such redistribution can compromise air traffic safety and manageability, particularly by introducing imbalances between sector capacity and traffic demand and increasing traffic complexity in specific regions, thereby raising concerns about the feasibility of climate-aware flight planning from the perspective of the ATM system. The current state of the literature lacks a framework for evaluating the climate benefits achievable through flight planning, one that extends beyond purely climate considerations to also account for the operational constraints and challenges associated with implementing such a measure in practice.

This thesis aims to fill this gap by developing frameworks that integrate air traffic manageability considerations into climate-optimal flight planning. To this end, we first introduce a sequential optimization framework. Within this strategy, individual flight plans are initially optimized considering climate impact and operational costs. Subsequently, these flight plans are integrated into the network traffic, and their collective impact on the manageability of air traffic is evaluated. Finally, resolution strategies are proposed to compensate for any adverse effects these trajectories may impose on the ATM system performance. Within this framework, several distinct approaches are presented, each characterized by its specific resolution strategy, performance indicators, and routing flexibility.

The initial implementation of the sequential approach is conducted in fully free-routing airspace. Individual flight plans are laterally optimized to mitigate climate impact, and the resulting potential conflicts associated with the adoption of these optimized trajectories are then evaluated. A centralized conflict resolution strategy based on the simulated annealing algorithm is proposed, employing speed adjustments to mitigate conflict probability while

preserving climate optimality. The developed strategy is evaluated through a regional case study over the Spanish airspace.

To overcome the scalability and adaptability limitations of the centralized heuristic approach, a distributed resolution strategy is subsequently proposed using a multi-agent reinforcement learning (MARL) framework. The proposed strategy leverages the twin delayed deep deterministic policy gradient algorithm to adjust aircraft speed during the flight planning phase to resolve the potential conflicts associated with climate-optimal trajectories. Given its scalability, the framework is benchmarked across the entire European airspace, considering three-dimensional optimized flight plans.

Building on the MARL framework, the study further incorporates traffic complexity as a more comprehensive and robust measure of air traffic manageability during the flight planning phase. It extends the analysis to realistic conditions within the currently structured airspace. To mitigate the traffic complexity of three-dimensional climate-optimized flight plans, a distributed MARL approach is proposed, modeling the problem as a partially observable Markov decision process and employing shared policy parameters to enhance scalability. The proximal policy optimization algorithm is used with decision variables extended to include both vertical maneuvers and speed adjustments.

Finally, to further enhance computational efficiency and enable large-scale analyses that support policy development, an integrated optimization framework is introduced. Unlike the sequential approach, the integrated framework simultaneously mitigates climate impact and traffic complexity in a single step. Climate hotspots are treated as constraints to be avoided, whereas traffic complexity is formulated as the objective function to be minimized. The problem is formulated as a constrained Markov decision process and solved using an adapted version of proximal policy optimization, incorporating Lagrangian techniques to enforce climate-related constraints. The decision space is further extended to allow lateral, vertical, and speed modifications.

The analyses conducted in this thesis using real traffic scenarios indicate that the climate benefits reported from individual flight planning may not be fully realizable in practice, primarily due to their adverse impact on the ATM system performance, particularly in terms of increased conflicts and traffic complexity. Through the implementation of the developed network-scale flight planning methodologies, we demonstrate that by accepting a modest reduction in the potential climate benefit achieved via individual flight planning, it is possible to maintain, and in some cases even improve, operational manageability at levels comparable to those of standard business-as-usual flight plans, albeit with a slight increase in the operational cost. Therefore, the proposed methodologies provide a practical means of balancing environmental, operational, and cost-efficiency objectives.

Overall, the work presented in this thesis represents an important advancement toward a more reliable evaluation of aviation's climate impact mitigation potential through flight planning and contributes to the development of practical and scalable frameworks supporting the future implementation of climate-optimal aircraft trajectories.

Contents

List of Figures

List of Tables

List of Symbols

t	Time
ϕ	Latitude
λ	Longitude
h	Altitude
ψ	Course
v_{gs}	Ground speed
v	True airspeed
m	Aircraft mass
v_{CAS}	Calibrated airspeed
s	Distance flown along the route
R_N	Ellipsoid radius of curvature in the prime vertical
R_M	Ellipsoid radius of curvature in the meridian
f_c	Fuel flow
M	Mach number
γ_p	Flight path angle
χ	Heading
C_T	Coefficient of thrust
C_L	Coefficient of lift
T	Thrust force
$\mathbf{x}$	Vector of states
$\mathbf{u}$	Vector of controls
D	Drag force
T	Temperature
w_x, w_y	Components of wind
GH	Geopotential
F_{in}	Incoming solar radiation
PV	Potential vorticity
OLR	Outgoing longwave radiation
CI	Cost index
EI	Environmental index
C_t	Time index
C_f	Fuel index
FT	Flight duration
FB	Fuel consumption
F	Flight plan
W	Set of meteorological variable
$\text{aCCF}_{\text{H}_2\text{O}}$	aCCF of water vapor
aCCF_{O_3}	aCCF of ozone

$\mathrm{aCCF}_{\mathrm{CH_4}}$	aCCF of methane
$\mathrm{aCCF}_{\mathrm{Cont.}}$	aCCF of contrails
$\mathrm{aCCF}_{\mathrm{CO_2}}$	aCCF of carbon dioxide
$\mathrm{ATR}_{\mathrm{H_2O}}$	ATR of water vapour
$\mathrm{ATR}_{\mathrm{O_3}}$	ATR of ozone
$\mathrm{ATR}_{\mathrm{CH_4}}$	ATR of methane
$\mathrm{ATR}_{\mathrm{Cont.}}$	ATR of contrails
$\mathrm{ATR}_{\mathrm{CO_2}}$	ATR of carbon dioxide
t_0	Initial time
t_f	Final time
$\mathrm{EI}_{\mathrm{NO_x}}$	Emission index of nitrogen oxides
N_{EPS}	Number of ensemble members
D_0	Minimum safe horizontal distance
H_0	Minimum safe vertical distance
$\mathrm{P}_{\mathrm{c},kl}$	Probability of conflict between aircraft k and l
J_R	Resolution objective
C_C	Cost of conflicts
Γ	Aircraft trajectory
$\mathcal{N}$	Set of agents
N	Number of agents
$\mathcal{S}$	State space
O	Observation space
γ	Discount factor
R	Reward function
C	Penalty function
$\mathcal{P}$	State transition function
s	State
a	Action
o	Observation
r	Reward
c	Cost (Penalty)
$\mathbf{a}$	Joint action
κ	Conflict index
C_V	Cost of speed deviation
$v_{\max}$	Maximum true airspeed
$v_{\min}$	Minimum true airspeed
V	Value function
Q	Action-value function
$\hat{Q}$	Target action-value function
η	Deterministic policy
$\hat{\eta}$	Target deterministic policy
π	Stochastic policy
θ	Actor network parameters
ϑ	Critic network parameters
α_a	Learning rate of the actor
α_c	Learning rate of the critic

$\hat{\alpha}_a$	Learning rate of the actor target
$\hat{\alpha}_c$	Learning rate of the critic target
$\mathcal{O}$	Computational/ space complexity
n_c	Number of critic parameters
n_a	Number of actor parameters
$\mathbb{E}$	Mathematical expectation
$\mathbf{O}$	Joint observations of all agents
$\mathbf{O}'$	Joint next observations of all agents
$\mathbf{R}$	Joint rewards of all agents
Λ	Joint actions of all agents
y	Target value
Ξ	Replay buffer
N_{batch}	Number of batches
A	Advantage function
Ψ	Complexity score
ν	Vertical different interacting flows
$\varkappa$	Horizontal different interacting flows
υ	Speed different interacting flows

List of Abbreviations

AIC	Aviation-induced cloudiness
aCCF	Algorithmic climate change function
ATC	Air traffic controller
ATM	Air traffic management
ATR	Average temperature response
ATS	Air traffic service
BAU	Business-as-usual
BFFM2	Boeing fuel flow method 2
CoCiP	Contrail cirrus prediction model
CCF	Climate change function
CO_2	Carbon dioxide
CH_4	Methane
DCB	Demand-capacity balancing
DDPG	Deep deterministic policy gradient
DRL	Deep reinforcement learning
DQN	Deep Q-network
ECAC	European civil aviation conference
EF	Energy forcing
ECMWF	European center for medium-range weather forecasts
EPS	Ensemble prediction system
GCN	Graph convolutional network
GNN	Graph neural network
GTP	Global temperature change potential
GWP	Global warming potential
GPU	Graphics processing unit
H_2O	Water vapor
ICAO	International civil aviation organization
IDQN	Independent deep Q-network
ISSR	Ice-supersaturated regions
IPOPT	Interior-point optimizer
LSTM	Long short-term memory
MARL	Multi-agent reinforcement learning
MAPPO	Multi-agent proximal policy optimization
MATD3	Multi-agent twin delayed deep deterministic policy gradient
MDP	Markov decision process
MIP	Mixed-integer programming
MILP	Mixed-integer linear programming
MINLP	Mixed-integer nonlinear programming
MPNNs	Message passing neural networks

NLP	Nonlinear programming
NM	Nautical mile
NO_x	Nitrogen oxides
O_3	Ozone
PCFA	Persistent contrail formation areas
POMDP	Partially observable Markov decision process
Ps-MADDPG	Policy-sharing multi-agent deep deterministic policy gradient
Ps-MATD3	Policy-sharing multi-agent twin-delayed deep deterministic policy gradient
PPO	Proximal policy optimization
RF	Radiative forcing
RL	Reinforcement learning
ROC	Robust optimal control for flight planning
ROOST	Robust optimization of structured trajectories
SA	Simulated annealing
SAC	Schmidt-Appleman criteria
SAF	Sustainable aviation fuel
SNOPT	Sparse nonlinear optimizer
SO_2	Sulfur oxide
SOC	Simple operating cost
SQP	Sequential quadratic programming
STD	Standard deviation
TD3	Twin delayed deep deterministic policy gradient
TH	Time horizon
TRPO	Trust region policy optimization
USD	US dollar

Chapter 1

Introduction

1.1 Motivation

The imperative to address climate change has reached a critical juncture, highlighted by the global target established in the Paris Agreement to limit the Earth's temperature increase to 2°C above pre-industrial levels, widely regarded as the upper and potentially irreversible threshold, while striving to remain below 1.5°C [1]. However, this more stringent target has already been breached. The year 2024 witnessed unprecedented global temperatures, continuing the record-breaking trend of 2023. For the first time, the annual global average temperature clearly exceeded 1.5°C above pre-industrial levels (see Figure 1.1). Multiple global records were shattered, including those for greenhouse gas concentrations, air temperature, and sea surface temperature, fueling extreme events such as floods, heatwaves, and wildfires, which are exacerbating risks to ecosystems, infrastructure, and human health. These data underscore the accelerating consequences of human-induced climate change and reinforce the urgent need for decisive action across all sectors. Without immediate and sustained mitigation efforts, the frequency and severity of these climate-related disasters will continue to rise, further threatening the stability of societies, economies, and natural systems worldwide.

The aviation industry has grown substantially over recent decades, becoming a key driver of global transportation and economic development [2]. However, this expansion has raised environmental concerns, particularly regarding aviation's contribution to climate change, which accounts for approximately 4% of total anthropogenic climate forcing since pre-industrial times [3]. This climate impact results not only from the direct emissions of carbon dioxide (CO_2), the most well-known and long-lived greenhouse gas, but also from a range of non-CO_2 effects. These include emissions of sulfur dioxide (SO_2) and water vapor (H_2O), as well as indirect effects such as the formation of contrail cirrus and chemically induced changes in atmospheric concentrations of ozone (O_3), methane (CH_4), and stratospheric water vapor due to nitrogen oxides (NO_x) emissions [4]. According to the latest estimates, these non-CO_2 effects contribute approximately twice the climate forcing of CO_2 emissions from aviation, although this estimate is subject to considerable uncertainty [3].

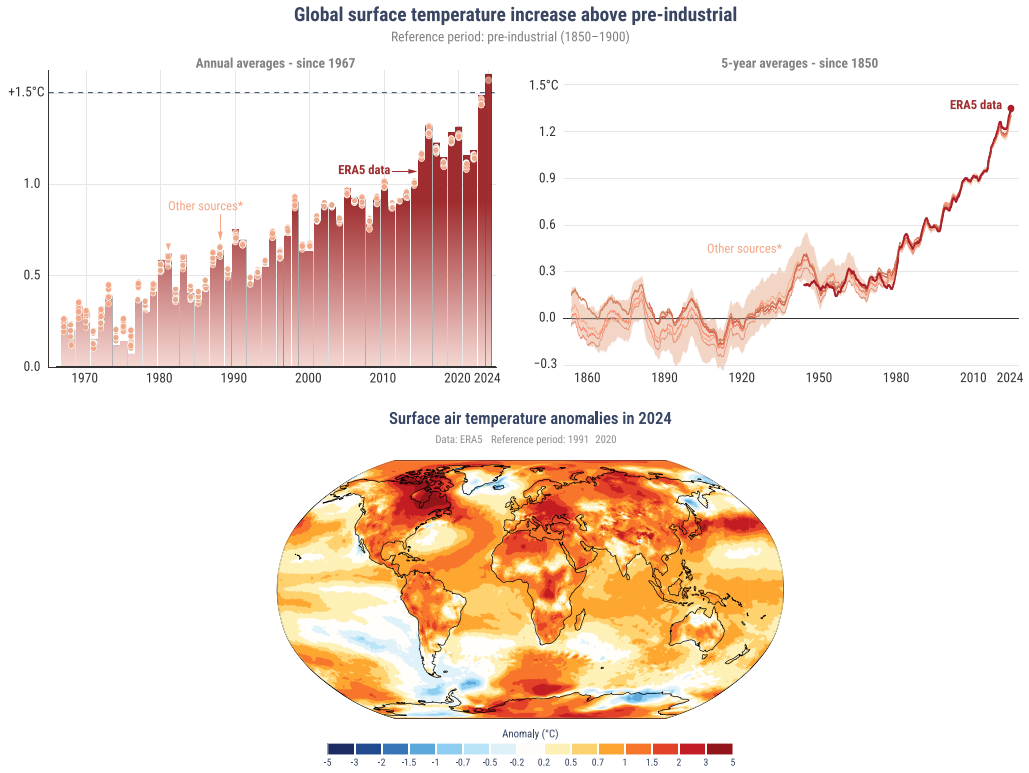


FIGURE 1.1: Global surface temperature trends and anomalies. Top-left panel: Annual global surface temperature increases relative to pre-industrial levels (1850–1900) since 1967, highlighting the recent breach of the critical 1.5°C threshold. Top-right panel: Five-year averaged temperature changes from 1850 to 2024, illustrating the accelerating warming trend. Bottom panel: Global map showing surface air temperature anomalies for 2024 relative to the reference period 1991–2020, indicating widespread warming with regional variations. Credit: C3S/ECMWF.

Emissions of both CO_2 and non-CO_2 pollutants change the Earth's radiation balance by altering the equilibrium between incoming solar energy and outgoing thermal radiation, quantified as radiative forcing (RF) [5]. For example, contrail cirrus clouds, as the most prominent non-CO_2 forcing agent and the primary contributor to net climate effects [3], trap outgoing thermal radiation and also reflect a portion of incoming solar radiation (during daylight hours). This radiative imbalance can push the climate system toward a new equilibrium state characterized by changes in temperature (and also precipitation and sea level). In the context of aviation, the net RF is positive, thereby contributing to global warming [3, 4]. In order to meet the set climate goals such as International Civil Aviation Organization's (ICAO) aspiration for net-zero carbon emissions by 2050 [6], along with the Paris Agreement's target to limit global warming to below 1.5°C [1], the aviation sector faces growing pressure to reduce its climate footprint [7, 8].

Multiple mitigation measures, ranging from policy-based and technological solutions to alternative fuels and operational improvements, are currently under consideration or

planned for implementation to address aviation's climate impact. Solutions such as sustainable aviation fuels (SAFs), electric aircraft, and hydrogen propulsion are among the most promising pathways to reduce both CO_2 and non-CO_2 effects [9]. However, these mitigation measures face several challenges that delay their widespread deployment. Electric aircraft are constrained by battery energy density, which limits range and payload, while fast turnarounds require high-power charging and clean-energy (i.e., low-carbon) infrastructure; hybrid propulsion mitigates energy demands but adds complexity, weight, and certification hurdles [10, 11]. Hydrogen propulsion, while offering zero CO_2 emissions, poses challenges related to green hydrogen production, cryogenic storage, and new infrastructure needs, along with concerns over increased contrail formation due to higher water vapor emissions [12, 13]. SAFs, although they can be compatible with current engines, are constrained by limited feedstock availability, high production costs, and the need for supportive policies to scale up supply chains [12, 14, 15]. All in all, new technologies and alternative fuels will require extended periods of research, development, certification, and gradual fleet replacement before full integration into the aviation sector [16, 17]. Thus, while these strategies offer substantial long-term potential, near-term complementary measures are needed to bridge the interim period.

Optimizing flight trajectories has emerged as a promising short-to-medium-term strategy to address non-CO_2 climate impacts [18]. CO_2 remains in the atmosphere for centuries and is well-mixed globally, meaning its radiative forcing largely depends on the cumulative mass released into the atmosphere over time [3]. In contrast, non-CO_2 species, such as nitrogen oxide-induced effects and contrail-cirrus, have significantly shorter atmospheric lifetimes and exhibit regionally and temporally variable effects. For example, the formation of persistent contrails can influence the climate within hours by altering high-altitude cloud coverage, while NO_x emissions modify atmospheric chemistry, affecting ozone and methane concentrations over months to years. As a result, the radiative forcing from non-CO_2 effects is highly sensitive to the time and location of emissions [19]. Consequently, rerouting aircraft to avoid regions where emissions exert large climate effects offers a practical and infrastructure-compatible approach to reducing aviation's climate footprint [20, 21].

In recent years, there has been growing interest in exploring the potential of flight planning to mitigate the climate impact of aviation (refer to [18] for a review of these studies). However, the existing literature has focused on assessing these opportunities through micro-scale flight planning, where flights are optimized individually and independently. This raises a natural question regarding the practicality of this measure and, consequently, the extent to which the anticipated climate mitigation benefits can be realized when applied to large-scale, real-world air traffic scenarios.

When aircraft trajectories are optimized for climate benefits, they are typically rerouted

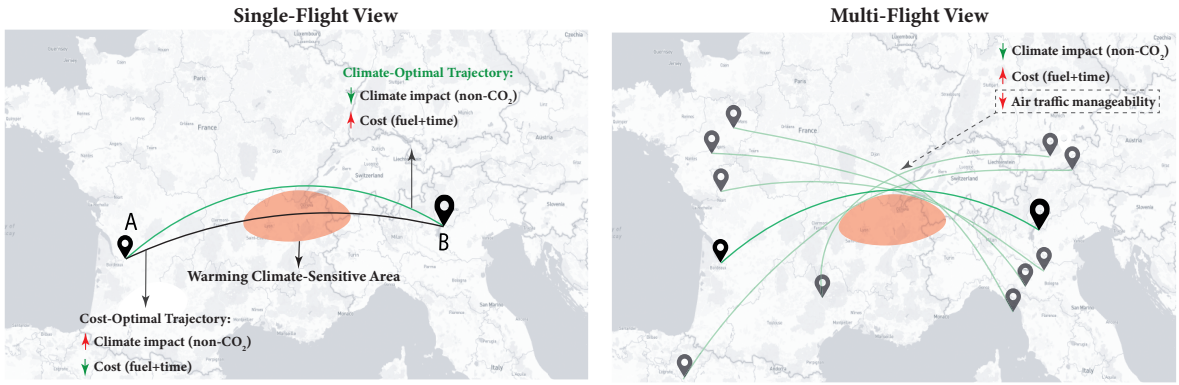


FIGURE 1.2: Rerouting aircraft to avoid regions where emissions have a strong warming effect can reduce the overall climate impact, though it often increases operational costs due to deviations from cost-optimal routes. Furthermore, extending this strategy from individual flights to multiple flights (as shown in the right panel) may result in localized congestion and operational challenges, thereby raising concerns about its operational feasibility.

to avoid airspace regions associated with warming effects (see Figure 1.2) and, when intended, planned to fly through areas favorable to cooling contrail formation. This climate-aware routing strategy can lead to a considerable shift in traffic distribution. Specifically, sectors associated with warming impacts tend to experience reduced traffic, while adjacent sectors or those linked to cooling effects may see increased congestion. Such redistribution can compromise the manageability of air traffic, particularly in high-density regions. This deterioration in manageability can be quantified using several indicators, such as traffic complexity (i.e., the level of difficulty in managing traffic), air traffic controller (ATC) workload, the balance between capacity and demand, and conflict likelihood (i.e., potential loss of separation). As manageability declines, the system's ability to maintain safe and efficient operational performance degrades, increasing the likelihood of safety-critical events.

Considering the projected growth of the aviation sector, any degradation in air traffic manageability resulting from the implementation of climate-optimal trajectories may be deemed unacceptable from the operational perspective. Therefore, ensuring the practicality of climatically optimized flight plans is paramount. Moreover, significant increases in fuel consumption, flight time, or associated operating costs resulting from deviations from standard routes can undermine stakeholder support for climate mitigation strategies. Therefore, a truly sustainable solution for aviation must simultaneously address climate impact mitigation, traffic manageability, and operational cost-effectiveness, as the intersection of these three factors is essential for incentivizing stakeholders to advance toward more climate-friendly operations (see Figure 1.3).

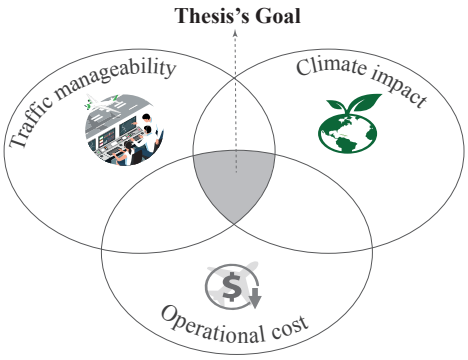


FIGURE 1.3: Key objectives in climate-aware trajectory planning. The intersection of climate mitigation, traffic manageability, and operational cost defines the solution space targeted in this thesis, illustrated by the shaded area.

This thesis is motivated by the need for a more realistic insight into the climate benefits achievable through flight planning, one that extends beyond purely climate considerations to also account for the operational constraints and challenges associated with implementing such a measure in practice. To this end, we first examine the operational implications of incorporating climate impact considerations into flight planning at the network scale and subsequently propose frameworks that enable the planning of climate-optimized flight trajectories while maintaining operational manageability when compared to business-as-usual operations.

1.2 Research gaps

A comprehensive review of the state of the art in climate-optimal flight planning, presented in Chapter 2, reveals that previous studies have primarily focused on optimizing individual trajectories [18, 22]. While these studies provide valuable insights into the potential of flight planning to mitigate aviation climate impact, they fall short of addressing the broader challenges associated with large-scale implementation in real-world air traffic operations. Given the vast number of aircraft operating in the airspace, averaging approximately 29,000 to 35,000 flights per day over Europe, depending on seasonal variations [23], analyzing isolated trajectories is insufficient to evaluate the mitigation potential achievable through the flight planning strategy. Therefore, it is essential to conduct analyses at the network scale that account for the collective behavior of all flights within the system, and to ensure that climate-optimal trajectories remain operationally feasible in practice without compromising air traffic manageability.

To the best of our knowledge, as of the beginning of this research in March 2022, no studies had addressed this challenge by developing climate-aware flight planning frameworks that explicitly consider the operational manageability of the traffic. Interest in this area, however, has grown in recent years. Concurrent with the work presented in this thesis, several

authors have published studies on related topics [24–26], which are reviewed in detail in Chapter 2. In the following section, we outline specific research gaps that emerge from this broader open problem.

Evaluating the operational manageability of climate-optimal trajectories at the network scale

Considering flight planning as one of the most immediately deployable measures for mitigating aviation-induced climate effects necessitates, first and foremost, a comprehensive evaluation of the impacts that implementing climate-aware trajectories may have on ATM system performance. While such routing strategies promise environmental benefits, their large-scale adoption may adversely affect air traffic manageability. Therefore, assessing their real-world viability must go beyond micro-scale trajectory planning and include a quantifiable analysis of their impact on ATM system performance through relevant indicators, such as air traffic complexity, conflict risk, capacity and demand, and system-wide resilience, to determine the extent to which these climate-optimized flight plans are operationally manageable. Despite the importance of this evaluation, at the outset of this PhD, no studies had evaluated the operational effects of climate-oriented routing strategies at the network scale, thus identified as an open problem.

Mitigating operational implications arising from the adoption of climate-optimal trajectories

Following the evaluation of the network-level impacts of independently optimized trajectories, it becomes necessary to mitigate any adverse effects these trajectories may impose on the ATM system performance. To deem such trajectories operationally feasible, their manageability must be at least preserved at levels comparable to those of business-as-usual operations. This necessitates adjustments to the initially optimized flight plans whenever detrimental consequences, such as increased traffic complexity, elevated conflict risk, or capacity-demand imbalances, are identified. In ATM studies, this strategy is commonly referred to as a resolution problem.

Addressing this problem is inherently complex due to its high dimensionality, the large number of aircraft involved, and the nonlinear nature of their dynamics [27]. Furthermore, the inclusion of operational objectives (e.g., conflict resolution, traffic complexity mitigation) introduces interdependencies among flights, reflecting the multi-agent structure of the ATM system. These inter-agent couplings transform the problem into a large-scale, interconnected optimization, where no single flight can be optimized in isolation. To address this problem, a sophisticated framework is needed that can handle high-dimensional, multi-agent interactions and ensure the manageability of air traffic in realistic, large-scale scenarios. Yet, such

a resolution framework is currently absent from the literature, highlighting an open gap in existing research.

Developing a scalable framework to support large-scale study of climate-optimal flight planning toward guiding policy actions

To incentivize stakeholders and policymakers to take the necessary steps toward implementing climate-optimized flight planning, a systematic evaluation of both its climate benefits and its operational feasibility is required. Such insight can only be gained through large-scale analysis under real traffic scenarios. Large-scale scenarios refer to analyses conducted over a wide range of weather patterns (e.g., an entire year) to capture the strong dependence of non-CO_2 climate effects on meteorological variables across different potential weather conditions, which can largely affect both the achievable climate benefits and the resulting traffic distribution.

Addressing this need requires a scalable framework to enable more realistic and operationally informed analysis of climate-optimized flight planning. The approach discussed in Scientific Gap 2 addresses this problem by sequentially optimizing individual trajectories, assessing their manageability, and applying adjustments to restore stability in the ATM system. While conceptually effective and suitable for planning daily operations, this multi-step approach may not be computationally efficient for large-scale analysis (e.g., year-long flight planning scenarios). In particular, the need to repeatedly perform micro-scale optimizations, followed by ATM performance evaluation and trajectory adjustment, results in considerable computational and data-handling burdens. To overcome these limitations, a unified, fast-time framework is required, one that can simultaneously generate operationally manageable climate-optimized trajectories in a single step. Such a framework would eliminate the inefficiencies of the sequential pipeline and enable real-time or near-real-time analysis of large-scale traffic scenarios. Despite its importance, no such integrated approach currently exists. Bridging this gap is essential to support large-scale analyses required for developing currently missing incentivizing indicators toward policy actions.

1.3 Objectives

Building on the identified scientific gaps, the main objective of this thesis is to propose frameworks for network-level climate-optimized flight planning that explicitly account for air traffic manageability. To achieve this, the research pursues the following specific objectives:

- **Goal 1:** Quantify the operational implications of adopting individually optimized climate-aware trajectories on air traffic manageability by evaluating key performance metrics, including air traffic complexity and potential conflicts.

- **Goal 2:** Develop resolution strategies to restore air traffic manageability in scenarios where climate-optimized trajectories compromise it, thus ensuring operational feasibility while preserving climate benefits.
- **Goal 3:** Develop an integrated and scalable framework that jointly addresses climate objectives and operational manageability, enabling fast-time, large-scale flight planning to generate realistic, policy-relevant insights into the climate mitigation potential of climate-aware operations.

This set of research goals follows a logical progression from micro-scale analysis to ATM system-wide application. First, the thesis aims to examine how individually optimized trajectories affect the performance of the ATM system, establishing the need for more holistic evaluation beyond climate goals alone. Second, to address any operational issues that arise, we propose resolution frameworks capable of adjusting these flight plans while preserving their environmental benefits. Finally, we move toward developing an integrated and scalable optimization framework that jointly considers environmental and operational objectives, suitable for large-scale, policy-relevant applications. This layered approach reflects the incremental research required to bridge the gap between micro-scale flight planning and its practical implementation in the real-world ATM system.

1.4 Contributions

In line with the defined objectives, this thesis makes several contributions to the field of climate-aware flight planning. These contributions address the aforementioned research gaps, including the lack of network-level evaluation of climate-optimized trajectories, the absence of resolution strategies to ensure operational manageability, and the need for scalable frameworks capable of supporting large-scale analysis. The main contributions are, therefore, as follows:

1. **Quantifying the impact of climate-optimized trajectories on air traffic manageability (Goal 1)**

   This thesis investigates how adopting individually optimized climate-aware trajectories influences the manageability of air traffic. Manageability is analyzed through a set of ATM-relevant performance indicators, including conflict count and complexity score. The analyses encompass a range of scenarios, including different airspace structures (free-routing and structured), spatial domains, and levels of flight planning flexibility (lateral only (2D) and full trajectories (3D)). This evaluation provides insight into how (micro-scale) climate-optimized flight planning may affect air traffic manageability and highlights the potential operational challenges that could arise when such trajectories are implemented across real traffic scenarios.

The findings related to this contribution have been documented in several publications. Analyses of potential conflicts and traffic complexity in structured airspace were presented in [28]. Evaluations of conflicts under free-routing conditions were published in [29] and [30] for 2D and 3D optimized trajectories, respectively.

2. **Development of a conflict resolution strategy based on a heuristic approach (Goal 2)**

   To address the operational challenges posed by adopting individually optimized trajectories in a climate-friendly manner, as the first attempt, this thesis introduces a heuristic conflict resolution strategy. The proposed algorithm aims to reduce the number of potential conflicts while preserving the environmental benefits of the optimized trajectories. This method is applied and evaluated in a case study over Spanish airspace (see the related publication in [29]).

3. **Development of a distributed multi-agent reinforcement learning framework for large-scale conflict resolution (Goal 2)**

   To overcome the scalability limitations of heuristic-based approaches, this thesis proposes a novel cooperative framework based on multi-agent deep reinforcement learning (MARL) to resolve conflicts arising from the adoption of climate-optimal trajectories during the flight planning phase. This contribution was published in [30]. The proposed strategy, referred to as the policy-sharing multi-agent twin-delayed deep deterministic policy gradient (Ps-MATD3) algorithm, is designed to handle the high-dimensional, continuous action space characteristic of fully free-routing airspace. The framework follows a centralized training and decentralized execution scheme. In this setup, the policy is trained in a centralized manner, but during execution, each aircraft applies the policy independently. To ensure scalability in large-scale scenarios involving varying numbers of agents, shared policy parameters are implemented, providing flexibility to accommodate diverse air traffic conditions and ensuring the framework can adapt to dynamic operational environments. Such computational enhancements allow the case study to be extended to European airspace.

4. **Development of multi-agent reinforcement learning framework for complexity management (Goal 2)**

   Extending the previous contributions focused on optimizing flights in free-routing airspace and on conflict resolution, this contribution further enhances the realism and applicability of the proposed developments by extending the analysis to structured airspace and incorporating air traffic complexity as a more suitable indicator of traffic manageability during the flight planning phase. To address the increased complexity resulting from the adoption of individually optimized trajectories, we propose a cooperative framework based on multi-agent proximal policy optimization (MAPPO),

well-suited for discrete action spaces in structured airspace. This approach derives an optimal policy that provides aircraft-specific altitude and speed modifications to mitigate air traffic complexity. Similar to the Ps-MATD3 algorithm, this approach follows the centralized training and decentralized execution paradigm and leverages parameter sharing to ensure generalizability and scalability to diverse air traffic scenarios. The performance of the developed methodology is evaluated using a real case study in European airspace within the structured airspace. This contribution was documented in a submitted manuscript currently under revision [31].

5. **Development of a scalable single-step framework for network-scale climate-optimized flight planning using constrained multi-agent reinforcement learning (Goal 3)**

   Contributions 2, 3, and 4 adopt a multi-step approach for network-level climate-optimal routing: first performing micro-scale trajectory optimization, then integrating those trajectories into the network traffic, and finally resolving any adverse effects that may arise from their adoption. Although the developed MARL-based conflict-resolution and complexity-management algorithm scales well, the prerequisite of micro-scale flight planning imposes computational constraints for large-scale studies required to develop policy-relevant indicators for stakeholders. The final contribution of this thesis is, therefore, to introduce a fast-time, scalable framework built on constrained multi-agent reinforcement learning, which enables single-step network-level flight planning by treating climate-sensitive area avoidance as constraints to be satisfied, and operational manageability as an objective to be optimized. The proposed method employs the MAPPO algorithm and adapts it to handle constraints related to climate hotspot avoidance using the Lagrangian technique. The framework enables direct optimization over business-as-usual trajectories to avoid climate-sensitive regions while mitigating air traffic complexity. By eliminating the need for individual trajectory optimization, the framework significantly reduces computational burden and enables large-scale analyses. The preliminary development of this approach was presented at the SESAR Innovation Days conference [32], and a journal article is currently under preparation.

1.5 Outline of the Thesis

This thesis is structured as follows:

- **Chapter 1**, the current chapter, presents the motivation for this research, outlines the identified research gaps, and defines the main objectives and contributions of the thesis.

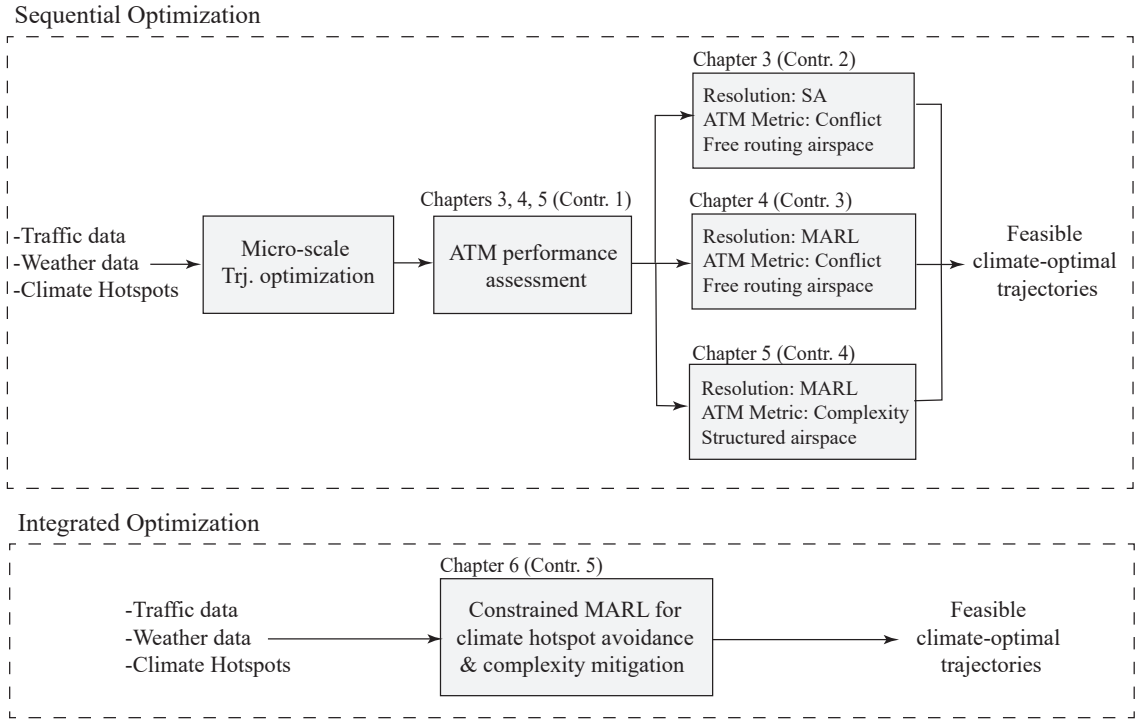


FIGURE 1.4: Structure of the chapters in this dissertation and their correlation with research contributions.

- **Chapter 2** reviews the state-of-the-art in climate-optimized trajectory planning, highlights current limitations, and frames the specific open problems that this thesis aims to address.
- **Chapter 3** investigates climate-aware flight planning within a free-routing airspace context. We utilize the number of conflicts as an indicator of traffic manageability and propose a heuristic-based resolution strategy to address conflicts arising from the adoption of independently optimized trajectories.
- **Chapter 4** enhances the framework from Chapter 3 by introducing a distributed MARL-based approach for conflict resolution. This method is designed to improve scalability and performance in high-density traffic scenarios, overcoming the limitations of heuristic-based strategies.
- **Chapter 5** extends the analysis to structured airspace and considers air traffic complexity as the performance indicator. To mitigate the potential complexity increase resulting from the adoption of climate-optimal trajectories, a scalable resolution framework based on MARL is proposed.

- **Chapter 6** introduces an integrated framework for climate-optimized flight planning at the network level. This framework enables large-scale scenario analysis without the need for individual trajectory optimization.
- **Chapter 7** summarizes the key findings of the thesis, discusses their broader implications, and outlines potential directions for future research.

Chapter 2

State of the art

The problem of climate-optimized flight planning at the network scale involves several components: models predicting the spatiotemporal climate sensitivity to aircraft emissions, the problem formulation and solution approaches for micro-scale trajectory optimization, performance indicators for evaluating operational manageability, and resolution strategies to address the operational challenges introduced by climate-optimized trajectories. This chapter provides an overview of these components and reviews the most recent studies in each area, ultimately identifying the existing gaps that this thesis aims to address.

2.1 Climate optimal flight planning

Flight planning is a fundamental operational process in aviation that defines an aircraft's intended trajectory, including its lateral path, altitude profile, and speed schedule. Conventionally, flight trajectories are planned mainly to minimize operational costs (e.g., considering fuel consumption and/or flight time) while satisfying safety constraints, airspace structure, airline preferences, and air traffic control requirements. However, in response to growing concerns about global warming, the aviation sector's contribution to climate change, and the projected increase in air traffic demand, there is a pressing need to incorporate environmental considerations, particularly climate impact, into flight planning [3].

While CO_2 emissions have long been recognized as the primary contributor to aviation-induced climate change, they represent only a fraction of the sector's total climate impact. According to a recent estimate by Lee et al. [3], non-CO_2 effects, primarily contrail cirrus and NO_x-induced ozone formation, account for approximately two-thirds of aviation's net effective radiative forcing. Unlike CO_2, which has a long-term warming effect, non-CO_2 impacts are highly sensitive to the atmospheric conditions at the time and location of emissions. Consequently, rerouting flights to avoid areas where emissions lead to high climate impact, often referred to as climate-sensitive regions, climate hotspots, or ECHO areas, offers a practical and infrastructure-compatible approach to climate impact mitigation [20, 21].

2.1.1 Climate impact estimation models

Mitigating the climate impacts of non-CO_2 forcing agents through flight planning requires, as a first step, accurate information on the location and intensity of climate sensitive areas. However, not all climate impact modeling approaches used in the literature for flight planning provide information on the spatiotemporal dependency of non-CO_2 climate effects, which is necessary to achieve the full potential of climate-optimized flight planning.

In early efforts to mitigate the climate effects of aviation through flight planning, researchers explored a range of strategies, including reducing emissions [33], avoiding persistent contrail formation areas [34, 35], reducing radiative forcing (RF) [36], and reducing global warming potential (GWP) [37, 38]. For the climate effects of different species, except for contrails, the typical approach involves first estimating emissions, commonly using methods such as the Boeing fuel flow method 2 (BFFM2), to compute emission indices for the most relevant species, particularly NO_x [39,40]. These indices, along with fuel consumption, are then used to calculate the corresponding RF, from which different climate metrics (e.g., GWP) are calculated. Each of these steps, such as emissions estimation, RF calculation, and climate metric evaluation, was incorporated in early studies to represent climate effects in the objective function of the flight planning problem.

For contrail formation, which cannot be simply attributed to a single emission quantity, assessments have typically been based on the distance traveled through regions conducive to the formation of persistent contrails. Therefore, instead of emissions, persistent contrail-forming areas are considered for calculating the corresponding RF or climate impact using different metrics. In some cases, in addition to the ice-supersaturated regions (ISSR), the Schmidt-Appleman criterion (SAC), which states that contrails form at sufficiently low temperatures and sufficiently moist (relative to liquid water) environments, was adopted [41–43]. In the literature, the consideration of both ISSR and SAC is referred to as persistent contrail formation areas (PCFA) [44]. For a detailed classification of these approaches, interested readers are referred to [18].

The aforementioned approaches to incorporating climate impact into flight planning, whether through emissions estimation, radiative forcing calculations, or derived climate metrics, are represented by fuel consumption in the end. In this type of modeling framework, when applied to flight planning, the optimization process typically aims to minimize fuel consumption, often favoring higher cruise altitudes within vertical constraints. Nevertheless, due to high weather and spatial dependencies of non-CO_2 climate effects, there is a risk of increasing these impacts and potentially offsetting the intended benefits of reduced fuel burn. [19,45]. In addition, in the case of avoiding ice-supersaturated areas, not all contrails have a warming impact (or significant warming impact) that needs to be avoided. Therefore, even if such modeling approaches result in a mitigation potential, the full potential of flight planning in the mitigation of climate effects is not realized.

To more reliably account for climate effects in aircraft trajectory planning, models that capture the spatiotemporal dependencies of non-CO_2 impacts are essential. Currently, only two state-of-the-art modeling approaches address this need:

- **Algorithmic Climate Change Functions (aCCFs)**: The first attempt to develop a model capable of providing spatiotemporal information on climate sensitivity to aircraft emissions was made under the EU-project REACT4C[1], which led to the development of climate change functions (CCFs). These functions estimate the impact of aviation emissions in terms of average temperature response (ATR) per flown distance (in terms of contrails) and per emitted mass (in terms of CO_2, NO_x, and H_2O), represented as five-dimensional datasets (longitude, latitude, altitude, time, and emission type) [46–48]. ATR quantifies the average change in global temperature over a defined time horizon (typically 20, 50, and 100 years). Due to their computational demands, original CCFs were not suitable for real-time applications. To address this, algorithmic climate change functions (aCCFs) were introduced, offering real-time evaluation based on simplified mathematical formulas using meteorological inputs [19]. These aCCFs are well-suited for trajectory optimization due to their computational efficiency [49]. An enhanced and harmonized version (in terms of climate metric), aCCF-V1.0a, has recently been introduced [50–52], which quantifies the effects of NO_x, water vapor, and persistent contrails. An open-source Python implementation is available via the CLIMaCCF library[2].

- **Contrail Cirrus Prediction Model (CoCiP)**: CoCiP quantifies the climate impact of contrails by simulating their properties (e.g., optical depth), evolution, and resulting energy forcing (EF), based on meteorological variables such as wind, temperature, and (relative or specific) humidity [53]. EF is computed by integrating radiative forcing per unit length and width of the contrail over its lifetime, representing the net energy added to or removed from the atmosphere. A recent open-source implementation is provided via the pycontrails library[3]. While the original Lagrangian (trajectory-based) version is not compatible with standard flight planning tools, a domain-filling (gridded) version has been proposed to support integration into such applications [54].

In addition to CoCiP and aCCFs, other models have been developed for estimating contrail climate effects. The two most well-known models are the Ames contrail simulation model (ACSM) [55] and the aircraft plume chemistry emission and microphysics model (APCEMM) [56]. However, due to extensive computational requirements, these models, under their current implementation frameworks, are not feasible for fast function evaluation

[1] `http://www.react4c.eu`
[2] CLIMaCCF library: `https://doi.org/10.5281/zenodo.6977272`
[3] pycontrails library: `https://doi.org/10.5281/zenodo.7877538`

required in climate-optimized flight planning. Nonetheless, ongoing efforts aim to adapt them to produce outputs that are computationally efficient and suitable for integration into flight planning frameworks.

With climate impact estimation models available, there is a need to develop flight planning methodologies that integrate this information in order to plan climate-optimized routes. Climate-optimal aircraft trajectory optimization problems can be generally categorized into two levels of analysis: micro-scale, which focuses on individual trajectory optimization, and macro-scale, which addresses network-level or multi-aircraft optimization. The following sections review the state-of-the-art literature in both domains.

2.1.2 Micro-scale trajectory optimization

Micro-scale climate-optimal trajectory optimization focuses on planning flight trajectories for individual aircraft with the primary objective of minimizing their climate impact, while also accounting for operational cost factors such as flight time and fuel consumption. In this approach, each flight is treated as an isolated entity, and its trajectory is optimized independently of surrounding traffic and the operational constraints of the ATM system.

Numerous studies have investigated the potential of micro-scale trajectory optimization for mitigating climate effects. A comprehensive review of this body of work is provided in [18], which surveys research conducted between 2000 and 2022. Table 2.1 summarizes the most recent contributions in this domain (including post-2022), highlighting trends and research directions within the field. One trend is the increasing use of models that provide spatiotemporal information on climate sensitivity to aircraft emissions, most commonly aCCFs (e.g., see [22, 57–60]). In parallel, recent efforts have moved toward more realistic operational scenarios by incorporating structured airspace constraints, as opposed to relying on fully free-routing conditions [22]. Indeed, optimization within structured airspace is more complex due to the hybrid nature of the decision space, which involves both discrete and continuous variables [18, 61].

The other recent direction in the field is the incorporation of uncertainty into the trajectory optimization problem in order to plan robust climate-optimized flight plans. This is particularly important because the climate impact of non-CO_2 forcing agents is highly complex to model and sensitive to many sources [3]. One of the main sources introducing direct uncertainty to climate impact estimates is the meteorological conditions. For flight planning, we rely on weather forecasts, which are inevitably uncertain, especially for variables required to estimate climate effects, such as humidity fields and radiation parameters [62, 63]. If the weather data used in flight planning fails to accurately capture future atmospheric states, the resulting trajectories may not only fail to mitigate climate impacts but also incur higher operational costs due to deviations from business-as-usual routes, leading to increased fuel consumption, longer flight times, and elevated CO_2 emissions, ultimately

exacerbating the climate impact rather than mitigating it. Therefore, improving the reliability and confidence of climate-optimized flight planning by accounting for uncertainty has become an increasingly important research focus in recent years.

TABLE 2.1: Overview of recent studies on climate-aware flight planning to mitigate aviation-induced climate effects. ISSR: ice-supersaturated regions; SAC: Schmidt-Appleman criterion; FFRA: fully free-routing airspace; GWP: Global Warming Potential; CG: Column Generation.

Study	*Forcing agents*	*Model*	*Routing*	*Uncertainty*	*Opt. scale*	*ATM KPI*	*Resolution*
Yamashita et al. (2020) [57]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale	–	–
Lührs et al. (2021) [64]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale	–	–
Yamashita et al. (2021) [58]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale	–	–
Yin et al. (2022) [50]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale	–	–
Castino et al. (2024) [65]	CO_2 and non-CO_2	aCCFs	FFRA	–	Micro-scale	–	–
Sausen et al. (2023) [66]	Contrails	ISSR	Structured	–	Micro-scale	–	–
Frias et al. (2024) [21]	Contrails	CoCiP	Structured	–	Micro-scale	–	–
Simorgh et al. (2022) [22]	CO_2 and non-CO_2	aCCFs	Structured	MET	Micro-scale	–	–
Simorgh et al. (2024a) [60]	CO_2 and non-CO_2	aCCFs	FFRA	MET	Micro-scale	–	–
Simorgh et al. (2024b) [20]	CO_2 and non-CO_2	aCCFs	Structured	MET	Micro-scale	–	–
Roosenbrand et al. (2023) [25]	Contrails	SAC, ISSR	Structured	–	Network-scale	Conflict risk	–
Zengerling et al. (2024) [24]	CO_2 and non-CO_2	aCCFs	Structured	–	Network-scale	Demand	–
Demouge et al. (2024) [26]	CO_2 and Contrails	GWP, aCCFs	FFRA	–	Network-scale	Demand	CG

To date, only two state-of-the-art methods have been proposed for robust flight planning under meteorological uncertainty, presented in [60] and [22]. These methods address climate-optimized trajectory planning in free routing and structured airspace, respectively. Meteorological uncertainty in these works is characterized using ensemble weather forecasts derived from the ensemble prediction system (EPS), a state-of-the-art numerical weather prediction framework that provides probabilistic forecasts (typically comprising 50 ensemble members each representing a probable realization of the atmospheric state) by running multiple simulations with slightly perturbed initial conditions and model parameters [67]. Both studies have released open-source tools implementing their proposed algorithms. The

tool for free-routing airspace is called ROC[4] (robust optimal control for flight planning), and the tool for structured airspace is called ROOST[5] (robust optimization of structured trajectories). These tools represent the most advanced frameworks currently available for robust climate-optimized flight planning under meteorological uncertainty.

Another direction of the research concerns the scale and representativeness of the traffic scenario under exploration. While many studies focus on the optimization of a single flight, only a few have considered broader traffic samples or representative subsets of flights to more reliably assess the achievable climate mitigation potential [20, 21, 64, 66]. For instance, the authors in [64] presented a large-scale study involving approximately 13,000 intra-European flights on a single day, optimizing each trajectory to reduce both CO_2 and non-CO_2 climate effects. A more recent study in [20], conducted a year-long analysis applying climate-aware trajectory optimization to 150 representative European flights per day within structured airspace. This study examined multiple aspects, including monthly and seasonal variability of climate effects and potential climate benefits, the proportion of flights requiring rerouting, changes in flight profiles, and the development of indicators for identifying "big-hit" scenarios, cases where substantial climate benefits can be achieved.

In a related study, Frias et al. [21] conducted a large-scale simulation to evaluate the feasibility of contrail avoidance within a commercial flight planning system. The study employed a forecast-based contrail avoidance model that integrated the CoCiP model. Unlike the conventional approach, which considers hotspot avoidance as a soft constraint in the objective function, this study implemented contrail avoidance as a hard constraint within the trajectory planning process. In addition, the work in [66] conducted an operational trial of contrail avoidance. The trial, conducted in structured airspace and involving over 200 flights, used satellite observations to confirm the outcomes under controlled conditions.

Nevertheless, the analyses performed in all the aforementioned studies, although extended to a larger number of flights, remain micro-scale in nature, as each trajectory was optimized independently to mitigate its climate impact, and the results were then aggregated [20, 64, 66]. This narrow scope overlooks interactions between flights and the resulting implications on air traffic manageability, thereby raising questions about the reliability of the reported climate benefit [28].

Climate-optimal trajectories are strategically planned to avoid regions where aircraft emissions have significant warming effects and, where possible and intended, target areas with the potential to generate cooling effects [60]. However, this shift in traffic flow increases congestion within climate-sensitive regions that are favorable for cooling effects while evacuating warming climate-sensitive areas. Such redistribution of traffic density can increase traffic congestion in particular areas and trigger new operational challenges. These

[4] https://github.com/Aircraft-Operations-Lab/roc
[5] https://github.com/Aircraft-Operations-Lab/roost

emergent effects may degrade the performance of the ATM system, ultimately raising concerns about the feasibility of implementing such a routing strategy. Therefore, to reliably understand the climate impact mitigation potential achievable through flight planning, it is essential to conduct analyses at the network scale, considering the collective behavior of all flights within the system.

2.1.3 Network-scale trajectory optimization

Network-scale (macro-level) climate-optimal trajectory optimization refers to the coordinated planning of multiple aircraft trajectories within a shared airspace, with the dual objective of minimizing environmental impact and ensuring the safety and efficiency of air traffic operations. Unlike micro-scale approaches, which optimize each flight independently, network-scale flight planning explicitly considers the interdependencies among flights and their collective impact on the ATM system performance. Such a network-level perspective is essential, as the air transportation system operates within a complex socio-technical and regulatory framework. It functions as a highly interconnected network whose overall behavior cannot be reliably predicted by analyzing individual components in isolation but rather from the dynamic interactions among them. Thus, the evolution toward climatically oriented air traffic requires a network-scale framework that can jointly address both environmental and operational objectives.

The current state of the literature includes limited studies that have partially or fully investigated climate-aware flight planning at the network scale (see Table 2.1 for an overview of these studies). Notably, all of these efforts have been conducted in parallel with this thesis, reflecting a broader and timely shift in the research community toward integrating environmental considerations into the operational framework.

Among these efforts, two studies have partially addressed the problem by evaluating the impact of optimized trajectories on ATM system performance. Roosenbrand et al. [25] proposed a methodology to minimize climate impact through tactical altitude adjustments and subsequently evaluated the resulting increase in potential conflicts. The study focused exclusively on contrail formation, identifying contrail-inducing flights and determining the altitude changes required to avoid ice-supersaturated regions. While this work provides insights into tactical contrail avoidance, it remains limited to performance assessment and does not incorporate a resolution framework to address the operational consequences of trajectory modifications to ensure the operational feasibility of the resulting trajectories.

Similarly, the study in [24] examined the potential for mitigating aviation-induced climate effects through adjustments to individual flight trajectories and subsequently evaluated their performance at the network level. Focusing on high-impact flights, the authors applied three modification strategies, vertical, temporal, and integrated, to reduce the climate impact of both CO_2 and non-CO_2 forcing agents. A network-level assessment was then

conducted using sector entry counts and nominal capacity thresholds to quantify the impact of the proposed adjustments on the performance of the ATM system. However, similar to the work by Roosenbrand et al. [25], this study was limited to performance assessment and did not include an optimization framework to address potential demand–capacity imbalances resulting from the trajectory modifications.

All in all, although network-level performance evaluation is an important step, it must be complemented by resolution strategies to mitigate the potential degradation in system performance resulting from the adoption of individually optimized trajectories. This includes addressing adverse effects such as increased conflicts, elevated traffic complexity, higher controller workload, and demand-capacity imbalances to ensure the operational feasibility of the climate-optimized flight planning.

Approaches to fully address climate-optimized flight planning at the network scale can be categorized into two classes: sequential optimization and integrated optimization, as illustrated in Figure 2.1.

2.1.3.1 Sequential optimization

Within the framework of sequential optimization, the overall process is decomposed into multiple stages. First, all individual flight trajectories are independently optimized with respect to climate impact using micro-scale aircraft trajectory optimization techniques. These climate-optimal trajectories are then integrated into the network traffic, where their collective effects on ATM system performance are evaluated using various indicators (presented in Section 2.2). To address any adverse operational consequences resulting from the adoption of individually optimized trajectories, an additional optimization stage is introduced. This step aims to ensure the operational feasibility of climate-aware trajectories by modifying them in a manner that preserves their environmental benefits while enhancing traffic manageability. In this thesis, this adjustment process is referred to as a *resolution strategy*, which is described in detail in Section 3.2.2.

To date, the only study that has applied this framework for network-scale flight planning is the work presented in [26], conducted in parallel with this thesis. The study aims to minimize the climate impact of aircraft trajectories, specifically contrail formation, while respecting sector capacity constraints. The model operates in a free-routing airspace environment and employs a column generation algorithm to determine feasible, climate-optimal paths for a set of flights. Each flight selects one trajectory from a set of candidates generated through a cost-minimizing subproblem, while the master problem ensures that no sector exceeds its predefined capacity. Climate impact is modeled using either the GWP metric or aCCFs, depending on the scenario.

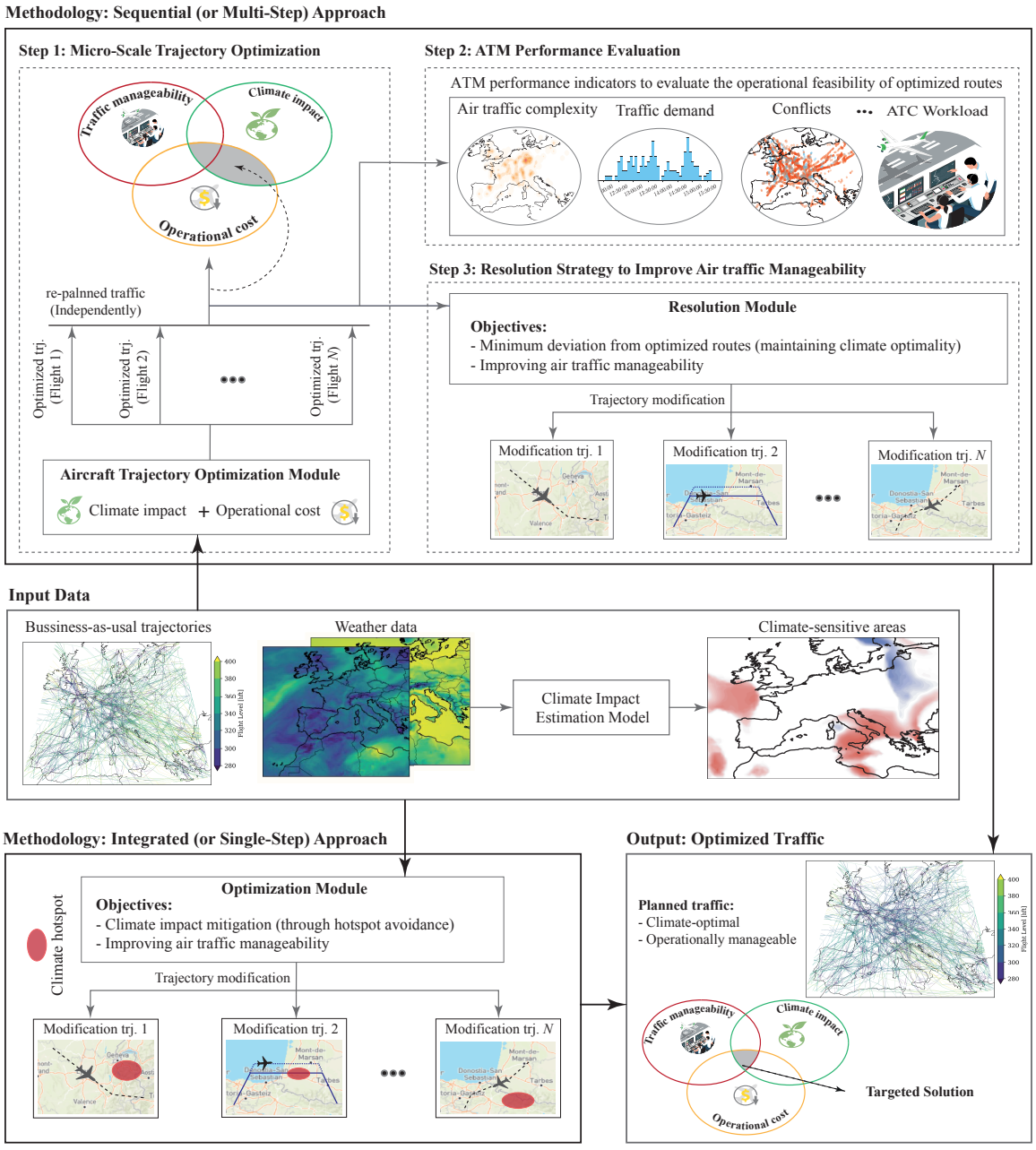


FIGURE 2.1: Frameworks for climate-optimal flight planning at the network-scale.

2.1.3.2 Integrated optimization

The integrated optimization framework optimizes environmental and operational objectives within a unified flight planning process. Instead of optimizing each flight independently and subsequently applying performance assessment and resolution strategy, all trajectories are planned dynamically in a single step, with the problem formulated either as a multi-objective or a constrained optimization problem. In this approach, climate impact metrics and ATM performance indicators are embedded within an integrated framework to directly generate trajectories that are both environmentally friendly and operationally feasible.

To the best of our knowledge, no prior studies have proposed or implemented such an integrated framework for network-scale climate-optimal trajectory planning.

2.2 Air traffic performance assessment

To determine the operational feasibility of climate-optimized trajectories, it is necessary to evaluate their collective performance when implemented across the entire air traffic network. This involves assessing the extent to which such trajectories impact the ATM system's ability to manage traffic effectively using key performance indicators. In the literature, although not related to climate impact consideration, a wide range of indicators has been proposed to evaluate the ATM system performance. These indicators can be broadly categorized into four main groups: demand and capacity, workload, safety, and complexity indicators. While these categories are widely recognized, they are not strictly independent; they may overlap or exhibit interdependencies. The following subsections provide a detailed examination of each category.

2.2.1 Capacity and demand indicators

This category quantifies the balance between the available resources and the demand placed on these resources by aircraft operations. They help determine whether the existing ATM infrastructure can accommodate the planned or actual traffic without compromising safety or efficiency [68].

Demand-capacity balance (DCB) in ATM is governed primarily by two components: aircraft trajectories and airspace sectors [69–71]. The airspace is divided into sectors, each defined by specific lateral boundaries and vertical limits (i.e., minimum and maximum flight levels). Sector configurations may be static or dynamically adjusted based on traffic density and operational requirements. Sector capacity refers to the maximum number of aircraft that can be safely accommodated within a sector during a given time interval. It is typically constrained by physical characteristics and operational factors, including airspace geometry, staffing levels, controller workload, coordination complexity, and procedural limitations [72]. In contrast, sector demand is directly determined by the planned trajectories of

aircraft. Demand indicators quantify the number of aircraft expected or scheduled to operate within a sector over a defined time horizon [73]. Several metrics are commonly employed to monitor demand dynamics. Among them, the occupancy count (i.e., the number of aircraft present in a sector at a given moment) and the sector entry rate (i.e., the number of aircraft entering a sector per unit of time) are widely used in both research and operational contexts [71,73].

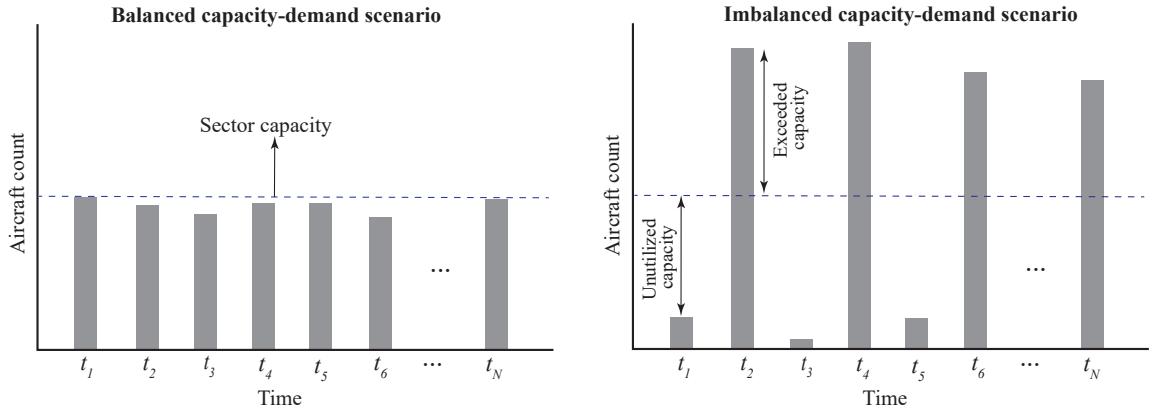


FIGURE 2.2: Illustration of sector capacity and demand for two different scenarios. The left figure shows a uniform distribution of traffic over time, where sector demand remains within capacity limits, ensuring manageable operations. The right figure presents an imbalanced distribution, with demand exceeding sector capacity at several time steps and significant underutilization at others.

In the context of climate-aware flight planning, the focus shifts toward demand-related indicators derived from aircraft trajectories. Climate-optimized routing often deviates aircraft away from regions associated with high warming potential toward regions where emissions have relatively lower or even cooling effects. This selective avoidance and preference mechanism alters standard traffic flows, leading to concentrated demand in specific sectors and generating non-uniform flow patterns. Assuming a fixed sector configuration, demand-related indicators offer an objective means to quantify traffic redistribution resulting from the adoption of climate-optimized trajectories. These indicators are essential for identifying potential bottlenecks or sector overloads. In this context, studies such as [26] and [24] have employed sector entry counts as performance metrics to evaluate the operational feasibility of climate-aware flight planning at the network scale.

2.2.2 Workload indicators

Controller workload is an aspect of ATM system performance, reflecting the cognitive and physical effort required by controllers to maintain safe and efficient operations. While demand can be objectively derived from aircraft trajectories, controller workload is less straightforward to quantify and is often assessed subjectively through task analysis [74]. Numerous studies have evaluated workload by examining specific controller responsibilities such as

communication, coordination, conflict resolution, and sector scanning. These evaluations typically rely on subjective measures (e.g., self-assessment questionnaires), physiological indicators such as heart rate variability [75], or operational metrics like task frequency and average service time [76, 77]. Although these methods offer useful information about controller workload, their dependence on subjective inputs and controller-specific data limits their direct applicability to strategic planning phases, where trajectories are determined well in advance.

2.2.3 Safety indicators

Safety is one of the key performance objectives in aviation, focused on ensuring secure aircraft operations and preventing hazardous events [78, 79]. Over the past five decades, the measurement and modeling of safety have been extensively studied, leading to the development of a wide range of methodologies [80–82]. These approaches can be broadly categorized into four groups: models that analyze the underlying causes of accidents, models that estimate aggregate collision risks, frameworks that assess the likelihood and impact of human error, and approaches that evaluate risks to ground populations, particularly near airports [83–85]. While these models provide valuable insights, they are predominantly retrospective in nature, relying on historical incident data, accident reports, or human-factor evaluations [86,87]. As such, their applicability is limited in strategic planning contexts, particularly during early design phases, where future traffic scenarios must be assessed prior to implementation [88].

Nevertheless, relative safety performance can be evaluated using simulation-based methods and trajectory-driven indicators. One widely used example is the number of potential conflicts (see Figure 2.3), defined as predicted losses of separation based on planned or simulated trajectories [89, 90]. Although such indicators do not replace empirical safety validation, they provide insight into the safety implications of new operational concepts, such as climate-aware flight planning, and contribute to more informed, risk-aware decision-making in the strategic planning phase.

2.2.4 Complexity indicators

Complexity is a multifaceted concept that reflects the intrinsic effort required to manage a given traffic situation [91]. The study of complexity in ATM emerged alongside a growing recognition that conventional capacity–demand metrics are insufficient to fully characterize operational challenges in dynamic airspace environments [92]. Importantly, high complexity is not solely a function of traffic volume; rather, it depends on factors such as traffic flow patterns, the spatial configuration of routes, aircraft interactions, and other dynamic factors that influence airspace manageability (see Figure 2.4). Unlike simpler metrics such as traffic count or density, complexity captures qualitative dimensions of traffic scenarios, including

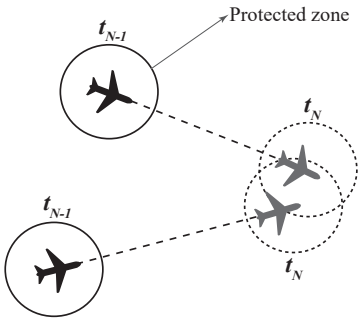


FIGURE 2.3: Potential conflict between two aircraft.

trajectory interactions, maneuvering demands, and spatiotemporal variability [93]. As such, it has been recognized as a comprehensive performance indicator for the ATM system.

Air traffic complexity is closely interconnected with other ATM performance indicators. As a vital determinant of workload, complexity reflects the cognitive demands placed on controllers more effectively than basic traffic volume metrics. For example, a well-structured traffic scenario with a high number of aircraft may be easier to manage than a disordered scenario with fewer flights, as organized and predictable flows reduce cognitive load and coordination effort. From a safety perspective, elevated complexity can be associated with an increased likelihood of loss of separation, particularly under peak workload conditions. Unlike static capacity thresholds, complexity-based evaluations provide dynamic, context-sensitive assessments of sector load, making them especially valuable for proactive traffic flow management.

Although a universally accepted definition of air traffic complexity has not been established, a wide range of metrics has been proposed to capture its various dimensions (see [68] for a comprehensive review). A widely recognized classification scheme groups these metrics into three main categories: structural characteristics, external constraints, and flow characteristics [94].

Structural characteristics refer to the fixed spatial and organizational features of the airspace that influence traffic patterns and controller workload. These include sector geometry and volume, route network design, and crossing points. Although static by nature, these structural factors play a critical role in shaping aircraft interactions and operational complexity. Studies such as [95] and [94] highlight that fragmented or irregular sector geometries, characterized by narrow corridors or multiple intersecting routes, can elevate complexity by increasing coordination demands and reducing maneuverability. To quantify such effects, metrics like the fractal dimension have been introduced to assess geometric complexity within route networks [96].

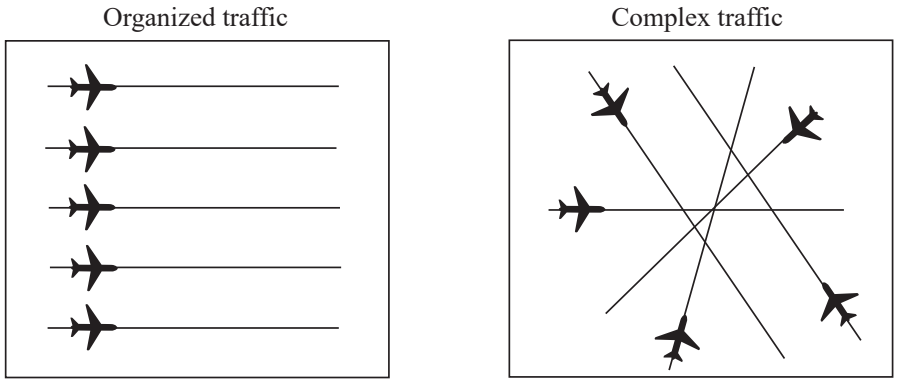


FIGURE 2.4: Illustration of traffic complexity levels. The left figure depicts an organized traffic configuration with parallel and coordinated trajectories, representing a scenario with low traffic complexity. The right figure shows a disorganized scenario with intersecting and multidirectional flows, representative of high-complexity traffic conditions.

External constraints refer to dynamic and often unpredictable factors that limit the usable airspace or restrict traffic flows. These include adverse weather events, temporary activation of military zones, reserved or restricted areas, and airspace closures due to emergencies or events. Such constraints reduce the available maneuvering space and may force rerouting, thereby increasing traffic density and the likelihood of interactions in surrounding sectors. For example, Perera et al. [97] demonstrated how the presence of hazardous weather significantly alters flow distributions and contributes to increased complexity, particularly when aircraft must deviate from nominal routes and cluster in confined corridors. External constraints are generally incorporated indirectly into complexity assessments using adjusted density metrics or scenario-based simulations [94, 98]. Their real-time impacts are critical, particularly when sudden restrictions cause localized surges in complexity, posing significant challenges for operational management.

Flow characteristics refer to trajectory-dependent features that influence the complexity of managing traffic. Unlike static structural factors, flow characteristics directly affect the potential interactions and the cognitive workload imposed on air traffic controllers. They are particularly relevant in trajectory-based operations (TBO), where aircraft are no longer constrained by fixed route networks but instead follow optimized, flexible 4D trajectories.

This dimension of complexity has been extensively studied through metrics such as traffic density [68, 93, 99] and dynamic density [100–102]. Traffic density, typically defined as the number of aircraft within a sector, is one of the most widely used indicators of air traffic complexity due to its simplicity and availability [68]. However, it has been criticized for overlooking key aspects such as traffic direction, spatial organization, and interaction potential. For instance, high-density flows with well-aligned trajectories may be easier to manage than lower-density but disordered traffic (see Figure 2.4). To address these limitations, Masalonis et al. [99] proposed using peak aircraft count relative to acceptable sector

thresholds, while Prandini et al. [93] introduced probabilistic models that assess complexity based on the likelihood of encountering nearby aircraft within spatial buffer zones. Despite its limitations, traffic density remains a foundational metric for complexity assessment.

To better capture spatial-temporal interaction potential, adjusted density has been proposed in several studies [94, 103]. In this approach, the airspace is divided into uniform cells, and the ratio between the cumulative time that multiple aircraft occupy the same cell and the total flight time within that cell is computed. This ratio reflects the duration-based probability of interaction, where a potential interaction is defined as two aircraft sharing the same cell from each other's perspective.

In a similar vein, the complexity score was developed as a composite metric based on three primary indicators that represent key traffic characteristics: vertical, horizontal, and speed-related differences among interacting flows [94]. Since the mere presence of two aircraft in the same airspace volume provides limited insight into the severity or duration of the associated hazards, these indicators aim to capture more nuanced features of disordered and difficult-to-manage traffic situations.

Dynamic density is a composite air traffic complexity metric designed to represent the collective influence of multiple interrelated factors. Dynamic density incorporates both static features, such as sector volume and route structure, and dynamic features that evolve over time, including aircraft count, speed, heading variations, and separation distances [102,104]. These variables are combined through linear models using empirically derived weights [100,101,104], nonlinear formulations [100], or machine learning approaches such as neural networks [105]. The weighting schemes are typically obtained through human-in-the-loop simulations, subjective controller workload ratings, or regression analysis. While this metric integrates some structural variables, it is primarily classified as a flow characteristic because it aims to capture the evolving interaction patterns and real-time control demands associated with dynamic traffic flows.

Several studies have introduced alternative methods for assessing air traffic complexity by developing intrinsic, sector-free metrics that do not rely on predefined sector boundaries or controller interventions [106–108]. For instance, nonlinear dynamical system approaches, including the use of Lyapunov exponents and topological entropy, assess the predictability and divergence of aircraft motion over time, thereby capturing the inherent disorder in traffic flows [106–108]. These models are well-suited for autonomous or decentralized systems, where trajectory flexibility is high and sector-based assumptions are less applicable.

Further advancements include graph-theoretical approaches that model air traffic as interaction networks, representing aircraft as nodes and potential conflicts as edges. Metrics such as network connectivity, clustering coefficients, and centrality have been employed to infer structural tensions and predict conflict propagation within these networks [109].

Although conflict metrics have been commonly classified as safety indicators representing separation risk [81,110–113], they have also served as complexity indicators [68,114,115].

Predicted conflicts can indicate regions with high interaction density, thus supporting the evaluation of traffic manageability. For instance, Roosenbrand et al. [25] employed conflict counts to evaluate the implications of climate-optimized flight paths on air traffic complexity. However, given inherent uncertainties in strategic planning, conflict-based metrics should be cautiously interpreted as indirect indicators of complexity.

2.3 Resolution strategies

Following the evaluation of ATM system performance, any degradation caused by climate-optimized trajectories necessitates the implementation of corrective measures to compensate for adverse effects and restore system manageability. In the literature, a variety of strategies have been developed to improve traffic manageability, including sectorization, dynamic airspace configuration, time-based management, and flight plan modifications [68]. The latter, in particular, focuses on modifying flight plans at the strategic level to enhance the overall manageability of the ATM system.

Since this thesis is grounded in the context of flight planning, the focus is placed on trajectory-based solutions. Specifically, when adverse effects are observed, additional optimization steps are introduced to adjust aircraft trajectories in order to ensure traffic manageability. These adjustments, referred to in this thesis as resolution strategies, are applied during the planning phase, i.e., from one day to a few hours prior to the estimated off-block time, and serve to mitigate the unintended operational impacts of climate-optimized routing on the ATM system performance.

Resolution strategies, which can be formulated as optimization problems, are typically classified along several dimensions, including hierarchical structure (centralized or decentralized decision-making), resolution objectives (e.g., minimizing conflicts, complexity, or controller workload), maneuver types (lateral, vertical, or speed adjustments), and the optimization methods employed [116, 117].

2.3.1 Hierarchical structure

A common criterion for classifying resolution strategies is the level of centralization in the decision-making process [71,116] (see Figure 2.5). In centralized resolution, a single planning system holds authority over trajectory modifications and coordinates adjustments across multiple flights based on a global view of the traffic situation [118]. Centralized methods reduce uncertainty and facilitate globally efficient solutions, but their performance may degrade with increasing traffic density due to the high computational burden and extensive data requirements [117]. Additionally, they rely on the availability of complete and accurate data for all aircraft.

In contrast, distributed[6] resolution strategies delegate decision-making to individual aircraft, allowing them to independently modify their planned trajectories using local or partially shared information [119–121]. These approaches support scalable and adaptive traffic management, particularly in high-density or unpredictable environments where centralized coordination may be infeasible. However, the lack of global coordination can lead to suboptimal system-wide outcomes and increased sensitivity to communication or sensing limitations.

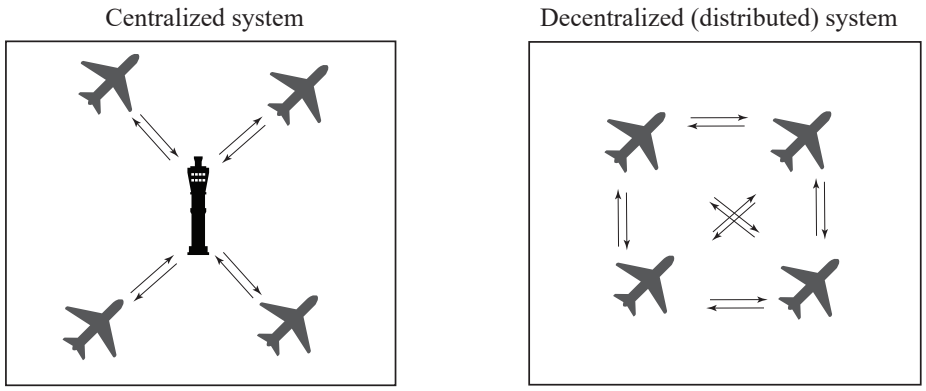


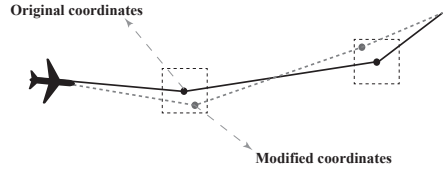
FIGURE 2.5: Difference between a centralized and a distributed resolution framework.

2.3.2 Maneuver type

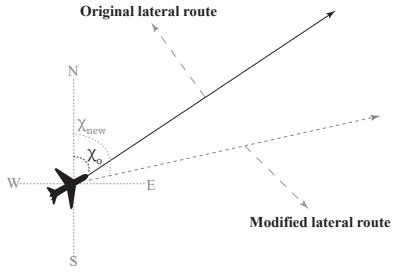
Resolution strategies can also be classified according to the type of maneuver used to achieve the set objectives. The objective should generally target traffic manageability by incorporating ATM performance metrics presented in Section 2.2, either individually or in combination. In parallel, it can also account for climate optimality, for example, by minimizing deviations from climate-optimized trajectories (for the sequential framework) or directly reducing climate impact (for the integrated framework).

In the literature, resolution maneuvers are typically categorized into four types: lateral deviations, vertical maneuvers, speed adjustments, and departure time modifications (see Figures 2.6) [92, 118, 122]. These maneuvers may be implemented either as continuous or discrete adjustments and are typically subject to constraints imposed by aircraft performance limitations and the requirement to maintain smooth and operationally feasible trajectories. These maneuver types can be applied either individually or in combination, depending on the specific resolution objective and operational context. For instance, speed regulation has been commonly used for conflict resolution and separation assurance, offering a subtle and efficient mechanism that avoids large deviations from the planned trajectory [33, 117, 121,

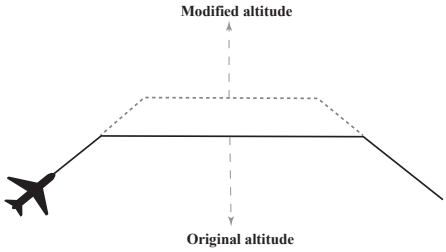
[6]Throughout this thesis, the terms decentralized and distributed are used interchangeably



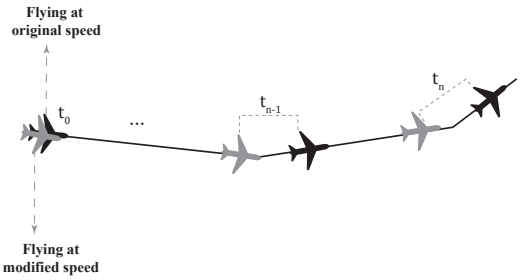
(A) Modifying coordination.



(B) Modifying heading angle.



(C) Altitude adjustment.



(D) Speed regulation.

FIGURE 2.6: Different types of maneuvers used for resolution.

123]. Nonetheless, certain head-on or converging conflicts may not be solvable through speed adjustments alone and could necessitate the use of additional maneuver types.

2.3.3 Optimization method

From the optimization perspective, four primary approaches have been widely employed in the literature to implement resolution strategies: mathematical programming [124], gradient-based methods [116], heuristic and metaheuristic algorithms [118], and learning-based techniques [125]. The first three share conceptual similarities, although they differ in how the optimization problem is formulated and solved. In contrast, learning-based techniques represent a distinct and emerging paradigm. The following subsections provide a more detailed overview of these approaches.

2.3.3.1 Mathematical programming

Mathematical programming formulates the resolution problem using mathematical models with well-defined constraints and objectives. Depending on the problem formulation, various solution techniques can be applied, such as mixed-integer programming [126–128], linear programming [129], or semi-definite programming [130,131]. Mathematical programming has been employed in several studies to address resolution problems [26, 132, 133]. In

[132], the problem was formulated as a mixed-integer linear programming problem, where the maneuvers were considered to be discrete speed and heading adjustments. The proposed solution was implemented within a centralized framework and solved using standard MILP solvers. Similarly, [133] formulated the conflict avoidance problem as a mixed-integer nonlinear programming model, aiming to minimize the magnitude of maneuvering required to ensure safe separation. Their formulation incorporated both discrete altitude changes and continuous velocity adjustments, with geometric and probabilistic constraints derived from relative motion analysis. Like the previous study, their approach was also centralized.

In related work, Demouge et al. [26] formulated the resolution problem as a mathematical programming model for demand–capacity balancing in air traffic management. The problem was defined over a set of precomputed trajectories, with binary decision variables used to assign one trajectory per flight. The resolution strategy was implemented in a centralized framework and solved using a column generation approach, which decomposed the problem into a master problem that selects trajectories while satisfying sector capacity constraints and a pricing subproblem that generates new feasible trajectory options. The maneuver space included both lateral and vertical deviations.

All in all, mathematical programming approaches, widely applied to resolution problems, are capable of providing deterministic (global) optimal solutions, though for small-scale and over-simplified problem instances. Indeed, to ensure tractability, the original problem must often be reformulated to fit the structure required by mathematical programming methods. For example, complex decision spaces are discretized into binary or integer variables, and nonlinear objectives or constraints are approximated or linearized.

2.3.3.2 Gradient-based methods

Gradient-based methods are particularly well suited for solving smooth and continuously differentiable optimization problems [134]. These problems are typically modeled using nonlinear programming (NLP), where both the objective function and constraints are differentiable. Solutions are obtained through iterative algorithms such as interior-point methods (IPOPT) or sequential quadratic programming (SQP), both of which are supported by mature solvers (see [135–137]).

A gradient-based approach was employed in [116] to address the resolution problem under wind forecast uncertainty, aiming to keep the probability of conflict below a predefined safety threshold while minimizing trajectory deviation. The maneuvers were modeled as lateral adjustments, implemented by modifying the geographic coordinates along predefined multi-segment trajectories. The problem was formulated as a nonlinear programming problem and solved in a centralized framework using the SQP algorithm.

When applied to resolution tasks, gradient-based methods can yield accurate solutions in small- to medium-scale settings with well-behaved and convex cost functions [116, 122].

However, their applicability becomes limited in real-world scenarios involving discrete decision variables (e.g., sector assignments or route selections), non-convex landscapes with multiple local minima, or non-differentiable dynamics due to switching behavior or piecewise models. Moreover, gradient-based methods are sensitive to initial conditions and may converge to suboptimal solutions, especially in high-dimensional or tightly constrained search spaces, which poses challenges for complex air traffic management problems.

2.3.3.3 (Meta-)Heuristic methods

Heuristic and metaheuristic methods are widely applied in resolution problems due to their flexibility in handling complex, high-dimensional, or combinatorial optimization problems where exact approaches, such as mathematical programming or gradient-based methods, may be unsuitable or computationally intractable [116, 118]. Rather than guaranteeing globally optimal solutions, these methods aim to find sufficiently optimal solutions (i.e., approximate global solutions) within a reasonable computation time. Common techniques include simulated annealing [118], genetic algorithms [138], and ant colony optimization [139], all of which employ stochastic or nature-inspired search strategies to explore the solution space.

An application of simulated annealing (SA) to strategic conflict resolution under uncertainty has been demonstrated in [118]. In this study, maneuvers were limited to speed adjustments along predefined flight plans, maintaining both lateral route and altitude profiles. The problem was formulated as a discrete optimization task aimed at conflict minimization within a centralized decision-making framework. Similarly, in [140], the SA algorithm was employed for conflict resolution using speed and lateral modification.

Similar to mathematical programming and gradient-based methods, most heuristic techniques are commonly implemented within a centralized decision-making framework, wherein a single agent or authority is responsible for coordinating resolution actions across all flights. While such centralized strategies facilitate globally consistent and highly coordinated solutions, their computational demands increase with problem size. Furthermore, in dynamic or high-density traffic environments, the need for frequent re-optimization in response to evolving traffic conditions or updated information further constrains their real-time applicability. These limitations have prompted growing interest in learning-based approaches, which offer the potential for decentralized, adaptive, and scalable decision-making.

2.3.3.4 Learning-based approaches

Learning-based approaches have emerged as a promising alternative for addressing large-scale, high-dimensional resolution problems [120]. Notably, deep reinforcement learning (DRL) has gained prominence as a powerful framework for sequential decision-making in dynamic and uncertain environments [141]. By integrating the representational capacity of deep neural networks with the trial-and-error learning paradigm of reinforcement learning,

DRL methods are capable of learning policies that map complex traffic states to effective trajectory adjustments [142]. This enables greater flexibility and adaptability compared to conventional optimization techniques, which often require re-solving each problem instance independently.

One of the key advantages of DRL lies in its capacity to generalize from prior experience. Instead of optimizing from scratch for every new scenario, RL-based models are trained offline, typically using simulated or historical data, to derive reusable policies that can be rapidly deployed in real-time operations [143]. This allows for efficient decision-making with minimal computational overhead during execution. Furthermore, DRL frameworks are inherently versatile: they can be implemented in both centralized and decentralized architectures, making them well-suited for a wide range of ATM applications, from ground-based coordination systems to decentralized, aircraft-level autonomy [144].

In recent years, DRL methods have gained growing application across various domains of aviation, particularly in air traffic management. Table 2.2 summarizes the most recent contributions in this domain. For a comprehensive review of these developments, see the survey by Razzaghi et al. [145].

In the specific context of conflict resolution, several DRL-based strategies have been proposed within single-agent frameworks, addressing both en-route and urban airspace operations [146–148]. For instance, Li et al. [147] introduced a DRL-based strategy for collision avoidance in unmanned aerial vehicle (UAV) operations, utilizing the deep Q-network (DQN) algorithm. The proposed framework employed discrete horizontal maneuvers and was tested in a two-dimensional airspace in a decentralized manner, where each agent made decisions independently based on local observations. Using the same algorithm, Mollinga et al. [123] applied DQN to structured en-route airspace. To enhance situational awareness, the study incorporated graph neural networks (GNNs), enabling the encoding of multi-level traffic information. The action space was expanded to include altitude, speed, and heading adjustments, allowing for more flexible conflict resolution.

Further studies on the applications of DRL in conflict resolution have investigated the use of the deep deterministic policy gradient (DDPG) algorithm. For instance, Wen et al. [149] implemented DDPG to adjust aircraft heading angles, ensuring conflict-free trajectories. In a similar vein, Pham et al. [146,150] applied DDPG to train agents capable of recommending heading adjustments for conflict avoidance in free-routing airspace. Extending this approach to structured environments, Ribeiro et al. [117] integrated both heading and speed changes into their resolution strategy using DDPG.

While these studies highlighted the effectiveness of DRL in conflict resolution, they were limited to single-agent frameworks. In single-agent settings, each agent optimizes its actions independently without explicitly considering the behaviors or intentions of other agents. However, real-world air traffic operations are inherently multi-agent in nature, as aircraft interact within shared airspace and make interdependent decisions. This interactivity gives

rise to complex learning dynamics that cannot be adequately captured by isolated, single-agent models. Consequently, there is a growing need to adopt multi-agent reinforcement learning (MARL) techniques to develop more realistic resolution frameworks that reflect the interactions among multiple aircraft.

In a basic implementation, Sui et al. [151] proposed a conflict resolution framework for multi-aircraft scenarios using an independent deep Q-Network (IDQN) algorithm, which combines DQN with the independent learning framework. In this setup, each aircraft functions as an independent agent, making discrete maneuver decisions, such as speed, altitude, and heading changes, to resolve conflicts. While the framework offers computational efficiency, it overlooks the interdependencies among agents, which can lead to suboptimal decisions and unstable behaviors in dynamic, interactive environments. This limitation highlights the challenge of the non-stationarity of the learning environment, where the actions of one agent affect the state transitions and reward functions experienced by others, making the environment inherently non-stationary from each agent's perspective [152]. This violates the Markov property assumed in single-agent reinforcement learning and significantly complicates the convergence of policies, particularly in high-dimensional or tightly coupled systems. Addressing this issue requires more sophisticated learning architectures and the incorporation of explicit coordination mechanisms to ensure stable and effective multi-agent decision-making.

The application of MARL to conflict resolution was initially pioneered by Brittain et al. [120] and subsequently advanced through a series of follow-up studies [121, 153]. These works employed the proximal policy optimization (PPO) algorithm within a multi-agent framework, where conflict resolution was handled through discrete speed adjustments in structured en-route airspace. In later developments, fully connected neural networks were replaced with long short-term memory (LSTM) architectures to enhance the model's ability to capture temporal dependencies in sequential decision-making tasks [121]. Additionally, attention mechanisms were introduced in [153] to allow agents to dynamically focus on relevant neighboring aircraft.

Building on these foundations, Dalmau et al. [154] proposed a PPO-based resolution framework that integrates MARL with message-passing neural networks (MPNNs), facilitating inter-agent communication prior to decision-making. Each agent selects discrete speed and heading adjustments to resolve conflicts, and the system was evaluated in a two-dimensional free-routing airspace. In a complementary effort, Zhao et al. [155] developed a physics-informed DRL framework for conflict resolution, which incorporates domain knowledge into the learning process to enhance both policy interpretability and training efficiency. Their strategy involved a combination of discrete and continuous modifications to heading angle and speed, evaluated in a 2D structured airspace.

More recent contributions have further explored architectural variations. Chen et al.

[156] introduced a general-purpose MARL approach for real-time conflict resolution in free-routing airspace, utilizing a Rainbow DQN algorithm. Maneuvers were represented as discrete combinations of speed and heading adjustments, applied through an adaptive strategy designed to ensure operational feasibility and support trajectory recovery. In a similar line of research, Papadopoulos et al. [157] developed a conflict resolution framework that integrates MARL with a graph convolutional neural network (GCN). In their formulation, each aircraft is modeled as an agent capable of selecting from a discrete set of maneuvers, including lateral deviations, speed changes, and altitude adjustments. Their approach was evaluated in a free-routing environment and further extended through a DQN-based variant that focused on discrete heading and speed modifications.

Considering complexity as the primary resolution objective, Ghosh et al. [158] proposed a deep ensemble MARL framework that integrates the PPO algorithm to reduce congestion and improve schedule adherence by regulating aircraft speed in two-dimensional structured airspace. To enhance policy stability under varying traffic conditions, the framework employed ensemble learning techniques. In a related effort, Guo et al. [159] developed a safety-aware MARL framework, also based on PPO, which incorporates dropout and data augmentation techniques to increase policy robustness under uncertainty and exposure to previously unseen scenarios. This model employed discrete speed adjustments and was evaluated in a similar 2D structured airspace environment.

To address demand–capacity imbalances, Kravaris et al. [160] proposed a MARL framework in which each aircraft agent selects from a predefined set of trajectory modifications, including ground delays and vertical maneuvers. The approach employed a DQN under a decentralized execution paradigm, with agents operating in the two-dimensional structured airspace encompassing Spanish sectors. The framework aimed to mitigate sector congestion and improve overall traffic flow efficiency through coordinated decision-making. In a related effort, Spatharis et al. [161] introduced a hierarchical MARL algorithm based on Q-learning to address demand–capacity balancing. Their method focused on strategic ground delay decisions, with agents interacting within a realistic 2D structured airspace. Learning was performed under both independent and collaborative settings, depending on the level of the hierarchical control structure employed.

Overall, deep reinforcement learning–based approaches have shown promise, particularly in terms of scalability and computational efficiency, and are increasingly emerging as a trend in ATM applications.

TABLE 2.2: Overview of recent DRL-based resolution studies. D: Discrete, C: Continuous. FFRA: Fully free-routing airspace, V: Vertical maneuver, H: Horizontal maneuver, S: Speed regulation. DEC: Decentralized, DCB: Demand-capacity balancing, DDPG: Deep deterministic policy gradient, DQN: Deep Q-network, PPO: Proximal policy optimization.

Study	*Maneuvers*	*Objective*	*Algorithm*	*Env.*	*Framework*	*Hierarchy*
Pham et al. (2019) [150]	H (C)	Conflict	DDPG	FFRA (2D)	Single	DEC
Tran et al. (2019) [125]	H (C)	Conflict	DDPG	FFRA (2D)	Single	DEC
Wang et al. (2019) [162]	H (C)	Conflict	Actor-Critic	FFRA (2D)	Single	DEC
Ribeiro et al. (2020) [119]	H, S (C)	Conflict	DDPG	Structured (2D)	Single	DEC
Hermans et al. (2020) [163]	H (D)	Conflict	DQN	Structured (2D)	Single	DEC
Wen et al. (2019) [149]	H (C)	Conflict	DDPG	FFRA (2D)	Single	DEC
Li et al. (2019) [147]	H (D)	Conflict	DQN	FFRA (2D)	Single	DEC
Mollinga et al. (2020) [123]	V, S, H (D)	Conflict	DQN	Structured (3D)	Single	DEC
Brittain et al. (2019) [120]	S (D)	Conflict	PPO	Structured (2D)	Multi	DEC
Brittain et al. (2021) [121]	S (D)	Conflict	PPO	Structured (2D)	Multi	DEC
Brittain et al. (2020) [153]	S (D)	Conflict	PPO	Structured (2D)	Multi	DEC
Guo et al. (2021) [159]	S (D)	Conflict	PPO	Structured (2D)	Multi	DEC
Zhao et al. (2021) [155]	H, S (C/D)	Conflict	PPO	Structured (2D)	Multi	DEC
Sui et al. (2021) [151]	H, S, V (D)	Conflict	DQN	Structured (3D)	Multi	DEC
Dalmau et al. (2020) [154]	H, S, V (D)	Conflict	Actor-Critic	FFRA (2D)	Multi	DEC
Ghosh et al. (2021) [158]	S (D)	Complexity	Model-Based	Structured (2D)	Multi	DEC
Checn et al. (2023) [156]	H, S (D)	Conflict	DQN	FFRA (2D)	Multi	DEC
Padpolos et al. (2024) [157]	H, S, V (D)	Conflict	DQN	FFRA (3D)	Multi	DEC
Kravaris et al. (2024) [160]	Delay, V (D)	DCB	DQN	Structured (3D)	Multi	DEC
Spatharis et al. (2024) [161]	Delay	DCB	Q-learning	Structured (2D)	Multi	DEC

2.4 Discussion and open problems

In the domain of climate-aware flight planning, several open problems exist. One important aspect is related to the current scientific understanding of aviation-induced climate effects, particularly those related to non-CO_2 forcing agents, which remain highly uncertain according to the latest estimates by Lee et al. [3]. Addressing this gap could improve the reliability of existing models or support the development of new ones, which is crucial for increasing confidence in the mitigation potential of climate-optimized flight planning. Another important problem is the need to move beyond the exclusive focus on conventional kerosene-powered aircraft in flight planning to evaluate the operational performance and environmental benefits of next-generation technologies, such as hydrogen-powered aircraft and sustainable aviation fuels (SAFs). Such evaluations are essential to support informed strategic decision-making regarding policies and investments in emerging technologies and alternative fuels to meet climate goals. These research directions, among others in this field, are actively being pursued by the scientific community and are contributing to a more comprehensive understanding of the aviation sector's potential for mitigating climate impact.

One particularly important yet underexplored research direction concerns climate-optimal flight planning at the network level, which is the primary scope of this thesis. This research direction can be approached from several perspectives, which are discussed in the following subsections.

Network-level evaluation of the feasibility of climate-optimized flight plans

While several studies have advanced micro-scale trajectory optimization techniques, the extent to which flight planning can reliably function as a climate mitigation measure in real-world operations remains underexplored. As discussed in Section 2.1.2, the majority of existing research has focused on optimizing individual trajectories in isolation, thus overlooking the impact of aggregated optimized flight plans on the overall manageability of air traffic and the performance of the ATM system. Without a network-level perspective, analyses of climate impact mitigation are incomplete or even misleading.

Only two studies conducted concurrently with this PhD research [24, 25] have examined the implications of climate-optimized flight plans on the manageability of traffic, focusing on metrics such as traffic demand and the number of conflicts. However, these studies focus on evaluating network-level impacts and do not incorporate any resolution strategies to compensate for the potential adverse effects. Moreover, performance indicators that are especially relevant for strategic planning, such as airspace complexity, are overlooked.

Strategies to address the operational challenges of climate-optimized trajectories

Once the impact of climate-optimized flight plans on traffic manageability has been quantified, an additional optimization step is required to mitigate any adverse operational effects in order to ensure the feasibility of these trajectories within the ATM system.

To date, only one study, conducted in parallel with the present thesis, has addressed this challenge through a mathematical programming framework that selects, from a predefined set of trajectories, those with lower climate impact while also satisfying capacity constraints [26]. This study represents a step toward addressing the operational manageability of climate-optimized flight plans. Nonetheless, conventional approaches of this kind face notable limitations in terms of scalability and adaptability, particularly when applied to large-scale or dynamic air traffic scenarios. Specifically, they rely on extensive pre-processing to generate multiple trajectory options for each flight, followed by static selection through a centralized optimization process. Due to the computational burden and limited flexibility of this approach, it is typically feasible only for regional-scale applications, as was the case in the referenced study. Moreover, the entire process must be re-executed whenever traffic scenarios or environmental inputs change, which restricts its utility in the real-world, dynamic ATM environment.

An advanced, scalable, and adaptive approach, capable of dynamically adjusting aircraft trajectories to restore operational feasibility, applicable across a broad range of scenarios without requiring extensive data preparation, scenario-specific assumptions, or full re-optimization, remains an open problem.

Integrated, single-step optimization frameworks for large-scale policy studies

While sequential optimization frameworks, where individual trajectories are first optimized and subsequently adjusted through a resolution strategy to ensure operational feasibility, can offer valuable insights into the practicality of climate-aware flight planning and are well-suited for planning daily operations, their dependence on the initial individual trajectory optimization step renders them computationally inefficient for large-scale analyses (e.g., multi-year flight planning to develop comprehensive incentivizing indicators for stakeholders). This holds even when advanced micro-scale planning algorithms and scalable resolution strategies are employed.

To enable large-scale analyses, integrated optimization frameworks that simultaneously address climate impact mitigation and operational feasibility within a single-step, fast-time process are needed. To the best of the author's knowledge, no such integrated approach has yet been proposed in the literature.

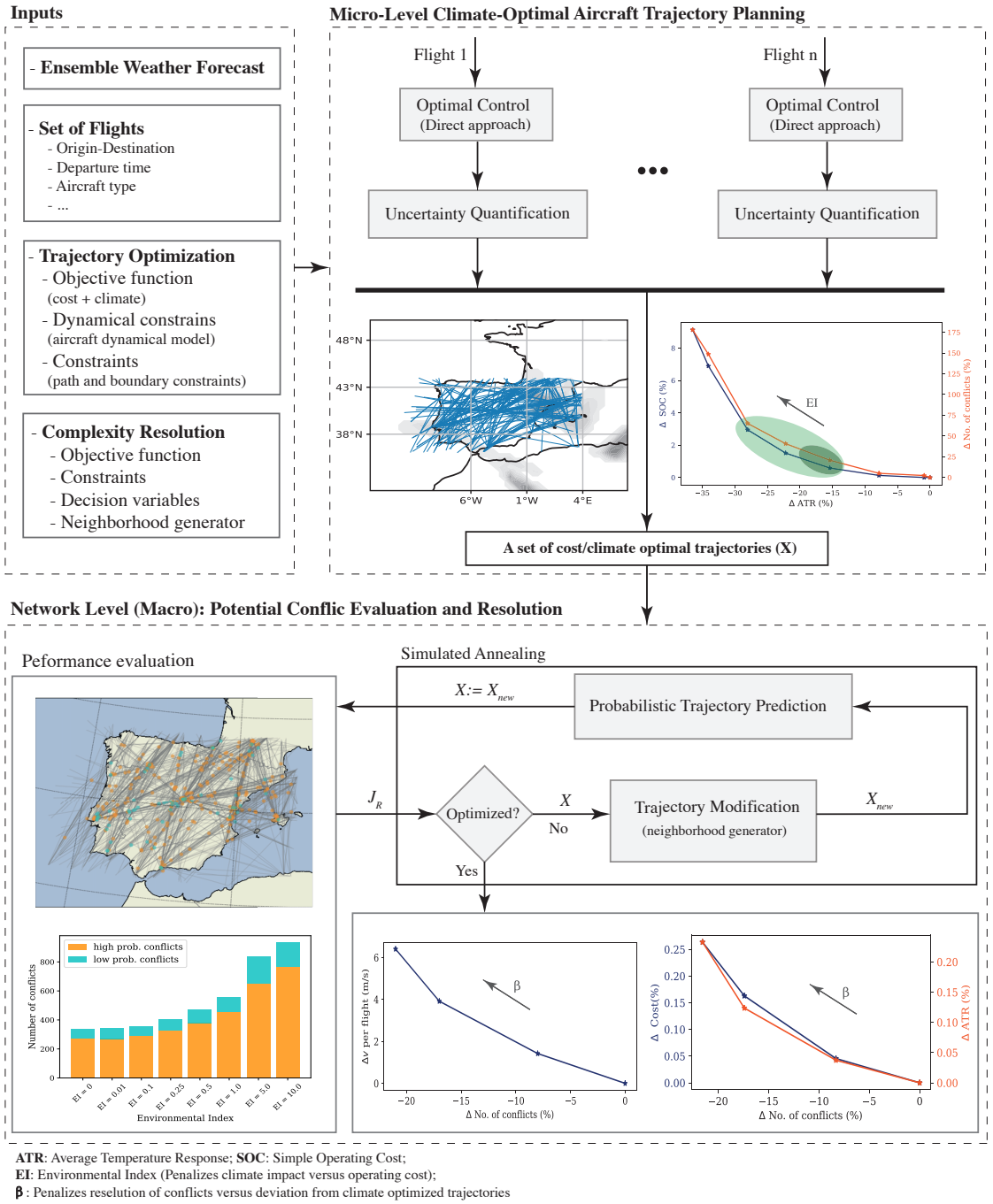
Chapter 3

Network-scale climate-optimized flight planning in free-routing airspace: Heuristic-based approach for conflict management

This chapter introduces an optimization framework for network-scale climate-aware trajectory planning within free-routing airspace. The proposed framework, depicted in Figure 3.1, follows a sequential structure. First, individual aircraft trajectories are independently optimized to minimize their climate impact. The integration of these climate-optimal trajectories into the air traffic management system is evaluated, with particular emphasis on the potential conflicts. To address conflicts arising from this integration, a resolution strategy is proposed. The resolution process is formulated as a multi-objective optimization problem with two primary goals: resolving the arising conflicts due to the consideration of climate impact, and minimizing deviations from the climate-optimal trajectories to preserve environmental benefits. To solve this problem, a heuristic method based on simulated annealing (SA) is employed, using aircraft speed profiles as decision variables. The analysis assumes that aircraft fly at a fixed cruising altitude, resulting in a two-dimensional (2D) optimization.

The remainder of this chapter is organized as follows: Section 3.1 details the micro-scale trajectory optimization method, Section 3.2 outlines the conflict assessment approach and the proposed conflict-resolution algorithm, and Section 3.3 presents and discusses the simulation results.

The content of this chapter is adapted from the work published in [164].



ATR: Average Temperature Response; **SOC**: Simple Operating Cost;
EI: Environmental Index (Penalizes climate impact versus operating cost);
β : Penalizes reselution of conflicts versus deviation from climate optimized trajectories

FIGURE 3.1: Workflow of the proposed heuristic-based framework for 2D climate-aware flight planning considering conflicts at the network scale.

3.1 Micro-scale climate-optimal flight planning

This section presents a framework for micro-scale trajectory optimization aimed at minimizing the climate impact of individual flights. The optimization problem is formulated within the framework of optimal control theory and is constrained to two-dimensional (2D) space, assuming that aircraft cruise at a fixed altitude. To address this problem, the essential components required for modeling are first introduced, including the aircraft dynamical model and the climate impact estimation model. These elements are then used to formulate an optimal control problem, which is solved using the direct optimal control approach to obtain climate-optimal trajectories. The proposed methodology consists of two steps: trajectory optimization and, then, trajectory prediction considering meteorological uncertainty (to quantify the associated uncertainty in the aircraft trajectory and its climate impact).

3.1.1 Modeling

Formulating the flight planning problem as an optimal control problem requires several components. These components are briefly presented in the following.

Aircraft dynamical model

To optimize the flight plan of an aircraft, it is essential to employ a dynamical model that enables the evaluation of key performance variables, such as flight time and fuel consumption, while ensuring the feasibility of the resulting flight profiles. Within the field of ATM research, a widely adopted modeling approach is the three-degree-of-freedom (3-DoF) point-mass model, which captures the principal dynamic characteristics necessary for ATM-related analyses by representing the aircraft as a point mass navigating through three-dimensional space [165]. In this chapter, a simplified 2D version of the point-mass model is employed, assuming flight at a constant cruising altitude:

$$\begin{bmatrix} \dot{\phi} \\ \dot{\lambda} \\ \dot{v} \\ \dot{m} \end{bmatrix} = \begin{bmatrix} \big(v\cos\chi + w_y\big)\big(R_M(\phi)+h\big)^{-1} \\ \big(v\sin\chi + w_x\big)\big((R_N(\phi)+h)\cos\phi\big)^{-1} \\ \big(\mathrm{T}(C_T) - \mathrm{D}(C_L)\big)m^{-1} \\ -f_c(C_T) \end{bmatrix}, \tag{3.1}$$

where φ is the latitude, λ is the longitude, h is the altitude, m is mass, R_M and R_N represent the ellipsoid radius of curvature in the meridian and prime vertical, respectively, f_c is the rate of fuel burn, T is the magnitude of thrust force, D is the magnitude of drag force, C_L is the lift coefficient, C_T is the thrust coefficient, v is the true speed, w_x and w_y are the wind

components, and χ is the heading angle. The aircraft's aerodynamic and propulsive performance is characterized using the BADA 4.2 model [166]. In this representation, the state and control vectors are defined as

$$\begin{aligned} \mathbf{x} &= \begin{bmatrix} \varphi & \lambda & v & m \end{bmatrix}, \\ \mathbf{u} &= \begin{bmatrix} \chi & C_T \end{bmatrix}. \end{aligned} \tag{3.2}$$

In order to account for both physical and operational factors, a set of path constraints is imposed on thrust coefficient, the Mach number (M), and calibrated airspeed (v_{CAS}):

$$\begin{aligned} v_{CAS,stall} \leq v_{CAS}(v) \leq v_{CAS,\max}, \\ C_{T,\min} \leq C_T \leq C_{T,\max}, \\ M(v_{tas}) \leq M_{\max}. \end{aligned} \tag{3.3}$$

Additionally, the following boundary conditions are imposed on the initial and final values of the aircraft's states:

$$\begin{aligned} \big[\varphi, \lambda, v\big](0) &= \big[\varphi_0, \lambda_0, v_0\big], \\ \big[\varphi, \lambda, v\big](t_f) &= \big[\varphi_f, \lambda_f, v_f\big], \\ m(0) &= m_0. \end{aligned} \tag{3.4}$$

It should be noted that the final mass and arrival time of the aircraft are not specified and will be optimized during the optimization.

Climate impact estimation model

To plan aircraft trajectories with minimal climate effects, we need to include spatiotemporal-dependent models in the objective function of the trajectory optimizer [18]. The state-of-the-art approach is to use the aCCFs V1.0a developed within EU-Projects FlyATM4E[1] and ALARM[2]. These functions estimate climate effects in temperature change computationally in real-time, thus enabling very fast function evaluation within the optimization.

The aCCFs provide spatiotemporally resolved information on aviation-induced non-CO_2 climate effects. They take as inputs specific weather variables and provide information on the climate impact of water vapor emissions, NO_x-induced methane, NO_x-induced ozone, and persistent contrails using average temperature response (ATR) over a 20 years time horizon (ATR20) as the climate metric. Note that the formulations express the relationship between climate effects and meteorological variables exhibiting the highest correlations, as identified through detailed simulations with a global climate–chemistry model. Depending

[1] https://flyatm4e.eu/
[2] https://alarm-project.eu/

on the aCCFs version, emission scenario, and time horizon considered, the parameters $\rho_{(\cdot)}$ are selected accordingly.

- **NO_x emissions**: The net climate impact of NO_x emissions is represented as the sum of warming from NO_x-induced ozone and cooling from NO_x-induced methane reduction. These effects are formulated as functions of local temperature (T), geopotential height (GH), and incoming solar radiation (F_{in}):

$$\text{aCCF}_{\text{O}_3} = \max\left(0,\, \rho_{11} + \rho_{12}T + \rho_{13}GH + \rho_{14}T \cdot GH\right), \tag{3.5}$$

$$\text{aCCF}_{\text{CH}_4} = \min\left(0,\, \rho_{21} + \rho_{22}GH + \rho_{23}F_{in} + \rho_{24}F_{in} \cdot GH\right). \tag{3.6}$$

- **Water vapor emissions**: The warming effect of emitted H_2O is modeled based on potential vorticity (PV):

$$\text{aCCF}_{\text{H}_2\text{O}} = \rho_{31} + \rho_{32}\left|PV\right|. \tag{3.7}$$

- **Contrail cirrus**: The climate impact of contrail cirrus is modeled separately for daytime and nighttime, as they exhibit distinct radiative effects: daytime contrails can contribute to both warming, by trapping outgoing longwave radiation, and cooling, by reflecting incoming solar radiation. In contrast, nighttime contrails cause warming, as the absence of solar radiation eliminates the cooling effect.

  For daytime contrails, outgoing longwave radiation (OLR) is identified as the most representative variable based on correlation analysis:

$$\text{aCCF}_{\text{day}} = \rho_{41}(\rho_{42} + \rho_{43} \cdot OLR). \tag{3.8}$$

  For nighttime contrails, a temperature-dependent formulation is used:

$$\text{aCCF}_{\text{night}} = \max\left(0,\, \rho_{51}\left(\rho_{52} \cdot 10^{\rho_{53} \cdot T} - \rho_{54}\right)\right). \tag{3.9}$$

- **Carbon dioxide (CO_2) emissions**: Due to its long atmospheric lifetime, the climate impact of CO_2 is considered independent of the emission location and is therefore represented by a constant factor:

$$\text{aCCF}_{\text{CO}_2} = \rho_{16}. \tag{3.10}$$

As evident from the aCCF formulations, they allow for real-time evaluation of climate impacts, making them efficient for use in flight planning (interested readers are referred to [167] for a detailed description of the aCCFs). Recently, a Python library called CLIMaCCF has been developed, implementing different versions of aCCFs with various emission scenarios, time horizons, and other user-defined settings. It is publicly available at DOI: `https://doi.org/10.5281/zenodo.7074582`.

Uncertainty

The climate impact of non-CO_2 emissions highly relies on meteorological conditions, including temperature, relative humidity, and outgoing longwave radiation [57, 168]. Therefore, a factor that can affect the reliability of the quantified climate impacts is the quality of the weather forecast [49]. The weather forecast is inevitably uncertain, which can also affect aircraft performance variables (e.g., flight time and fuel consumption) due to dependency on wind and temperature [169]. Ensemble prediction systems (EPS) have been introduced to quantify the uncertainties associated with weather forecasts, providing a collection of N_{EPS} probable realizations of weather situations, called ensemble members [170]. Forecasts within EPS can be obtained using different techniques. For instance, in one approach called ensemble data assimilation, the initial conditions and/or parameters of models used in producing forecasts are perturbed.

3.1.2 Trajectory optimization

For optimizing aircraft trajectory, we take the mean values of the ensemble weather variables and solve the problem in a deterministic manner using the aircraft dynamical model (Equation (3.1)), path constraints (Equation (3.3)), boundary constraints (Equation (3.4)), and the following objective function:

$$J = \mathrm{CI}\Big[\underbrace{\mathrm{C}_t \cdot \big[t_f - t_0\big] + \mathrm{C}_f \cdot \big[m(t_0) - m(t_f)\big]}_{\mathrm{SOC}}\Big] + \mathrm{C} \cdot \mathrm{EI} \underbrace{\int_{t_0}^{t_f} \sum_{i=1}^{5} \mathrm{ATR}_i^{\mathrm{mean}}\big(t, \mathbf{x}(t), \mathbf{u}(t)\big)\,\mathrm{d}t}_{\mathrm{ATR}}, \tag{3.11}$$

for $i \in \{\mathrm{CH_4}, \mathrm{Contrails}, \mathrm{O_3}, \mathrm{H_2O}, \mathrm{CO_2}\}$, where $\mathbf{x}(t)$ and $\mathbf{u}(t)$ are the state and control vectors, respectively, and SOC is the simple operating cost [57]. $\mathrm{ATR}_i^{\mathrm{mean}}$ is the ATR for the agent i calculated from considering mean values of ensemble members. The ATR for different non-CO_2 species are calculated as follows:

$$\begin{aligned}
&\mathrm{ATR}_{\mathrm{O_3}}(t, \mathbf{x}, \mathbf{u}) = 10^{-3} \times \mathrm{aCCF}_{\mathrm{O_3}}\big(t, \mathbf{x}\big) \cdot \mathrm{EI}_{\mathrm{NO_x}}(t, \mathbf{x}, \mathbf{u}) \cdot f_c(\mathbf{u}),\\
&\mathrm{ATR}_{\mathrm{CH_4}}(t, \mathbf{x}, \mathbf{u}) = 10^{-3} \times \mathrm{aCCF}_{\mathrm{CH_4}}\big(t, \mathbf{x}\big) \cdot \mathrm{EI}_{\mathrm{NO_x}}(t, \mathbf{x}, \mathbf{u}) \cdot f_c(\mathbf{u}),\\
&\mathrm{ATR}_{\mathrm{Contrails}}(t, \mathbf{x}) = 10^{-3} \times \mathrm{aCCF}_{\mathrm{Contrails}}(t, \mathbf{x}) \cdot v_{gs}(t),\\
&\mathrm{ATR}_{\mathrm{H_2O}}(t, \mathbf{x}, \mathbf{u}) = \mathrm{aCCF}_{\mathrm{H_2O}}(t, \mathbf{x}) \cdot f_c(\mathbf{u}),\\
&\mathrm{ATR}_{\mathrm{CO_2}}(t, \mathbf{x}, \mathbf{u}) = \mathrm{aCCF}_{\mathrm{CO_2}} \cdot f_c(\mathbf{u}),
\end{aligned} \tag{3.12}$$

where v_{gs} is the ground speed and $\mathrm{EI}_{\mathrm{NO_x}}(t, \mathbf{x}, \mathbf{u})$ is the $\mathrm{NO_x}$ emission index calculated using Boeing Fuel Flow Method 2 [40]. To explain how the $\mathrm{ATR}_i^{\mathrm{mean}}$ for the agent i is calculated, let us consider aCCF of water vapor. According to [57], the aCCF of water vapor depends on

the meteorological variable potential vorticity. Since within EPS, we are provided with N_{EPS} forecasts, N_{EPS} different aCCF values can be calculated for water vapor. Thus, $\mathrm{ATR}_{\mathrm{H_2O}}^{\mathrm{mean}}(\cdot)$ indicates that the ATR is calculated from considering the average of aCCF values obtained from different ensemble members.

As can be seen in Equation (3.11), the defined objective function considers both climate impact and cost. Normally, there exists a trade-off between these two objectives. Weighting parameters environmental index (EI) and cost index (CI) determine the importance of these objectives compared to each other. Moreover, constant parameters C_t and C_f are used to allow different explanations of cost, and C adjusts the order of climate impact with cost.

The direct optimal control approach, as an efficient technique for solving highly nonlinear dynamical optimization problems with different sets of constraints, is employed to solve the formulated trajectory optimization problem [171]. In this approach, the dynamical optimization problem (trajectory optimization problem in our case) with the objective function, dynamical model, and different types of constraints (e.g., path and boundary constraints) is transcribed to a nonlinear programming problem (NLP) represented with the following general form [172]:

$$\min_{\Theta} \quad J_{NLP}(\Theta) \tag{3.13}$$

$$\text{s.t} \quad \Phi_i(\Pi) = 0 \qquad i = 1, \cdots, n_\omega \tag{3.14}$$

$$\Xi_i(\Theta) \leq 0 \qquad i = 1, \cdots, n_\xi \tag{3.15}$$

where the vector $\Theta \in \mathbb{R}^{n_\Theta}$ includes the decision (or NLP) variables, and equality and inequality constraints are represented by Equation (3.14) and Equation (3.15), respectively. The optimal solution to the original dynamical optimization problem is obtained by solving the resulting discretized problem. The NLP problems can be solved using various techniques, such as gradient-based and meta-heuristic methods. Moreover, several efficient software packages, including IPOPT [173] and SNOPT [174], exist for solving such a class of optimization problems. Interested readers are referred to [172] for more details on the direct optimal control approach.

After solving the trajectory optimization problem, an optimal trajectory for the aircraft i-th is received and denoted as $\Gamma_i^o := (\varphi_i^o, \lambda_i^o, v_i^o)$.

3.1.3 Ensemble trajectory prediction

To assess the associated uncertainty in the aircraft trajectory and its climate impact, we evaluate the performance of the optimized trajectory $\Gamma_i^o := (\varphi_i^o, \lambda_i^o, v_i^o)$ for all ensemble members. First of all, we assume a unique lateral path and speed profile (obtained from trajectory optimization) for all ensemble members. The consideration of a unique lateral path implies a

constant aircraft course (ψ^o). In this case, the uncertainty will affect ground speed as:

$$v_{gs}^j = \sqrt{v^2 + w_c^{j^2}} + w_a^j, \tag{3.16}$$

where w_a and w_c are the decomposition of wind along the along-track and cross-track directions, respectively, and $(\cdot)^j$ denotes the ensemble member j. The uncertainty in ground speed affects the time aircraft flies the trajectory through:

$$\frac{\mathrm{d}t^j}{\mathrm{d}s} = (v_{gs}^j)^{-1}, \tag{3.17}$$

where t^j is the flight time associated with the ensemble member j and s is the distance flown along route. The fuel consumption is also predicted in a similar way. Now, we consider the range of uncertainty in the climate impact. For the ensemble member j, one has:

$$\mathrm{ATR}^j = \int_{t_0^j}^{t_f^j} \sum_{i=1}^{5} \mathrm{ATR}_i^j\big(t^j, \mathbf{x}^j(t^j), \mathbf{u}(t^j)\big)\mathrm{d}t^j, \tag{3.18}$$

where the superscript j in $\mathbf{x}^j$ is due to the calculated ensemble values for aircraft mass within the ensemble trajectory prediction. For the unique profile determined in the optimization step, the effects of uncertainty are quantified and reflected on the flight time, fuel consumption, and climate impacts, which are included in:

$$\mathrm{P}_i := \{(t_i^1, \cdots t_i^{N_{EPS}}), (m_i^1, \cdots m_i^{N_{EPS}}), (v_{gs_i}{}^1, \cdots v_{gs_i}{}^{N_{EPS}}), (\mathrm{ATR}_i^1, \cdots \mathrm{ATR}_i^{N_{EPS}})\}. \tag{3.19}$$

3.2 Conflict assessment and resolution

The implementation of independently climate-optimized trajectories may lead to the emergence of traffic bottlenecks, as multiple flights tend to avoid warming climate-sensitive regions. This can increase the likelihood of congestion in certain areas, raising operational concerns regarding the feasibility of such trajectories. In this respect, a preliminary step toward planning operationally feasible trajectories within the network scale is to evaluate the operational challenges arising from the adoption of climate-optimized flight plans on the ATM system performance and subsequently develop resolution strategies to mitigate any adverse effects.

In this chapter, the impact of employing climate-optimal trajectories at the network scale is assessed in terms of the number of conflicts (presented in Section 3.2.1). Then, a resolution strategy is proposed to mitigate the encountered conflicts (Section 3.2.2).

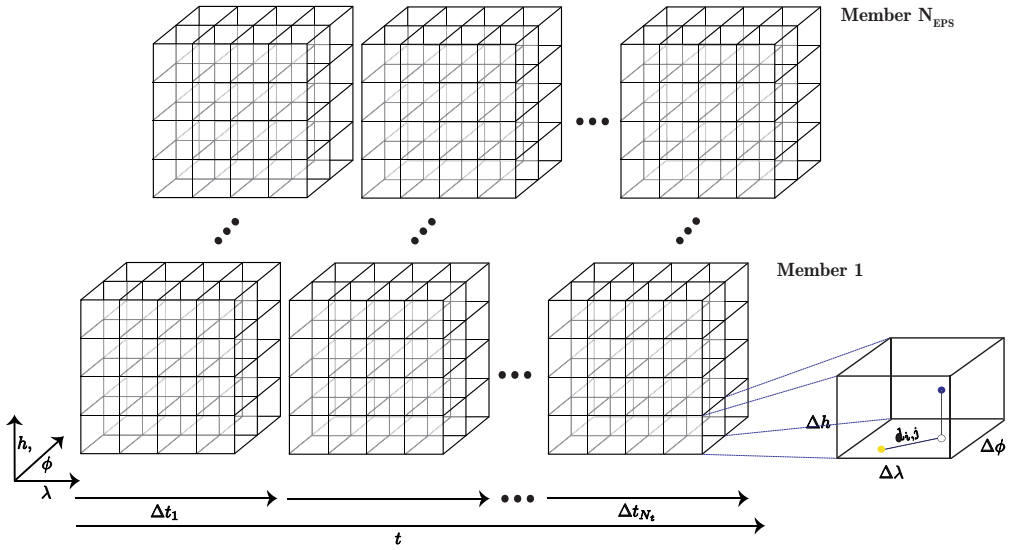


FIGURE 3.2: Grid-based representation of space duplicated with the number of ensemble members [140].

3.2.1 Probabilistic conflict assessment

A conflict is defined as a predicted situation in which an aircraft is expected to lose its minimum separation from other aircraft or airspace constraint within a specified look-ahead time, if no corrective action is taken. One criterion of standard separation is D_0 = 5 NM of horizontal distance and H_0 = 1000 ft vertical distance. In other words, a cylinder with a 5 NM diameter and 1000 ft height can represent a protected zone for each aircraft that should not be intersected with other aircraft-protected zones. The aim of conflict detection is to predict the probable loss of separation that may occur in the future at a specific time and position. Conflict detection for large-scale scenarios generally suffers from computational time since most algorithms compute conflicts pairwise. Moreover, the input trajectories for conflict detection may have some perturbations due to the consideration of uncertainty, such as the ones proposed in Section 3.1. To deal with the mentioned concerns (i.e., computational time and uncertainty), we employ the probabilistic grid-based technique firstly proposed in [175]. Within this technique, aircraft trajectories are saved into grid cells (the size of each cell should be greater than or equal to the minimum standard separation) as depicted in Figure 3.2. The conflict between aircraft k and l is evaluated if their assigned cells are the same or are neighboring. In this case, their distance corresponding to the ensemble member j is computed at the time instant Δt_n. If the computed distance is less than the standard separation, there exists a conflict, i.e.,

$$p_{kl}^{j} = \begin{cases} 1 & \text{if} \quad d_{kl}^{j} < D_0 \quad \text{and} \quad h_{kl}^{j} < H_0 \\ 0 & \text{else} \end{cases} \tag{3.20}$$

where d_{kl} and h_{kl} are their relative horizontal and vertical distances, respectively. The probability of conflict between aircraft k and aircraft l is then calculated over all ensemble members as:

$$\mathrm{P}_{\mathrm{c},kl} = \frac{1}{N_{EPS}} \sum_{j=0}^{N_{EPS}} p_{kl}^{j}, \tag{3.21}$$

where N_{EPS} represents the number of ensemble members.

One challenge associated with the implementation of this method is the required memory to store all grids. To deal with this issue, the hash table structure is employed [176]. The hash table extensively reduces the required memory as we neglect empty cells.

3.2.2 Conflict resolution

The resolution strategy can be formulated as a nonlinear programming problem with the following two objectives, subject to a set of operational constraints:

- Minimize the number of high-probability conflicts.
- Minimize the deviation of the modified trajectories from the climate-optimal ones.

These objectives are mathematically modeled and included in the following cost function:

$$J_R = \beta.K.C_C + (1-\beta) \sum_{i=1}^{N} (\Delta v_i), \tag{3.22}$$

where N represents the number of flights. $\beta \in [0,1]$ is a weighing parameter that determines the importance of the mentioned objectives compared to each other, K is the scaling parameter, and C_C is the cost of conflicts defined as:

$$C_C = \sum_{k=1}^{N} \sum_{l=1, l\neq k}^{N} \mathrm{Y}_{kl}, \qquad \mathrm{Y}_{kl} = \begin{cases} 1 & \text{if} \quad \mathrm{P}_{\mathrm{c},kl} \geq P_\tau, \\ 0 & \text{else}, \end{cases} \tag{3.23}$$

where the parameter P_τ is a user-defined probability threshold that allows users to adjust the level of conservativeness. For instance, $P_\tau = 0$ considers all possible conflicts caused by ensemble members (both high and low probability conflicts) to be resolved, while $P_\tau \geq 0.5$ only considers those conflicts that occur at least within half of the ensemble members.

The decision variables of the optimization problem are the true airspeed of each flight. The speed change (deviation (Δv) from climate-optimal speed (v^o)) is constrained as:

$$-0.04v^o \leq \Delta v \leq 0.04v^o. \tag{3.24}$$

Due to the complexity of the optimization problem (i.e., a large number of decision variables), we rely on the SA algorithm as the solution approach [118]. The SA algorithm,

inspired by the annealing process, aims at reaching the minimum energy of the metal by rearranging its particles. This algorithm is based on temperature change through the heating and cooling process. In the heating process, the particles have enough freedom to move around in random directions. By reducing the temperature, they tend to find a new stable configuration to achieve minimum energy [177]. The SA in optimization applications uses the same strategy to minimize a defined objective function. A neighbor solution is generated at each iteration, and its associated cost is compared with the current solution. The generated solution at each iteration is accepted with a probability that is related to the temperature and the difference between the current and new costs. At the beginning of the process, when the temperature is high, the worst solutions (i.e., the solutions that yield higher costs compared to the previous step) are more likely to be accepted. The acceptance rate of the worst solutions is reduced within the cooling process. The property to accept worse solutions provides the ability to avoid getting stuck in local minima and is one of the main advantages of SA over the gradient-based solvers, such as interior-point, and successive quadratic programming [177].

For the considered resolution problem, at each iteration, a neighbor solution is generated from the neighborhood function presented in Subsection 3.2.2 and the objective function Equation (3.22) is re-evaluated (see Figure 3.1). The new cost $J_{R,n}$ is compared with the current cost $J_{R,c}$ and accepted with the following probability:

$$P_{acc} = \begin{cases} 1 & \text{if} \quad J_{R,n} \leq J_{R,c} \\ e^{-(J_{R,n} - J_{R,c})T^{-1}} & \text{else.} \end{cases} \tag{3.25}$$

According to Equation (3.25), if the cost associated with the neighbor solution is improved, it is accepted, and in the case of degradation, it is accepted with the defined probability (obtained from the Boltzmann distribution).

Neighborhood function

At each iteration within SA, new solutions need to be generated, called neighbor solutions. Here, we propose a neighborhood function to generate candidate solutions by modifying the aircraft's true airspeed. To avoid blindly exploring within a large search space, we inform the algorithm only to modify the trajectories of aircraft in conflict. The proposed algorithm starts from a set of climate-optimal trajectories (determined in section 3.1) as the initial solution S_0 with the cost $J_{R,0}$. At each iteration, the neighborhood function generates a neighbor solution S_n with the cost $J_{R,n}$. The steps to generate a neighbor solution within the neighborhood function are as follows:

1. Two sets of aircraft are selected independently: Ω_1, comprising 20 aircraft with the highest number of conflicts, and Ω_2, comprising 20 aircraft with the largest deviation

from optimal speed.

2. An aircraft $\mathrm{Ac_m}$ is selected for trajectory modification with the probability β from the set Ω_1, and $1-\beta$ from the set Ω_2. By this setting, with increasing β, the aircraft with a higher number of conflicts are more likely to be selected for the resolution.

3. The airspeed v_m of $\mathrm{Ac_m}$ is modified within the defined permissible range, and its associated performance Γ_m is re-computed employing the ensemble trajectory prediction method presented in Section 3.1.3.

The SA algorithm and the proposed neighborhood function are presented in Algorithm 1.

Algorithm 1 SA Algorithm

Require: N_{tr}, α, T_{in}, T_f, β
Require: A set of climate-optimal trajectories S_0
  $T_c \leftarrow T_{in}$, $J_{R,c} \leftarrow J_{R,0}$, $S_c \leftarrow S_0$
  while $T_c < T_f$ **do**
    $N_c \leftarrow 0$
    while $N_c < N_{tr}$ **do**
      From Sc get subsets Ω_1 and Ω_2.
      Get a random number $p \in [0, 1]$
      if $p \leq \beta$ **then**
        Get random aircraft $Ac_m \in \Omega_1$
      else
        Get random aircraft $Ac_m \in \Omega_2$
      end if
      Get random number $\gamma \in [0.96, 1.04]$
      $v_m^o \times \gamma$
      Predict the aircraft performance Γ_m
      Compute $J_{R,n}$
      if $J_{R,n} \leq J_{R,c}$ **then**
        Replace the new profile of Ac_m in S_c
        $J_{R,c} \leftarrow J_{R,n}$
      else
        Get a random number $\sigma \in [0, 1]$
        $\mathrm{P}_{acc} \leftarrow \mathrm{e}^{-(J_{R,n} - J_{R,c})T^{-1}}$
        if $\sigma < P_{acc}$ **then**
          Replace the new profile of Ac_m in S_c
          $J_{R,c} \leftarrow J_{R,n}$
        end if
      end if
      $N_c \leftarrow N_c + 1$
    end while
    $T_c \leftarrow T_c \times \alpha$
  end while
  $S^* \leftarrow \mathrm{S}_c$
  Return S^*

3.3 Simulation results

A case study including 1006 flights is presented to illustrate the applicability of the proposed methodology. In Section 3.3.1, the independently optimized trajectories are determined for different routing options. The implications of adopting these optimized trajectories at the network scale are evaluated in Section 3.3.2, focusing on the resulting number of conflicts. Finally, the identified conflicts are addressed through the resolution strategy presented in Section 3.3.3.

3.3.1 Trajectory optimization

The flight data have been extracted from Eurocontrol's Demand Data Repository (DDR2) dataset[3] by limiting airspace to an area mainly covering Spain and Portugal on May 6^{th} 2018, from 12:00 to 16:00. The information regarding the origin, destination, cruise altitude, and flight time has been provided within the DDR2 dataset. The aircraft models are all considered to be A330-341 with an initial mass of 200 tons.

To model the uncertainty in meteorological variables, the ERA5 reanalysis data, containing ten ensemble members, is adopted and used in trajectory optimization and prediction steps. Due to ease of availability, we employ this dataset in this study; however, forecast data with a different number of ensemble members can be employed in a similar manner.

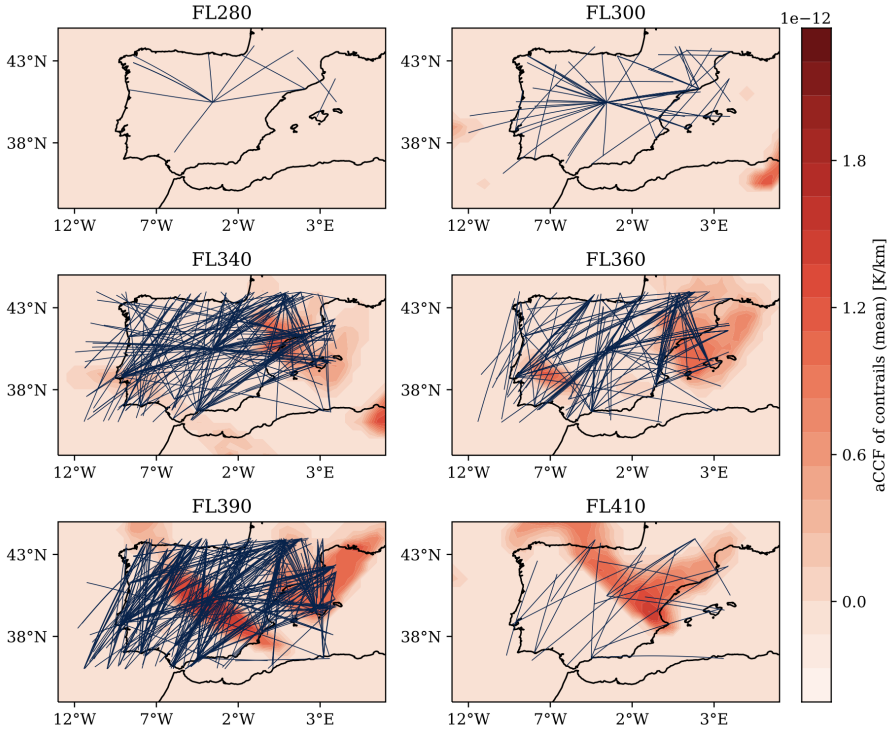
The extracted flights are optimized with the direct optimal control approach for the problem formulated in Section 3.1, considering the mean values of ensemble members for weather variables. We use the Trapezoidal rule to transform the original optimal control problem into a nonlinear programming problem. The resulting NLP problem is then solved using the IPOPT solver in Python, employing the interior-point method to find the optimal solution. The number of discretization nodes is N_d=100. The discretized shortest path is considered the initial guess for the optimizer.

The trajectories are optimized in 2D airspace for eight different environmental indices (i.e., EI [-]), considering CI = 1.0 [-], C_t = 0.75 [USD/s], C_f = 0.51 [USD/kg], and C = 10^{14} [USD/K]. To quantify the uncertainty in the flight performances due to the ensemble members, the ensemble trajectory prediction proposed in Section 3.1.3 is implemented. For each weighting coefficient, the computational time to solve the trajectory optimization and prediction for all flights is approximately 100 minutes (i.e., 6s per flight). The optimized trajectories for two environmental indices are depicted in Figure 3.3. According to the recent studies employing aCCFs to quantify climate impact ([57,58]), the climate effect of contrails outweighs the impacts caused by other species. To this end, the optimized trajectories are plotted with the aCCF of contrails on different flight levels.

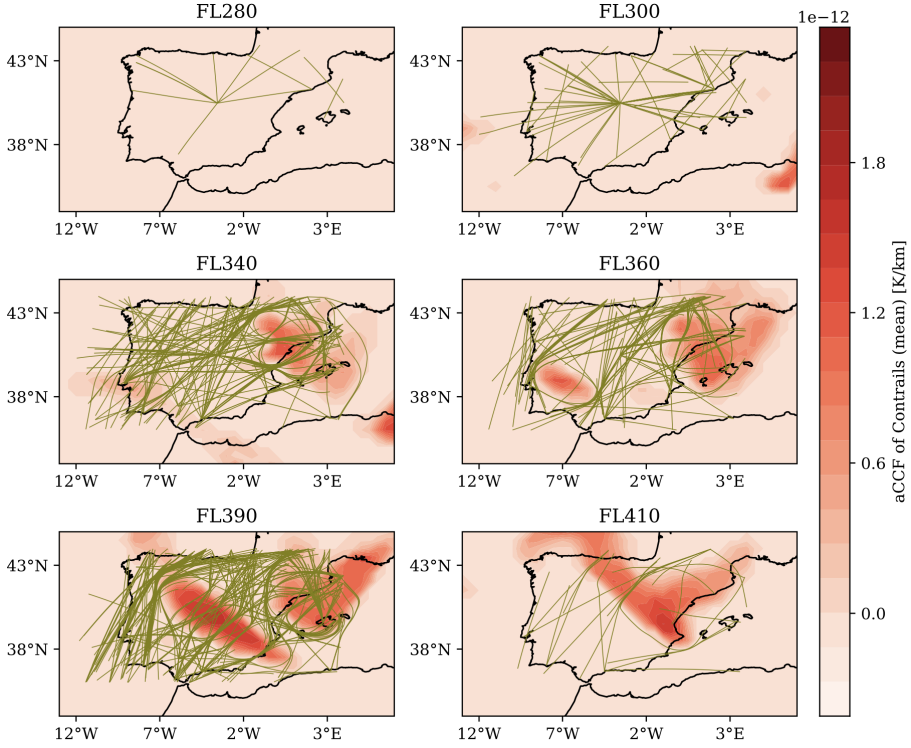
Figure 3.3a shows the pure cost-optimal trajectories (i.e., EI = 0.0). It can be seen that contrail-sensitive regions are crossed by aircraft. Whereas, as can be observed in Figure 3.3b, these regions are mostly avoided with the climate-optimal routing option, associated with EI = 1.0. For the flight levels with no persistent contrail formation, such as FL280 and FL300, both routing strategies result in almost similar trajectories.

The effects of increasing EI on ATR and simple operating cost are shown in Figure 3.4a, and 3.4b, respectively. As can be seen, higher EI values mitigate the climate impact, albeit with an associated increase in operational cost. The error bars in Figure 3.4a and 3.4b show

[3] `https://www.eurocontrol.int/ddr`



(A) cost-optimal routing option (EI = 0.0)



(B) climate-optimal routing option (EI = 1.0)

FIGURE 3.3: Lateral paths depicted with the contrail-sensitive regions (aCCF of contrails) as colormaps for different flight levels.

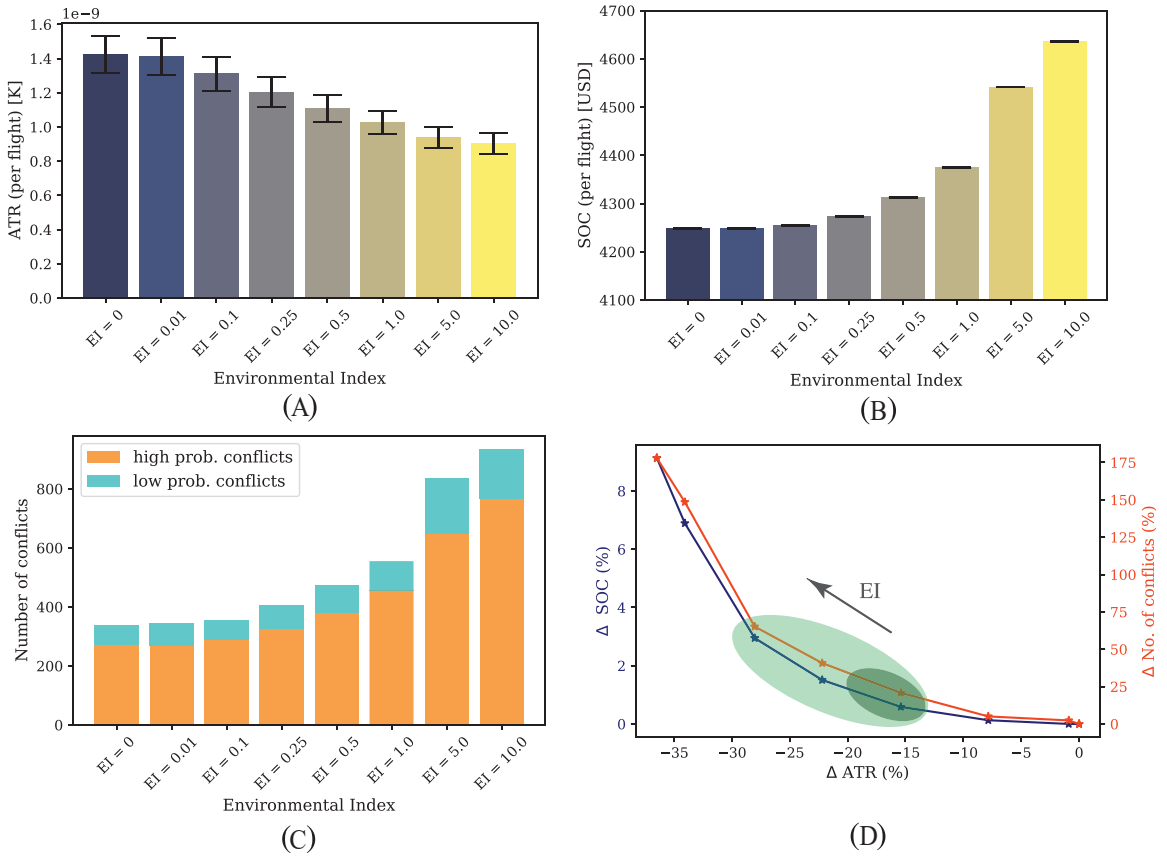


FIGURE 3.4: The obtained performances for different EI values. SOC and ATR have been divided by the number of flights.

the ranges of uncertainty (standard deviation) due to the ensemble forecast. The Pareto-frontier, showing the trade-off between climate impact mitigation potential and relative increase in cost, is given in Figure 3.4d. It can be concluded from Figure 3.4d that there is a potential to reduce the climate impact by 28% at the expense of accepting a 3% increase in operating cost.

3.3.2 Conflict assessment

To see how the adoption of climate-optimized routes affects the complexity of the traffic, the number of conflicts is calculated using the approach presented in Section 2.2. This approach is particularly suitable for evaluating potential conflicts given its relatively efficient computation and its compatibility with uncertainty-aware modeling [116]. Nevertheless, alternative conflict detection techniques, such as analytical closest-point-of-approach [122], can also be used, particularly when higher geometric precision or explicit risk estimation is required. These methods, however, tend to face scalability limitations in high-density traffic

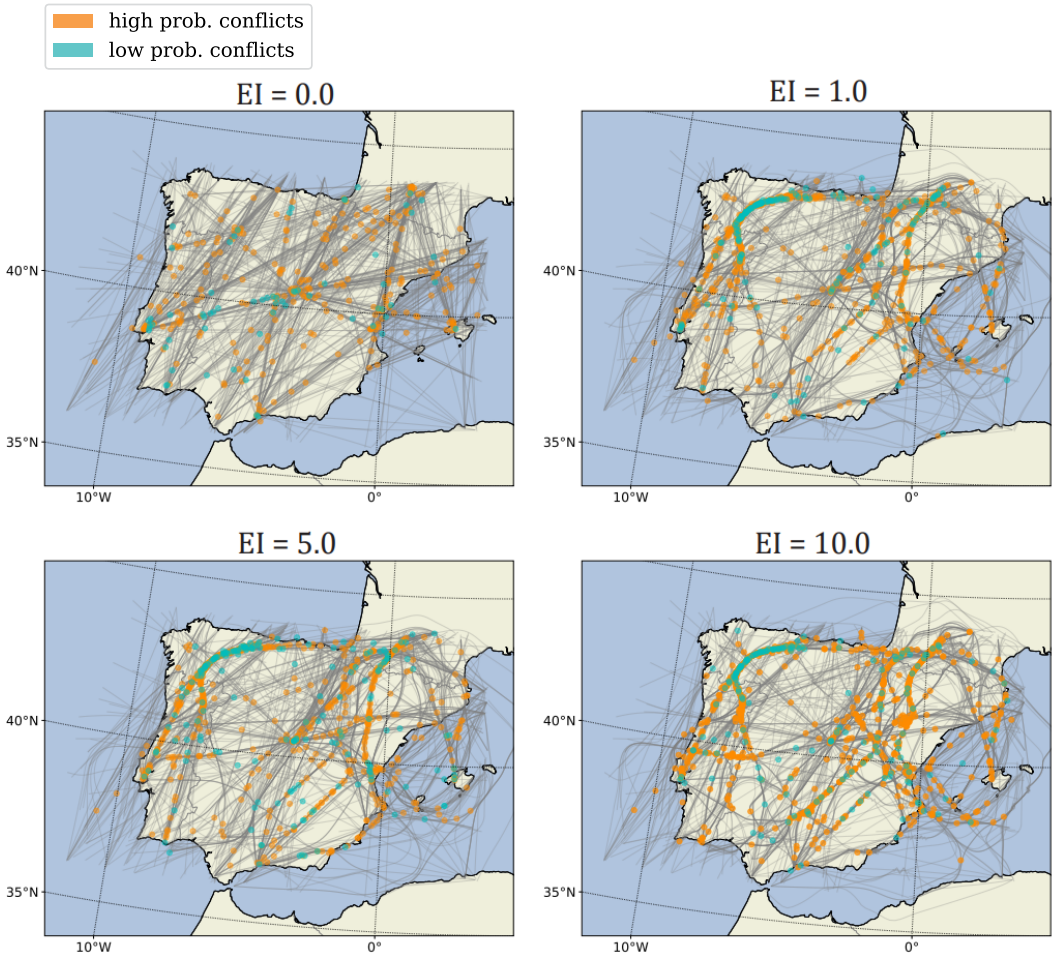


FIGURE 3.5: Lateral paths with the location of high and low probability conflicts for different EIs.

scenarios. While their performance can be enhanced through spatial or temporal filtering, this often comes at the cost of increased implementation complexity.

The airspace is divided into a four-dimensional grid spanning latitude, longitude, altitude, and time. Each grid cell has dimensions of $0.1^\circ \times 0.1^\circ \times 1000ft \times 1s$, which is small enough that no conflict is missed. We define a probability threshold ($P_\tau = 0.5$ in this study) to differentiate between low and high probable conflicts. The conflicts with a probability less than 0.5 are labeled as low probability conflicts, and vice versa for high probability ones.

The number of conflicts for the considered EI values is shown in Figure 3.4c, illustrating that climate impact reduction is achieved at the expense of higher potential conflict occurrence. Therefore, in addition to the cost, the increase in the number of conflicts is another challenge that arises when considering climate impact. To depict this, the geographical locations of conflicts for different routing options are plotted in Figure 3.5. From Figure 3.5, it can be seen that regions with low or negligible contrail-induced climate impact become more congested with increasing EI. This is because the aircraft tend to fly in these areas to mitigate

climate impact, resulting in a predicted loss of standard separation between aircraft.

Therefore, there exists a trade-off between mitigating the climate impact and the number of conflicts. Figure 3.4d shows that a 28% reduction in climate impact is achieved at the expense of a 65% increase in the number of conflicts, for EI = 1.0. It is worth mentioning that the range of uncertainty in the obtained cost (Figure 3.4b) is negligible. Moreover, most of the conflicts are high probable ones. One justification for such a low impact of uncertainty is the usage of reanalysis data. For weather forecast data, the ranges of uncertainty for the same scenario are expected to be higher.

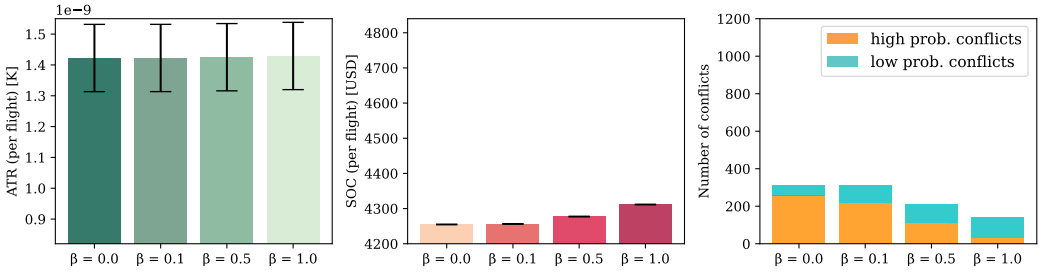
3.3.3 Conflict resolution

As shown, the mitigation of climate impact is achieved at the expense of an increase in the number of potential conflicts for higher EI values. To strategically resolve the encountered conflicts, we employ the resolution algorithm presented in Section 3.2.2. To this end, four sets of optimized trajectories associated with different EIs (i.e., EI $\in [0.0, 1.0, 5.0, 10.0]$) are selected for the resolution process.

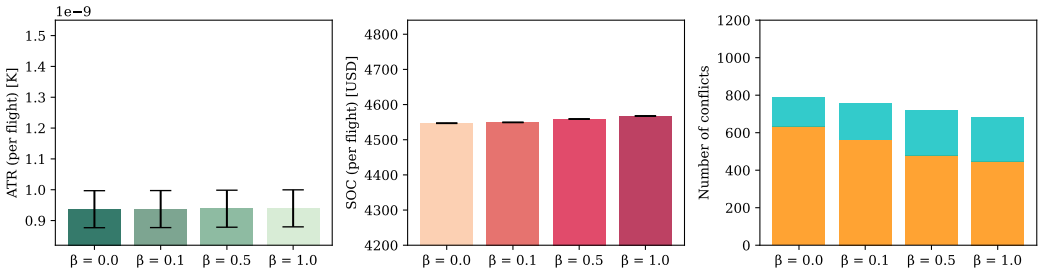
For each set of trajectories, the proposed SA Algorithm 1 is implemented in Python for four different weighting coefficients β's (i.e., $\beta \in [0.0, 0.1, 0.5, 1.0]$), considering $K = 200$, and $P_\tau = 0.5$. At each iteration within the resolution algorithm, after modifying the speed of aircraft in conflict, the total number of conflicts between all aircraft is re-evaluated to consider the domino effect of de-conflicting one pair of aircraft on the others. The performance of the proposed resolution strategy in terms of climate impact, cost, and the number of conflicts is depicted in Figure 3.6 for different EI and β values.

As can be seen in Figure 3.6, for all sets of trajectories corresponding to different EI values, by increasing β, the number of high probability conflicts is reduced by a slight increase in cost and climate impact. The trade-off between reducing conflicts and relative increases in cost and climate impact is depicted in Figure 3.7 for different sets of optimized trajectories. Figure 3.7 indicates that a reduction in conflicts can be achieved at the cost of a marginal increase in both operational cost and climate impact. In all cases, by accepting less than 0.5% increase in cost and climate impact, the conflicts are reduced at least by 20%. It is worth mentioning that by increasing EI, the potential to resolve conflicts is reduced.

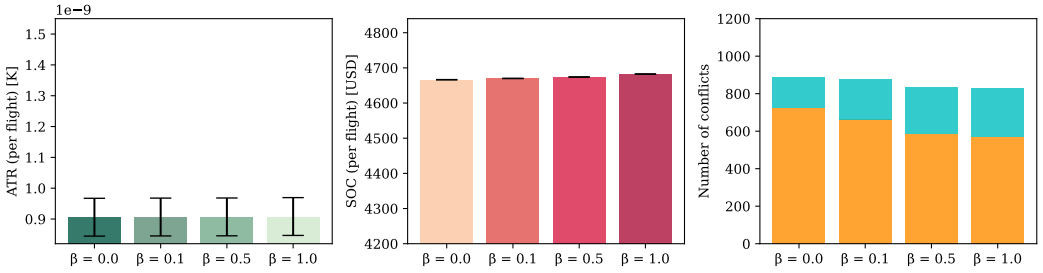
The limited effectiveness of the proposed conflict resolution algorithm, particularly for trajectories corresponding to higher EIs (i.e., lower climate impact), stems from the assumptions adopted in this analysis. In the first stage of the framework, trajectory optimization was performed only in the lateral dimension, with flight levels kept fixed. This constraint restricts flexibility, as the only available option to reduce climate impact is to reroute aircraft around climate-sensitive regions, causing redistribution of traffic and increased concentration of trajectories in neighboring regions, which in turn elevates the likelihood of conflicts.



(A) EI = 0.0



(B) EI = 5.0



(C) EI = 10.0

FIGURE 3.6: The obtained results from the resolution method for different routing options. For each routing option, the effects of increasing β (i.e., the weighing parameter in the cost function Equation (3.22)) on the number of conflicts, simple operating cost, and climate impact are depicted.

Since all conflicts occur within the horizontal plane and the resolution algorithm can only adjust speed profiles to modify temporal separation, its ability to generate fully conflict-free solutions is inherently limited. In particular, head-on conflicts cannot be resolved using speed adjustments alone. Furthermore, because the permissible range of speed variations is small, the achievable separation distances remain insufficient to eliminate all crossing conflicts. These results highlight the need to expand the decision space to incorporate vertical and lateral maneuvers, thereby improving conflict resolution effectiveness.

Finally, the trade-off between reducing the number of conflicts and deviation from optimal speeds (obtained from trajectory optimization) is depicted in Figure 3.8 by means of Pareto frontiers. The results indicate that the reduction of conflicts results in an increase in deviation from climate-optimal speed.

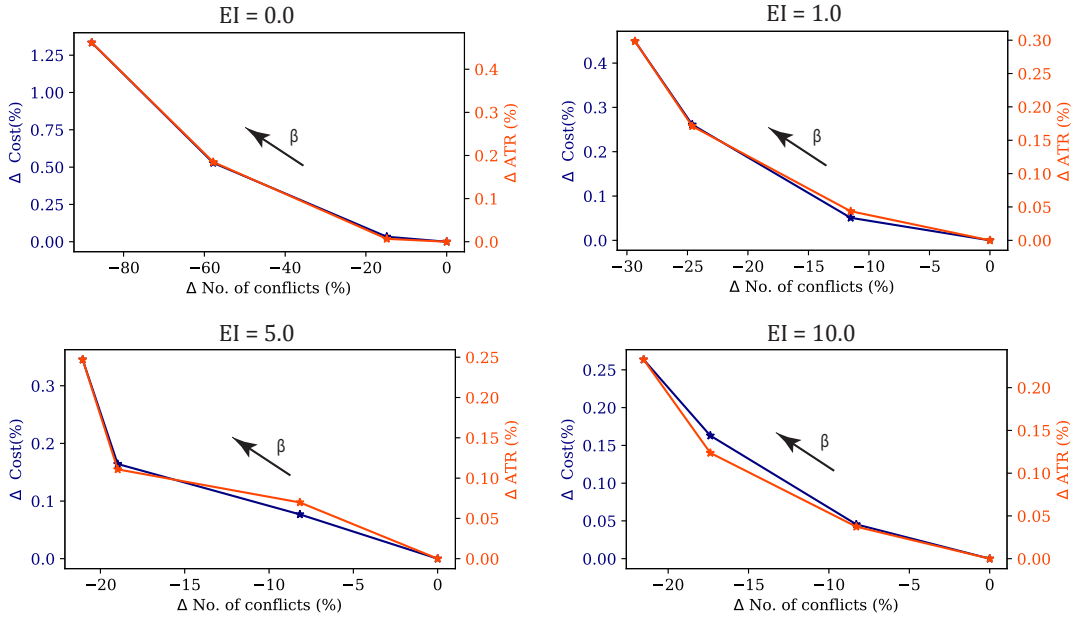


FIGURE 3.7: Pareto-frontiers obtained by employing the proposed probabilistic resolution method for different sets of optimized trajectories (i.e., different EI values). The Pareto frontiers show the trade-off between decreasing the number of high-probability conflicts and relative increases in cost and climate impact obtained using different β values.

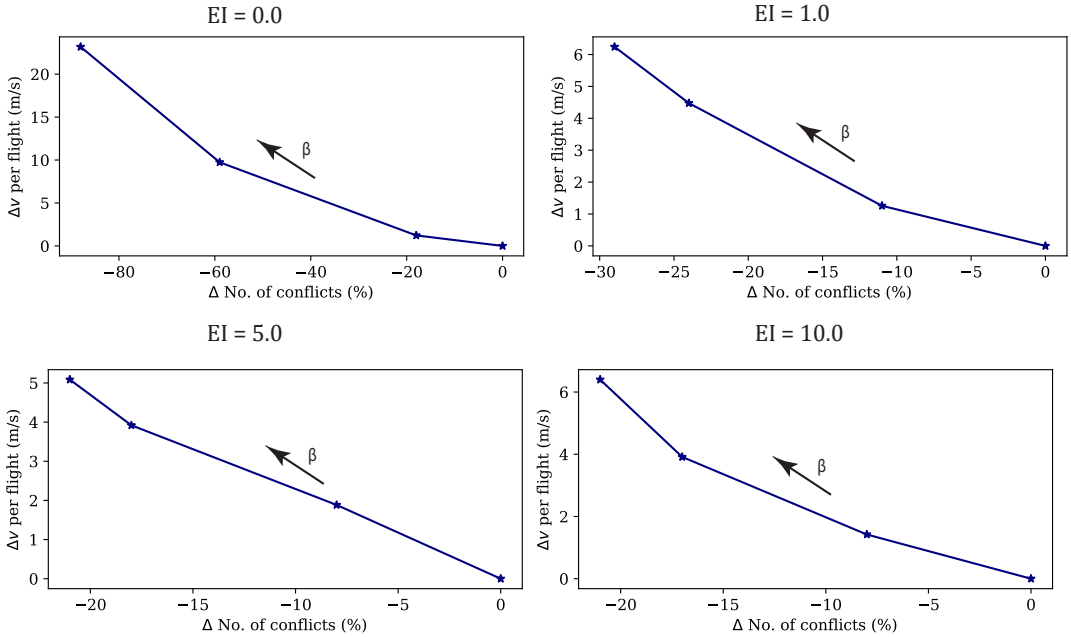


FIGURE 3.8: Pareto-frontiers obtained by employing the proposed probabilistic resolution method for different sets of optimized trajectories (i.e., different EI values). The Pareto frontiers illustrate the trade-off between reducing the number of high-probability conflicts and the increase in deviations of speeds from the optimized ones obtained using different β values.

TABLE 3.1: Summary of the obtained results from the proposed resolution method across different climate-optimal trajectory sets (EI values). Metrics reflect how resolving high-probability conflicts for various β values impacts climate, cost, and trajectory alterations.

Environmental Index	Performance	β = 0.0	β = 0.1	β = 0.5	β = 1.0
EI = 0.0	**ATR [K]**	$14.224e^{-8}$	$14.225e^{-8}$	$14.250e^{-8}$	$14.290e^{-8}$
	(relative increase w.r.t. non-resolved trjs. [%])	(0%)	(0.006%)	(0.184%)	(0.461%)
	SOC [tUSD]	4.254	4.256	4.277	4.311
	(relative increase w.r.t. non-resolved trjs. [%])	(0%)	(0.033%)	(0.528%)	(1.333%)
	High prob. conflicts	266	218	108	31
	(relative decrease w.r.t. non-resolved trjs. [%])	(0%)	(18%)	(59%)	(88%)
	Low prob. conflicts	65	92	99	110
	Average speed variation [m/s]	0	1.223	9.738	23.177
	Number of speeded up aircraft	0	29	70	132
	Number of slowed down aircraft	0	28	81	150
EI = 5.0	**ATR [K]**	$9.362e^{-8}$	$9.375e^{-8}$	$9.381e^{-8}$	$9.393e^{-8}$
	(relative increase w.r.t. non-resolved trjs. [%])	(0%)	(0.043%)	(0.171%)	(0.298%)
	SOC [tUSD]	4.547	4.549	4.558	4.567
	(relative increase w.r.t. non-resolved trjs. [%])	(0%)	(0.050%)	(0.261%)	(0.449%)
	High prob. conflicts	634	561	478	448
	(relative decrease w.r.t. non-resolved trjs. [%])	(0%)	(11%)	(24%)	(29%)
	Low prob. conflicts	157	197	221	235
	Average speed variation [m/s]	0	1.257	4.476	6.241
	Number of speeded up aircraft	0	20	33	31
	Number of slowed down aircraft	0	24	37	49
EI = 10.0	**ATR [K]**	$9.059e^{-8}$	$9.065e^{-8}$	$9.069e^{-8}$	$9.081e^{-8}$
	(relative increase w.r.t. non-resolved trjs. [%])	(0%)	(0.069%)	(0.110%)	(0.247%)
	SOC [tUSD]	4.666	4.669	4.674	4.682
	(relative increase w.r.t. non-resolved trjs. [%])	(0%)	(0.076%)	(0.164%)	(0.345%)
	High prob. conflicts	722	663	585	564
	(relative decrease w.r.t. non-resolved trjs. [%])	(0%)	(8%)	(18%)	(21%)
	Low prob. conflicts	167	215	246	259
	Average speed variation [m/s]	0	1.884	3.919	5.085
	Number of speeded up aircraft	0	31	35	23
	Number of slowed down aircraft	0	17	35	36

A brief summary of the performance of the proposed resolution algorithm is presented in Table 3.1. An observation from the number of modified aircraft trajectories implies that in almost all scenarios, less than 10% of aircraft trajectories are modified. Moreover, the number of seeded-up aircraft is almost similar to the slowed-down ones. Since the focus of resolution is on reducing high-probability conflicts, an increase in the number of low-probability conflicts is observed across all cases.

3.4 Summary

This chapter introduced a framework for climate-optimal flight planning at the network scale. A realistic traffic scenario over Spanish airspace, comprising 1006 flights, was used to evaluate the performance of the proposed approach. Initially, all flights were individually optimized under various routing objectives, ranging from cost-optimal to climate-optimal trajectories. The effects of adopting these climate-optimized routes were subsequently assessed in terms of operational cost and the number of conflicts. Pareto-frontiers were provided to study the existing trade-offs between climate impacts, operating cost, and the number of conflicts.

The results demonstrated that the number of conflicts increases as we employ trajectories with lower climate impact. For example, in the considered case study, an overall 28% reduction in climate impact increased the number of conflicts by 65%, while a 7% more reduction in climate impact was achieved at the expense of a 110% additional increase in the number of conflicts, which must be crucially taken into consideration.

Subsequently, a strategic conflict-resolution procedure was implemented using the SA algorithm to see how many conflicts can be reduced with only speed modifications and different weights penalizing deviation from climate-optimal trajectories. The results showed that by accepting less than 0.5% increase in cost and climate impact, at least 20% of conflicts could be resolved for all routing options. However, the potential to reduce conflicts using speed change as the only decision variable is decreased by accepting trajectories with lower climate impacts. For instance, for the cost-optimal routing option, 80% of the conflicts are resolvable, whereas, for the trajectories with less climate impact, the reduction is limited to 25%. Including other decision variables in the resolution, such as lateral path and altitude, are alternatives that can increase this potentiality by resolving encountered conflicts in a more efficient manner (having more degrees of freedom).

It should be noted that in this study, the main goal was to establish an overall understanding of the potential mitigation of climate impact by trajectory planning and the associated effects on traffic patterns. In this respect, the same aircraft type (i.e., A320-214) was considered to create an initial picture of these aspects while avoiding introducing additional

elements that could bias and complicate the analysis. However, incorporating a heterogeneous fleet would introduce variation in optimal trajectories, particularly in altitude selection under identical weather conditions. Such differences in traffic patterns could affect several key performance metrics discussed earlier, including spatial distribution and number of conflicts, the magnitude of climate impact mitigation, and the corresponding operational cost. Therefore, the proposed frameworks should be extended to accommodate mixed-fleet operations to align more closely with operational practices.

Chapter 4

Network-scale climate-optimized flight planning in free-routing airspace: Multi-agent RL approach for conflict management

The previous chapter introduced a sequential framework for network-scale climate-optimal flight planning, in which the impact of climate-aware trajectories on ATM system performance was first evaluated, focusing on the number of potential conflicts, followed by the development of a resolution strategy to mitigate the resulting conflicts. While the approach demonstrated the feasibility of reducing climate impact in an operationally manageable manner, it exhibited limitations related to the scope of the scenario considered and the resolution method. With regard to the scenario, the analysis was restricted to flight planning with constant flight levels and was confined to a regional-scale scenario, thereby limiting the generalizability of the findings to real-world scenarios. As for the resolution strategy, the proposed approach relied on a heuristic algorithm, which struggles with scalability in large-scale air traffic scenarios. The execution time for these methods further escalates as air traffic density increases. Furthermore, such conventional approaches to tackle the optimization problem lack adaptability, requiring complete recalculation whenever the traffic scenarios change or new trajectory sets are introduced.

To address these challenges, this chapter advances the framework in the previous chapter in two directions: 1) considering full 3D trajectory optimization (i.e., both lateral and vertical optimization) to mitigate climate impacts and extending the traffic scenario to the entire European airspace, and 2) developing a scalable and adaptable algorithm for conflict mitigation capable of managing large-scale air traffic scenarios.

Meeting the demands of such a large-scale, high-dimensional optimization problem necessitates a paradigm shift from conventional heuristics toward more scalable and adaptive

methodologies. In this context, deep reinforcement learning (DRL) offers a promising solution. By enabling agents to learn optimal behaviors through interaction with their environment, DRL supports dynamic, data-driven decision-making that can generalize across scenarios [141]. The DRL, in particular, has proven effective in decentralized, high-dimensional control tasks and provides the flexibility required to model the dynamic nature of air traffic management systems [145]. Importantly, DRL approaches allow for the derivation of optimal policies through offline training, which can then be deployed across varied scenarios with reduced computational overhead.

Building on these capabilities, this chapter introduces a novel cooperative framework based on multi-agent DRL for network-scale climate-optimal flight planning. The framework is designed to strategically adjust individually optimized, climate-aware trajectories in order to enhance overall ATM system manageability by reducing potential conflicts, while preserving the environmental benefits of the original flight plans. The proposed algorithm, termed the policy-sharing multi-agent twin-delayed deep deterministic policy gradient (Ps-MATD3) algorithm, extends the well-established twin-delayed deep deterministic policy gradient (TD3) algorithm to a multi-agent setting. Each aircraft is modeled as an autonomous agent operating within a shared environment, and trained with two objectives: 1) to minimize conflicts with other aircraft, and 2) to maintain close adherence to its initial climate-optimal trajectory.

To overcome the challenge of non-stationarity inherent in multi-agent environments, fully observable critic networks are employed. These critics have access to the states and the actions of all agents during training, thereby providing each agent with a stable learning environment and improving the reliability of the training process. Finally, to ensure adaptability across varying traffic scenarios, the proposed algorithm incorporates a policy-sharing mechanism, allowing a single trained policy to be deployed across an arbitrary number of aircraft. This feature enhances the generalizability and scalability of the framework, making it suitable for large-scale air traffic management applications.

The overall view of the proposed framework in this chapter is illustrated in Figure 4.1. Initially, flight plans are optimized using the framework presented in Section 4.1 for a real traffic scenario considering the mitigation of the climate impact induced by non-CO_2 forcing agents and operating cost as the flight planning objectives. By adjusting the weight of climate impact in the objective function, we generate various alternative trajectories, ranging from cost-optimal to climate-optimal routing options. Then, we investigate the effects of adopting climate-optimized trajectories on traffic complexity in terms of the number of conflicts. To mitigate the arisen conflicts, a conflict resolution is proposed using the framework presented in Section 4.2. The proposed resolution strategy presented in Section 4.3 is then implemented to drive an optimal policy that takes as input aircraft information and provides speed modifications to resolve conflicts of climatically optimized trajectories at the planning phase (Section 4.4).

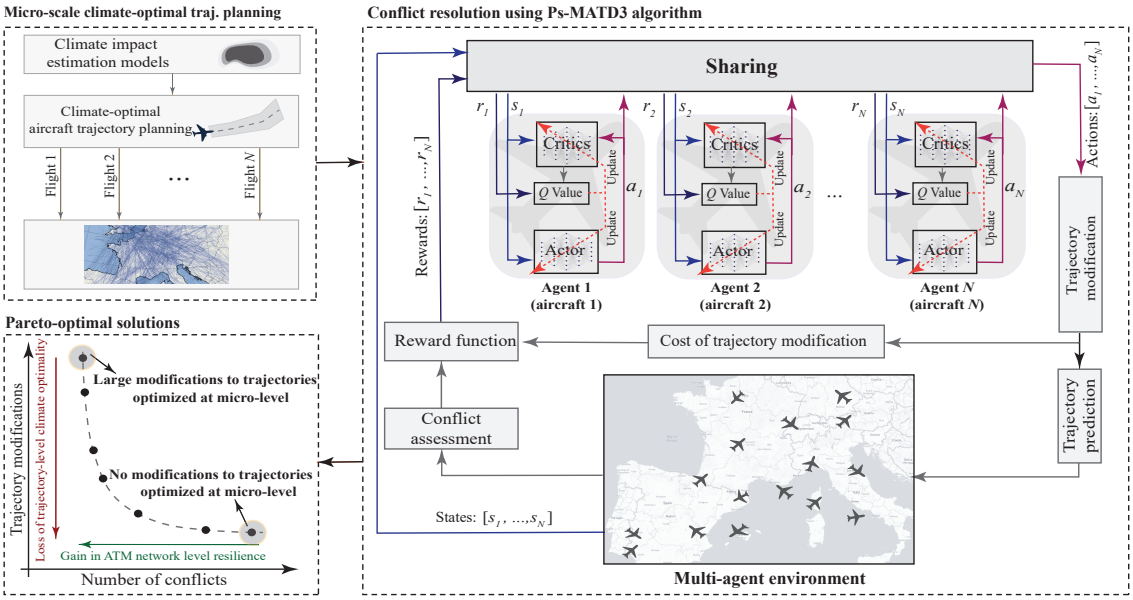


FIGURE 4.1: Workflow of the proposed MARL-based framework for climate-aware flight planning considering conflicts.

The methodology and results presented in this chapter are extracted from the published work in [30].

4.1 Micro-scale 3D climate-optimal flight planning

This section extends the trajectory optimization framework introduced in Chapter 3 by incorporating altitude as an additional decision variable, thereby enabling full three-dimensional (3D) optimization. Similar to Chapter 3, the optimization problem is formulated within the framework of optimal control theory. The key components required for this formulation, such as the aircraft dynamical model, the climate impact estimation method, and the objective function, remain consistent with those introduced in Chapter 3 (i.e., Section (3.1.1)). However, to account for the vertical dimension, the aircraft's dynamical behavior is now modeled using a three-degree-of-freedom point-mass model as:

$$\begin{bmatrix} \dot{\phi} \\ \dot{\lambda} \\ \dot{h} \\ \dot{v} \\ \dot{m} \end{bmatrix} = \begin{bmatrix} \left(v\cos\gamma_p\cos\chi + w_y\right)\left(R_M(\phi)+h\right)^{-1} \\ \left(v\cos\gamma_p\sin\chi + w_x\right)\left(\left(R_N(\phi)+h\right)\cos\phi\right)^{-1} \\ v\sin\gamma_p \\ \left(\mathrm{T}(C_T) - \mathrm{D}(C_L)\right)m^{-1} - g\sin\gamma_p \\ -f_c(C_T) \end{bmatrix}, \quad \begin{array}{l} \text{Latitude, Longitude} : \phi, \lambda \\ \text{Altitude, True airspeed} : h, v \\ \text{Mass, Fuel flow} : m, f_c \\ \text{Thrust \& Drag forces} : \mathrm{T}, \mathrm{D} \\ \text{Thrust \& Lift coefficients} : C_T, C_L \\ \text{Components of wind} : w_x, w_y \\ \text{Heading \& Path angles} : \chi, \gamma_p \end{array} \tag{4.1}$$

This extension introduces additional state and control variables to account for altitude variations. Accordingly, the state and control vectors are defined as follows:

$$\mathbf{x}(t) = \begin{bmatrix} \varphi & \lambda & h & v & m \end{bmatrix}^T, \qquad \mathbf{u}(t) = \begin{bmatrix} \chi & C_T & \gamma_p \end{bmatrix}^T.$$

Incorporating altitude as a variable also necessitates the introduction of altitude-specific constraints to ensure compliance with physical limitations and operational requirements. The complete set of extended constraints is given below:

$$\begin{aligned} v_{CAS,stall} \leq v_{CAS}(v) \leq v_{CAS,\max}, \\ C_{T,\min} \leq C_T \leq C_{T,\max}, \\ M(v) \leq M_{\max}, \\ h \leq h_{\max}. \end{aligned} \tag{4.2}$$

In addition to the physical and operational constraints defined above, the trajectory optimization problem is also subject to initial and final boundary conditions:

$$\begin{aligned} t(0) &= t_0, \\ \big[\varphi, \lambda, h, v, m\big](0) &= \big[\varphi_0, \lambda_0, h_0, v_0, m_0\big], \\ \big[\varphi, \lambda, h, v\big](t_f) &= \big[\varphi_f, \lambda_f, h_f, v_f\big]. \end{aligned} \tag{4.3}$$

The objective function used in this chapter is identical to that defined in Chapter 3 (Equation (3.11)), which combines operational cost and climate impact through a weighted sum (i.e., J = CI . SOC + C . EI . ATR). The solution method also follows the approach introduced in the previous chapter, i.e., the direct optimal control. All in all, the optimized trajectory for the flight i-th (i.e., trajectory-level optimization) is obtained as:

$$\Gamma_i^o := (\varphi_i^o, \lambda_i^o, h_i^o, v_i^o, t_i^o, m_i^o).$$

Having planned the climate-optimal trajectories for individual aircraft through the presented methodology, we integrate them into network-scale analysis to assess their effects on the manageability of the ATM system. To address any adverse impacts, resolution strategies are required aimed at re-stabilizing the ATM system without compromising climate optimality. In this respect, the next section will introduce our proposed framework, specifically designed to resolve conflicts that arise as a consequence of adopting climate-optimized trajectories.

4.2 Conflict resolution modeling using multi-agent reinforcement learning

As introduced in the previous chapter, a conflict arises when two aircraft are predicted to violate a predefined minimum safe distance, either horizontally or vertically. To prevent such scenarios, a safety buffer is considered around each aircraft, conceptualized as a virtual protected zone. This zone adheres to specific horizontal and vertical separation standards, establishing clear spatial safety boundaries. If another aircraft is projected to enter this zone, a conflict is identified. Resolving such conflicts involves a sequence of interdependent decisions, which can be effectively addressed using multi-agent reinforcement learning (MARL) approaches [178]. In decentralized settings, the MARL framework is often formalized as a decentralized partially observable Markov decision process (POMDP), defined by the tuple $\langle \mathcal{N}, \mathcal{S}, \mathcal{A}, \mathcal{P}, O, \gamma, R \rangle$, where:

- $\mathcal{N} = \{1, \ldots, N\}$ represents the set of agents (decision makers) within the interactive environment.
- $\mathcal{S}$ denotes the state space. The state $s \in \mathcal{S}$ encapsulates the current situation of the environment and can be viewed as the aggregate state of all agents, expressed as $s := s^1 \times s^2 \times \ldots \times s^N$.
- O is the observation space. o^i is the observation received by agent i, based on the state s.
- $\mathcal{A}$ is the joint action space, described as $\mathcal{A} := \mathcal{A}^1 \times \mathcal{A}^2 \times ... \times \mathcal{A}^N$. It includes all possible action combinations that agents can execute, where $\mathcal{A}^i$ represents the action set available to the agent i.
- $\mathcal{P}$ is the state transition function, $\mathcal{P} : \mathcal{S} \times \mathcal{A} \to \Delta(\mathcal{S})$, determining the probability of moving from the current state $s \in \mathcal{S}$ to the next state $s' \in \mathcal{S}$, given the joint action $\mathbf{a} \in \mathcal{A}$.
- $R : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ is the reward function stated as $R := R^1 \times R^2 \times ... \times R^N$.
- $\gamma \in [0, 1]$ represents the discount factor. This parameter balances the emphasis between immediate and long-term rewards. A lower γ leads the agent to prioritize immediate rewards, while a higher γ encourages a focus on long-term benefits. The specific value of γ is chosen based on the requirements of the application at hand.

At each time step t, agent i receives a local observation o_t^i and takes an action a_t^i according to its policy. The joint action of all agents $\mathbf{a}_t = (a_t^1, \ldots, a_t^N)$ is then applied to the environment, resulting in the transition of the environment to a new state $s_{t+1} \sim \mathcal{P}(\cdot|s_t, \mathbf{a}_t)$. Each

agent receives a reward r_t^i. In the cooperative setting, the goal of the agents is to maximize the expected total reward:

$$J \triangleq \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t, \mathbf{a}_t)\right],$$

In this section, the problem of conflict resolution is formulated as a MARL problem, where each aircraft in the airspace is considered an agent. Consequently, the airspace, encompassing all aircraft, is treated as an environment with N decision-makers involved. Within this framework, every aircraft is tasked with making critical decisions about its flight profile, aiming to cooperatively avoid conflicts with other aircraft in the airspace. The formulated problem to represent the decision-making processes of the aircraft will be described in detail in the following.

State space

The state s_t at time t is constructed by concatenating the local observations from all aircraft:

$$s_t = \left[o_t^1, \cdots, o_t^N\right].$$

Observation space

The observation of each aircraft is defined based on the required information for decision-making. At any given time t, for each aircraft operating in the airspace, the observation encompasses the heading, flight path angle, and current speed of the aircraft. These parameters can capture the intentions of the aircraft, thereby facilitating strategic decision-making within this dynamic environment. Given the multi-agent nature of this environment, which includes multiple interacting decision-makers, effective communication between aircraft is vital. This communication informs each aircraft about its surrounding traffic, forming an essential part of its observation representation. To achieve this, the observation of each aircraft includes information about neighboring aircraft. This information includes their heading, path angles, speed profiles, and the minimum distance relative to the observing (ownship) aircraft over the next time interval, i.e., until time step $t' = t + \Delta t$. Therefore, at time t for aircraft i, the observation is described as follows:

$$o_t^i = \{\chi_t^i, \gamma_{pt}^i, v_t^i, I_t^{i,1}, ..., I_t^{i,m}\}, \tag{4.4}$$

where χ_t^i, $\gamma_{p_t}^i$, and v_t^i represent the heading, path angle, and speed of aircraft i at time t, respectively. The term $I_t^{i,m}$ details the information of m^{th} neighboring aircraft as:

$$I_t^{i,m} = (\chi_t^m, \gamma_{p_t}^m, v_t^m, L_{ht}(m,i), L_{vt}(m,i)). \tag{4.5}$$

The parameters $L_{ht}(m,i), L_{vt}(m,i)$ are the minimum horizontal and vertical distance of aircraft m to aircraft i in the interval $[t,t']$, serving as a critical metric to assess proximity and a key component of the state representation.

Action space

In this study, the modification of aircraft speed profiles is considered as the action space for agents to mitigate the complexity in terms of potential conflicts. In addition to conflict resolution, agents are also tasked with maintaining their modified speed profiles as close as possible to the original climate-optimal trajectories. To support this objective, a continuous action space is adopted to allow for fine-grained control over speed adjustments. The preference for a continuous action space is based on the limitations associated with discrete action selection, particularly in terms of flexibility and precision in trajectory adjustments. Discrete actions can lead to deviations from the original, climate-optimized profiles due to their limited range of options. For instance, an aircraft might resolve a conflict effectively with a small speed adjustment, possibly only half the maximum value considered in a discrete system. However, without such intermediate options, an agent might be forced to select a larger, less optimal adjustment, resulting in an unnecessary deviation from the desired climate-optimal path. Therefore, for each aircraft i, we define a continuous action space A^i as follows:

$$A^i = \{\zeta | \zeta \in [v_{\min}, v_{\max}]\}. \tag{4.6}$$

At each time step t, the chosen action value ζ by agent i is added to the aircraft's optimal speed (determined during the optimization phase) for the interval $[t,t']$. Here, $v_{\min}$ and $v_{\max}$ are the minimum and maximum permissible speed changes, meeting physical and operational constraints.

Reward function

Modifying an aircraft's speed results in the transition of the environment to a new state through complex interactions among various aircraft, providing feedback in the form of a reward to each agent. In the context of our proposed methodology, the reward function is designed to fulfill a dual objective: mitigating the potential loss of separation and deviations from the climate-optimized trajectory. The reward for aircraft i at time t is expressed as:

$$r_t^i = \kappa \cdot C_{C_t}^i + C_{V_t}^i. \tag{4.7}$$

where κ is a weighting parameter that serves to regulate the relative importance of conflicts over the speed modifications. $C_{C_t}^i$ represents the cost of conflicts and can be quantified as

follows:

$$C^i_{C_t} = \sum_{\tau=t}^{t'} \sum_{j=1, j\neq i}^{N} -c^{ij}_{\tau}, \qquad c^{ij}_{\tau} = \begin{cases} 1 & \text{if} \quad d^{ij}_{\tau} < D_0 \quad \text{and} \quad h^{ij}_{\tau} < H_0 \\ 0 & \text{else} \end{cases} \tag{4.8}$$

Here, d^{ij}_{τ} and h^{ij}_{τ} represent the horizontal and vertical distances between aircraft i and j at time τ, with D_0 and H_0 being the respective minimum required separation standards. This component of the reward function prioritizes the maintenance of safe distances between aircraft, thereby reducing the likelihood of conflict situations.

Given that deviations from the reference speed can compromise the optimal performance originally established at the trajectory level, we incorporate a term in the reward function that penalizes speed adjustments. To address this, our model introduces a negative term within the reward function, explicitly designed to discourage excessive speed changes as:

$$C^i_{V_t} = -\sum_{\tau=t}^{t'} \omega^{i2}_{\tau}, \qquad \omega^i_t := ({v_{ct}}^i - {v_{ot}}^i)/(v_{max} - v_{min}), \tag{4.9}$$

where ${v_{ct}}^i$ and ${v_{ot}}^i$ are the modified and the original speed of aircraft i at time t, respectively, and $v_{\min}$ and $v_{\max}$ are the minimum and maximum allowable speed changes, respectively, given in Equation 4.6.

This criterion is designed to incentivize agents to prioritize actions that minimize potential conflicts by maintaining adequate separation distances while also adhering to optimized trajectories, ensuring the climate optimality of the flight plans.

4.3 Multi-agent deep reinforcement learning

The problem of conflict resolution in air traffic can be modeled as a cooperative game, where the efforts of individual agents contribute to the overall goal of the system. In recent years, MARL algorithms have emerged as powerful tools for addressing decision-making problems modeled as Markov games. Building on this foundation, we develop a MARL framework based on the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm, specifically tailored to address the complexities of the conflict resolution problem.

4.3.1 TD3 algorithm

The TD3 algorithm was originally developed for continuous control tasks in single-agent reinforcement learning settings. It represents a significant advancement over its predecessor, the deep deterministic policy gradient (DDPG) method [179], particularly in addressing the overestimation of Q-values [180].

At its core, DDPG integrates an actor and a critic network. The critic network employs a parameterized action-value function (Q_ϑ) to estimate the expected rewards associated with

the deterministic policy $\eta_\theta : S \to A$. The actor network selects deterministic actions based on the current state received from the environment. The critic network evaluates these actions taken by the actor and provides gradient information to guide the actor in updating its parameters θ. To stabilize the training, two target networks $\hat{Q}_{\hat{\vartheta}}$ and $\hat{\eta}_{\hat{\theta}}$ identical to the original networks, called target critic and target actor, respectively, are created, and their parameters are periodically updated by copying ϑ and θ from the original networks.

Building upon the foundation laid by DDPG, TD3 introduces several enhancements that improve its performance [181]. A key improvement in TD3, drawing inspiration from the double Q-learning approach, is the use of a dual-critic network. This approach reduces the likelihood of overestimating Q-values, a challenge often encountered in single-critic models like DDPG [180]. It involves the utilization of two distinct critic networks, Q_{ϑ_1} and Q_{ϑ_2}, each paired with a corresponding target network, $\hat{Q}_{\hat{\vartheta}1}$ and $\hat{Q}_{\hat{\vartheta}2}$. This dual-critic structure allows TD3 to compute two separate action-value estimates for the subsequent state:

$$\begin{aligned} \hat{Q}_{\hat{\vartheta}_1}(s', a') &= \hat{Q}_{\hat{\vartheta}_1}(s', \hat{\eta}_{\hat{\theta}}(s')), \\ \hat{Q}_{\hat{\vartheta}_2}(s', a') &= \hat{Q}_{\hat{\vartheta}_2}(s', \hat{\eta}_{\hat{\theta}}(s')). \end{aligned} \tag{4.10}$$

To address the issue of Q-value overestimation, TD3 selects the minimum of these two values to compute the target Q-value corresponding to the next state action pair:

$$y = r + \gamma \min_{j=1,2} \{\hat{Q}_{\hat{\vartheta}_j}(s', a')\}. \tag{4.11}$$

Here, r denotes the immediate reward for a single step and γ stands for the discount factor. Subsequently, the minimum is inserted into the Bellman equation to compute the loss function as:

$$L(\vartheta_j) = \mathbb{E}_{(s,a,s',r)}\left\{ (Q_{\vartheta_j}(s, a) - y)^2 \right\}, \quad j = 1, 2, \tag{4.12}$$

where the operator $\mathbb{E}\{\cdot\}$ signifies the mathematical expectation. The parameter ϑ is updated to minimize the expectation of the loss function concerning the actual and target Q-value using gradient descent as:

$$\vartheta_j = \vartheta_j - \alpha_c \nabla_{\vartheta_j} L(\vartheta_j), \tag{4.13}$$

where α_c is the learning rate of the critic and $\nabla_{\vartheta_j} L(\vartheta_j)$ is the gradient of loss function calculated as:

$$\nabla_{\vartheta_j} L(\vartheta_j) = \mathbb{E}_{(s,a,s',r)}\left\{ (Q_{\vartheta_j}(s, a) - y) \nabla_{\vartheta_j} Q_{\vartheta_j}(s, a) \right\}. \tag{4.14}$$

Similar to DDPG [141], the objective of the actor network in TD3 is to maximize the Q-value defined as:

$$J(\theta) = \mathbb{E}_{(s)}\left\{ Q_{\vartheta_1}(s, \eta_\theta(s)) \right\}. \tag{4.15}$$

Accordingly, the parameter θ is updated to maximize the Q value as:

$$\theta = \theta + \alpha_a \nabla_\theta J(\theta), \tag{4.16}$$

where α_a is the learning rate of actor network and $\nabla_\theta J(\theta)$ computed as:

$$\nabla_\theta J(\theta) = \mathbb{E}_{(s)} \Big\{ \nabla_\theta \eta_\theta(s) \nabla_a Q_{\vartheta_1}(s,a)|_{a=\eta_\theta(s)} \Big\}. \tag{4.17}$$

Another strategic modification in TD3 is the implementation of delayed policy updates. This approach differs from traditional reinforcement learning paradigms by updating the actor and target networks less frequently. Specifically, it introduces a fixed delay between these updates. In contrast to conventional methods, TD3 prioritizes more frequent updates to the critic (Q-value) networks, only adjusting the actor and target networks after this fixed delay interval. The deliberate delay in actor and target networks updates is designed to enhance overall stability, mitigate potential divergence issues, and improve the accuracy of Q-value estimation [180].

To mitigate over-fitting, which can result from discrepancies in estimation and the instability inherent in deterministic policy methods, it is essential to smooth the evaluation of the Q-value. This smoothing process involves a balance between minimizing estimation variations and controlling fluctuations. To achieve this equilibrium, one can incorporate clipped normal distribution noise into the output actions generated by the target actor network:

$$\overline{a} = \hat{\eta}_{\hat{\theta}}(s') + \sigma, \sigma \sim clip(\text{Normal}(0, \epsilon), -c, c). \tag{4.18}$$

With these insights into the TD3 algorithm, we propose a framework leveraging TD3 within a multi-agent environment. In the subsequent sections, we will delve into the details of our proposed framework, explaining the adaptation of TD3, its expected benefits, and our experimental findings, demonstrating its effectiveness in real-world multi-agent scenarios.

4.3.2 Policy-sharing multi-agent TD3 algorithm

In this section, we introduce the policy-sharing multi-agent TD3 (Ps-MATD3) algorithm, an adaptation of the TD3 framework designed to mitigate the potential conflict in real-world air traffic scenarios. The proposed framework is tailored for multi-agent cooperative environments and can handle an arbitrary number of concurrently operating agents. Within this framework, each aircraft is considered as an individual TD3 agent operating within a shared environment. To address the challenge of scalability in multi-agent systems, a crucial aspect in scenarios involving hundreds or thousands of agents, we propose a unique policy shared

across all agents. This shared policy approach not only enhances scalability but also simplifies the computational complexity typically encountered in optimizing policies for thousands of agents. Moreover, policy sharing facilitates deployment in dynamic environments with a variable number of agents, ensuring consistent and coordinated decision-making regardless of system scale.

Delving into the mechanism of the Ps-MATD3 algorithm, we consider that each agent i in the multi-agent system utilizes two centralized critic networks, $Q^i_{\vartheta_1}$, and $Q^i_{\vartheta_2}$, which are parameterized by ϑ_1 and ϑ_2, respectively. The critic networks are paired with corresponding target networks $\hat{Q}^i_{\hat{\vartheta}_1}$, and$\hat{Q}^i_{\hat{\vartheta}_2}$. In addition, the Ps-MATD3 algorithm incorporates a shared actor network, denoted as η_θ, parameterized by θ. The shared policy η facilitates a unified approach in decision-making, ensuring that each agent, despite operating individually, aligns with the collective objective of the system. Furthermore, we introduce a target actor network $\hat{\eta}_{\hat{\theta}}$, with the parameters of the target actor and target critics networks being periodically updated by copying parameters from the original networks.

At each time step t, for the agent $i \in [1, 2, ..., N]$, the observation o^i_t is input to the actor network to generate action $a^i_t = \eta_\theta(o^i_t)$. The actions generated by all agents are applied to the environment, resulting in the transition to new states and obtaining feedback rewards. These transitions are stored in a reply buffer Ξ as tuple $(\mathbf{O}, \Lambda, \mathbf{O}', \mathbf{R})$, where $\mathbf{O} : \{o^1, o^2, ..., o^N\}$, $\Lambda : \{a^1, a^2, ..., a^N\}$, $\mathbf{O}' : \{o'^1, o'^2, ..., o'^N\}$, $\mathbf{R} : \{r^1, r^2, ..., r^N\}$ represent the states, actions, next states and obtained rewards of all agents, respectively. Once one episode is done (i.e., all agents approach their terminal states), for each agent, the networks are trained using N_{batch} samples of Ξ.

However, one challenge in training networks is the non-stationarity of the environment [152]. In multi-agent systems, the actions taken by one agent not only affect its rewards but also influence the rewards and state dynamics experienced by other agents [182]. This causes instability in learning as it violates the stationary assumption required for the convergence of single-agent reinforcement learning algorithms. The assumption states that the environment follows a stationary Markovian property if the reward and current state depend solely on the previous state and action [182]. To mitigate this issue, fully observable critic networks are employed, wherein the critic is provided with access to the state and joint actions of all agents. Thus, the environment is stationary from the perspective of each agent. Even if the policy of agents changes during the training (i.e., $\pi^k \neq \pi'^k$), the environment is still stationary since, regardless of the change in the policy of others, the environment returns the same state as:

$$\mathcal{P}(s'|s, a^1, a^2, ..., a^N, \pi^1, \pi^2, ..., \pi^N) = \mathcal{P}(s'|s, a^1, a^2, ..., a^N, \pi'^1, \pi'^2, ..., \pi'^N). \tag{4.19}$$

Suppose that the tuple $(\mathbf{O}, \Lambda, \mathbf{O}', \mathbf{R})$ is one sample from Ξ. Each agent updates the parameters of its centralized critic networks as

$$\vartheta_j^i = \vartheta_j^i - \alpha_c \nabla_{\vartheta_j^i} L(\vartheta_j^i), \quad j = 1, 2. \tag{4.20}$$

Here, $L(\vartheta_j^i)$ represents the critic loss function, which is to be minimized. This function is defined as follows:

$$L(\vartheta_j^i) = \mathbb{E}_{(\mathbf{O},\Lambda,\mathbf{O}',\mathbf{R})\sim\Xi}\left\{ (Q_j^i(\mathbf{O}, a^1, a^2, ..., a^N) - y^i)^2 \right\}, \tag{4.21}$$

where $Q_j(\mathbf{O}, \Lambda)$ is the predicted value of the j critic network, and y^i is the target value obtained using:

$$y^i = r^i + \lambda \min_{j=1,2} \left\{ \hat{Q}_j^i \left(o'^1, o'^2, ..., o'^N, \hat{\eta}_{\hat{\theta}}(o'^1) + \sigma, \hat{\eta}_{\hat{\theta}}(o'^2) + \sigma, ..., \hat{\eta}_{\hat{\theta}}(o'^N) + \sigma \right) \right\}. \tag{4.22}$$

In Equation (4.22), $\hat{\eta}_{\hat{\theta}}(o'^i) + \sigma$ is the predicted action by target actor for agent i with a gaussian noise $\sigma \sim clip(\text{Normal}(0, \epsilon), -c, c)$ that is clipped and added for policy smoothing, and $\hat{Q}_j^i(o'^1, ..., o'^N, \hat{\eta}_{\hat{\theta}}(o'^1) + \sigma, ..., \hat{\eta}_{\hat{\theta}}(o'^N) + \sigma)$ is the Q-value predicted by target critic j.

Once the parameters of the critic networks are updated, the parameters of the actor network are updated as follows:

$$\theta = \theta + \alpha_a \nabla_\theta J(\theta), \tag{4.23}$$

where $\nabla_\theta J(\theta)$ is the gradient of the expected return computed as:

$$\nabla_\theta J(\eta_\theta) = \mathbb{E}_{(\mathbf{O},\Lambda)} \left\{ \nabla_\theta \eta_\theta(o^i) \nabla_{a_i} Q_1^i(\mathbf{O}, a^1, a^2, ..., a^N)|_{a^i = \eta_\theta(o^i)} \right\}, \tag{4.24}$$

After training, the parameters of the critic target networks are updated with a fixed delay for all agents as follows:

$$\hat{\vartheta}_j^{\,i} \leftarrow \hat{\alpha}_c \vartheta_j^i + (1 - \hat{\alpha}_c) \hat{\vartheta}_j^i, \tag{4.25}$$

where $\hat{\alpha}_c$ is the learning rate for critic targets. Finally, the parameters of the shared actor target are periodically updated after a predefined number of iterations as:

$$\hat{\theta} \leftarrow \hat{\alpha}_a \theta + (1 - \hat{\alpha}_a) \hat{\theta}, \tag{4.26}$$

where $\hat{\alpha}_a$ is the learning rate for the actor target. Once the training process is done, the trained critic networks are removed. Then, at each time step t, only the local observation of the agent i is required to be input into the trained actor model η^* to obtain the optimal action $a_t^i = \eta^*(o_t^i)$. The algorithm showing the training process based on the presented steps is demonstrated in Algorithm 2.

Algorithm 2 Ps-MATD3 algorithm.

Initialize the environment
Initialize the replay buffer Ξ
Initialize the actor network parameterized by θ
Initialize the actor target parameterized by $\hat{\theta}$
for agent i, $i = 1, ..., N$ **do**
  Initialize the critic networks parameterized by ϑ_1^i and ϑ_2^i
  Initialize the critic targets parameterized by $\hat{\vartheta_1}^i$ and $\hat{\vartheta_2}^i$
end for
for s = 1 to S_{max} **do**
  Reset the environment
  for time step = 0 to t_{max} **do**
    Receive the current observation of all agents $[o_t^1, ..., o_t^N]$
    Select actions a_t^i for $i = 1, ..., N$:

$$a_t^i = \eta_\theta(o_t^i) + \sigma, \quad \sigma \sim \text{Normal}(0, \epsilon)$$

    Execute joint action $\mathbf{a}_t = a_t^1, ..., a_t^N$ in the environment.
    Update the environment and get new observations $[o_t'^1, ..., o_t'^N]$ and reward r_t
    Store transition (s_t, a_t, r_t, s_t') in Ξ
  end for
  Randomly sample a mini-batch of l samples
  for agent i, $i = 1, ..., N$ **do**
    Compute target action for agent i:

$$a'^i = \hat{\eta}_{\hat{\theta}}(o'^i) + \sigma, \quad \sigma \sim clip(\text{Normal}(0, \epsilon), -c, c)$$

    Compute y^i according to Equation (4.22)
    Update ϑ_1, ϑ_2 using Equation (4.20)
  end for
  if s mod τ = 0 **then**
    Update θ using Equation (4.23)
    Update $\hat{\vartheta}$ and $\hat{\theta}_j$ from Equation (4.25) and Equation (4.26)
  end if
end for

4.3.3 Computational complexity analysis of Ps-MATD3

The complexity of the proposed Ps-MATD3 algorithm is analyzed in terms of computational and space complexity. This analysis is divided into two stages: the training phase and the execution phase.

Computational complexity: The computational complexity during the training phase primarily arises from two components: updating the critic networks and updating the actor network. To update the critic networks, each agent requires the target actions of all N agents. Specifically, for each of the N agents, we compute the actions of all N agents, resulting in a computational complexity of $\mathcal{O}(N^2 \times n_a)$, where n_a is the number of parameters in the actor network. Next, we need to calculate the Q-values based on the joint actions and observations of all agents. Each agent has two critic networks, as per the TD3 architecture, but the factor of two does not affect the order of complexity. Thus, the computational complexity of updating all critic networks is $\mathcal{O}(N \times n_c)$, where n_c represents the number of parameters in each critic network.

Second, for the actor network update, Ps-MATD3 requires updating only one shared actor network, with gradients computed from the critic networks of all N agents. Therefore, the total complexity for the actor update phase is $\mathcal{O}(n_a + N \times n_c)$, which includes the complexity of updating the actor network parameters $\mathcal{O}(n_a)$ and computing gradients from the N critic networks $\mathcal{O}(N \times n_c)$. Thus, the overall computational complexity per training step for Ps-MATD3 is:

$$\mathcal{O}(N^2 \times n_a + N \times n_c).$$

The primary difference between Ps-MATD3 and standard MATD3 in terms of computational complexity lies in the actor update phase. In standard MATD3, each of the N agents updates its own actor network, resulting in a complexity of $\mathcal{O}(N \times n_a + N \times n_c)$. Ps-MATD3 reduces this complexity by utilizing a single shared actor network for all agents, lowering the actor update complexity to $\mathcal{O}(n_a + N \times n_c)$. However, despite this reduction, the total computational complexity for both Ps-MATD3 and standard MATD3 is dominated by the term $\mathcal{O}(N^2 \times n_a)$, due to the necessity of computing the target actions for all agents during the critic updates. Consequently, the overall complexity of both algorithms for the training phase remains of the same order:

$$\mathcal{O}(N^2).$$

The computational complexity in the execution phase of Ps-MATD3 is similar to MATD3, where each agent independently selects its action, resulting in a computational complexity of:

$$\mathcal{O}(N \times n_a).$$

Space Complexity: Each agent in Ps-MATD3 has two critic networks, resulting in a

space requirement of $\mathcal{O}(N \times n_c)$ for the critics. Since there is only one shared actor network, the space required for the actor network is $\mathcal{O}(n_a)$. Thus, the total space complexity for Ps-MATD3 is:

$$\mathcal{O}(N \times n_c + n_a).$$

In contrast, standard MATD3 requires $\mathcal{O}(N \times n_a)$ space for the actor networks, as each agent has its own actor network. Therefore, the total space complexity for standard MATD3 is $\mathcal{O}(N \times n_c + N \times n_a)$.

During the execution phase of Ps-MATD3, only the actor network is required, resulting in a space complexity of $\mathcal{O}(n_a)$. By comparison, MATD3 requires each agent to have its own actor network, leading to a higher space complexity of $\mathcal{O}(N \times n_a)$.

4.4 Simulation results

This section presents a case study to demonstrate the effectiveness of the proposed methodology. The components of the simulation framework are outlined, including the experimental setup, scenario definition, neural network architecture, parameters used for trajectory planning and conflict resolution, and the results obtained.

4.4.1 Experimental details

This experiment employs a real large-scale scenario representing the traffic on December 20, 2018, within the European Civil Aviation Conference (ECAC) airspace. The analysis focuses specifically on flights operating between 12:00 and 14:00 UTC. Flight data, including boundary conditions required for aircraft trajectory optimization (e.g., initial flight time and coordination of initial and final waypoints), is extracted from the DDR2 dataset[1]. In cases where flights depart or arrive outside the ECAC airspace, only the flight segments within this airspace during the specified time frame are considered.

The weather information necessary for the proposed optimizations (i.e., trajectory planning and conflict resolution using Ps-MATD3), including wind, temperature, humidity, radiation, potential vorticity, and geopotential, is obtained from the ERA5 reanalysis data products. These data are publicly available at the Copernicus Data Store and offer a resolution of $0.5^\circ \times 0.5^\circ$ with a temporal granularity of three hours (for the ensemble data).

The flight plans are optimized using the methodology presented in Section 4.1, with weighting parameters in the objective function Equation (3.11) set to $C = 10^{10}$ [USD/K], $C_t = 0.75$, and $C_f = 0.51$. The Trapezoidal rule is applied to convert the formulated optimal control problem into an NLP problem. Subsequently, the NLP problem is solved using the interior-point method by employing the IPOPT solver in Python.

[1] `https://www.eurocontrol.int/ddr`

A range of routing options is generated by varying the parameter penalizing climate effects (EI $\in [0.0, 0.001, 0.01, 0.1, 1.0, 10.0]$) in Equation (3.11), ranging from cost-optimized (EI = 0.0) to climate-optimized routes (EI = 10).

We implemented the Ps-MATD3 algorithm for the conflict resolution problem with continuous action spaces. The algorithm is applied in a custom environment where each aircraft is considered as an agent, and the airspace encompassing all agents is the MARL environment. A single actor network is shared among all agents. This actor network is responsible for generating the actions for each agent based on its local observations. The actor network is configured as a fully connected neural network comprising three layers. These layers consist of 300, 200, and 80 nodes, respectively. The design of the input layer is specifically tailored to align with the observation dimension of each agent. For the output of the actor network, we employ a tangent activation function, which generates a one-dimensional continuous value for speed adjustment.

Each agent has two critic networks to estimate the Q-values of state-action pairs. These networks are structured as fully connected neural networks, each containing two hidden layers with 80 nodes. As the proposed Ps-MATD3 algorithm utilizes fully observable critic networks to avoid non-stationarity during training, critic networks take the joint actions and observations of all agents. The critics follow the clipped double Q-learning strategy, where the minimum Q-value between two critics is used to reduce overestimation bias. The ReLU activation function is applied to all hidden layers, and the outputs of the critics employ linear activation functions. To enhance the stability of the training process, target networks, identical to the original actor and critics, are created, with their parameters periodically updated from the original networks. The Adam optimizer is utilized for both actor and critic loss functions.

A centralized replay buffer is implemented to store agents' experiences in the form of (state, action, reward, next state) tuples. The replay buffer size is set to 10^6. Experiences are sampled from this buffer during training to update the networks. To encourage exploration during training, Gaussian noise with a standard deviation of 5.0 is added to the output of the action by the actor network. The standard deviation of the noise is decayed over time to allow for more stable policies as training progresses. Additionally, noise is added to the target action during the critic update, and this noise is clipped to the range [-5, 5] to prevent excessive deviations and ensure smoother value updates. After every step, the critic networks are updated using the mean squared error between the predicted Q-values and the target Q-values, which are computed using the target networks. The actor network is updated less frequently (every two critic updates), following the delayed policy update strategy from TD3. Both the actor and critic networks use target networks, which are updated using a soft update factor 10^{-3}. A detailed summary of the hyperparameters used in the training process is provided in Table 4.1.

TABLE 4.1: The set of training parameters and their values.

Parameter	Value
Time step (Δt)	600 s
Discount factor for future rewards (γ)	0.95
Learning rate of updating target networks $(\hat{\alpha}_a, \hat{\alpha}_c)$	10^{-2}
Size of mini-batch (N_{batch})	20
Exploration noise std (σ)	5.0
Maximum time step in each episode	13
Learning rate of actor network (α_a)	10^{-3}
Learning rate of critic networks (α_c)	10^{-3}

To train the proposed Ps-MATD3 algorithm, we select the trajectories associated with EI = 0.0 (i.e., cost-optimal trajectories). These trajectories serve as the baseline, representing standard flight paths, and provide a reference point for our algorithm's training process. Within the proposed framework, aircraft not part of any conflict maintain their original flight profiles, thereby assigning the responsibility of conflict resolution exclusively to the aircraft directly involved in the conflict. This is because allowing all aircraft to modify their profiles can significantly increase the complexity of the problem. Furthermore, from the operational perspective, it is preferable for aircraft not engaged in conflicts to adhere to their predetermined optimal flight plans to maintain overall system efficiency and stability. This approach ensures that the algorithm focuses on conflict resolution where it is most needed while minimizing unnecessary adjustments to non-conflicting aircraft.

Each agent's state includes its current position (i.e., latitude, longitude, altitude), velocity, heading angle, and the relative positions and velocities of nearby aircraft within a radius of 25 nautical miles (NM) and a height of 3000 feet (ft). At each time step $\Delta t = 10$ minutes, agents receive their state and take actions in the range $v_{\min} = -30$ knots to $v_{\max} = 30$ knots. The selected action is applied uniformly to all segments of the trajectory during the time step. For instance, if an agent chooses to increase its speed, the selected action is added to the current speed at that interval, and the entire flight profile is updated to reflect the new speed. Accordingly, all information related to the trajectory, such as data on neighboring aircraft, is updated. Once all agents select their actions, the environment is updated, and conflicts between aircraft are computed. The separation criterion for having a conflict is defined as a horizontal distance of $D_0 = 5$ NM and a vertical distance of $H_0 = 1000$ ft, consistent with the definition used in Chapter 3. Therefore, any pair of aircraft that is predicted to violate this separation receives a penalty of -1 (negative reward). The total penalty for each agent increases based on the number of potential conflicts with other agents. Additionally, a negative reward proportional to the magnitude of the speed change is applied to discourage unnecessary deviations from the original trajectory. Specifically, the penalty is calculated

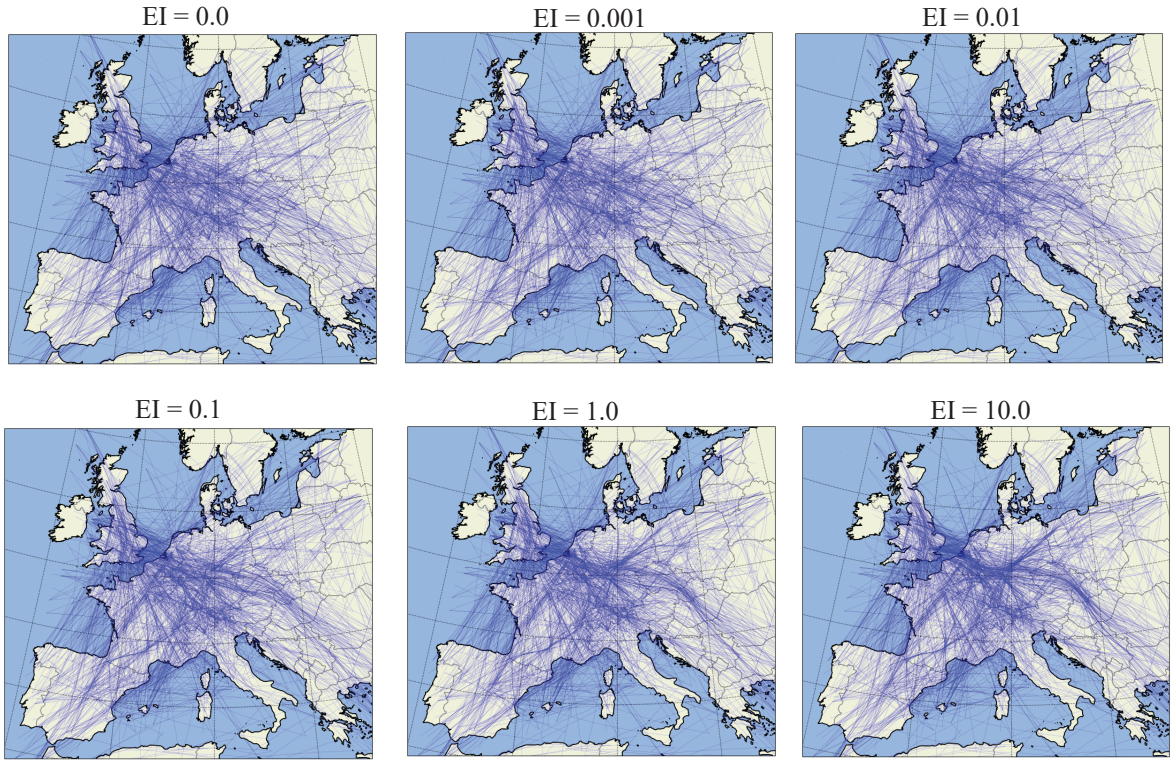


FIGURE 4.2: Lateral routes of optimized trajectories associated with different EI's.

using Equation (4.9). The parameter κ in Equation (4.7) is set to 0.58, determined through the tuning process to balance effective conflict resolution with minimal trajectory deviation.

4.4.2 Climate-optimal flight planning

To assess the impact of climate considerations on flight trajectories, we generate six sets of alternative trajectories. Each set is associated with a specific EI value, covering a range from cost-optimized trajectories to trajectories with an increased level of climate optimality. After optimizing flight plans, they are integrated into the ATM system for performance evaluation. The optimized lateral paths for different routing options (i.e., different EI values) are shown in Figure 4.2. The figure highlights the impact of varying EI values on the traffic pattern. It can be seen that for lower EI values, the distribution of traffic flow is relatively uniform across the airspace. However, as EI values increase, a noticeable concentration of traffic flow emerges in specific areas.

Figure 4.3 illustrates the climate impact of the optimized trajectories, evaluated both in terms of the net climate effect and the contributions of individual species. The results show that increasing EI leads to a decrease in the net ATR. Notably, optimizing flight trajectories can contribute to more than just mitigating warming effects; it also causes cooling effects

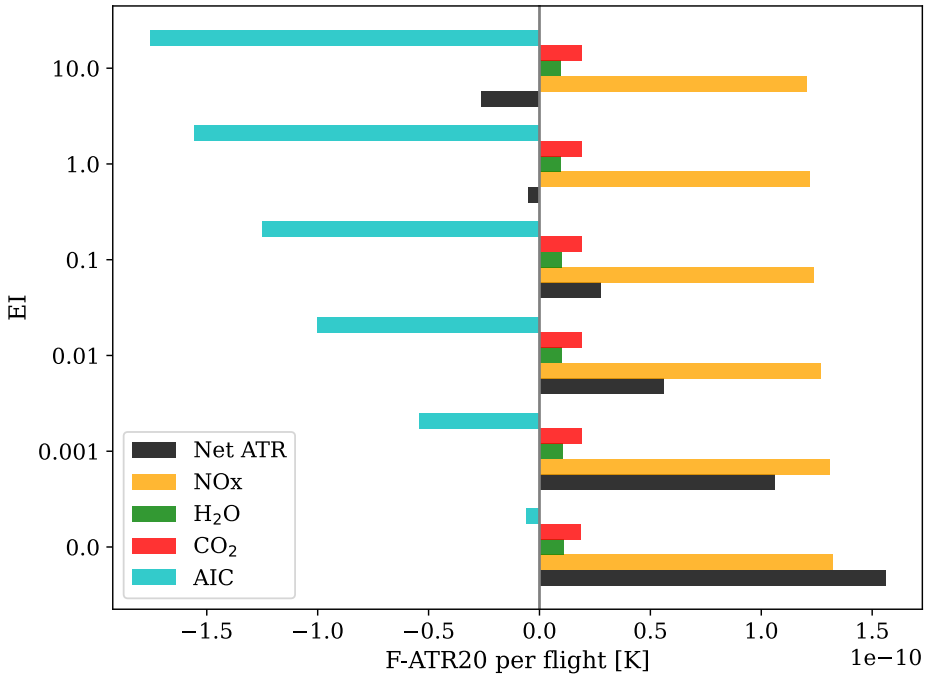


FIGURE 4.3: Climate effects associated with individual species for different sets of optimized trajectories.

(i.e., EI $\in$ $[1.0, 10.0]$). This outcome arises from forming cooling contrails while simultaneously minimizing the generation of warming contrails (see Aviation-Induced Cloudiness (AIC) in Figure 4.3). Such behavior is in line with related studies employing aCCFs, stating the climate impact mitigation potential is generally achieved through reducing the impact of contrails (see [183]). Figure 4.4 shows the geographical distribution of formed persistent contrails for different sets of flight trajectories. It is evident that trajectories associated with higher EI values generate more cooling contrails while simultaneously reducing the formation of warming contrails. One can see in Figure 4.2 that locations where the traffic density is increased match the areas where cooling contrails are formed (refer to Figure 4.4). Indeed, these specific areas have a heightened potential for cooling contrail formation, leading aircraft to fly through these regions.

While trajectory planning offers opportunities to reduce aviation's climate impact (as shown in Figure 4.3), it is also crucial to evaluate its effects on other performance metrics. In particular, in this chapter, we focus on the operational cost and potential conflicts. To assess the operational cost, we utilize the SOC, which is calculated using a weighted sum of flight time and fuel consumption (given in Equation (3.11)). The results are detailed in Figure 4.5. The findings indicate that adopting flight trajectories with lower climate impact increases operating costs. This trend is due to the behavior of aircraft in response to different climate conditions: they tend to avoid areas contributing to warming effects while seeking to cross

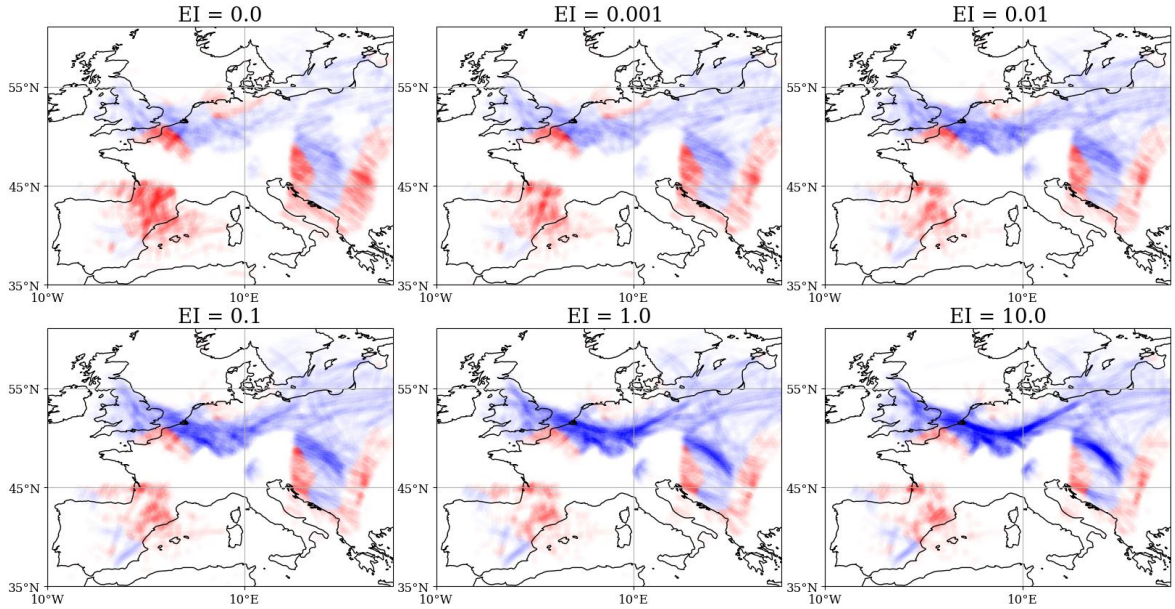


FIGURE 4.4: Geographical distribution of persistent warming and cooling contrails across various sets of optimized flight plans. Formed warming contrails are indicated in red, while cooling contrails are depicted in blue.

regions conducive to cooling impacts. While these optimized flight plans are effective for climate mitigation, they generally result in longer routes, thus increasing operational costs.

In addition to the operating cost, the traffic concentration within/around climate hotspots increases traffic complexity. The congestion arises as multiple aircraft fly through the areas with high potential to form cooling and less warming contrails, resulting in proximity and potential conflicts between flights (see Figure 4.5). A comparison of different trajectory sets reveals that adopting trajectories with lower climate impacts is associated with an increase in the number of potential conflicts.

The locations of conflicts for different sets of trajectories are illustrated in Figure 4.9, providing a visual understanding of areas where climate-optimized trajectories potentially compromise air traffic manageability. Notably, Figure 4.9 illustrates an increase in the number of potential conflicts at lower altitudes for trajectories with higher climate impact mitigation. One justification for this behavior is the possibility of generating cooling contrails at lower flight levels for this scenario. As a result, these specific flight levels experience an increase in traffic flow, leading to a rise in potential conflicts.

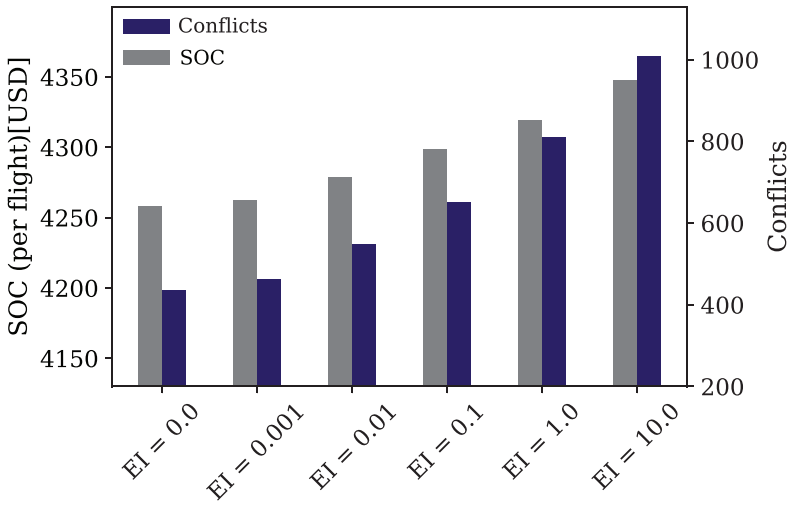


FIGURE 4.5: Operating cost and number of potential conflicts for different sets of optimized flight plans (different EI's).

4.4.3 Mitigating potential conflicts using Ps-MATD3

The results indicate that efforts to mitigate climate impact through trajectory optimization lead to increased traffic concentration in specific regions. This redistribution of traffic flow contributes to a heightened number of potential conflicts, which may compromise the operational feasibility of the optimized flight plans. Addressing this challenge, it is imperative to strategically (i.e., at the planning phase) manage traffic complexity. To this end, we employ the method detailed in Section 4.3 to resolve the encountered conflicts, ensuring the feasibility of climate-optimal trajectories. Our objective is to create a balance between the climate benefits achieved through trajectory optimization and the operational requirements of the ATM system.

The MARL environment is structured according to the state space, action space, and reward function presented in Section 4.2. The proposed Ps-MATD3 is compared against three algorithms: the original MATD3, multi-agent deep deterministic policy gradient (MADDPG), and policy-sharing MADDPG (Ps-MADDPG). Figure 4.6 illustrates the reward performance of these four algorithms, where the vertical axis represents the obtained rewards, and the curves are averaged over 10000 episodes.

As shown in Figure 4.6, Ps-MATD3 achieves consistently higher rewards compared to the other methods. When compared with the original MATD3, which does not employ policy sharing, Ps-MATD3 consistently outperforms MATD3 in terms of total reward. This improvement can be attributed to the ability of policy sharing in Ps-MATD3, which accelerates learning by enabling agents to benefit from shared experiences.

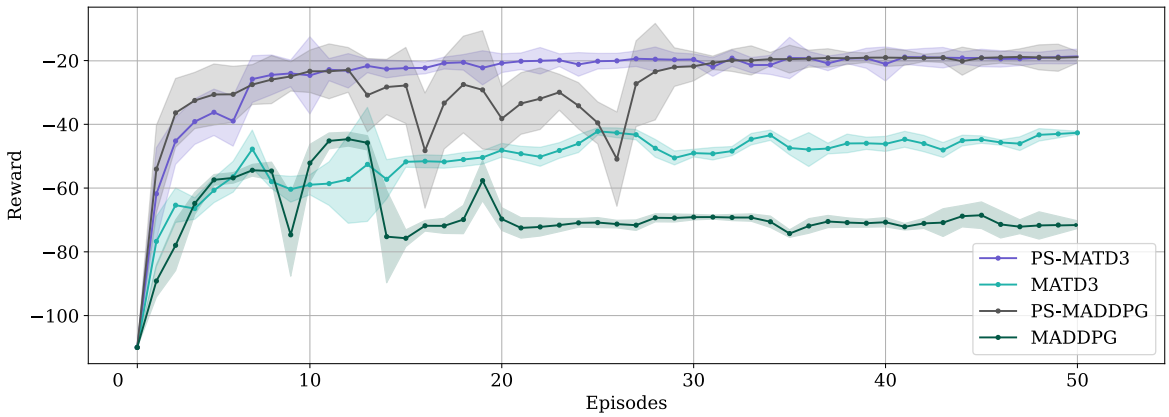


FIGURE 4.6: Comparison of reward performance across four multi-agent reinforcement learning algorithms: Ps-MATD3, MATD3, Ps-MADDPG, and MADDPG. Each point corresponds to the mean reward over 10,000 episodes, and shaded regions show the standard deviation.

Ps-MATD3 demonstrates stable reward curves, likely due to the twin-critic structure and delayed policy updates of the TD3 algorithm, which help stabilize the learning process. In contrast, Ps-MADDPG, while competitive in terms of reward, exhibits higher fluctuations in its reward curve. Finally, the original MADDPG algorithm shows comparatively inferior performance in terms of rewards and exhibits high variability in the learning curve. All in all, the results indicate that the policy-sharing mechanism (i.e., Ps-MATD3 and Ps-MADDPG) improves performance and enhances the stability of learning compared to their counterparts that do not utilize policy-sharing.

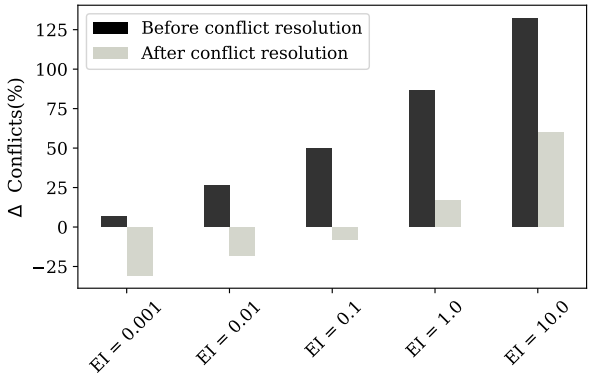
Table 4.2 summarizes the performance metrics of the proposed Ps-MATD3 algorithm compared to other methods. The results demonstrate that Ps-MATD3 converges faster to higher rewards, achieving an optimal performance after approximately 110k episodes. In contrast, MATD3 requires more training steps (around 350k episodes) to converge and still fails to reach the same reward level as Ps-MATD3. This faster convergence reduces computational complexity, as fewer iterations are needed to achieve optimal performance. Additionally, Ps-MATD3 has a shorter training time per episode (0.24 seconds vs. 0.26 seconds for MATD3), further decreasing the computational burden. This reduction is attributed to the policy-sharing mechanism, which decreases the number of networks that need to be trained, streamlining the learning process.

When comparing Ps-MATD3 to MADDPG and Ps-MADDPG, it is evident that Ps-MATD3 outperforms both in terms of stability and reward performance. Although Ps-MADDPG demonstrates improvement in training times per episode, it requires more iterations to achieve the same level of stability and reward as Ps-MATD3. Consequently, despite the shorter per-episode training time, the overall time to converge is longer for Ps-MADDPG. MADDPG, without policy sharing, experiences slower convergence and lower reward levels, indicating

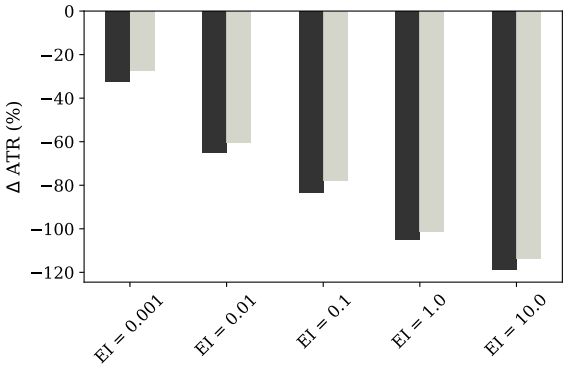
TABLE 4.2: Performance comparison of the Ps-MATD3 algorithm with MATD3, MADDPG, and Ps-MADDPG.

	Convergence pint (Episode)	Training time per episode (s)	Final reward
Ps-MATD3	$\sim 110k$	0.24	-19
MATD3	$\sim 350k$	0.26	-41
Ps-MADDPG	$\sim 290k$	0.19	-19
MADDPG	$\sim 210k$	0.21	-72

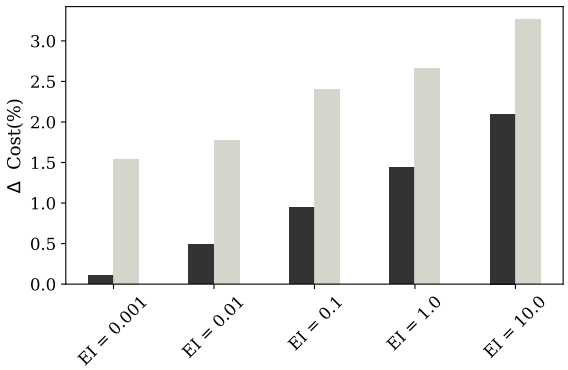
that policy sharing contributes to both enhanced learning efficiency and reduced computational complexity.



(A) Relative differences of the number of potential conflicts from cost-optimal trajectories (EI = 0.0) for different sets of climate-optimized trajectories.



(B) Relative differences of climate impact from cost-optimal trajectories (EI = 0.0) for different sets of climate-optimized trajectories.



(C) Relative differences of the operating cost from cost-optimal trajectories (EI = 0.0) for different sets of climate-optimized trajectories.

FIGURE 4.7: Comparison between the climate impact, operating cost, and number of conflicts before and after applying the proposed resolution methodology.

The policies derived from PS-MATD3 are tested across different sets of climate-optimal trajectories, i.e., EI $\in [0.001, 0.01, 0.1, 1.0, 10.0]$, to modify the speeds of conflicting aircraft.

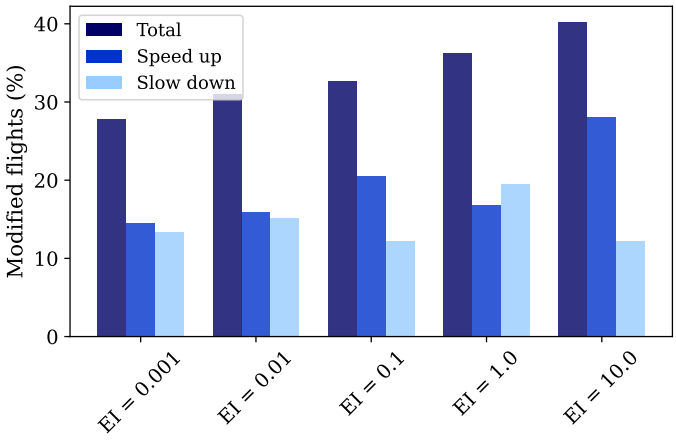


FIGURE 4.8: Percentage of flights subject to speed adjustment (speed up, slow down, and total modifications) after implementing conflict mitigation strategies for different sets of climate-optimized flight plans.

For each aircraft in conflict, the speed profile is modified according to the action suggested by the agent at each time step, and the flight trajectory is updated based on the new velocity profiles. Subsequently, we reassess the performance of the modified flight plans in terms of ATR, SOC, and the number of conflicts to analyze the effectiveness of the proposed conflict resolution strategy.

A comparative analysis of flight performance, assessed before and after the implementation of the resolution strategy, is presented in Figure 4.7. This figure highlights the differences in potential conflicts, ATR, and SOC compared to cost-optimal trajectories as the reference. The findings reveal that the policy developed through this study reduces traffic complexity in all cases. For most sets of optimized flight plans, it successfully lowers the number of conflicts to levels close to or even below those encountered in scenarios with cost-optimal trajectories. However, as indicated in Figure 4.7, achieving this conflict reduction is associated with increased operational costs, which rise by less than 1.5% in all scenarios. Additionally, there is a compromise in climate impact mitigation, decreasing by a maximum of 5% over all sets of flight plans. Such a trade-off is expected as we deviate from the optimized routes determined in the trajectory optimization phase.

All in all, the presented results indicate that the proposed conflict-resolution strategy contributes to improving the balance of the ATM system by enhancing manageability, albeit with a slight compromise in operating cost and climate performance.

Figure 4.10 illustrates the conflict locations following the implementation of the proposed strategy, indicating a substantial reduction in traffic conflicts, especially in high-density areas (i.e., areas where cooling contrails are formed in Figure 4.4). This decrease in potential conflicts demonstrates the strategy's efficiency in improving overall traffic manageability. The percentage of flights that adjusted their speed to minimize conflict occurrences is illustrated

in Figure 4.8. This analysis shows a clear trend: with an increase in EI value, the percentage of modified flights also increases. This indicates that trajectories with lower climate impact require more modifications to re-stabilize the ATM system.

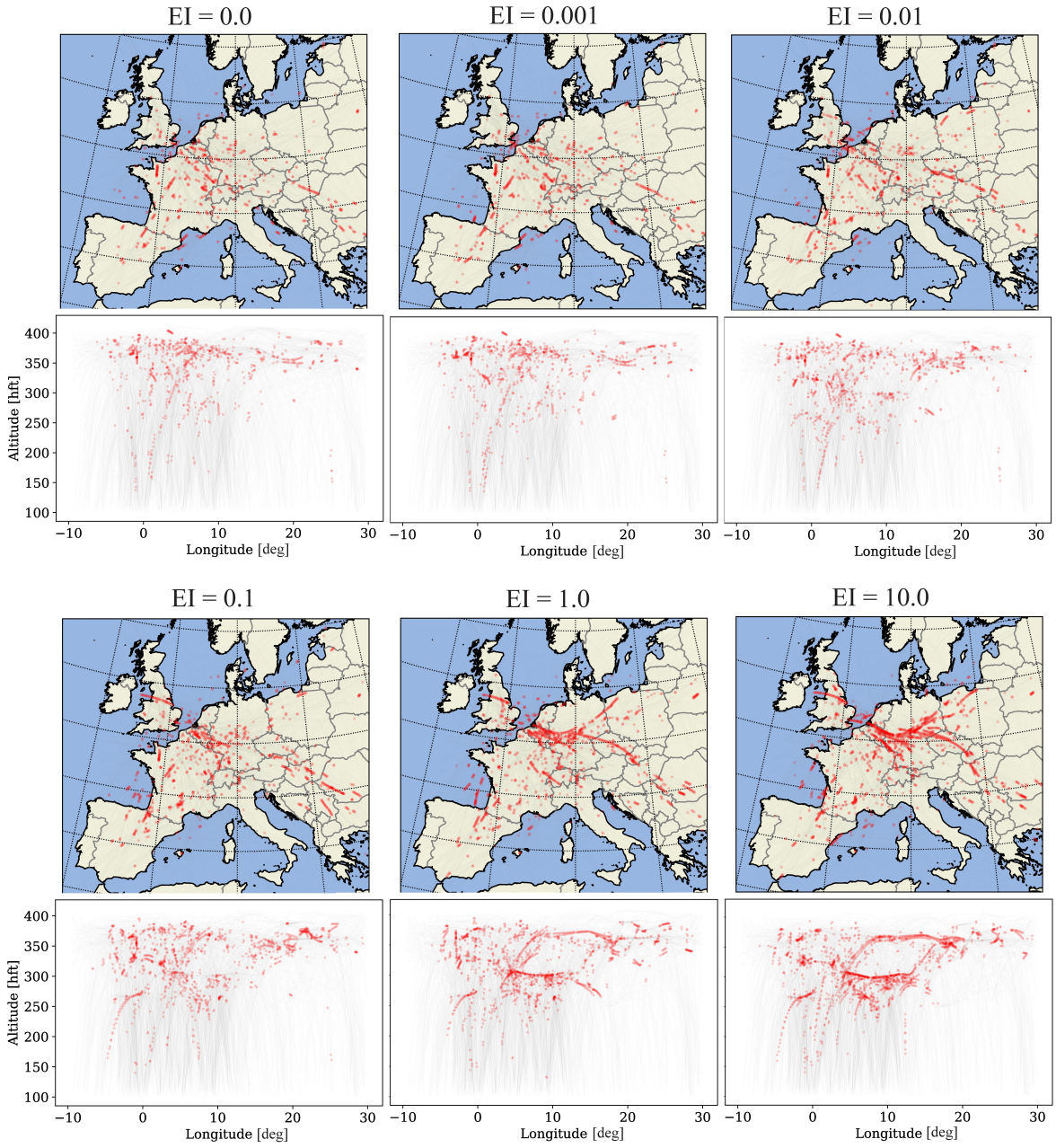


FIGURE 4.9: Locations of potential conflicts for different sets of optimized trajectories pre-resolution implementation: red circles indicate potential loss of separation between aircraft, increasing as trajectories with lower climate impact are adopted.

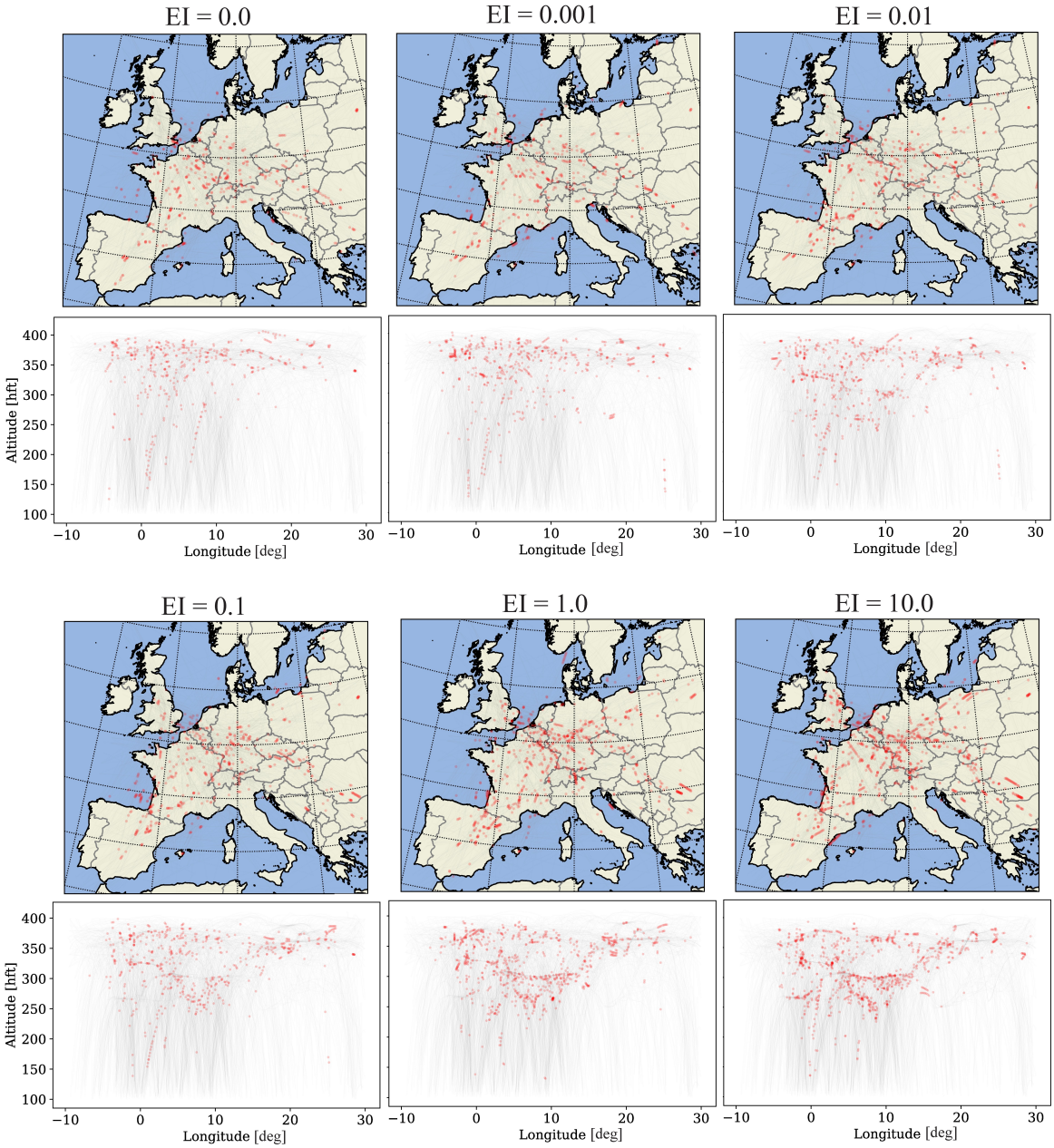


FIGURE 4.10: Locations of potential conflicts for different sets of optimized trajectories post-resolution implementation. Red circles indicate potential loss of separation between aircraft.

4.5 Summary

This chapter presented a cooperative decision-making framework employing MARL to plan operationally feasible climate-friendly routes from the perspective of the ATM system. The proposed strategy leveraged the TD3 algorithm to adjust flight trajectories during the planning phase to resolve the potential conflicts associated with climate-optimal trajectories.

The performance of the framework was evaluated using a real-world traffic scenario over the ECAC airspace. A set of alternative trajectories, spanning from cost-optimal to climate-optimal routing options, was generated using a micro-scale trajectory optimization approach. The results indicated that the climate benefits reported from individual trajectory optimization approaches might not be fully achievable in practice due to their potential impact on ATM system performance, particularly through an increased number of conflicts. For instance, achieving an 85% reduction in climate impact was associated with a 50% increase in the number of conflicts.

To address this challenge, the Ps-MATD3 algorithm was introduced as a resolution strategy aimed at balancing environmental objectives with traffic complexity. Although not all conflicts were resolved, mainly due to the limited action space, the proposed approach demonstrated the potential to improve traffic manageability by reducing the number of predicted conflicts. Restricting decision space to only speed changes, up to 80% climate impact reduction was achievable while decreasing potential conflicts by 10% compared to standard business-as-usual trajectories.

One of the key findings of this study is the advantage of policy sharing between agents. Sharing policy parameters across aircraft improves the learning process by enabling more efficient use of data, as agents benefit from collective experience, reducing the time and computational resources required for convergence. Another advantage of policy sharing is that it allows training in smaller traffic scenarios while still being applicable to larger, more complex environments. This is possible because conflict resolution typically depends on interactions with neighboring aircraft, making the shared policy effective regardless of the overall size of the scenario. Without policy sharing, training policies on smaller subsets may not generalize well to larger scenarios. In such cases, scaling these individually trained policies to handle a significantly larger number of agents becomes unclear, and the entire system may require retraining. Shared policies avoid this issue by being adaptable, regardless of the number of agents in the environment.

Chapter 5

Network-scale climate-optimized flight planning in structured airspace: Multi-agent RL approach for complexity management

The previous chapter introduced a sequential framework for network-scale climate-optimal flight planning within fully free-routing airspace. While the proposed method demonstrated potential in balancing environmental benefits with operational feasibility, measured in terms of conflict reduction, it might reveal limitations from the operational perspective. In particular, relying on the number of conflicts as a proxy for air traffic manageability during the planning phase may be insufficient, as conflicts are typically assessed over short time horizons and depend on precise trajectory information, which is inherently uncertain at this stage. To enable a more robust evaluation, it is essential to incorporate performance indicators that operate over longer time horizons and are more closely aligned with the characteristics of trajectory-based operations. Moreover, to assess the real-world applicability of climate-aware flight planning, it is necessary to move beyond the context of futuristic fully free-routing scenarios and consider the constraints imposed by current structured airspace.

Motivated by these needs, this chapter introduces a MARL framework for network-scale climate-optimal flight planning within the constraints of current structured airspace, using air traffic complexity as the primary performance metric to evaluate air traffic manageability. Building upon the sequential optimization framework presented in the previous chapters, the proposed approach follows a similar multi-step structure (see Figure 5.1). First, individual trajectories are optimized to mitigate the climate effects. Next, the collective performance of these optimized trajectories is evaluated in terms of air traffic complexity. Finally, to address the unintended increase in complexity that arises from adopting independently optimized trajectories, we introduce a cooperative resolution strategy based on Multi-Agent Proximal Policy Optimization (MAPPO).

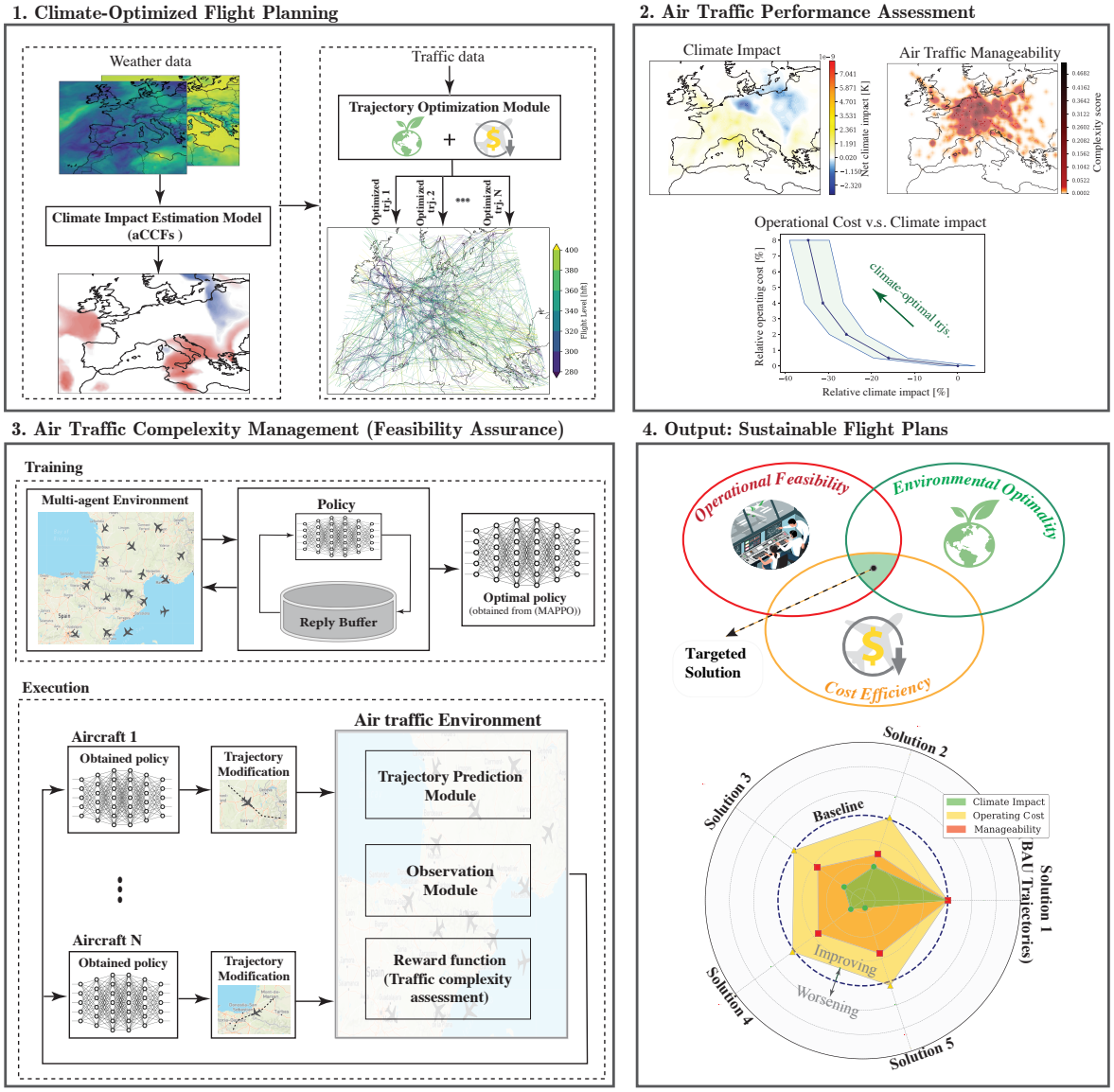


FIGURE 5.1: Workflow of the proposed MARL-based framework for climate-aware flight planning considering traffic complexity.

The resolution problem is formulated as a decision-making process in which each aircraft adjusts its speed and altitude profiles to mitigate traffic complexity while preserving climate benefits. Similar to the previous chapter, the proposed framework follows a centralized learning and decentralized execution scheme. In this setup, the policy is trained in a centralized manner, but during execution, each aircraft applies the policy independently based on the available information. To ensure scalability in scenarios involving varying numbers of agents, shared policy parameters are implemented, providing flexibility to accommodate diverse air traffic conditions and ensuring the framework can adapt to dynamic operational environments.

The remainder of this chapter is organized as follows. Section 5.1 introduces the micro-scale trajectory optimization framework within structured airspace. Section 5.2 outlines the methodology used to evaluate air traffic complexity. Section 5.3 presents the decision-making strategy to mitigate the increased complexity associated with climate-optimized routes. Finally, section 5.4 details the simulation setup, including scenario definitions and the configuration of the optimization tool, and presents and discusses the findings of this study.

We note that the content in this chapter is mainly extracted from papers [28] and [31].

5.1 Micro-scale trajectory optimization within structured airspace

The trajectory planning approach outlined in the previous chapters assumes that the lateral routes and altitude profiles of the aircraft can be freely determined between predefined departure and arrival points to minimize a given objective function. While this framework aligns with the flexibility envisioned in future fully free-routing airspace, it does not fully reflect the constraints imposed by current operational practices. In reality, aircraft must follow structured route networks and operate at designated flight levels, which restricts trajectory planning flexibility.

Transitioning from free routing to structured airspace introduces a fundamental shift in the optimization formulation. Specifically, decision variables in the problem shift from a continuous decision space to a hybrid decision space comprising both continuous and discrete variables. This shift significantly increases the complexity of the problem, placing it within the domain of mixed-integer nonlinear programming (MINLP). To tackle this optimization problem more efficiently, instead of optimizing a sequence of control inputs ($\mathbf{u}^o$), we directly

optimize the flight plan, defined as $F : \{\overline{\mathrm{M}}, R, \overline{\mathrm{FL}}, C, D, d_D\}$ [27], where

$\overline{\mathrm{M}}$: Mach schedule
R : Lateral path
$\overline{\mathrm{FL}}$: Flight level of the cruise phase
C : Climb profile
D : Descent profile
d_D : Distance-to-go to the destination node.

The direct optimization of flight plans, rather than control and state variables, enables more efficient handling of structured airspace constraints. This is achieved by directly searching over admissible flight plan variables that comply with the air traffic services (ATS) route network, rather than imposing these requirements as path constraints, which would otherwise lead to a more complex optimization problem. The goal is to determine a flight plan that minimizes the following objective function, which includes both operating cost and ATR, while satisfying the aircraft dynamical model (i.e., Equation 4.1) and complying with the underlying constraints, such as those imposed by structured airspace and the flight envelope:

$$J(F) = \mathrm{CI} \cdot \Big[\mathrm{C}_t \cdot \mathbb{E}\{\mathrm{FT}\} + \mathrm{C}_f \cdot \mathbb{E}\{\mathrm{FB}\} \Big] + \mathrm{EI} \cdot \mathrm{C} \cdot \mathbb{E}\{\mathrm{ATR}\}. \tag{5.1}$$

Here, FT is the flight time, and FB is the fuel consumption, and C_t, C_f, CI, and EI are constant parameters weighting flight time, consumed fuel, operating cost, and climate effects, respectively, and C adjusts the order of climate impact with cost. The expectation over flight performance variables is used to capture the effects of meteorological uncertainty on aircraft dynamical behavior and climate impact estimates, which are quantified using ensemble prediction weather forecasts [22].

To evaluate the objective function defined in Equation (5.1), we need to calculate flight time, fuel consumption, and ATR. In this respect, the dynamical model given in Equation (4.1) needs to be integrated for a given flight plan (F), a set of meteorological variables (W), and initial time and mass (t_0, m_0), denoted with $\mathrm{TI}(\cdot)$. By performing trajectory integration, we can compute performance variables of interest:

$$\Big[\mathbb{E}\{\mathrm{FT}\},\ \mathbb{E}\{\mathrm{FB}\},\ \mathbb{E}\{\mathrm{ATR}\} \Big] = \mathbb{E}\Big\{ \mathrm{TI}(F, W, t_0, m_0) \Big\}, \tag{5.2}$$

allowing for evaluating the performance of a flight plan with the objective function given in Equation (5.1). Due to the discrete distribution of meteorological uncertainty characterized by EPS, the expectation operators in Equation (5.2) can be calculated with the simple average of different trajectory integrations, each corresponding to one member of EPS and sampled

initial flight time and initial fuel burn, with equal weights.

Trajectory optimization tool

The robust climate-optimal flight planning problem is solved using the recently developed tool ROOST, which is appropriate for the current structured airspace and capable of optimizing flight plans, taking into account meteorological uncertainty and uncertainty corresponding to the initial flight conditions. ROOST employs a heuristic algorithm called augmented random search as the optimization method, and due to parallelization on GPUs, it delivers the optimized robust flight plan in seconds. This library is publicly available and can be accessed using DOI: `https://doi.org/10.5281/zenodo.7495472`.

It is important to point out that the optimization problem formulated and presented in this chapter involves a combination of both discrete and continuous decision variables, i.e., resulting in a MINLP problem. This is an NP-hard problem class, and the associated theoretical difficulties are indeed borne out in practice: both the mathematical programming approaches to MINLP and metaheuristics are considered fairly computationally expensive. To cope with such a complex optimization problem efficiently, ROOST adopts the method proposed in [27], which defines a continuous search space over probability distributions of flight plans, rather than searching directly for a single optimal trajectory. In the context of this work, this formulation results in an optimization problem with hundreds of variables, depending on the size of the selected airspace subgraph and the discretization of Mach number and flight level values.

After performing optimization with ROOST for individual flights, we will receive N optimized flight plans. We denote the optimized flight plan for aircraft i as:

$$F^{i^o} = (R^{i^o}, \overline{\mathrm{FL}}^{i^o}, \overline{\mathrm{M}}^{i^o}, C^{i^o}, D^{i^o}, d_D^{i^o}).$$

5.2 Air traffic complexity evaluation

Accounting for network effects and operational requirements is crucial when planning climate-optimized trajectories for individual flights in order to ensure feasibility and enhance the realism of the achievable climate benefits. In this context, once individual flight trajectories are optimized, they are integrated into the overall traffic network to assess their collective impact on ATM system manageability, particularly regarding air traffic complexity.

This study adopts the complexity score as a quantitative measure to evaluate air traffic complexity. The complexity score consists of three different indicators, each being used to show a certain traffic characteristic: vertical, horizontal, and speed, different interacting flows [94]. As the presence of two aircraft in the same volume does not provide information on the severity and period of the potential hazards, the three mentioned indicators

are exploited to find an insight into the duration of having a disordered and difficult-to-handle traffic situation. To evaluate complexity, the airspace is partitioned into uniform four-dimensional grids. Two aircraft are considered to be interacting at any given time if they occupy the same grid from each aircraft's perspective. Once an interaction is detected, the above indicators are computed for each aircraft pair.

Vertical interaction is used as a measure of traffic in evolution. This indicator concerns not only the vertical and horizontal distance but also the flight phases. Basically, managing aircraft with different flight phases in a cell, such as a mix of cruising, descending, or climbing, is more complicated than flights in similar phases (e.g., only cruise phases). Indeed, the interactions are considered bilaterally. Following this logic, for two flights with different phases being simultaneously present in a similar grid, two interactions are counted. In this respect, for two aircraft k and i in the same cell, the vertical different interacting flows ν is computed as:

$$\nu^{i,k} = \begin{cases} 2\kappa^2 / ((t_x{}^i - t_e{}^i) + (t_x{}^k - t_e{}^k)) & \text{if} \quad (\kappa \neq \varnothing) \quad \textit{and} \quad (p^i \neq p^k), \\ 0 & \text{else,} \end{cases} \tag{5.3}$$

where t_e^i and t_x^i are the entering and exit times of aircraft within the cell, respectively, p^i is the flight phase of the aircraft i, and κ is defined as:

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

Horizontal different interacting flow is a complexity indicator that accounts for the duration of potential horizontal interaction between flights. It is assumed that controlling a structured traffic situation, such as parallel flows, requires less effort than a disordered situation with crossing flows. The term horizontal interaction is used to refer to the scenario where two aircraft with different heading angles are present in the same unit at the same time. In this respect, for two aircraft k and i in the same cell, the horizontal different interacting flows $\varkappa$ is computed as:

$$\varkappa^{i,k} = \begin{cases} 2\kappa^2 / ((t_x{}^i - t_e{}^i) + (t_x{}^k - t_e{}^k)) & \text{if} \quad (\kappa \neq \varnothing) \quad \textit{and} \quad (|\chi^i - \chi^k| > 20^\circ), \\ 0 & \text{else,} \end{cases} \tag{5.4}$$

where χ^i is the heading angle of the aircraft i.

In general, air traffic situations are perceived to be less complex when aircraft within a controller's area of responsibility maintain similar speeds [94]. Accordingly, significant speed variations among interacting flows are considered a contributing factor to increased complexity. The relative speed between interacting aircraft flows, denoted by v, is computed

as:

$$v^{i,k} = \begin{cases} 2\kappa^2 / ((t_x{}^i - t_e{}^i) + (t_x{}^k - t_e{}^k)) & \text{if} \quad (\kappa \neq \varnothing) \quad and \quad (|v^i - v^k| > 35kts), \\ 0 & \text{else} \end{cases} \tag{5.5}$$

where v^i is the speed of the aircraft i.

Once all indicators are computed, the complexity score is calculated as

$$\Psi = \nu + \varkappa + \upsilon.$$

By computing the complexity score for each individual cell, a complexity map is constructed as a function of the position and time.

5.3 Multi-agent reinforcement learning to manage air traffic complexity

The integration of individually optimized, climate-aware trajectories into the air traffic network can unintentionally increase traffic complexity, thereby challenging the operational feasibility of such flight plans. To ensure that traffic remains manageable, it is necessary to adjust flight profiles so that the resulting complexity does not exceed levels observed under business-as-usual operations. This adjustment process requires each aircraft to make decisions based on surrounding traffic conditions, dynamically modifying its trajectory to mitigate complexity. Such a decentralized decision-making problem can be modeled as a partially observable Markov decision process (POMDP).

5.3.1 Modeling complexity management as a MARL problem

As introduced in Chapter 4, the POMDP is defined by the tuple $\langle \mathcal{N}, \mathcal{S}, \mathcal{A}, \mathcal{P}, O, \gamma, R \rangle$. The components retain the same definitions used in that chapter. We model the problem as a fully cooperative framework where the reward function $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is shared across all agents.

In the following, we define the specific components of the POMDP framework used to address the problem of mitigating air traffic complexity, including the observation space, action space, and reward function.

State space

In this study, the state s_t at time t is defined as the concatenation of the local observations from all aircraft:

$$s_t = \left[o_t^1, \cdots, o_t^N \right].$$

where each $o_t^i = O(s_t, i)$ represents the local observation of aircraft i, derived from the full environment state via the observation function O.

This formulation assumes that the information of all agents is accessible during training, consistent with the centralized training and decentralized execution paradigm, while during execution, each aircraft relies solely on its own observation to make decisions.

Observation space

The observation space comprises two types of information: trajectory data and information about neighboring aircraft. Trajectory data consists of the true airspeed (v_t^i), heading angle (χ_t^i), flight phase (p_t^i) (i.e., climb, cruise, descent), and the duration (τ_t^i) over which the aircraft fly the most complex grid segment of its trajectory during $[t, t + \Delta t]$. This is complemented by information on neighboring aircraft within a defined vicinity, which aids the agent in assessing the complexity of the surrounding traffic and making strategic decisions to mitigate traffic complexity. We denote (v_r^j, χ_r^j, p_r^j) as the relative speed, heading, and phase differences of neighboring aircraft j to the aircraft i. The term I_t aggregates this detailed information for the m-th neighboring aircraft:

$$I_t = \left((v_r^1, \chi_r^1, p_r^1)_t, \ldots, (v_r^m, \chi_r^m, p_r^m)_t \right).$$

Based on the available data for each aircraft, the state observed by agent i at time t is defined as:

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

Action space

The decision space for the agents consists of admissible modifications to the aircraft's trajectory. Specifically, agents can perform predefined maneuvers to reduce traffic complexity, categorized into two primary types: speed adjustments and altitude changes. Agents can select modifications in one or both of the categories:

- Speed Adjustments:
  - Increase Mach by 0.03
  - Decrease Mach by 0.03
  - Maintain current Mach
- Altitude Changes:
  - Decrease by 2000 ft
  - Increase by 2000 ft
  - Maintain current altitude.

The chosen action is consistently applied across the entire trajectory. For example, if an agent opts to increase altitude, the entire flight profile is adjusted to reflect this change. Consequently, all trajectory-related data, including information on neighboring aircraft, is also updated.

Reward Function

The reward function R quantifies the immediate reward obtained by all agents for transitioning from state s to state s' due to the joint action $\mathbf{a}$. In this study, the agents cooperate with the objective of enhancing air traffic manageability. Accordingly, the complexity score is employed as the performance indicator, which is presented in Section 5.2.

The reward function is formulated to promote network-level improvements by minimizing aggregated traffic complexity. It is defined as the difference between the initial complexity, Ψ_0, and the current complexity at each decision step:

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

Accordingly, agents are positively rewarded only when their collective actions contribute to lowering overall traffic complexity, directly aligning the learning objective with the strategic goal of enhancing network manageability.

5.3.2 MAPPO algorithm

Several approaches exist for finding the optimal policy in reinforcement learning, such as value-based methods, policy gradient techniques, and actor-critic frameworks [141]. Policy gradient methods, which directly optimize the policy, have demonstrated significant potential in addressing various reinforcement learning problems [143]. Within this category, trust region learning is particularly notable, with implementations such as trust region policy optimization (TRPO) and proximal policy optimization (PPO) standing out [184, 185]. TRPO updates the policy within a trust region to ensure stable and reliable policy improvements, preventing drastic changes that could degrade performance [184]. However, TRPO is computationally demanding to implement due to its reliance on second-order optimization techniques, such as the Fisher information matrix. This complexity makes it less practical for many applications [185].

To overcome these limitations, PPO has been introduced in [185]. PPO uses a surrogate objective function to optimize the policy using first-order optimization methods while retaining the stability benefits of trust region approaches. We define the state value function as:

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

which estimates the expected reward obtained by starting from state s and following the policy π. The policy and value function are approximated using neural networks and parameterized using finite-dimensional parameters. Specifically, π_θ denotes the actor-network used to approximate the policy, and V_ϑ represents the critic network utilized for approximating the value function. The surrogate objective function is formulated as follows:

$$L(\theta) = \mathbb{E}_{\mathbf{a}\sim\pi_\theta, s\sim\mathcal{P}}\left[\min\left(\frac{\pi_\theta(a|s)}{\pi_{\theta_{old}}(a|s)}A_{\pi_\theta}(s,a), \text{clip}\left(\frac{\pi_\theta(a|s)}{\pi_{\theta_{old}}(a|s)}, 1\pm\epsilon\right)A_{\pi_\theta}(s,a)\right).\right]$$

The clipping mechanism ensures that the policy ratio $\frac{\pi_\theta(a|s)}{\pi_{\theta_{old}}(a|s)}$ is constrained between $1+\epsilon$ and $1-\epsilon$. This effectively allows PPO to regulate the extent of policy updates, ensuring that the new policy π_θ does not deviate excessively from the old policy $\pi_{\theta_{old}}$. By doing so, PPO retains the advantages of the TRPO algorithm, such as stable and reliable policy improvements, while being significantly easier to implement.

The advantage function A_π is computed using generalized advantage estimation as [141]:

$$A_{\pi_\theta}(s,\mathbf{a}) := \mathbb{E}_{\mathbf{a}_t\sim\pi_\theta, s_t\sim\mathcal{P}}[\sum_{t=0}^{\infty}(\gamma)^t\delta_{t+1}|s_0=s,\mathbf{a}_0=\mathbf{a}],$$

$$\delta_t = R(s_t,\mathbf{a}_t) + \gamma V(s_{t+1}) - V(s_t).$$

This function evaluates the benefit of taking action a in the state s relative to the baseline value $V(s)$.

Multi-agent PPO (MAPPO) extends the PPO algorithm to multi-agent environments, enabling multiple agents to interact within the same environment and learn coordinated policies for efficient decision-making [186]. MAPPO operates through two distinct phases: centralized training and decentralized execution. During the decentralized execution phase, each agent follows a policy $\pi^i_{\theta^i}$, which is parameterized by θ^i.

In this study, we assume that all agents are homogeneous, sharing identical capabilities, observation spaces, action spaces, and a common reward function that aligns them toward the shared goal of minimizing air traffic complexity [187]. This homogeneity implies that the optimal behavior for one agent is also optimal for all others, allowing agents to serve interchangeable roles within the system and supporting the use of a common policy framework. With homogeneous agents, policies can be trained more efficiently through parameter sharing (i.e., $\theta^1 = \theta^2 = ... = \theta^N = \theta$). Following this assumption, we parameterize the policy π using a finite-dimensional parameter θ shared among all agents. In this setup, while the policy is trained using the collective experience of all agents, each agent can still take individual actions, $\pi_\theta(\cdot|o^i_t)$, based on individual local observations o^i_t [188].

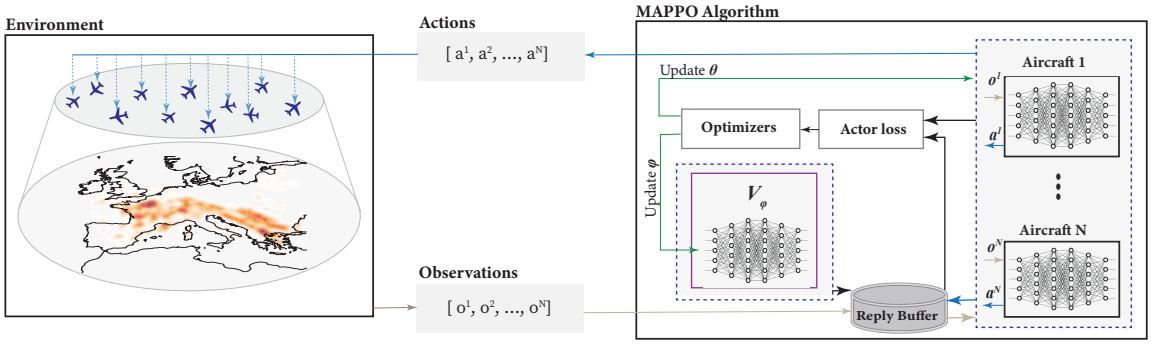


FIGURE 5.2: The proposed framework for managing the complexity of climate-optimized routes based on MAPPO.

The policy parameters θ are optimized by maximizing the following surrogate objective function:

$$L(\theta) = \sum_{i=1}^{N} \mathbb{E}_{\mathbf{a}\sim\pi_\theta, s\sim p}\left[\min\left(\frac{\pi_\theta(a^i|o^i)}{\pi_{\theta_{\text{old}}}(a^i|o^i)} A_\pi(s,\mathbf{a}),\ \text{clip}\left(\frac{\pi_\theta(a^i|o^i)}{\pi_{\theta_{\text{old}}}(a^i|o^i)}, 1-\epsilon, 1+\epsilon\right) A_\pi(s,\mathbf{a})\right)\right]. \tag{5.6}$$

In the centralized training step, a critic network V_ϑ is employed as a central coordinator to facilitate variance reduction. The critic network leverages global state information formed by combining the observations of all agents, represented as $s = [o^1, ..., o^N]$. This representation allows the critic to effectively estimate the overall system value, incorporating information that is not available to individual agents during decentralized execution.

The critic network is updated by minimizing the temporal difference error between the estimated value and the target value. The loss function for updating the critic network is defined as:

$$L(\vartheta) = \mathbb{E}_{\mathbf{a}\sim\pi_\theta, s\sim\mathcal{P}}\left[\sum_{t=0}^{\infty}(V_\vartheta(s_t) - y_t)^2\right],$$

where the target value y_t is given by

$$y_t = R(s_t, \mathbf{a_t}) + \gamma V_\vartheta(s_{t+1}).$$

The proposed MAPPO framework is depicted in Figure 5.2, providing an overview of its structure and components. The training procedure, following the outlined steps, is detailed in Algorithm 3.

Algorithm 3 Multi-Agent proximal policy optimization to mitigate complexity.

Input: α_θ, Number of agents N, Batch size B, episodes K, steps per episode T, discount factor γ, and GAE parameter λ
Initialize policy parameters θ, value function parameters ϑ
Initialize optimizers for policy and value function
for $k = 0, 1, ..., K-1$ **do**
    Collect a set of trajectories using policy π_θ
    Compute rewards $r_t = R(s_t, a_t)$, and log-probabilities $\log \pi_\theta(a_t^i | o_t^i)$
    Push transition $\{o_t^i, a_t^i, o_{t+1}^i, r_t\}$ into the replay buffer
    for each mini-batch B **do**
        Calculate advantage function $\hat{A}(s, a)$ using critic network

$$\delta_t = R(s_t, \mathbf{a}_t) + \gamma V_\vartheta(s_{t+1}) - V_\vartheta(s_t)$$

$$\hat{A}(s_t, \mathbf{a}_t) = \sum_{l=1}^{T} (\gamma\lambda)^l \delta_{t+l}$$

        Compute

$$V_{\text{target}} = \hat{A}(s_t, a_t) + V_\vartheta(s_t)$$

        Compute surrogate objective $L(\theta, \lambda)$ according to Equation (5.6)
        Update policy parameters θ via gradient ascent on $L(\theta, \lambda)$
        Update value function parameters ϑ via gradient descent on L_V

$$L_V = \frac{1}{BT} \sum_{j=1}^{B} \sum_{t=1}^{T} \left[\left(V_\vartheta(s_t) - V_{\text{target}} \right)^2 \right]$$

    end for
end for

5.4 Simulation results

In this section, we assess the effectiveness of the proposed approach in planning practical climate-optimized routes. First, the scenario definitions and simulation settings are outlined. Subsequently, the simulation results are analyzed and discussed.

5.4.1 Scenario definition

The scenario considered in this chapter builds upon that introduced in Chapter 4, with an extended temporal scope. Specifically, it includes approximately 6,000 flights operating within ECAC airspace on December 20, 2018, between 12:00 and 16:00 UTC.

The flight data, including the time and altitude of the first crossed waypoint in the considered time window, are extracted from the DDR2 dataset. For flights that start or land outside ECAC airspace, we model only the segments of the flights that take place within ECAC airspace. The first and last crossed waypoints within the considered areas and time window are selected as the origin and destination for the simulation. For all flights that are in the climb, cruise, or descent phases at the considered spatiotemporal frame, trajectory optimization is performed.

The weather data required for the optimization are sourced from the ERA5 reanalysis data product, including wind components, humidity (i.e., relative and specific), top net thermal radiation, potential vorticity, and geopotential. These variables are needed for representing aircraft dynamics and estimating the climate impacts of non-CO_2 forcing agents using aCCFs. The reported values in this chapter represent the mean outcome across 10 ensemble members of ERA5 weather data used in the flight planning process.

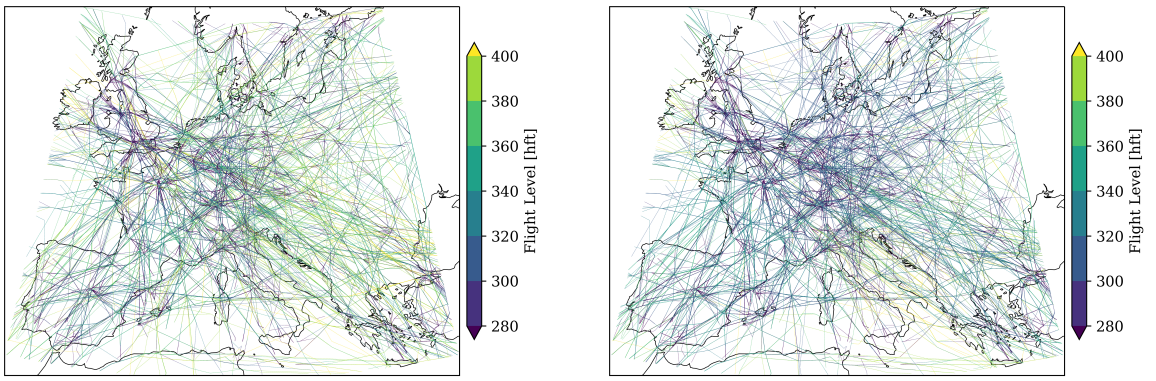
Since flight planning is performed within structured airspace, admissible route graphs are required for the optimization. In this study, the route graphs for the considered scenario are derived from the DDR2 environment data of June 2018.

5.4.2 Micro-scale trajectory optimization

Once we have all flight data and weather variables, we employ ROOST to plan climate-optimal trajectories. To explore different routing options, we follow a similar approach as in the previous chapters, varying only EI while keeping the remaining weighting coefficients in the objective function (i.e., Equation (5.1)) fixed as follows:

$$\mathrm{CI} = 1[-], \quad \mathrm{C}_t = 0.75, \quad \mathrm{C}_f = 0.51.$$

Five different routing options are generated, corresponding to EI values of $[0.0, 0.01, 0.1, 1.0, 10.0]$. The parameter EI, as defined in Equation (5.1), serves as a weighting parameter that prioritizes climate effects relative to operational costs. For illustration, the



(A) Cost-optimized routes (EI = 0.0) (B) Climate-optimized routes (EI = 10.0)

FIGURE 5.3: Individually optimized flight trajectories for two different EI values, color-coded based on flight levels. The color bar indicates flight altitudes.

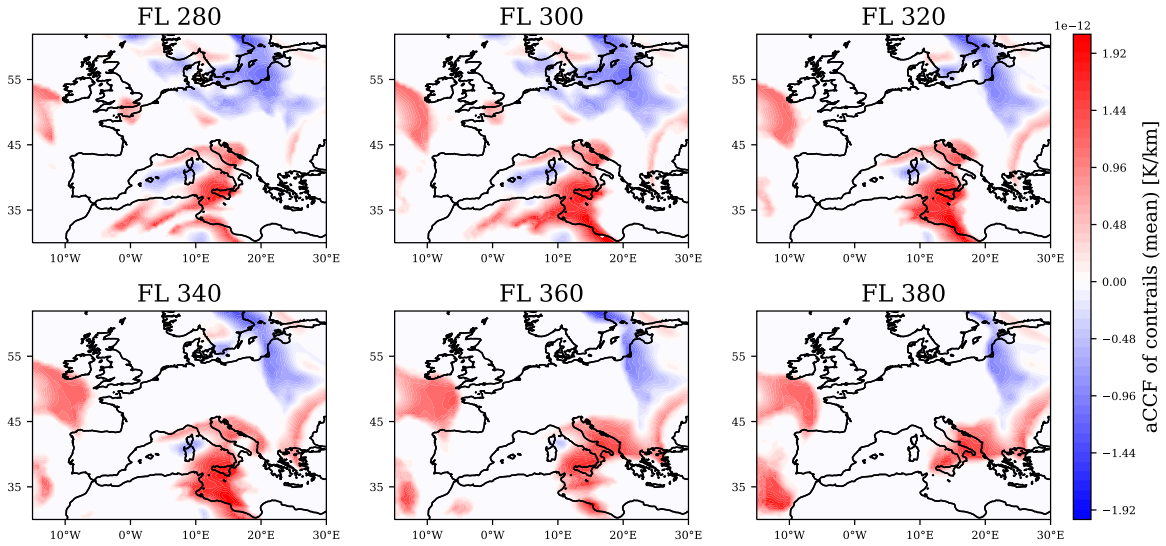
trajectories associated with EI $= 0.0$ and EI $= 10.0$ are shown in Figure 5.3, with color coding indicating the flight levels at which aircraft fly. The figure shows that as the weight assigned to climate impact in the objective function increases (prioritizing climate benefits), the aircraft tend to fly at lower altitudes. This trend is likely due to the greater potential for cooling contrail formation at lower flight levels, as shown in Figure 5.4a. The computational time for solving each trajectory optimization problem is approximately three seconds, with the maximum number of iterations set to 4000 and each iteration taking about 0.8 milliseconds. The trajectory planning simulations were performed using an NVIDIA GeForce RTX 3090 GPU.

5.4.3 Policy training setup

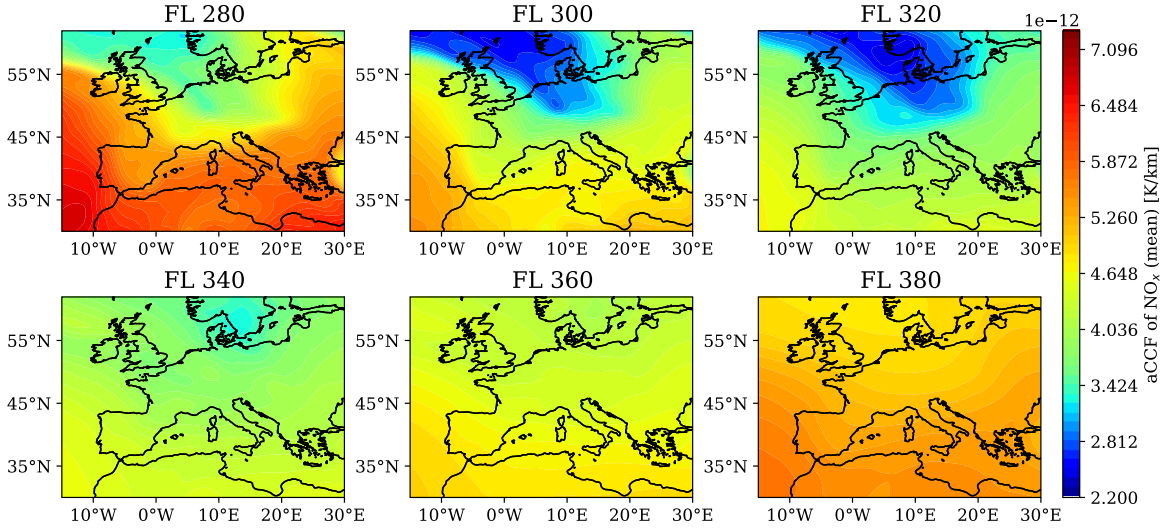
Following the micro-scale trajectory optimization, the proposed MARL algorithm described in Section 5.3.2 is implemented in a custom environment. For training, trajectories corresponding to EI = 0.0 (i.e., cost-optimal trajectories) are used. The experimental setup includes the state space, action space, and reward function detailed in Section 5.3.1. All aircraft are modeled as homogeneous agents with identical performance characteristics, sharing the actor network.

The actor is a fully connected multi-layer perceptron (MLP) with two hidden layers of 264 neurons and *Tanh* activations. This architecture was selected based on empirical tuning. It maps local observations to a categorical distribution over the action set via a softmax transformation. Actions are sampled stochastically during training and chosen greedily during evaluation. If an aircraft selects an action that results in a constraint violation, the action is invalidated and replaced with a no-modification, meaning the aircraft continues along its original trajectory for that decision step. The chosen action is applied uniformly to the aircraft's trajectory over the considered interval. For example, if an agent selects an increase

in speed, the adjustment is applied to the current speed, and the entire flight profile is updated accordingly. This update propagates to all trajectory-related information, including interactions with neighboring aircraft. Once all agents have selected their actions, the environment is updated, and the overall traffic complexity is recomputed. The average execution time per aircraft is approximately 2 μs (i.e., 2×10^{-6} seconds). The critic network is also a fully connected MLP with two hidden layers of 264 neurons, taking as input the state $s_t = [o_t^1, \cdots, o_t^N]$, and outputs a scalar value $V_\phi(s_t)$, which provides stable feedback under the shared reward and mitigates non-stationarity in multi-agent learning.



(A) The aCCF of contrails.



(B) The aCCF of NO_x.

FIGURE 5.4: The aCCF of contrails and NO_x emissions on December 20^{th}, 2018 (1200UTC).

For complexity evaluation, the airspace is discretized into cells of size 20 nm$\times$20 nm$\times$30 hft, and a time window of 60 minutes is considered [94]. Trajectories are sampled at 10-second intervals, which is sufficient to capture all cell transitions. The complexity is then computed, and rewards are assigned in proportion to the change in traffic complexity: positive values when complexity decreases, zero when it remains unchanged, and negative values when it increases. To stabilize learning, rewards are normalized during training.

Transitions are stored in a replay buffer and generalized advantage estimation is used to compute advantages, with $\lambda = 0.95$. Since the optimization objective concerns the total accumulated impact over a known, finite flight duration, we set the discount factor $\gamma = 1$ so that rewards at all time steps contribute equally to the return [141]. The PPO surrogate objective with clipping ($\epsilon = 0.2$) is optimized for 10 epochs per update using minibatches of size 32 sampled from batches of 124 episodes.

Several stabilization strategies are applied during training. Network weights are initialized orthogonally (gain = 1.0 for hidden layers; 0.01 for the output layer) to improve signal propagation. Entropy regularization with a coefficient of 0.01 is applied to encourage sufficient exploration. Gradient clipping is applied with a maximum global norm of 10.0 to prevent exploding gradients. The learning rate is linearly decayed from an initial value of 5×10^{-4} to zero over the course of training. Finally, both advantages and rewards are normalized to reduce variance and accelerate convergence.

Hyperparameters for MAPPO were determined partly from established MAPPO studies [186], including the GAE parameter ($\lambda = 0.95$), PPO clipping threshold ($\epsilon = 0.2$), optimizer epsilon, and network initialization. Other parameters were tuned empirically to ensure stable convergence in our domain, such as the entropy coefficient ($\beta = 0.01$), learning rate (5×10^{-4}), and batch size and minibatch size. Table 5.1 summarizes the hyperparameters used for the MAPPO algorithm.

To stabilize training and ensure computational tractability in large-scale scenarios ($\sim$ 6,000 flights), we adopt a sampling strategy during training. For the training stage, subsets of air traffic are created to divide the overall traffic into smaller groups for computational convenience. This approach is also justified by the fact that traffic complexity is primarily localized, where an aircraft is mainly influenced by its nearby flights. This is due to the requirement that both temporal and spatial distances between flights must be below certain thresholds for interactions to be relevant. To generate each subset, a random geographical location (defined by random latitude and longitude coordinates) is selected. From this central point, a square region with a side length of 10° (i.e., 600 nautical miles (NM)) is considered. Flights intersecting these regions during the specified timeframe are grouped into subsets. This method allows individual flights to appear in multiple subsets alongside varying combinations of other flights, ensuring the coverage of the airspace and introducing variability into the training data. Importantly, by limiting each training step to a subset of locally interacting agents, this approach reduces gradient variance and improves training stability.

TABLE 5.1: Hyperparameter settings for MAPPO training.

Parameter	Value
Batch size (episodes)	124
Minibatch size	32
Training epochs per update (K)	10
Discount factor (γ)	1.0
GAE parameter (λ)	0.95
PPO clipping parameter (ϵ)	0.2
Entropy coefficient (β)	0.01
Learning rate	5×10^{-4} (decayed linearly)
Optimizer	Adam
Gradient clipping	10.0
Hidden layer dimension (actor/critic MLP)	264
Activation function	*Tanh* (hidden layers), *Softmax* (actor output)
Orthogonal initialization	Gain = 1.0 (hidden), 0.01 (output layer)

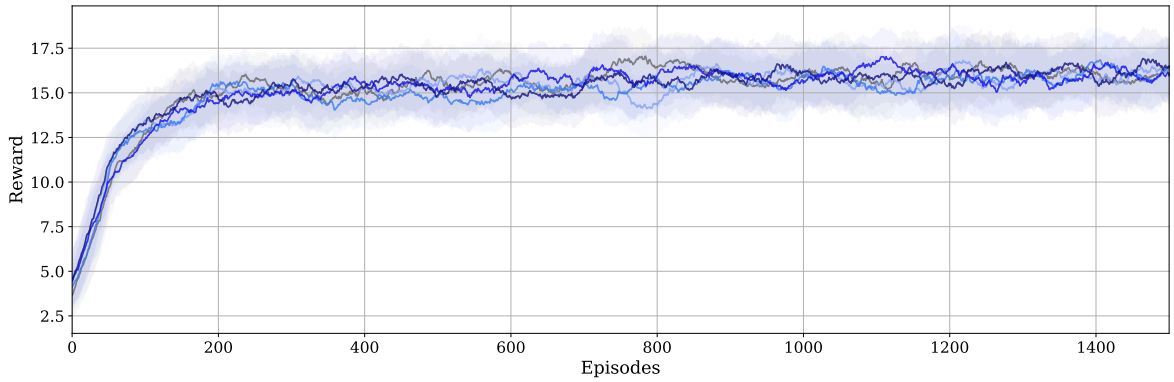
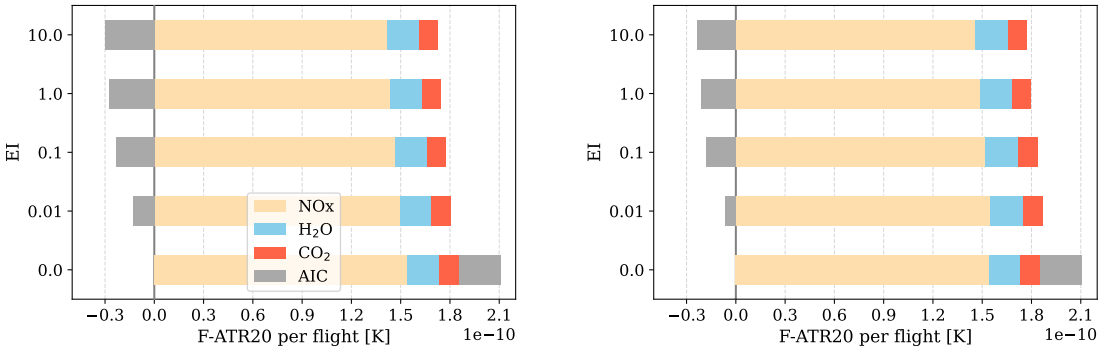


FIGURE 5.5: Learning curves from five independent training runs, with shaded regions indicating variability across runs.

Figure 5.5 shows the learning curves from five independent training runs, recorded every 500 training steps. A moving average with a window size of 50 was applied to improve visual clarity. Across all runs, the policies consistently converged to stable performance levels with only minor variations in the final reward values. These results indicate that the MAPPO framework is robust to initialization and training stochasticity.

5.4.4 Results

This section presents the simulation results, where the performance of the proposed algorithm is evaluated and compared with the micro-scale trajectory optimization framework.



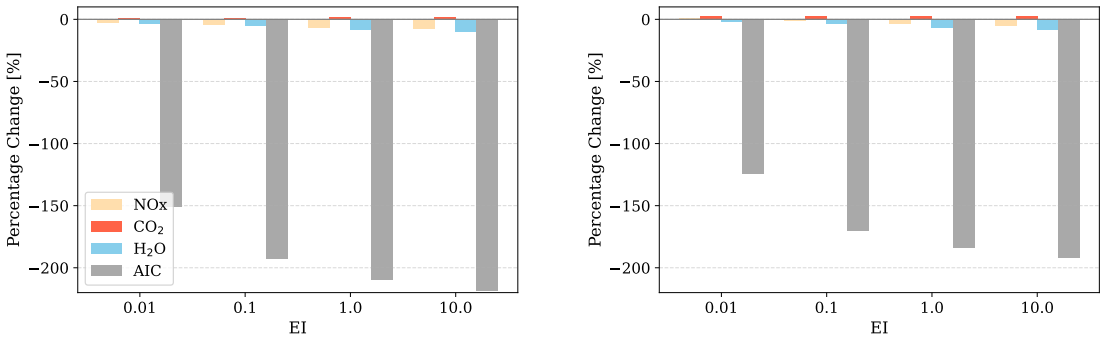
(A) Micro-scale trajectory optimization

(B) Network-scale trajectory optimization using MAPPO

FIGURE 5.6: Climate impact per flight for NO_x, CO_2, H_2O, and contrails (AIC) across different EI values.

The trajectories obtained from both approaches (i.e., micro-scale trajectory optimization and the proposed network-scale trajectory optimization using MAPPO) are analyzed to evaluate their climate impact. Key species such as NO_x, H_2O, CO_2, and aviation-induced cloudiness (AIC) are examined. Figure 5.6 illustrates the average temperature response based on future emission scenarios (F-ATR20) for different species across various EI values. The horizontal axis represents the climate impact per flight in Kelvin, and the vertical axis shows the EI values, ranging from 0.0 to 10.0. The results show that trajectory planning provides the potential to mitigate climate effects. It can be seen that for the set of cost-optimal trajectories (EI = 0.0), contrails predominantly have a warming impact. However, as the weight of climate impact is increased in the optimization process, the total climate impact is reduced, primarily due to the generation of cooling contrails. A similar trend is observed with our proposed approach, where the climate impact also decreases with increasing EI values. Nevertheless, the degree of mitigation is slightly diminished compared to the micro-scale approach, particularly in the generation of cooling contrails.

Figure 5.7 illustrates the percentage change in climate impact for each species compared to the baseline scenario (micro-scale trajectory optimization with EI = 0.0). It can be seen that contrails have the highest potential for climate impact mitigation. This mitigation is primarily achieved through cooling contrail generation. These findings align with the research in [20], which highlighted that flight planning is most effective in reducing contrails climate effects. This is mainly due to the sharp aCCF pattern (i.e., related to conditions for persistent contrail formation) of contrails, in which small trajectory changes can yield significant mitigation (see Figure 5.4a). However, for emissions such as NO_x, as shown in [60] and illustrated in Figure 5.4b, their climate sensitivity tends to be smoother, limiting the effectiveness of trajectory planning in mitigating their climate effects, considering the trade-offs with operating costs.



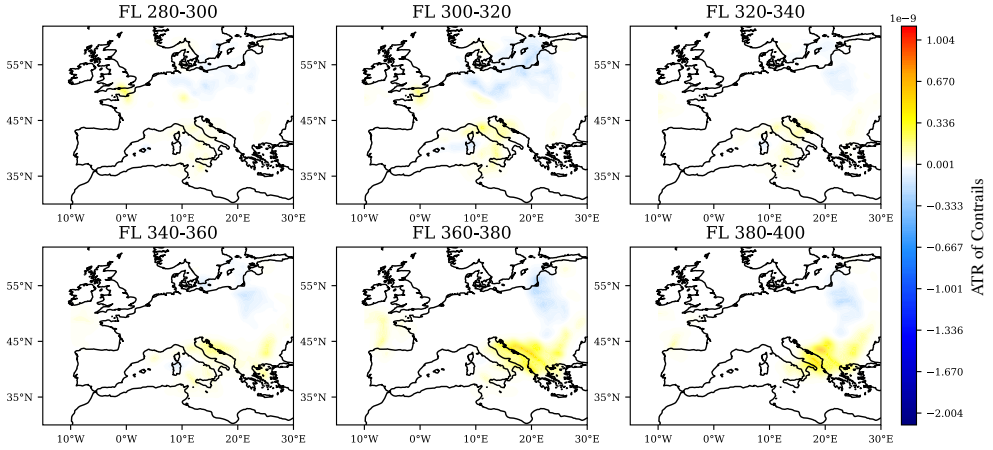
(A) Micro-scale trajectory optimization (B) Network-scale trajectory optimization

FIGURE 5.7: Percentage change in the climate impact for different species.

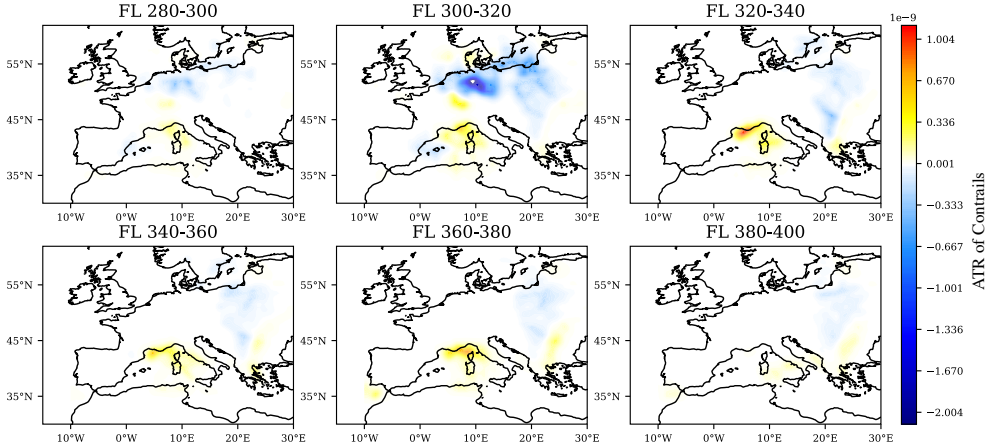
The spatial distribution of the climate effects of contrails for the extreme EI scenarios (0.0 and 10.0) is presented in Figure 5.8, providing insight into where persistent contrails are formed and how they vary across different flight levels and regions (for a complementary presentation of the overall climate impact (i.e., net ATR) see Figure 5.9). Figure 5.8a shows the distribution of persistent contrails for cost-optimal trajectories, which predominantly exhibit warming effects. In contrast, the maps corresponding to EI = 10.0, obtained from micro-scale trajectory optimization and network-scale optimization using MAPPO, demonstrate a reduction in warming contrails and an increase in cooling contrails, particularly at lower flight levels, compared to the cost-optimal scenario. This shift is attributed to the rerouting of climate-optimal flight paths from higher altitudes to lower altitudes (e.g., FL 280-300, 300-320, 320-340), where conditions are more favorable for generating cooling contrails (see Figure 5.4a). Comparing Figures 5.8b and 5.8c, we observe that while the overall pattern remains similar, the intensity and extent of cooling effects are slightly reduced when using the proposed approach. This indicates that the network-scale optimization with MAPPO effectively mitigates climate impacts, albeit achieving a slightly lower degree of mitigation due to the need to address air traffic manageability.

To evaluate the manageability of the ATM system, the complexity of all routing options is assessed. Figure 5.10 presents a comparison of aggregated and maximum air traffic complexity for trajectories obtained using micro-scale optimization and the proposed network-scale approach. The complexity of different routing options is reported to examine the manageability of air traffic under different levels of climate optimality. Aggregated complexity represents the cumulative measure of traffic complexity across the entire airspace and time period under analysis. The results reveal a clear trend: for trajectories generated through micro-scale optimization, increasing the weighting parameter EI to enhance climate impact mitigation results in a rise in air traffic complexity.

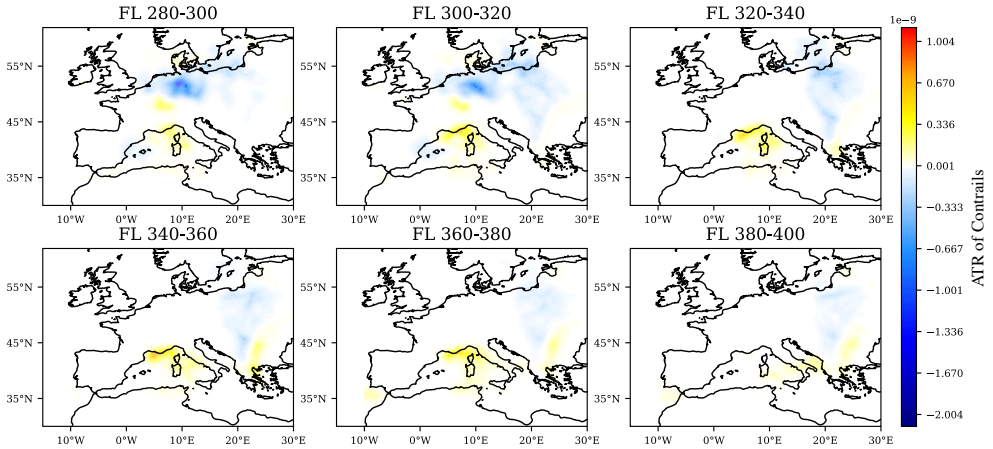
Notably, for the EI=10.0, the aggregated complexity increases by 68.4% compared to the



(A) cost-optimal trajectories (EI = 0.0).

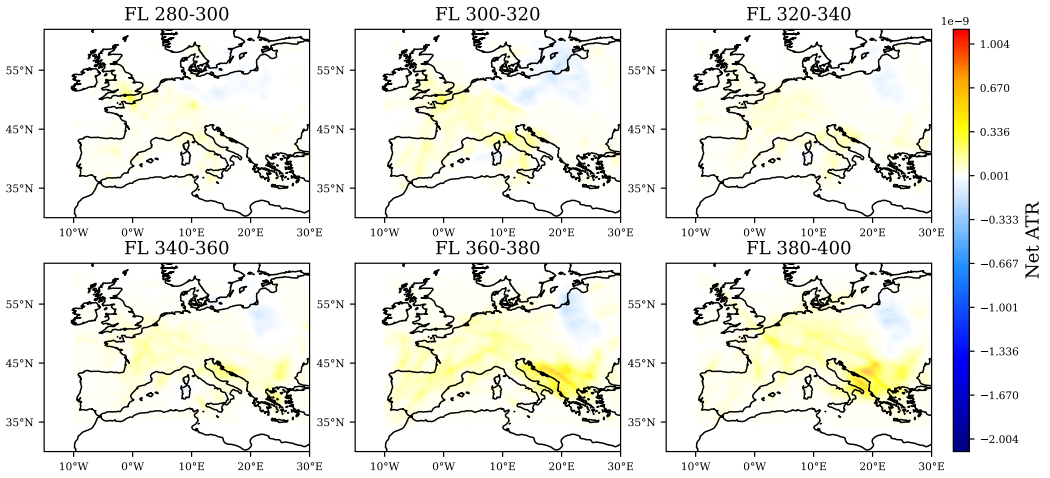


(B) Climate-optimal trajectories (EI = 10.0) obtained from micro-scale trajectory optimization.

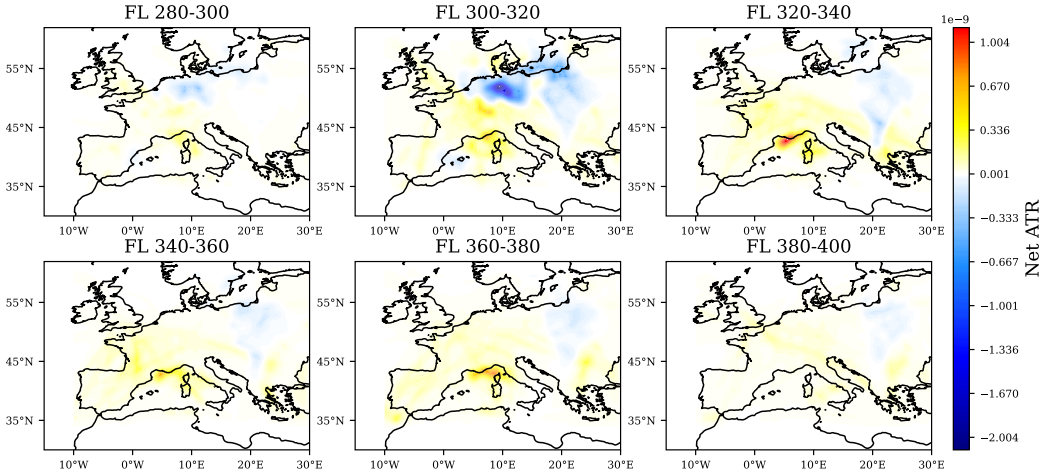


(C) Climate-optimal trajectories (EI = 10.0) obtained from network-scale trajectory optimization using MAPPO.

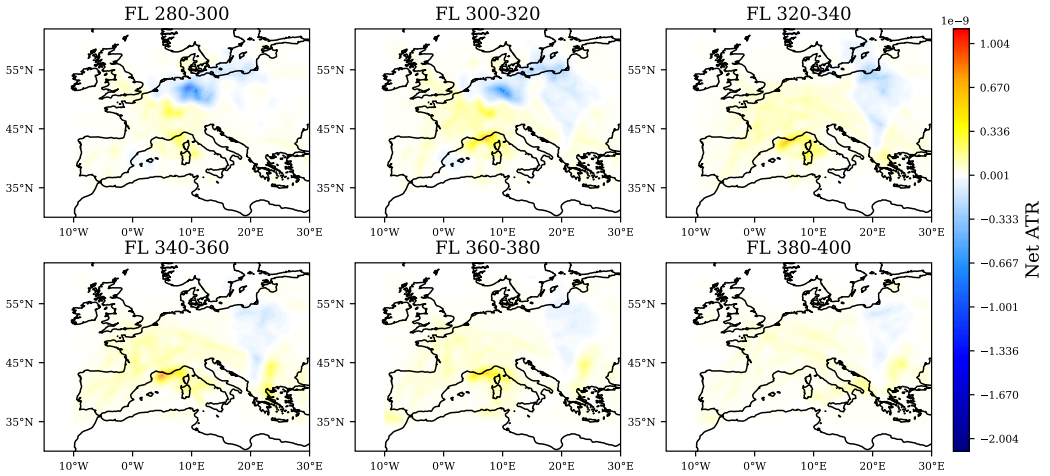
FIGURE 5.8: The spatial distribution of persistent contrails (both warming and cooling) formed along trajectories for different sets of optimized routes.



(A) cost-optimal trajectories (EI = 0.0).



(B) Climate-optimal trajectories (EI = 10.0) obtained from micro-scale trajectory optimization.



(C) Climate-optimal trajectories (EI = 10.0) obtained from network-scale trajectory optimization using MAPPO.

FIGURE 5.9: Net average temperature response (net ATR) maps.

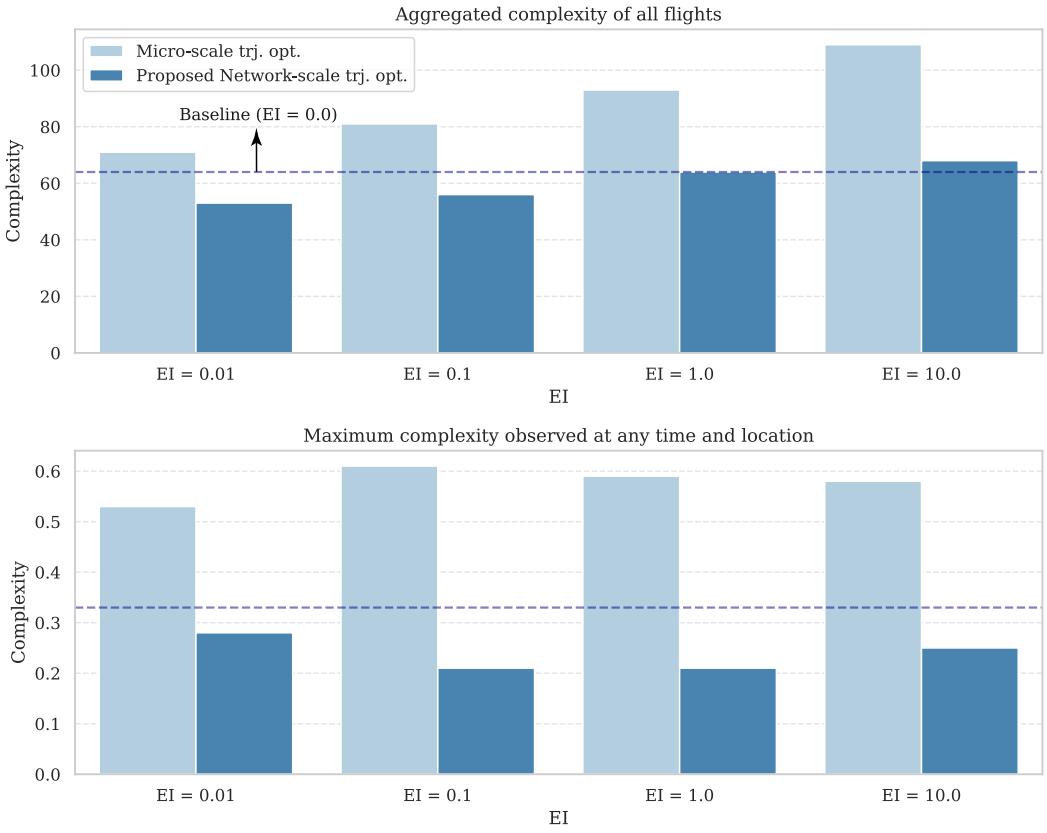


FIGURE 5.10: Comparison of aggregated and maximum air traffic complexity for micro-scale and network-scale trajectory optimization approaches for different EI values.

cost-optimal baseline. The majority of this increase (approximately 51.6%) stems from lateral interactions, as more flights are rerouted horizontally to avoid climate-sensitive regions, leading to a higher number of crossing and merging points (with heading differences often exceeding 20°). Additionally, speed interactions contribute 16.5% to the total increase. As the weight of climate impact increases, aircraft tend to adopt different combinations of flight levels and Mach numbers to minimize environmental impact and operational costs, which results in greater speed differences between aircraft pairs. In contrast, vertical interactions exhibit a negligible change (only 0.2%), since most flights remain in the cruise phase throughout their trajectory and continue to be classified as such at the cell level. These findings highlight that optimizing solely for climate benefits, without accounting for air traffic interactions, can challenge airspace manageability. In contrast, the proposed network-scale approach achieves a more balanced outcome; the aggregated complexity remains comparable to, or in most cases lower than, that of the cost-optimal trajectories.

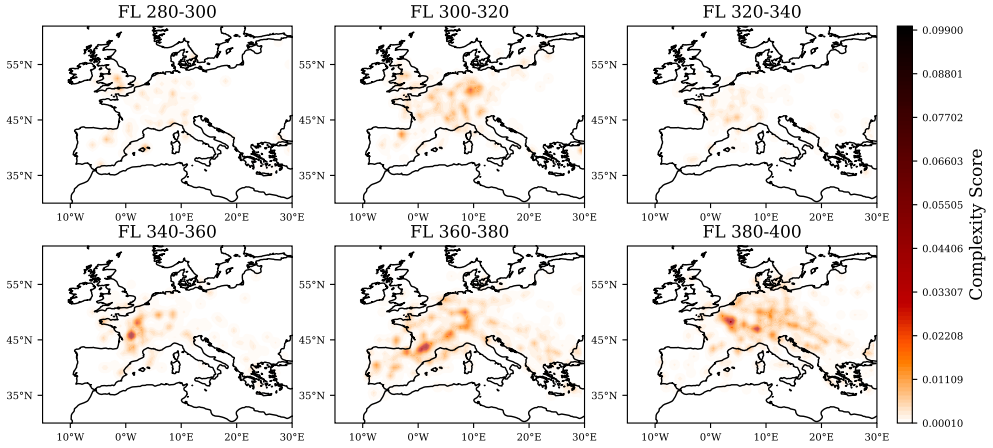
Figure 5.10 also presents the maximum complexity values for both approaches. Maximum complexity represents the highest level of traffic complexity observed at any specific

time and location within the airspace. This metric highlights the peak, or worst-case scenario, of traffic complexity. The results show that maximum complexity significantly increases when adopting climate-optimized trajectories from micro-scale optimization, particularly as EI increases. This can create hotspots of high congestion, posing safety and operational challenges for air traffic controllers. However, when using the proposed network-scale approach, the maximum complexity consistently remains below the levels observed in the cost-optimal scenario. This underscores the effectiveness of the introduced framework in maintaining operational feasibility and preventing the creation of high-risk areas in the airspace while still achieving reductions in climate impact. Overall, Figure 5.10 highlights the difference between the two optimization methods. While the micro-scale optimization approach exhibits a sharp increase in both metrics as we increase EI, the proposed approach maintains relatively stable complexity values.

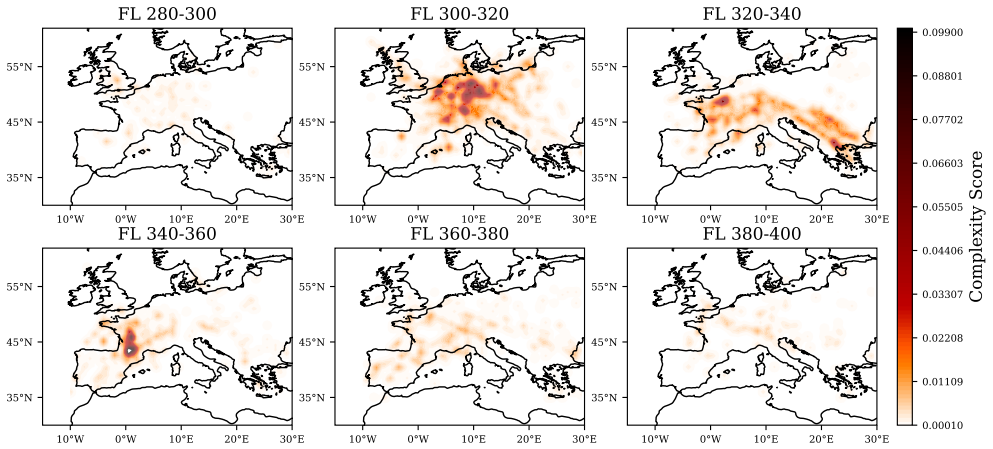
To provide visual insight into the locations where traffic complexity increases, complexity maps are generated. Figure 5.11 presents the spatial distribution of air traffic complexity for three scenarios: (a) cost-optimal routes (EI = 0.0), (b) climate-optimal routes (EI=10.0) obtained through micro-scale optimization, and (c) climate-optimal routes (EI=10.0) derived from network-scale optimization using MAPPO. Figure 5.11a illustrates the baseline scenario with cost-optimal trajectories (EI = 0.0). It can be seen that when the optimization focuses solely on cost, without accounting for climate impacts, air-traffic complexity remains lower than that observed in the climate-optimal scenarios (i.e., higher EI values) and exhibits a relatively more uniform spatial distribution.

Figure 5.11b illustrates the complexity for trajectories corresponding to EI = 10.0 obtained through the micro-scale trajectory optimization approach. A significant increase in complexity is observed, particularly at flight levels 300–320 and 320–340. Darker areas on the map highlight regions with high complexity. A comparison between Figure 5.8 and Figure 5.11 reveals that these regions of heightened complexity align with areas where cooling contrails are generated. This correlation suggests that, in regions with high potential for cooling contrail formation to mitigate overall climate impact, the concentration of flights increases, leading to traffic congestion and, consequently, higher complexity.

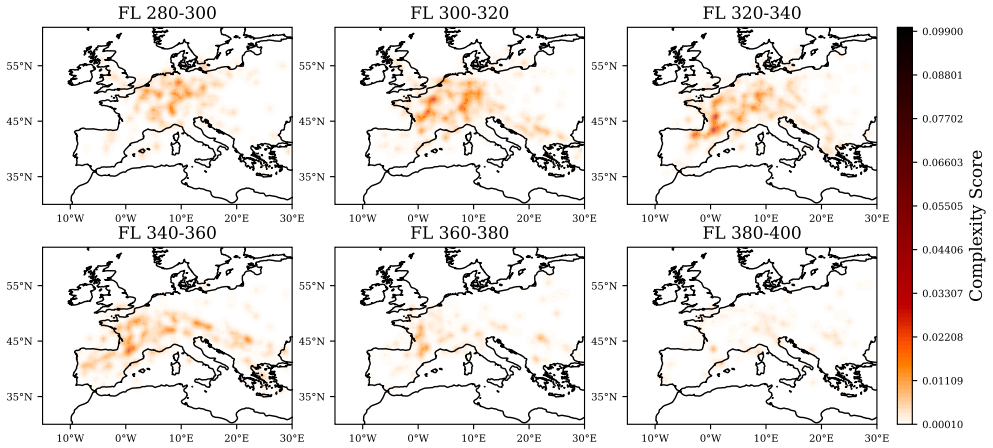
Figure 5.11c illustrates the trajectories corresponding to EI = 10.0 obtained from the network-scale optimization using MAPPO. While the spatial distribution of complexity differs from that in Figure 5.11a (i.e., the map associated with cost-optimal routes), the overall severity remains comparable. This highlights the ability of the proposed network-scale approach to achieve climate benefits without deteriorating air traffic complexity. In comparison to the micro-scale optimization, the proposed method provides a more balanced distribution of complexity across various flight levels, effectively reducing the congestion hotspots observed in Figure 5.11b. This represents a key advantage of the proposed network-scale approach, which delivers climate benefits while maintaining air traffic complexity at levels comparable to those of cost-optimal scenarios.



(A) Cost-optimal trajectories (EI = 0.0).



(B) Climate-optimal trajectories (EI = 10.0) obtained from micro-scale trajectory optimization.



(C) Climate-optimal trajectories (EI = 10.0) obtained from network-scale trajectory optimization using MAPPO.

FIGURE 5.11: The complexity maps obtained for different sets of optimized trajectories.

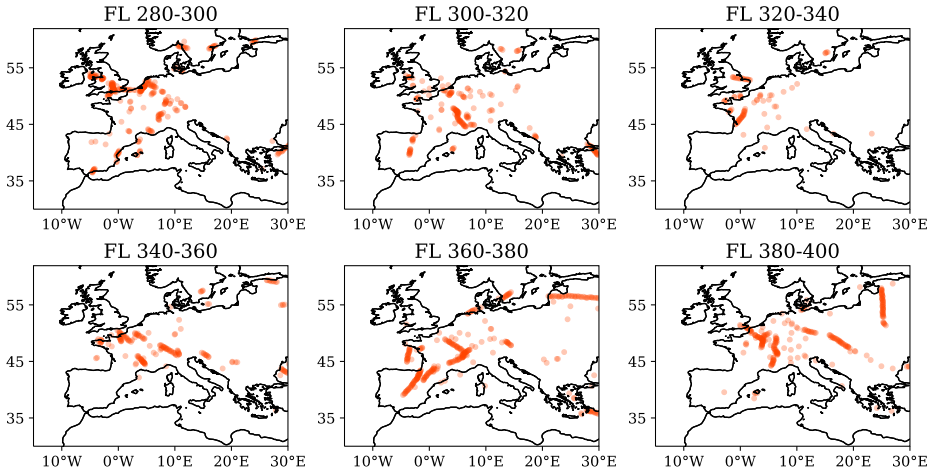
TABLE 5.2: Number of potential conflicts after micro-scale optimization and the proposed MAPPO-based framework across different EI values

	EI = 0.0	EI = 0.01	EI = 0.1	EI = 1.0	EI = 10.0
Micro-scale traj. opt.	1369.0	1503.0	1534.0	1551.0	1571.0
Proposed macro-scale traj. opt. (MAPPO)	825.0	905.0	963.0	1010.0	943.0

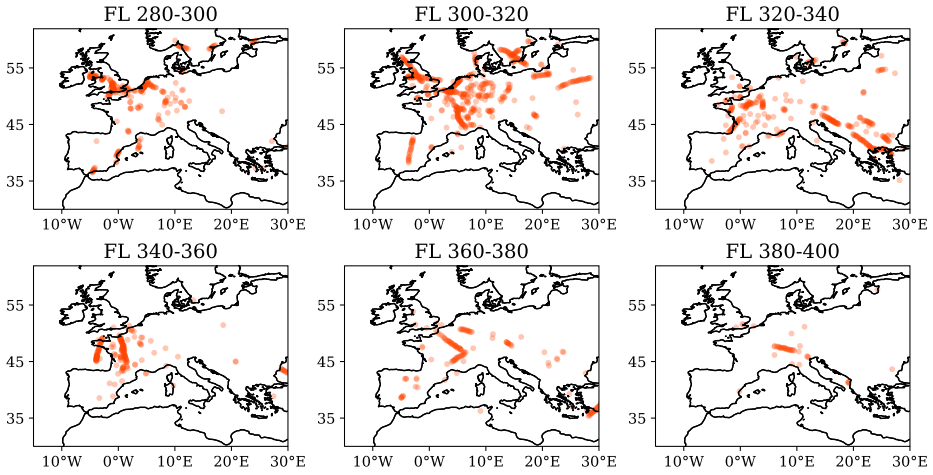
A post-optimization analysis was conducted to evaluate potential conflicts following both the micro-scale trajectory optimization and the proposed network-level framework based on MAPPO. Table 5.2 reports the total number of potential conflicts observed under different EI values. The results show that, under micro-scale optimization, which prioritizes climate impact mitigation without accounting for traffic complexity, higher EI values lead to an increase in potential conflicts. In contrast, the proposed MAPPO-based resolution reduces conflicts, with values consistently lower than those observed in the baseline business-as-usual trajectories. These findings suggest that addressing traffic complexity can also contribute to conflict mitigation.

To provide further insight, Figure 5.12 illustrates the spatial distribution of potential conflicts across flight levels. Under micro-scale optimization, conflicts not only increase in number but also tend to concentrate at specific altitudes (e.g., FL300–320), which aligns with the complexity patterns observed in Figure 5.11. In comparison, the MAPPO-based resolution reduces the overall number of conflicts and yields a more balanced distribution across altitude layers.

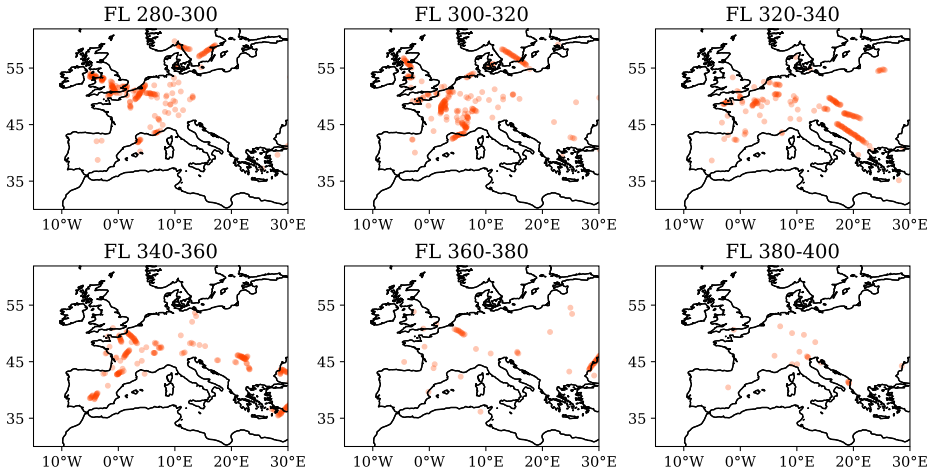
Figure 5.13 presents a comparison of airspace user preferences, including fuel consumption, flight time, and operating cost, across different EI values for both micro-scale and network-scale optimization approaches. The top and middle panels show that fuel consumption and flight time increase for higher EI values. This trend can be explained by the fact that climate-optimal trajectories typically deviate from cost-optimal routes to avoid (or fly through) warming (or cooling) climate-sensitive regions, leading to increased operational costs (see the bottom panel). While both approaches exhibit this trend, the proposed method results in higher fuel consumption compared to the micro-scale approach, while consistently achieving shorter flight times. This can likely be attributed to adjustments made to independently climate-optimized flight profiles to maintain traffic complexity at standard levels. Specifically, some flights are shifted from the most complex flight levels (FL300-320; see Figure 5.11b) to FL280-300 (compare Figures 5.11b and 5.11c), leading to shorter flight times. Overall, the figure demonstrates that while the proposed network-scale approach leads to slightly higher fuel consumption, it compensates for this with shorter flight times and maintains SOC values close to those of the micro-scale method.



(A) Cost-optimal trajectories (EI = 0.0).



(B) Climate-optimal trajectories (EI = 10.0) obtained from micro-scale trajectory optimization.



(C) Climate-optimal trajectories (EI = 10.0) obtained from network-scale trajectory optimization using MAPPO.

FIGURE 5.12: The conflict maps obtained for different sets of optimized trajectories.

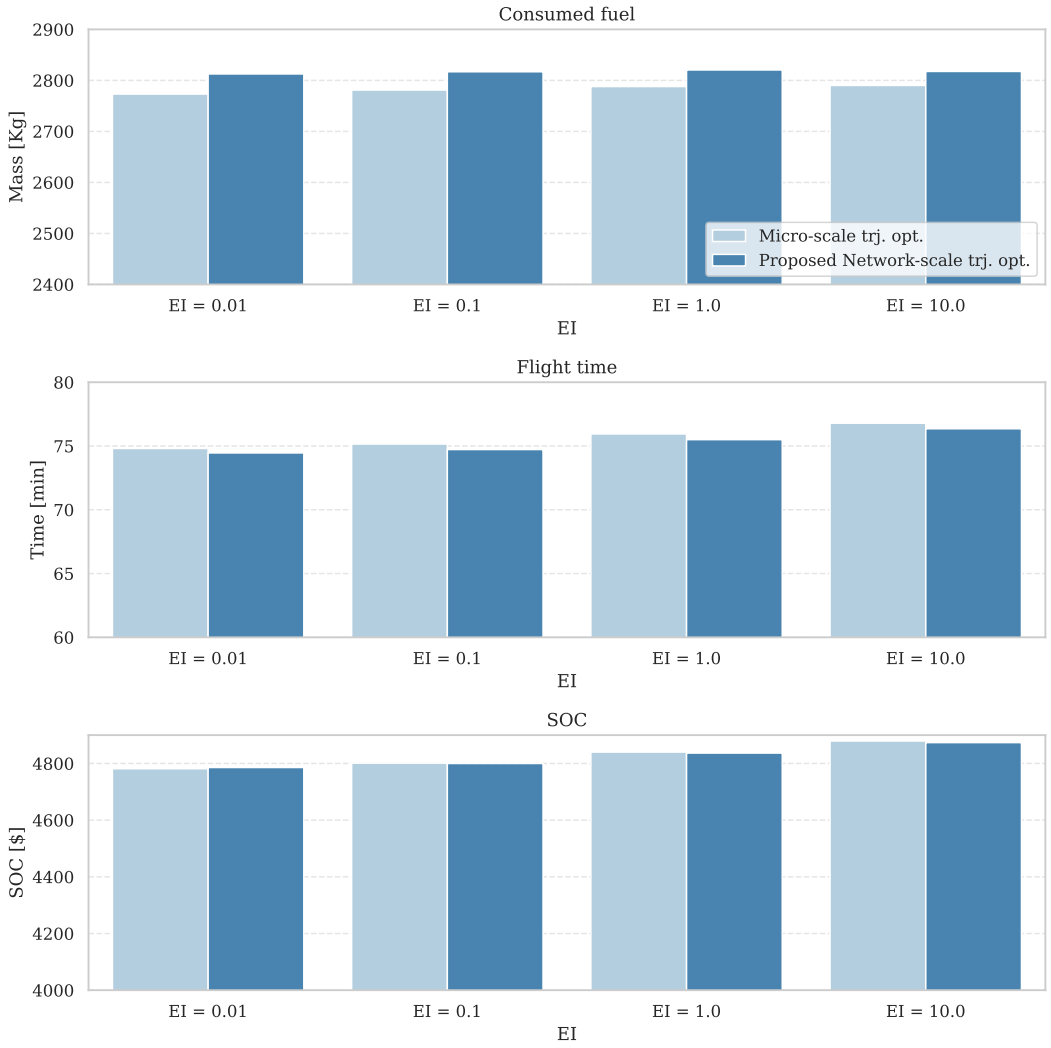
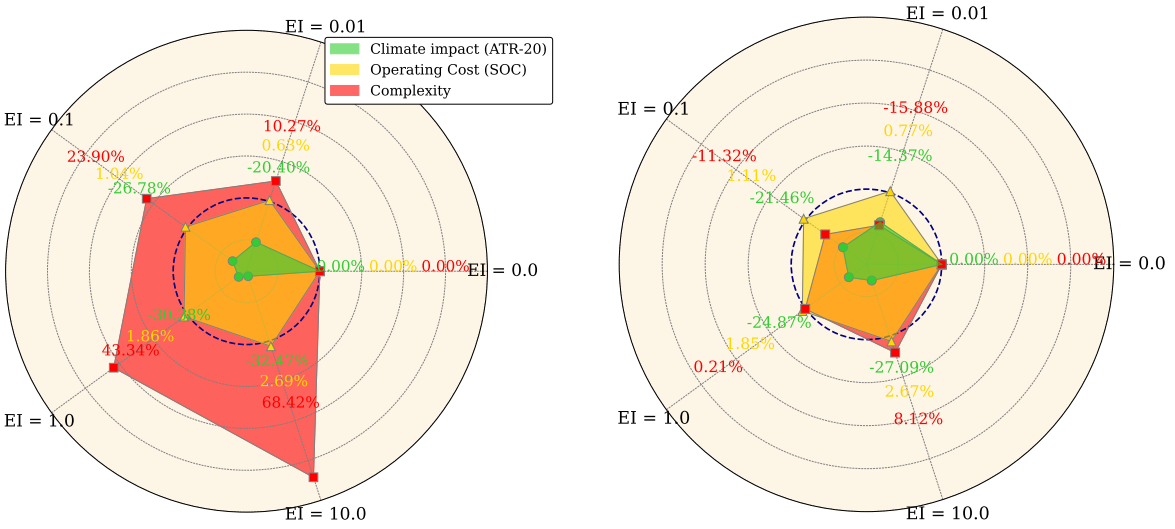


FIGURE 5.13: Comparison of consumed fuel, flight time, and operating cost for micro-scale and network-scale trajectory optimization approaches for different EI values.

Figure 5.14 presents a radar plot comparison of three key metrics, climate impact, operating cost, and air traffic complexity across different EI values for both approaches: micro-scale trajectory optimization and proposed network-scale trajectory optimization using MAPPO. Figure 5.14a illustrates the performance of the trajectories optimized solely for climate impact. As the EI value increases, there is a reduction in the climate impact, as indicated by the decreasing radial distance in the green zone. However, this is accompanied by increases in both operating costs and air traffic complexity, particularly when EI = 1.0 and EI = 10.0.



(A) Micro-scale trajectory optimization

(B) Network-scale trajectory optimizations using MAPPO

FIGURE 5.14: Performance comparison of micro-scale and proposed network-scale trajectory optimization approaches, illustrated using radar plots for three key metrics: climate impact, operating cost, and air traffic complexity, across different EI values. The dashed blue circle represents the baseline for all metrics.

Figure 5.14b illustrates the performance of the proposed network-scale trajectory optimization using MAPPO, demonstrating a more balanced outcome across the evaluated metrics. While there is a slight reduction in climate benefits in comparison with the micro-scale optimization method, the introduced framework outperforms in terms of air traffic complexity. The complexity remains consistently lower, even at higher EI values, highlighting the network-scale approach's ability to redistribute air traffic and mitigate the creation of complexity hotspots. Additionally, the operating cost remains comparable, highlighting the effectiveness of the proposed approach in planning operationally feasible climate-optimized flight plans.

All in all, the results highlighted a trade-off between climate impact mitigation, operational costs, and air traffic complexity, as prioritizing one objective often necessitates compromises in the others. The key challenge lies in finding an optimal balance between these competing objectives. Figure 5.14 illustrates this trade-off. Ideally, a feasible solution would position all objectives within the baseline radius (dashed blue circle) on the radar plot. However, due to the trade-offs involved, it is not possible to minimize all three simultaneously. For a solution to be considered feasible and optimal, it must maintain operational costs and air traffic complexity near the baseline while achieving reductions in climate impact that fall within the baseline range.

5.5 Summary

This chapter presented a framework for planning climate-friendly aircraft routes in an operationally feasible manner within the structured airspace. The proposed methodology comprised two key steps: first, optimizing individual trajectories to mitigate climate effects, and second, developing a decision-making strategy to mitigate the increased complexity associated with climate-optimized routes.

The effectiveness of the proposed framework is validated through a case study using real traffic data over European airspace. The results demonstrated that the proposed framework is capable of providing solutions that achieve a relatively balanced trade-off between climate impact, operational costs, and air traffic complexity. Notably, for EI = 0.1, the proposed framework achieved a 21.0% reduction in climate impact, accompanied by an 11.0% decrease in overall air traffic complexity compared to cost-optimal trajectories. This improvement was attained with a 1.0% increase in operational cost.

Compared to the results obtained in Chapter 4, the climate impact mitigation achieved in this chapter is lower, even for higher EI values. This difference stems from the underlying airspace structures considered in each case. In the previous chapter, the optimization was applied to a fully free-routing environment, where aircraft had greater flexibility to avoid warming regions or deliberately fly through cooling regions. In contrast, the structured airspace considered in this chapter imposes more routing constraints, limiting the potential for climate-aware rerouting and, consequently, reducing the achievable mitigation.

As discussed in previous chapters, such a sequential framework provides stakeholders with the flexibility to tailor flight planning according to their operational priorities. Those focused on minimizing operational costs may prefer cost-optimal trajectories (EI = 0.0), while stakeholders prioritizing climate impact reduction may opt for higher EI values, leading to the highest possible reductions in climate effects. Those seeking a more balanced routing strategy might select intermediate EI values, aiming to reduce climate impact while maintaining manageable traffic complexity and controlling operating costs.

One limitation of the proposed approach in this chapter concerns the definition of the decision space, which was restricted to vertical and speed adjustments for simplicity. Expanding it to include lateral modifications is expected to increase flexibility, as lateral deviations can introduce additional spatial separation between flights, offering alternative routing options to alleviate localized congestion and mitigate traffic complexity. From the climate perspective, any trajectory adjustment alters the climate impact due to the spatiotemporal dependency of non-CO_2 effects. However, these effects are generally more sensitive to altitude than to lateral deviations [183]. This suggests that a promising strategy would be to retain the climate-optimal altitude profiles obtained in the first stage and use lateral deviations for complexity reduction. Such an approach could enhance the framework's performance and support a more balanced trade-off between operational feasibility and climate objectives.

Chapter 6

Integrated framework for climate-optimal flight planning at the network-scale within structured airspace

In the previous chapters, we introduced sequential frameworks to address the problem of climate-optimal flight planning at the network scale. While these frameworks demonstrate effectiveness, they encounter computational limitations when applied to large-scale scenarios, such as yearly analysis. The primary challenge lies in the first phase of the process, the individual trajectory optimization. Despite the use of fast flight planning tools, optimizing each trajectory independently remains a time-consuming process. For instance, conducting a year-scale analysis involves millions of flights, each requiring optimization under different trade-offs between climate impact and operational cost, imposing significant computational burdens. Therefore, an integrated optimization framework that eliminates the need for a sequential process is essential to enable large-scale climate-aware flight planning, supporting the development of indicators toward policy actions. However, such a solution remains an open research gap.

This chapter introduces a scalable optimization framework for climate-optimal flight planning that simultaneously addresses climate impact mitigation and traffic manageability within a unified formulation. To mitigate climate effects, we identify specific airspace regions where aircraft emissions exert a strong warming impact, referred to as climate hotspots or ECHO areas. These regions are incorporated as avoidance constraints within the optimization problem. To ensure the operational feasibility of trajectories, traffic complexity is considered as the objective function to be minimized. Starting from business-as-usual (cost-optimal) trajectories, each aircraft adjusts its flight plan to avoid ECHO areas while collectively reducing overall traffic complexity (see Figure 6.1).

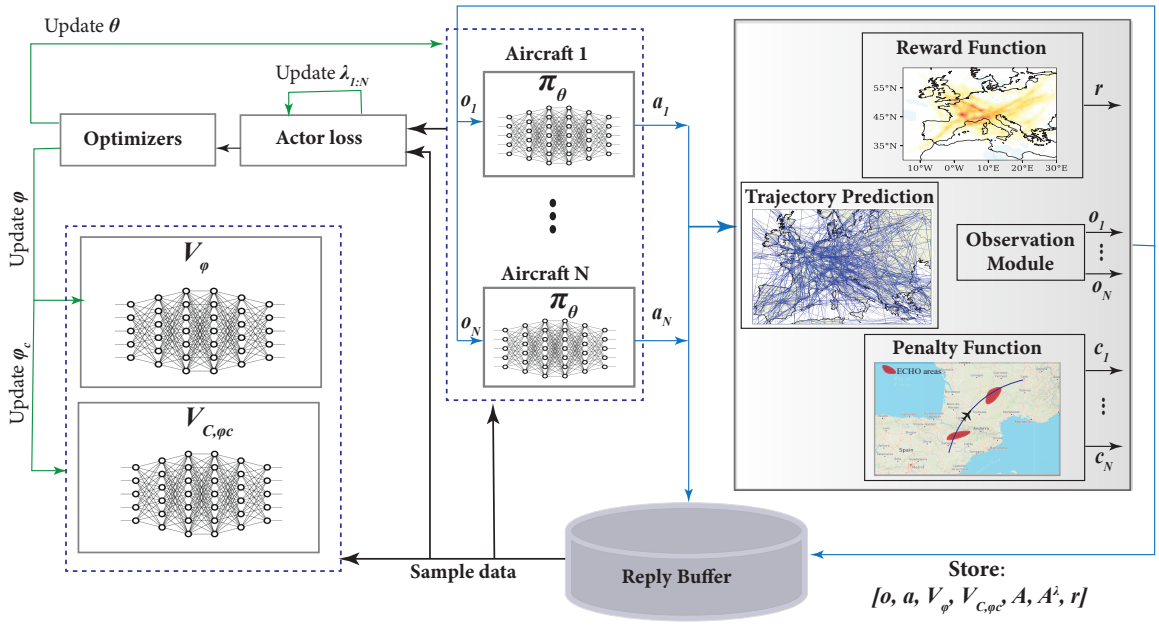


FIGURE 6.1: Overview of the proposed framework. Each aircraft receives a local observation about the surrounding traffic and the location of climate hotspots (ECHO areas). Based on these observations, the aircraft executes actions according to a trained policy aimed at avoiding hotspots while maintaining manageable aerial traffic.

The proposed approach employs a cooperative framework based on multi-agent proximal policy optimization (MAPPO) and adapts it to incorporate constraint handling for individual agents through the Lagrangian approach. Similar to the approach presented in Chapter 5, we employ a centralized training with decentralized execution paradigm, which facilitates efficient coordination among aircraft during training while allowing each agent to operate independently during execution [189]. Recognizing the need for scalability in scenarios involving varying numbers of agents, we implement shared policy parameters, ensuring flexibility across diverse air traffic situations.

The rest of this chapter is structured as follows. Section 6.1 defines the problem of network-scale aircraft trajectory optimization for climate impact mitigation, formulating it as a constrained optimization problem. Section 6.2 presents the proposed constrained multi-agent reinforcement learning (CMARL) algorithm to solve the stated problem. Section 6.3 then reformulates the network-scale optimization problem within the CMARL framework. Finally, Section 6.4 describes the simulation setup, outlines the implementation details of the proposed algorithm, and presents the experimental results.

Parts of the content presented in this chapter are adapted from the published work in [32].

6.1 Problem statement

This section formulates the climate-optimal flight planning problem at the network scale and presents its key components, including representation of air traffic, integration of climate impact, and consideration of air traffic manageability.

6.1.1 Air traffic

To formulate the problem of climate-optimal flight planning at the network scale, we consider a traffic consisting of N aircraft operating simultaneously in a shared airspace over a finite planning horizon T. Each aircraft $i \in \{1, \ldots, N\}$ follows a trajectory denoted by Γ^i which represents the sequence of its flight state variables over time and is defined as:

$$\Gamma^i = \left[(\varphi_0^i, \lambda_0^i, h_0^i, \chi_0^i, p_0^i, v_0^i, m_0^i, t_0^i), \cdots, (\varphi_k^i, \lambda_k^i, h_k^i, \chi_k^i, p_k^i, v_k^i, m_k^i, t_k^i) \right].$$

Each trajectory Γ^i is discretized into k time steps using a specified temporal resolution δt, such that $k = t_f^i / \delta t$, where t_f^i is the flight time of aircraft i. At each discretized point l, φ_l^i denotes latitude [degrees], λ_l^i longitude [degrees], h_l^i altitude [hft] (1 hft = 100 feet), χ_l^i heading angle [degrees], p_l^i flight phase (i.e., climb, cruise, descent), v_l^i speed [kts], m_l^i aircraft mass [Kg], and t_l^i time [s].

In this formulation, it is assumed that the initial flight plans of all aircraft are cost-optimal, and the objective is to iteratively modify these trajectories through a series of discrete actions in order to mitigate climate impact while maintaining traffic manageability. These actions, applied at each decision point, are defined as:

$$u_l^i = [\delta\varphi, \delta\lambda, \delta h, \delta v],$$

which represent adjustments to the aircraft's lateral position (latitude and longitude), altitude, and speed, respectively.

As a result of applying a control action u_l^i at step l, the flight plan is updated. The full trajectory is then recalculated to reflect the new profile, accounting for changes in mass, thrust, Mach number, and other performance-related variables. This update is performed using the trajectory prediction module based on the dynamical model presented in Chapter 4, Equation (4.1), in which the equations of motion are evaluated algebraically.

6.1.2 Climate impact

As presented in the previous chapters, the non-CO_2 climate impacts of aircraft emissions, including contrail formation and nitrogen oxides (NO_x)-induced effects, exhibit large spatiotemporal variability. Rather than accounting for the full spectrum of climate-sensitive areas with varying intensity levels in flight planning, a more practical alternative is to focus

exclusively on regions where emissions are expected to have a particularly strong climate impact. These regions are commonly referred to as climate hotspots or ECHO areas. Such areas are typically identified by applying predefined thresholds to filter climate-sensitive areas. In this framework, the aCCFs of all relevant species are first computed using the formulations presented in Chapter 3. These aCCFs are then merged to represent the overall climate sensitivity at each spatiotemporal location. A region is classified as a climate hotspot if its merged aCCF value exceeds a predefined threshold. This threshold is determined by computing the cumulative distribution function of the merged aCCF values and selecting a user-defined upper quantile (e.g., the top 5% corresponding to the 95^{th} percentile). For a detailed definition of climate hotspots and an approach to determine appropriate thresholds, the reader is referred to [190], which also describes the implementation in the CLIMaCCF library [190].

Once the hotspot areas are defined, a cost penalty is imposed on aircraft that intersect these regions in order to account for their associated climate impact.

6.1.3 Air traffic manageability

In addition to mitigating climate impact, ensuring the operational manageability of the resulting trajectories is a key objective. Air traffic manageability reflects the system's capacity to safely and efficiently accommodate new (updated) trajectories without overloading air traffic controllers or compromising safety. To incorporate manageability into the optimization framework, it is necessary to quantify it in a tractable and interpretable manner. In this study, air traffic complexity is used as a proxy for manageability.

Similar to Chapter 5, to quantify the air traffic complexity, we adopt the complexity score as the performance metric to provide insights into potential hazards within the airspace, focusing on the duration and severity of interactions rather than solely the presence of aircraft in the same volume. Two aircraft are considered to be interacting at a given time if, from each aircraft's perspective, the other is located within the same four-dimensional grid cell, defined over latitude, longitude, altitude, and time [94]. The complexity for each aircraft i with respect to aircraft k is computed as follows:

$$\Psi_t^{i,k} = \sum_{g_t}^{g_{t+\Delta t}} (v^{i,k} + \varkappa^{i,k} + v^{i,k}), \tag{6.1}$$

where g_t is the grid that aircraft i enters at time t, and $g_{t+\Delta t}$ is the cell that aircraft i exits at time $t + \Delta t$. The sum operator encompasses all cells that aircraft i crosses between g_t and $g_{t+\Delta t}$. The variables $v^{i,k}$, $\varkappa^{i,k}$ and $v^{i,k}$ are computed using the formulations presented in Chapter 5, Section 5.2.

6.1.4 Optimization problem formulation

The overall objective is to determine a set of trajectory modifications $\mathcal{U} = [u^0, \ldots, u^N]$, where $u^i = [u^i_0, \ldots, u^i_d]$ denotes the sequence of actions applied to the aircraft i over d decision steps, that jointly minimize traffic complexity while avoiding climate hotspot areas.

Formally, the problem can be stated as:

$$\min_{\mathcal{U}=[u^1,\ldots,u^N]} \quad \sum_{i=1}^{N} \sum_{\substack{k=1 \\ k\neq i}}^{N} \sum_{t=0}^{T} \Psi_t^{i,k}(u^i, u^k)$$

subject to

$$\Gamma^i(u^i) \cap \mathcal{H}_{\text{ECHO}} = \emptyset, \quad \forall i \in \{1, \ldots, N\}.$$

6.2 Constrained multi-agent reinforcement learning

To solve the above problem in a scalable and adaptive manner, we formulate it as a decentralized partially observable constrained Markov decision process, which is then addressed using a constrained multi-agent reinforcement learning (MARL) framework. In the following, we first introduce the decentralized partially observable constrained Markov decision process formulation and then present the proposed algorithm used to solve it.

6.2.1 Partially observable constrained Markov decision process

A partially observable constrained Markov decision process is defined by the tuple $\langle \mathcal{N}, \mathcal{S}, \mathcal{A}, \mathcal{O}, \mathcal{P}, \gamma, R, \{C^i\}_{i=1}^N, \{c^i\}_{i=1}^N, s_0 \rangle$, where, $\mathcal{N} = \{1, \ldots, N\}$ represents the set of agents, $\mathcal{S}$ is the state space, $\mathcal{A}^i$ denotes the action space for agent i, $\mathcal{A} = \mathcal{A}^1 \times \cdots \times \mathcal{A}^N$ is the joint action space for all agents, $o^i = \mathcal{O}(\mathcal{S}, i)$ represents the local observation for agent i at state s, $\mathcal{P} : \mathcal{S} \times \mathcal{A} \to \Delta(\mathcal{S})$ is the state transition probability function, $\gamma \in [0, 1)$ is the discount factor, $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is the shared reward function, $C^i : \mathcal{S} \times \mathcal{A}^i \to \mathbb{R}$ is the cost function (penalty function) for agent i, with a cost threshold c^i.

We denote the initial state of the environment by $s_0 \in \mathcal{S}$. At each time step t, agent i observes o^i_t and selects an action a^i_t according to a randomized stationary policy $\pi^i(\cdot|o^i_t) \in \Pi^i$. The joint action of all agents $\mathbf{a}_t = (a^1_t, \ldots, a^N_t)$ is then executed, transitioning the system to a new state $s_{t+1} \sim \mathcal{P}(\cdot|s_t, \mathbf{a}_t)$. Each agent i then receives a reward $R(s_t, \mathbf{a}_t)$ and incurs a penalty cost $C^i(s_t, a^i_t)$. In this fully cooperative setting, the reward function R depends on the joint actions of all agents, reflecting its coupling across agents. The constraints are decoupled, as each agent's cost C^i depends only on its own actions a^i.

The set of joint policies is denoted by $\pi = \{\pi_i\}_{i\in\mathcal{N}}$ and is represented as $\Pi := \Pi^1 \times \cdots \times \Pi^N$. For any joint policy $\pi \in \Pi$, we define the reward value function at state s as

$V_R^\pi(s) := \mathbb{E}_{\mathbf{a}_t \sim \pi, s_t \sim \mathcal{P}} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, \mathbf{a}_t) \mid s_0 = s \right]$ and cost value function at state s as $V_{C^i}^\pi(s) := \mathbb{E}_{\mathbf{a}_t \sim \pi, s_t \sim \mathcal{P}} \left[\sum_{t=0}^{\infty} \gamma^t C^i(s_t, a_t^i) \mid s_0 = s \right]$ for $i \in \mathcal{N}$.

The goal is to find a policy that maximizes the reward value function $V_R^\pi(s_0)$ while ensuring that the constraint $V_{C^i}^\pi(s_0) \leq c^i$ is satisfied for every agent i. Formally, this is expressed as:

$$\max_{\pi \in \Pi} V_R^\pi(s_0), \text{ s.t. } V_{C^i}^\pi(s_0) \leq c^i, \forall i \in \mathcal{N}. \tag{6.2}$$

In this study, we focus on a large population of agents that are assumed to be homogeneous. This assumption is justified by a common reward function that aligns all agents' interests toward minimizing air traffic complexity while avoiding climate-sensitive regions. Such homogeneity also implies that the agents play interchangeable roles in the system's evolution and are nearly indistinguishable from each other [191]. Due to the homogeneity of the agents, parameter sharing can be applied to enhance scalability and training efficiency [187]. This allows all agents to use a single shared policy [186,192]. The shared policy, denoted by π_θ and parameterized by θ, enables training to utilize the collective experience of all agents. Meanwhile, each agent i can still take its own actions, $\pi_\theta(\cdot|o_t^i)$, based on its observations o_t^i [188].

6.2.2 Constrained multi-agent proximal policy optimization

The constrained Markov decision process formulated above can be addressed using a constrained multi-agent reinforcement learning (MARL) framework. These approaches allow agents to learn policies that maximize cumulative rewards while satisfying individual or shared constraints. However, solving constrained MARL problems involves several challenges. These include non-stationarity, where the environment evolves dynamically in response to the actions of multiple agents; scalability, where computational complexity grows exponentially with the number of agents; and training stability, where large policy updates can lead to instability and performance collapse by erasing previously learned good behaviors. Additionally, balancing reward optimization with constraint satisfaction further complicates the problem.

In this study, we address these challenges by developing a constrained multi-agent reinforcement learning framework that builds on Multi-Agent Proximal Policy Optimization (MAPPO) [186]. We extend MAPPO to handle explicit constraints via a Lagrangian relaxation, embedding constraint costs directly into the learning objective to ensure reliable safety performance [193]. The resulting optimization is formulated as the following max–min problem:

$$\max_{\theta} \min_{\lambda^i \geq 0, i \in \mathcal{N}} V_R^{\pi_\theta}(s_0) - \sum_{i \in \mathcal{N}} \lambda^i (V_{C^i}^{\pi_\theta}(s_0) - c^i). \tag{6.3}$$

To solve this problem, we apply gradient descent on the Lagrange multipliers $\{\lambda^i\}_{i\in\mathcal{N}}$ and gradient ascent on the policy parameters θ. However, directly applying gradient ascent on θ can lead to large, unstable updates, potentially causing the policy to forget previously learned good behaviors, which results in performance collapse. The PPO framework mitigates this issue by employing trust region optimization, which constrains the magnitude of policy updates. PPO achieves this by clipping the probability ratio $\frac{\pi_\theta}{\pi_{\theta_{\text{old}}}}$ within $(1-\epsilon, 1+\epsilon)$, ensuring that the new policy π_θ remains close to the old policy $\pi_{\theta_{\text{old}}}$. This clipping mechanism enhances stability and enables more controlled updates.

To introduce the MAPPO method [186], we define the reward state-action value function and cost state-action value function as follows:

$$Q_R^{\pi_\theta}(s, \mathbf{a}) := \mathbb{E}_{\mathbf{a}_t\sim\pi_\theta, s_t\sim\mathcal{P}}[\sum_{t=0}^{\infty}\gamma^t R(s_t, \mathbf{a}_t)|s_0 = s, \mathbf{a}_0 = \mathbf{a}],$$

$$Q_{C^i}^{\pi_\theta}(s, a^i) := \mathbb{E}_{\mathbf{a}_t\sim\pi_\theta, s_t\sim\mathcal{P}}[\sum_{t=0}^{\infty}\gamma^t C^i(s_t, a_t^i)|s_0 = s, a_0^i = a^i],$$

for $i \in \mathcal{N}$. The advantage function is defined as:

$$A_u^{\pi_\theta}(s, a) := Q_u^{\pi_\theta}(s, a) - V_u^{\pi_\theta}(s),$$

for $u \in \{R\} \cup \{C^i | i \in \mathcal{N}\}$. This function evaluates the benefit of taking action a in state s relative to the baseline value $V_u^{\pi_\theta}(s)$. Using these definitions, the MAPPO objective is formulated as:

$$\begin{aligned} L\left(\theta, \{\lambda^i\}_{i\in\mathcal{N}}\right) := \mathbb{E}_{\mathbf{a}\sim\pi_\theta, s\sim p}\Bigg[\sum_{i=1}^{N}\min\Bigg(&\frac{\pi_\theta(a^i \mid o^i)}{\pi_{\theta_{\text{old}}}(a^i \mid o^i)}A_{\lambda^i}^{\pi_\theta}(s, \mathbf{a}), \\ &\text{clip}\left(\frac{\pi_\theta(a^i \mid o^i)}{\pi_{\theta_{\text{old}}}(a^i \mid o^i)}, 1-\epsilon, 1+\epsilon\right)A_{\lambda^i}^{\pi_\theta}(s, \mathbf{a})\Bigg)\Bigg], \end{aligned} \tag{6.4}$$

where $A_{\lambda^i}^{\pi_\theta}(s, \mathbf{a})$ is defined as:

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

The advantage functions are computed using generalized advantage estimation [141]. Here, λ^i penalizes constraint violations.

To solve problem (6.3), we iteratively apply the following update rules:

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

where α_λ and α_θ are the learning rates for updating $\{\lambda^i\}_{i\in\mathcal{N}}$ and θ, respectively. These updates balance constraint satisfaction with reward maximization at the individual agent level. In practice, automatic differentiation frameworks such as TensorFlow and PyTorch facilitate this gradient computation by calculating $\nabla_\theta \mathcal{L}(\theta)$ automatically once the objective function is defined; the optimizer subsequently applies this gradient to update parameters accordingly [189]. The details of CMAPPO are summarized in Algorithm 4.

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

In this section, we outline the key components of the MARL framework used to solve the network-scale flight planning problem for the benefit of the climate. Specifically, we define the observation space, action space, reward function, and cost (penalty) function.

To ensure better alignment with operational realities, we adopt a modeling assumption in which the MARL environment is formulated using distance-based decision points rather than discrete time steps. Specifically, each episode begins at the origin of a trajectory, and observations, actions, and updates are performed at fixed spatial intervals, for example, every Δd nautical miles flown. This formulation provides a consistent and operationally meaningful decision structure, as elapsed time can vary across aircraft due to differences in speed, whereas distance flown offers a uniform basis for control and coordination.

6.3.1 State

The global state s_d at decision step d is constructed by concatenating the local observations from all aircraft:

$$s_d = \left[o_d^1, \cdots, o_d^N\right].$$

6.3.2 Observation

The local observation o_d^i for each aircraft consists of the following components:

- **Trajectory information**
  This includes data on the aircraft's trajectory up to the current decision point:

  $$\Gamma_d^i = \left[(t_1^i, \chi_1^i, v_1^i, p_1^i), \ldots, (t_j^i, \chi_j^i, v_j^i, p_j^i)\right].$$

  The trajectory over the interval $[d, d + \Delta d]$ is discretized into j points (j is fixed for all flights to have the same observation dimension).

- **Information about neighboring aircraft**
  The information about neighboring aircraft within a certain vicinity is provided, which

Algorithm 4 Constrained Multi-Agent Proximal Policy Optimization (Constrained MAPPO)

Input: α_θ, α_λ, Batch size B, Number of agents N, episodes K, steps per episode T, discount factor γ, safety threshold c, and GAE parameter λ
Initialize policy parameters θ, value function parameters ϕ, cost value function parameters ϕ_c, and Lagrange multipliers λ_i for $i = 1, 2, ..., N$
Initialize optimizers for policy, value function, and cost value function
for $k = 0, 1, ..., K-1$ **do**
    Collect a set of trajectories using policy π_θ
    Compute rewards $r_t = R(s_t, a_t)$, costs $c_t^i = C(o_t^i, a_t^i)$, and log-probabilities $\log \pi_\theta(a_t^i | o_t^i)$
    Push transition $\{o_t^i, a_t^i, o_{t+1}^i, r_t, c_t^i\}$ into the replay buffer
    for each mini-batch B **do**
        Compute advantage function $A_R(s, a)$ based on value network with GAE

$$\delta_{R_t} = R(s_t, \mathbf{a}_t) + \gamma V_{R,\phi}(s_{t+1}) - V_{R,\phi}(s_t)$$

$$A_R(s_t, \mathbf{a}_t) = \sum_{l=0}^{T} (\gamma\lambda)^l \delta_{R_{t+l}}$$

        Compute

$$V_{R,\text{target}} = A_R(s_t, a_t) + V_{R,\phi}(s_t)$$

        for each agent i **do**
            Compute cost advantage functions $A_C^i(o^i, a^i)$, based on cost value networks with GAE

$$\delta_{C_t}^i = C(o_t^i, a_t^i) + \gamma V_{C,\phi_c}(o_{t+1}^i) - V_{C,\phi_c}(o_t^i)$$

$$A_C^i(o_t^i, a_t^i) = \sum_{l=0}^{T} (\gamma\lambda)^l \delta_{C_{t+l}}^i$$

            Compute

$$V_{C,\text{ target}}^i = A_C^i(o_t^i, a_t^i) + V_{C,\phi_c}(o_{t+1}^i)$$

            Calculate policy ratio $r_t^i = \frac{\pi_\theta(a_t^i | o_t^i)}{\pi_{\theta_{\text{old}}}(a_t^i | o_t^i)}$
            Compute average episode cost $\bar{C}^i = \frac{1}{T}\sum_{t=1}^{T}$

$$\Delta\lambda^i = -\frac{1}{BT}\sum_{j=1}^{B}\sum_{t=0}^{T}\Big((C(o_t^i, a_t^i) - c) + (r_t A_C^i(o_t^i, a_t^i))\Big)$$

            Update Lagrange multipliers $\lambda^i \leftarrow \text{ReLU}\left(\lambda^i - \alpha_\lambda \Delta\lambda^i\right)$
        end for
        Compute surrogate objective $L(\theta, \lambda)$ according to Equation (6.4)
        Update policy parameters θ via gradient ascent on $L(\theta, \lambda)$
        Update value function parameters ϕ and cost value function parameters ϕ_c via gradient descent on L_V and $L_{C,V}$

$$L_V = \frac{1}{BT}\sum_{j=1}^{B}\sum_{t=0}^{T}\left[\left(V_{R,\phi}(s_t) - V_{R,\text{target}}\right)^2\right]$$

$$L_{C,V} = \frac{1}{BTN}\sum_{j=1}^{B}\sum_{t=0}^{T}\sum_{i=1}^{N}\left[\left(V_{C,\phi_c}(o_t^i) - V_{C,\text{target}}^i\right)^2\right]$$

    end for
end for

aids in assessing local air traffic complexity. The information about neighboring aircraft for the agent i is given by:

$$I_d^i = \left[(t_d^1, \chi_d^1, v_d^1, p_d^1), \ldots, (t_d^m, \chi_d^m, v_d^m, p_d^m) \right],$$

where m is the number of neighboring aircraft.

- **Climate hotspot information**
  The hotspot information is defined as:

  $$E = [e_1, \ldots, e_r],$$

  where each element e_r represents the cumulative hotspot impact associated with the action combination r over the interval $[d, d + \Delta d]$. Thus, E captures the aggregated environmental effect across all available actions at the current decision step.

Based on the above, the local observation for agent i at step d is defined as:

$$o_d^i = \left[\tau_d^i, I_d^i, E \right].$$

6.3.3 Action

The action space for each agent is defined by potential modifications to the aircraft's trajectory, encompassing three types of maneuvers: lateral deviations, altitude changes, and speed adjustments. Specifically, agents can modify their flight path by altering (I) the position of trajectory nodes in latitude and longitude, (ii) the flight level, or (iii) the Mach number.

- **Lateral deviations** are implemented by modifying the geographic coordinates (latitude and longitude) of the next discretized trajectory points. Specifically, the position is adjusted by δl degrees in both latitude and longitude, where:

  $$\delta l = [-0.4, 0.2, 0.0, 0.2, 0.4].$$

- **Altitude adjustments** are performed by modifying the flight level by δh, expressed in units of 100 feet, where:

  $$\delta h = [-40.0, -20.0, 0.0, 20.0, 40.0].$$

- **Speed adjustments** involve modifying the Mach number by:

  $$\delta M = [-0.03, 0.0, 0.03].$$

6.3.4 Reward function

The reward function R provides the immediate reward received by agents for transitioning from state s to state s' due to the joint action $\mathbf{a}$. In this study, the reward is defined as follows:

$$R = -\sum_{i=1}^{N} \sum_{k=1, k\neq i}^{N} \Psi^{i,k},$$

where $\Psi^{i,k}$ is computed using Equation (6.1). In this formulation, the reward is always non-positive, as it is defined as the negative sum of complexity. The maximum attainable reward is zero, corresponding to the ideal scenario in which all complexity terms vanish, indicating a manageable traffic configuration.

6.3.5 Constraints

The avoidance of climate hotspots is modeled as a set of constraints that each agent must satisfy in order to ensure climate efficiency of planned routes. We define the following function to quantify the cost of constraint violation (i.e., the intersection of flight trajectories with the climate-sensitive areas):

$$C_d^i = \begin{cases} 0 & \text{if } \tau_d^i \cap E = \emptyset, \\ c_h & \text{otherwise.} \end{cases}$$

Here, τ_d^i represents the trajectory segment of aircraft i during interval $[d, d + \Delta d]$, c_h is a penalty term proportional to the number of hotspot violations, and E denotes ECHO areas assumed to be fixed during the considered interval. The equation implies that when an aircraft flies through climate-sensitive areas, a cost c_h is incurred.

6.4 Simulation results

This section evaluates the performance of the proposed approach through a real-world case study. First, the experimental setup and simulation details are described, followed by a presentation and analysis of the obtained results.

6.4.1 Experimental setup

The case study considered in this section is the same as that introduced in Chapter 5. It is based on a real traffic scenario over the ECAC airspace on December 20, 2018, encompassing all the flights operating between 12:00 UTC and 16:00 UTC. The weather data, including wind and temperature, is obtained from the ERA5 reanalysis data products. The initial flight trajectories are generated using our in-house trajectory planning tool, ROOST, as detailed in

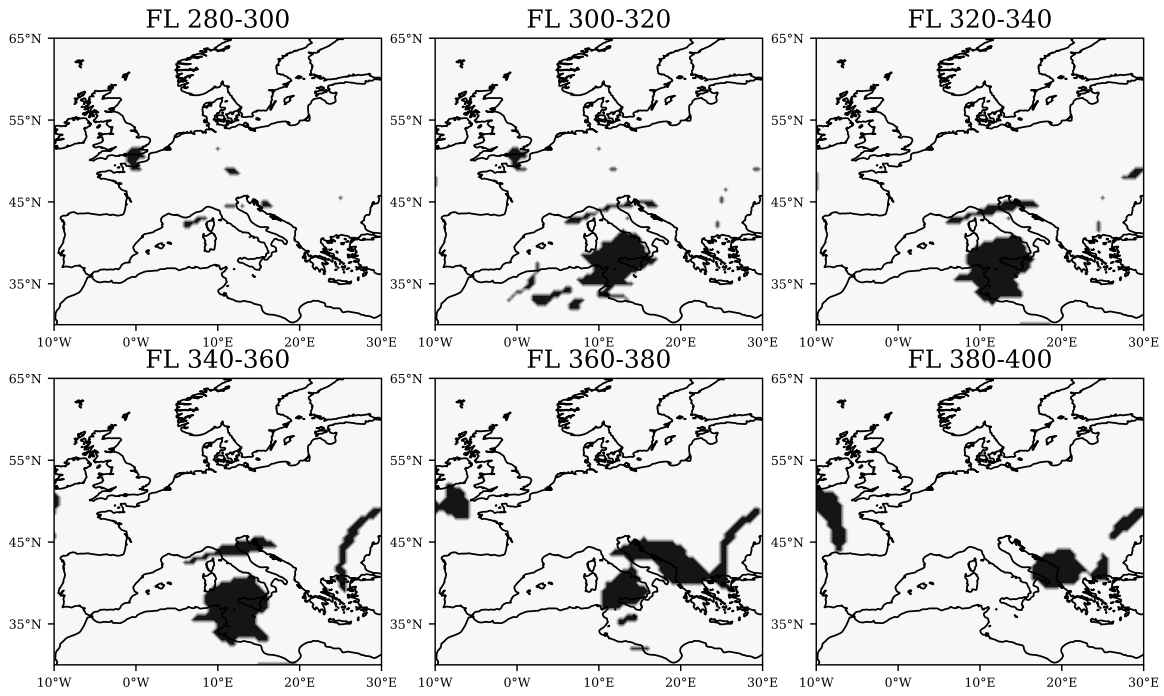


FIGURE 6.2: The identified climate hotspots for December 20, 2018, at 12:00 UTC.

Chapter 5. These trajectories correspond to the EI value of 0.0, representing the cost-optimal trajectories. Although the initial trajectories in this study are generated using ROOST, the framework is flexible and can use other planned trajectories, such as those available from the DDR2 or automatic dependent surveillance–broadcast (ADS-B) dataset.

By inputting meteorological data, the CLIMaCCF library provides climate hotspot regions, as detailed in Section 6.1.2. The output is a gridded dataset covering the European airspace, with a horizontal resolution of 0.5°× 0.5°and a vertical resolution of 20 hft. To ensure sufficient variability in hotspot distribution during training, one full year of climate hotspot data from 2018 is used. Figure 6.2 illustrates an example of the identified hotspots for December 20, 2018, at 12:00 UTC. In this study, only the aCCF associated with contrails is used to identify the hotspots, as contrails have been shown to contribute the most to climate impact and offer the greatest mitigation potential through flight planning [62]. Accordingly, the resulting hotspot map shows regions with a strong warming potential due to contrail formation. During trajectory planning, agents are encouraged to avoid these warming-sensitive areas. It should be noted that regions with potential cooling effects are not considered in the hotspot identification process.

Given the computational complexity of the traffic scenario, which involves approximately 6,000 flights, a graph-based clustering strategy is implemented to partition the traffic scenario into smaller subsets of interacting flights. In this approach, each flight is represented as a node in a graph, and edges are established between flights (nodes) if they are

within a defined spatial vicinity, within 2 degrees of latitude and longitude, and 50 hft of altitude.

For each flight, a subgraph is formed consisting of the flight itself and all directly connected neighboring aircraft. This grouping captures local interactions relevant to air traffic complexity while allowing the exclusion of distant flights that have no influence on the trajectory under consideration. Since air traffic complexity is a coupled objective, dependent on spatiotemporal interactions between flights, this method ensures that only relevant subsets of traffic are considered, significantly reducing the computational burden without compromising the fidelity of the complexity estimation.

Once all preprocessing steps, such as trajectory clustering and climate hotspot generation, are completed, the proposed strategy introduced in Section 6.2.2 is implemented. The experimental setup incorporates the state space, action space, reward, and cost formulations detailed in Section 6.3. The learning architecture consists of three neural modules: an actor network that approximates each agent's policy, a cost value network that estimates the penalty associated with climate hotspot violations, and a value network based on a graph neural network (GNN) that evaluates the overall complexity of the traffic configuration.

The actor network is implemented as a fully connected multi-layer perceptron (MLP). It receives the agent's observation as input and processes it through three hidden layers, each with 64 neurons, using the *Tanh* activation function. The final layer outputs action logits, which are passed through a softmax function to yield a probability distribution over the discrete action space. Orthogonal weight initialization is applied across all layers to improve training stability.

The cost network shares the same MLP architecture as the actor but serves to estimate the expected cost associated with an agent's trajectory, specifically the extent of hotspot violation. The network takes the agent's state as input and outputs a scalar value using a $softplus$ activation function, ensuring non-negativity. This cost estimate is integrated into the constrained optimization framework via Lagrangian relaxation, allowing the agent to learn a policy that respects the safety threshold. The safety threshold is set to 1 to prevent Lagrangian over-penalization in cases where hotspot avoidance is infeasible. Specifically, some trajectories may inevitably intersect with hotspot regions, regardless of the action chosen. Allowing a limited tolerance ensures that the learning process remains stable even when strict constraint satisfaction is unattainable.

The value network is used to estimate the complexity of the current traffic configuration and is implemented using a GNN to accommodate the variable number of aircraft in each traffic cluster. Unlike standard MLP that require fixed input sizes, the GNN architecture enables flexible reasoning over dynamic interaction graphs. Each node in the graph represents an aircraft, and edges encode spatial proximity relations. The network includes two stacked graph convolutional network layers, each with 64 hidden units, followed by a global mean pooling operation that aggregates the node-level features into a fixed-size graph embedding.

TABLE 6.1: Hyperparameters used in the experiment.

Hyperparameters	Values	Hyperparameters	Values
lr rate	1e-4	Batch size	124
Safety_bound	1	ϵ	0.1
Eval. episodes	1000	Lagrangian coef. rate	0.001
Optimiser	Adam	N mini-batch	24

This embedding is then passed through two fully connected layers to produce a scalar estimate of the value function, representing the complexity of the traffic from a network-level perspective. ReLU activations are used within all layers, enabling nonlinear transformations of the aggregated features.

To ensure stable and efficient training, several well-established reinforcement learning techniques are employed, including reward normalization, advantage normalization, entropy regularization (with a coefficient of 0.01), learning rate decay, and gradient clipping. Details of the hyperparameters used in the experiment are presented in Table 6.1.

6.4.2 Results

This section presents the experimental results obtained from implementing the proposed CMAPPO algorithm for climate-aware aircraft trajectory planning.

To illustrate the learning performance of the proposed algorithm, the evolution of the reward and cost during training is shown in Figure 6.3 and Figure 6.4, respectively. The results are recorded at evaluation intervals of 1000 training steps. For clarity of presentation, a moving average with a window size of 500 is applied to smooth the curve. Figure 6.3 illustrates the learning performance of the CMAPPO framework, showing the evolution of the average episodic reward over the training process. As depicted, the reward increases steadily throughout training, indicating consistent improvement in policy performance and successful adaptation to the trajectory planning task under constraint-aware conditions.

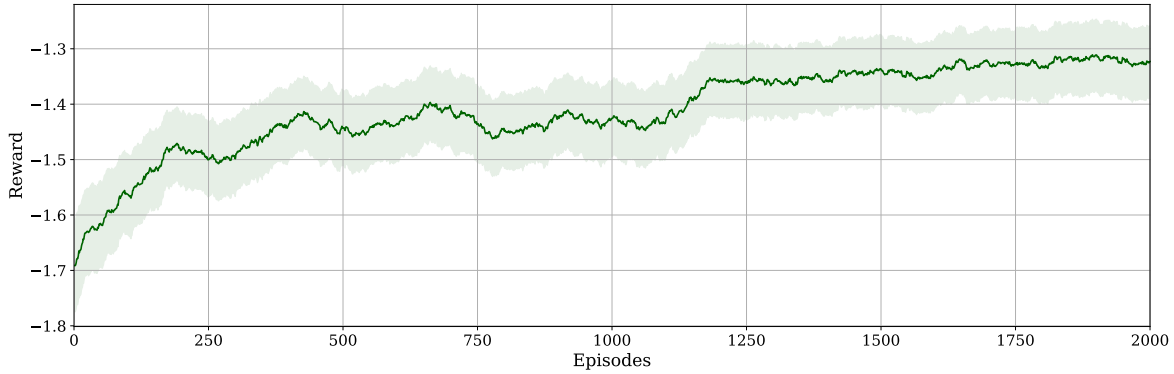


FIGURE 6.3: Evolution of the average episodic reward over training steps.

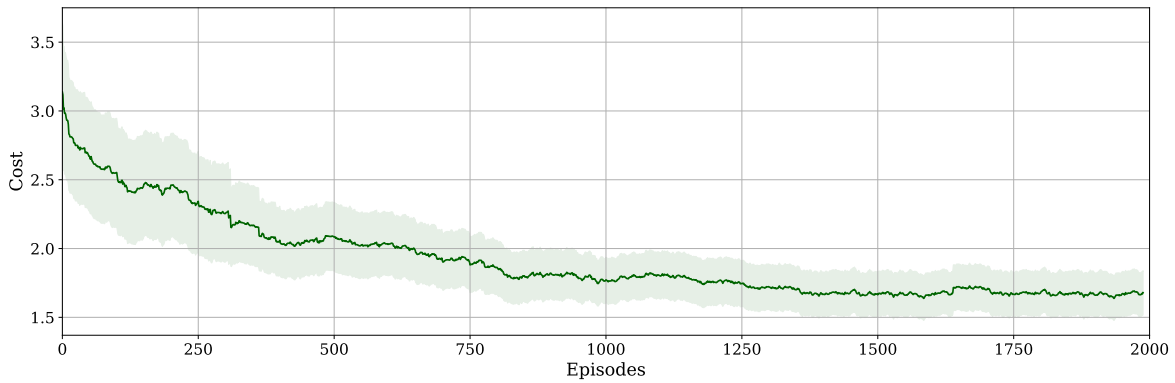


FIGURE 6.4: Evolution of the average episodic cost associated with hotspot violations over training steps.

Figure 6.4 shows the evolution of the average cost associated with hotspot violations. For each episode, this cost is computed as the average violation across all agents, providing a measure of how well the multi-agent system adheres to the constraints on a per-flight basis. A clear downward trend is observed during the training phase, indicating that the agents progressively learn to avoid climate-sensitive regions. This result demonstrates that the Lagrangian-based optimization framework effectively encourages constraint satisfaction while allowing for limited, unavoidable violations in scenarios where strict compliance is not feasible. It can be seen that the cost curve exhibits more stable behavior compared to the reward curve. This can be attributed to the nature of the optimization objectives: while the reward is coupled across agents, requiring coordinated behavior and mutual adaptation, the cost is computed on an individual basis, making it comparatively easier for each agent to satisfy the constraint independently.

The policy derived from the proposed CMAPPO algorithm is applied to the presented scenario (i.e., December 20, 2018, from 12:00 to 16:00 UTC). In this study, only aircraft crossing climate hotspots or associated with high traffic complexity are allowed to deviate from their original flight plans, while the rest follow their planned business-as-usual (cost-optimal) trajectories. Figure 6.5 compares the flight profiles of business-as-usual trajectories with those optimized using the proposed CMAPPO framework.

Figure 6.6 presents a comparative analysis between the performance of the business-as-usual trajectories and those optimized using the CMAPPO algorithm. As shown in the first figure, the optimized trajectories using CMAPPO achieve 87.2% reduction in contrail climate impact, measured in terms of ATR. This contributes to an overall 9.1% reduction in net climate impact when considering all relevant forcing agents. The third figure indicates a 5.2% reduction in aggregated air traffic complexity, indicating that the algorithm not only mitigates environmental impact but also improves traffic manageability by reducing traffic complexity. However, as illustrated in the bottom figure, the observed environmental and complexity-related benefits are accompanied by a 0.7% increase in operational cost relative

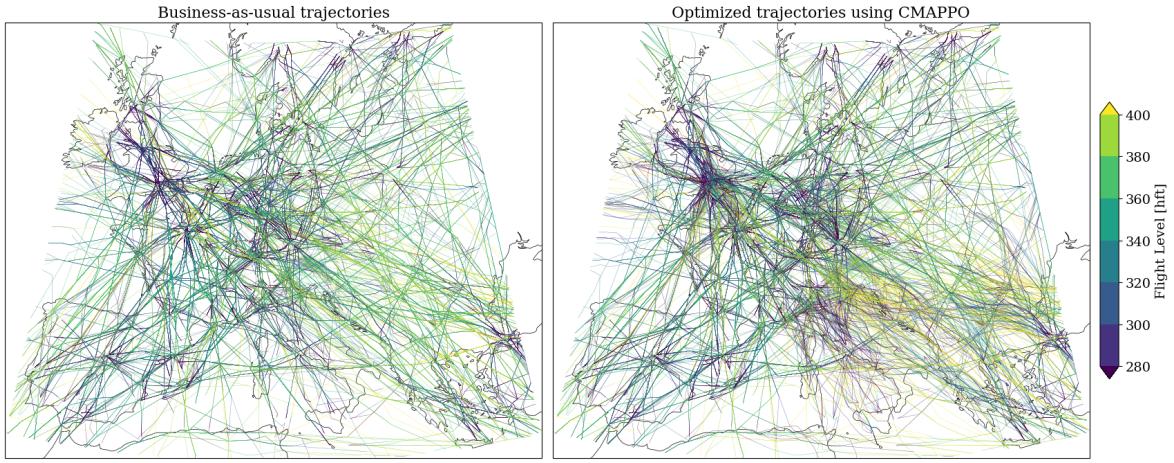


FIGURE 6.5: Traffic distribution of business-as-usual trajectories and those optimized using the proposed CMAPPO algorithms, color-coded based on flight levels. The color bar indicates flight altitudes.

to the cost-optimal baseline. This increase reflects a 0.3% rise in total flight time and a 1.7% increase in fuel consumption.

To visualize the mitigation of contrail-induced climate impact, Figure 6.7 presents the ATR of contrails formed along the original trajectories (Figure 6.7a), the trajectories optimized using the proposed CMAPPO algorithm (Figure 6.7b), and their difference (Figure 6.7c). As illustrated, the use of the proposed approach results in a reduction in warming contrail impacts, particularly at higher flight levels (e.g., FL 360–400), where intense warming impacts are significantly diminished. Nevertheless, residual warming impacts are observed in certain regions. This is expected, as certain climate-sensitive regions may be unavoidable due to trajectory feasibility constraints, or no feasible control action can completely avoid them. Moreover, Figure 6.7c reveals a few localized increases in warming, which are limited in both magnitude and spatial extent relative to the regions exhibiting reductions. A likely explanation is that the optimization process focuses solely on avoiding warming contrails and does not explicitly consider the preservation or increase of cooling contrails. As a result, some cooling contrails may be inadvertently reduced due to trajectory adjustments aimed at minimizing complexity, leading to increases in net warming in specific regions.

To evaluate the impact of the proposed optimization strategy on air traffic complexity, Figure 6.8 presents the spatial distribution of the complexity score across different flight levels. Figure 6.8a shows the complexity map for business-as-usual trajectories, while Figure 6.8b illustrates the corresponding map for trajectories optimized using the CMAPPO algorithm. A visual comparison between the two figures indicates a high degree of similarity, suggesting that the overall level of complexity remains comparable to the baseline scenario.

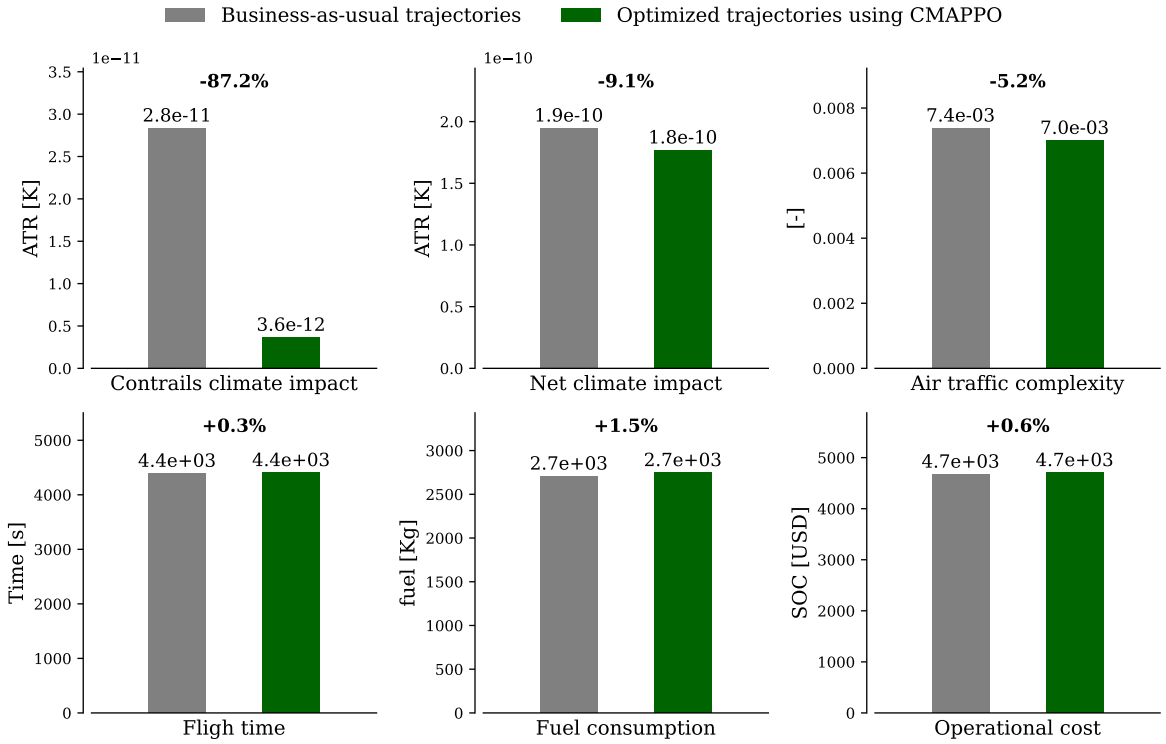
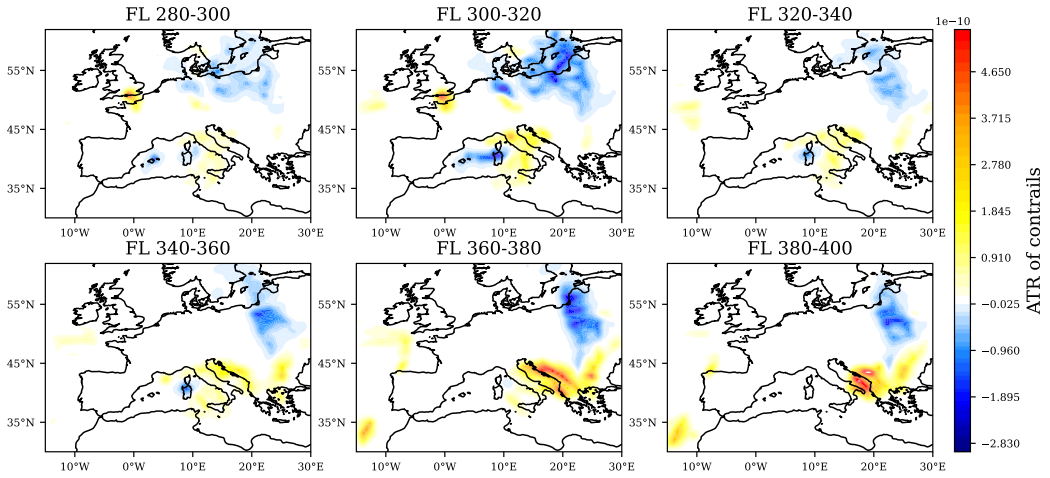
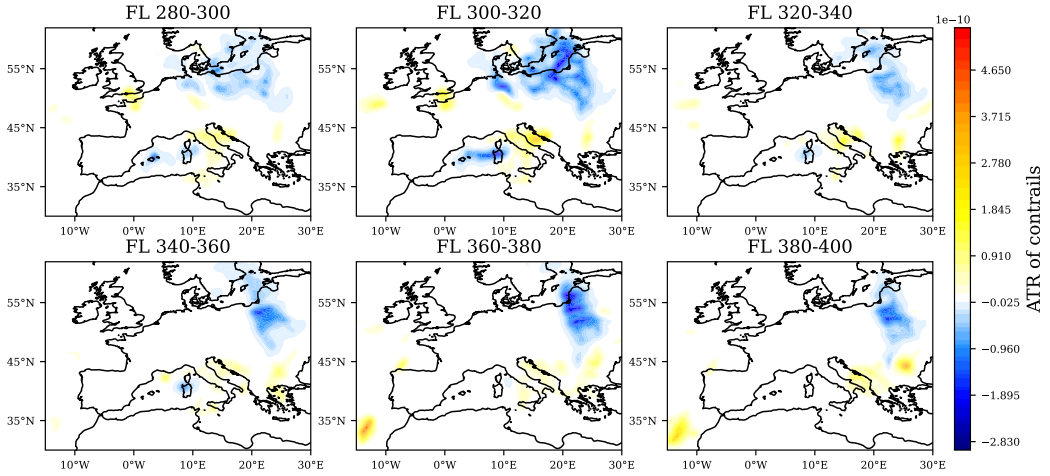


FIGURE 6.6: Performance comparison between business-as-usual trajectories and those optimized using the proposed constrained MAPPO algorithm.

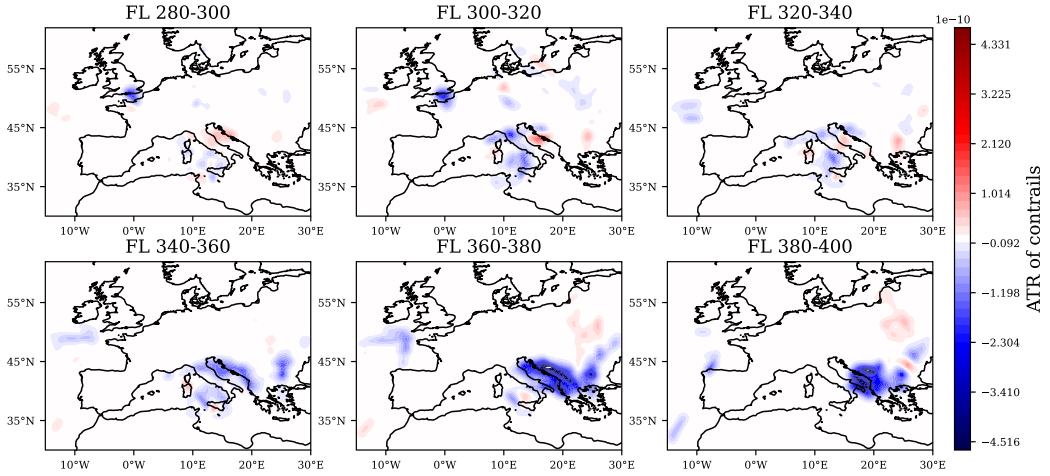
The difference between the two scenarios is illustrated in Figure 6.8. The figure highlights regions where complexity has either increased or decreased as a result of trajectory modification. As shown, the overall changes in complexity are relatively small in magnitude. However, regions with decreased complexity (in blue) are more frequent and spatially widespread than those with increased complexity (in red). This indicates that the proposed algorithm consistently reduces complexity across broader areas of the airspace. These findings confirm that the CMAPPO algorithm is capable of generating climate-aware trajectories while maintaining the manageability of air traffic relative to current operations.



(A) ATR of contrails along business-as-usual trajectories.

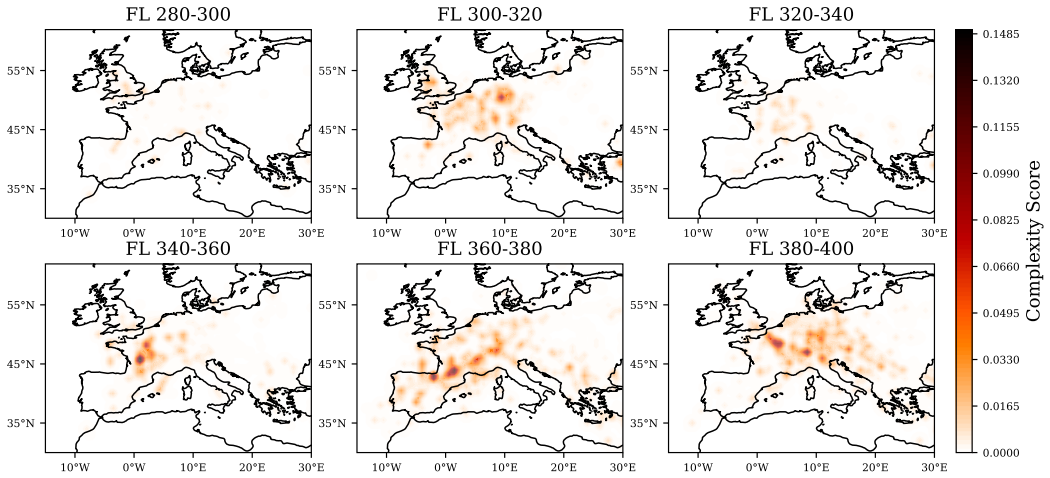


(B) ATR of contrails along trajectories optimized using the CMAPPO algorithm.

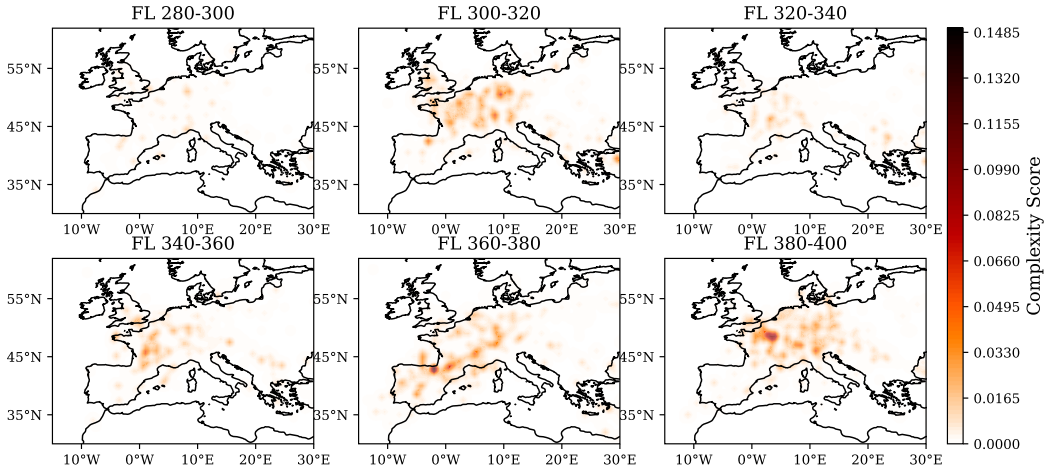


(C) Difference in ATR of contrails between the two scenarios.

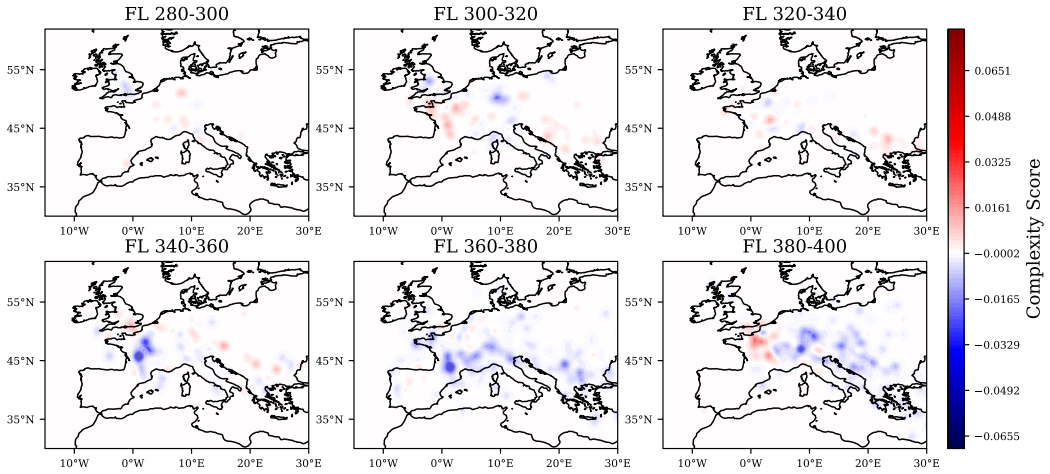
FIGURE 6.7: Comparison of the spatial distribution of persistent contrails formed along business-as-usual trajectories and trajectories optimized using the CMAPPO algorithm.



(A) Complexity map of business-as-usual trajectories.



(B) Complexity map of optimized trajectories using the CMAPPO algorithm.



(C) Difference in complexity between the two scenarios.

FIGURE 6.8: Comparison of air traffic complexity between business-as-usual trajectories and trajectories optimized using the CMAPPO algorithm. Blue regions indicate a reduction in complexity under the CMAPPO policy, while red regions denote localized increases.

6.5 Summary

This chapter introduced an integrated approach for climate-optimal flight planning considering air traffic complexity. The problem was formulated as a constrained MARL, where air traffic complexity served as the objective function, and climate hotspot avoidance was imposed as constraints. The proposed method employed the multi-agent proximal policy optimization algorithm and adapted it to handle constraints related to climate hotspot avoidance using the Lagrangian technique. The developed method was evaluated using a real-world traffic scenario over European airspace. The results demonstrated that, compared to business-as-usual trajectories, the integrated method effectively reduces both climate impact and traffic complexity, albeit with an increase in operational cost.

Compared to the sequential framework presented in Chapter 5, the integrated approach resulted in lower performance in terms of both climate impact mitigation and traffic complexity reduction under the same traffic scenario. Specifically, for an EI of 0.01, the sequential approach achieved a 14.37% reduction in climate impact, accompanied by a 0.7% increase in operational cost and a 15.8% reduction in overall traffic complexity. In contrast, the integrated CMAPPO-based approach resulted in a 9.1% reduction in climate impact, a 0.6% increase in operational cost, and a 5.2% reduction in complexity.

This difference in performance is primarily attributable to the optimization scope. In the sequential framework, all non-CO_2 forcing agents were considered in the optimization process, thereby enabling additional mitigation through the reduction of other non-CO_2 climate effects(e.g., NO_x) and formation of cooling contrails, which accounted for approximately two-thirds of the total mitigation achieved. In contrast, the integrated approach developed in this chapter focused exclusively on avoiding strong warming contrails and did not explicitly account for other non-CO_2 forcing agents or the climate benefits achievable through the formation of cooling contrails. This design choice inherently limits the achievable mitigation. Specifically, the 9.14% net climate mitigation achieved by the integrated framework resulted from a 12.18% reduction in contrail-related climate impact, offset by a 3.03% increase in the impact of other forcing agents (e.g., NO_x-induced effects), partially attributed to a 1.47% increase in fuel consumption. Of the mitigation achieved through contrails, 98.30% resulted from the avoidance of warming contrails, while the remaining 1.69% was due to an unintended increase in cooling effects. While there is ongoing debate in the literature about the extent to which cooling contrails should be considered, this limitation could be addressed by extending the current framework also to target cooling-sensitive regions and incentivize agents to pass through them when beneficial.

The sequential approach results in a more pronounced reduction in air traffic complexity compared to the integrated method. This can be attributed to the design of its optimization process, in which complexity reduction is the sole objective of the MARL framework. Moreover, during simulation, all aircraft contributing to complexity are allowed to execute

maneuvers, maximizing the resolution potential. In contrast, the integrated CMAPPO setup incorporates two simultaneous objectives, climate impact mitigation and complexity reduction, which naturally introduces trade-offs that can limit the extent of achievable complexity reduction. Furthermore, to minimize the number of rerouted flights from the BAU trajectories, only aircraft intersecting climate hotspots are permitted to take action during execution. This constraint further limits the system's ability to reduce overall complexity.

It is also worth noting that the integrated approach demonstrates slightly improved cost efficiency, achieving a lower increase in operational cost compared to the sequential method. Furthermore, the integrated approach offers advantages in terms of scalability and applicability to large-scale or real-time scenarios. Unlike the sequential method, which requires an initial trajectory optimization phase, the CMAPPO model can be applied directly to BAU trajectories once trained. While cost-optimal trajectories from ROOST were used in this study for consistency with Chapter 5, the framework can readily accept alternative input sources, such as ADS-B trajectories. This flexibility makes the integrated approach more suitable for large-scale evaluations and operational implementation, with reduced computational cost.

Chapter 7

Conclusions and future work

This thesis has proposed advanced optimization frameworks to align climate-optimal flight trajectories with the operational requirements of the ATM system. To support the concluding remarks, all quantitative results reported in this chapter refer to the representative case study conducted on 20 December 2018, from 12:00 to 16:00 UTC, within structured airspace. The results associated with the sequential framework specifically correspond to the case with an environmental index (EI) of 0.01.

Based on the analyses and findings presented in previous chapters, the following final conclusions are drawn:

On micro-scale climate-optimized flight planning

- Given the strong spatiotemporal dependency of aviation-induced non-CO_2 climate effects, flight planning can be used as a viable operational strategy to mitigate their impact. This strategy is particularly effective for contrails, as the climate effect of contrails exhibits substantial spatiotemporal variability, enabling effective mitigation through slight trajectory adjustments. In contrast, mitigating other non-CO_2 effects, such as those induced by NO_x emissions, was shown to be less effective due to their smooth spatiotemporal climate sensitivity patterns. For the scenario examined, the micro-scale trajectory optimization presented in Chapter 5 resulted in a 20.40% reduction in net climate impact, of which 18.54% was attributable to contrail mitigation.

- Micro-scale flight planning, while effective in reducing climate impact, introduces operational challenges by creating traffic imbalances due to the consideration of climate-sensitive areas, thereby increasing air traffic complexity and potential conflicts within (i.e., for cooling impact) or around (i.e., for warming impact) these areas. Additionally, deviating from cost-optimal routes results in higher operational costs. This highlights inherent trade-offs between environmental performance, operational cost, and ATM manageability. In the representative case study, the 20.40% reduction in climate impact was accompanied by a 0.63% increase in operational cost and a 10.27% increase in traffic complexity.

On network-scale climate-optimized flight planning

- A sequential framework, where individual trajectories are first optimized for cost or climate objectives, followed by a second optimization layer to ensure operational manageability, proved promising for network-scale climate-optimal flight planning. This framework enables trade-off analysis by generating a spectrum of solutions that reflect varying degrees of climate optimality, air traffic complexity, and cost efficiency. Applied to the considered traffic scenario, the framework presented in Chapter 5 achieved a 14.37% reduction in climate impact. Notably, this reduction in climate impact was predominantly attributable to contrail-related effects, approximately one-third due to a decrease in warming contrails and two-thirds due to an increase in cooling contrails. Compared to micro-scale trajectory optimization, a reduction of approximately one-quarter in climate benefit was made to ensure air traffic manageability, resulting in a 15.88% reduction in traffic complexity relative to cost-optimal trajectories. These results were achieved with a 0.77% increase in operational cost.

- The integrated framework that jointly optimizes climate impact and traffic manageability in a single step offers a scalable and computationally efficient alternative. For the same case study, this approach achieved a 9.14% reduction in net climate impact and a 5.27% decrease in traffic complexity, with a 0.64% increase in operational cost. The lower climate benefits of the integrated approach, compared to the sequential framework, are attributable to its optimization scope, focusing on avoiding strong warming contrails without explicitly considering other non-CO_2 effects or the contribution of cooling contrails. Specifically, 98.30% of the reduction in contrail climate impact was achieved through the mitigation of warming effects, with the remaining 1.69% due to an unintended increase in cooling effects. Given its computational efficiency and integrated structure, requiring approximately 35 minutes to process ≈ 6000 flights on a standard workstation, this approach is particularly well-suited for large-scale, policy-relevant applications, enabling near real-time planning of operationally manageable climate-optimized traffic.

- The selection between sequential and integrated frameworks should be guided by the intended application. The sequential approach is more suitable for high-fidelity scenarios, such as daily flight planning, as it offers greater flexibility, albeit at a higher computational cost. In contrast, integrated frameworks are better suited for applications requiring large-scale analysis of flight planning, such as the development of operational and environmental indicators to support stakeholder decision-making and policy guidance. Therefore, although integrated frameworks may offer slightly reduced optimality compared to sequential approaches, they provide a more scalable and efficient solution for high-level assessments.

On resolution strategies

- The findings indicate that the effectiveness of resolution strategies diminishes as trajectories become more climate-efficient, revealing a natural constraint in achieving feasible solutions under stringent climate objectives. This confirms a trade-off between climate impact mitigation, operational cost, and air traffic manageability, where improvements in one objective often necessitate compromises in the others. One of the main aims of this thesis was to identify a balanced solution that appropriately reconciles these competing objectives.

- Both centralized and distributed resolution frameworks offer viable approaches to address the operational challenges posed by climate-optimal trajectories, yet their suitability depends on the context and scale of application. Centralized approaches (i.e., simulated annealing approach in Chapter 3) are conceptually simpler, making them easier to prototype and evaluate in controlled or small-scale settings. However, they lack scalability and depend on centralized access to system-wide data, which may not always be available in real-world operations. In contrast, distributed MARL approaches, which rely on local information, have demonstrated greater effectiveness in large-scale, realistic scenarios due to their scalability and robustness.

On multi-agent reinforcement learning

- Multi-agent reinforcement learning provides a scalable and effective approach for trajectory planning in complex, high-dimensional air traffic environments. It enables decentralized decision-making while preserving global coordination and is particularly well-suited to partially observable and dynamic environments such as the ATM system. Parameter sharing within the MARL framework proved to be an efficient strategy for achieving scalability, allowing the system to manage varying numbers of aircraft across diverse traffic scenarios.

- Constrained multi-agent reinforcement learning proves effective in handling multiple, often competing objectives, such as minimizing climate impact while addressing traffic complexity, within a unified learning framework. Its ability to incorporate diverse objectives and constraints, combined with its scalability, fast execution time, and adaptability to new traffic scenarios, makes it a practical solution for aviation applications.

- A challenge identified in this thesis is the strong dependence of MARL performance on the formulation of the underlying optimization problem. The effectiveness of both standard and constrained MARL is highly sensitive to how states, actions, and reward

functions are mathematically defined. Identifying an appropriate problem representation is often non-trivial and may not yield an optimal abstraction of the operational environment, potentially limiting both learning efficiency and generalizability.

Future work

Some future lines of investigation can be pursued to expand and complement the work presented in this thesis.

- **Incorporating meteorological uncertainty into flight planning:** The climate impact of non-CO_2 emissions is highly dependent on meteorological variables (e.g., temperature, humidity, and radiation). As a result, the accuracy of climate impact estimates is closely tied to the quality of weather forecasts. However, weather forecasts are inevitably uncertain, which can highly affect the reliability of the identified climate-sensitive areas. Beyond climate impact, aircraft performance variables (e.g., fuel consumption and flight time) are also influenced by atmospheric conditions, particularly wind and temperature. If such weather forecast uncertainties are not accounted for during flight planning, there is a risk that the optimized trajectories may fail to deliver the intended climate benefits and may not satisfy ATM manageability performance metrics, such as air traffic complexity or the likelihood of conflicts. While Chapter 3 explicitly accounts for meteorological uncertainty using ensemble weather forecasts, the resolution strategies presented in the subsequent chapters were implemented deterministically. Future work should aim to extend flight planning frameworks by explicitly incorporating meteorological uncertainty in order to generate robust flight plans at the network scale with greater confidence in both climate benefits and traffic manageability.

- **Using multiple climate impact estimation models:** In this study, we relied on a single model to estimate climate effects. However, previous research has highlighted potentially large discrepancies between different climate impact estimation models [62]. This suggests that the climate optimality achieved using one model may not hold when evaluated with another, an aspect that requires careful consideration. Such discrepancies between models are generally referred to as modeling uncertainty. Currently, only two models, CoCiP and aCCFs, are capable of providing spatiotemporal estimates of non-CO_2 climate effects suitable for flight planning. To enhance confidence in the achieved climate benefits, one potential direction is to incorporate multiple climate impact estimation models (e.g., aCCFs, CoCiP, and other emerging models) into flight planning. This would support the identification of climate-optimal routes that yield consistent benefits across different models. Such a multi-model approach is particularly relevant from a macro-level perspective, as it allows for evaluating how the

inclusion of additional climate-sensitive regions identified by different models may affect traffic manageability.

- **Incorporating forecast data and evaluating associated uncertainty:** The weather data used in this study are based on a reanalysis product (i.e., ERA5). However, since reanalysis data are only available after the actual realization of atmospheric conditions, they are not suitable for the flight planning stage. Instead, forecast data should be employed for such applications. In this thesis, reanalysis was used due to open-source availability and ease of accessibility. However, the proposed framework is compatible with other types of meteorological datasets, including forecast data products (e.g., ensemble prediction system weather forecast). A relevant direction for future work is to utilize forecast data for flight planning and subsequently evaluate the performance of the planned trajectories under reanalysis data. This approach would allow for a systematic assessment of forecast-induced uncertainty, as well as the sensitivity and robustness of the optimized trajectories to such effects.

- **Investigating alternative metrics for evaluating ATM system performance:** In this thesis, the number of potential conflicts and complexity scores were used as primary performance indicators to evaluate the operational manageability of climate-optimized flight planning. However, as outlined in Chapter 2, other ATM performance metrics, such as capacity-demand balance, controller workload, and alternative complexity indicators, offer complementary perspectives on the system behavior. Future research should explore these metrics to enable a more comprehensive assessment of ATM performance under climate-aware trajectory planning. A promising direction would be to investigate the interdependencies among performance indicators, for example, by optimizing with respect to one and analyzing its effects on the others to identify which metrics are most effective and operationally relevant. Alternatively, multiple indicators could be incorporated jointly in the flight planning framework to support a more integrated and robust evaluation of ATM system performance.

- **Incorporating additional operational constraints into the constrained MARL framework:** In Chapter 6, a constrained MARL framework was proposed to jointly optimize climate impact and air traffic manageability. This framework assumed that the original trajectories were cost-optimal and that slight modifications would not significantly affect operational efficiency. A potential extension would be to incorporate operational cost as an explicit constraint, by limiting cost increases using a mechanism similar to the one employed for climate hotspot avoidance. Additional constraints could also be introduced to reflect other ATM considerations, such as sector capacity limits. For instance, a capacity threshold could be enforced to ensure that rerouting around climate-sensitive areas does not lead to excessive demand in already saturated regions. Such

extensions would enable a more comprehensive and operationally realistic application of the constrained MARL framework.

- **Extending the analysis to multiple aircraft types:** This thesis focused on a single aircraft type (i.e., A320-214), and the developed methodologies were built on that assumption of shared policy parameters. However, real-world operations involve a heterogeneous fleet composed of various aircraft types. Therefore, the proposed frameworks should be extended to accommodate multiple aircraft types to better reflect operational realities. A potential approach to maintaining scalability while considering different aircraft types is to cluster aircraft into categories based on their performance models and assign a shared policy to each group. This ensures scalability, as each group uses a single policy, and the number of categories remains limited while capturing the distinct operational constraints and capabilities of different aircraft types.

- **Exploring mean field reinforcement learning for large-scale trajectory optimization:** Mean-field reinforcement learning (MFRL) has been successfully applied in various large-scale multi-agent systems and offers certain advantages over conventional multi-agent reinforcement learning approaches. By approximating the interactions among agents through a representative population distribution, MFRL can reduce computational complexity and improve scalability. These properties make it a promising candidate for macro-scale trajectory optimization problems, where thousands of aircraft must coordinate under both climate and operational constraints. Future work could investigate the potential of MFRL to enable efficient network-scale climate-optimal flight planning.

- **Large-scale analysis of climate-aware trajectory planning:** The results presented in this thesis are based on specific case studies; therefore, the conclusions drawn may not be fully generalizable to other traffic scenarios or weather patterns. To comprehensively assess the mitigation potential of climate-aware flight planning, future research should extend the analysis to large-scale scenarios, for example, through full-year simulations under diverse traffic and meteorological situations. The framework proposed in Chapter 6 provides a foundation for conducting such large-scale analyses. This would facilitate the generation of robust, system-wide insights into the effectiveness of climate-optimized flight planning for climate impact mitigation.

- **Extending the analysis to mixed-fleet operations:** This thesis focused on conventional aircraft powered by kerosene fuel. However, future airspace will comprise a mixed fleet, including emerging aircraft types and fuels. Mixed-fleet operations, along with novel operational concepts, require more targeted and strategic allocation of future

resources and technologies, an approach that is referred to as smart mixed-fleet operations. Emerging fuels and aircraft technologies typically involve significant costs (including production and infrastructure), and thus their integration into fleets will likely occur gradually. Therefore, maximizing their climate benefits requires strategically allocating these limited resources to flights where the greatest mitigation potential can be achieved. Crucially, this targeted resource allocation approach must be integrated with macro-scale flight planning. It cannot effectively function as two separate optimization processes, one that allocates technologies to different flights, followed by another performing macro-scale operational optimization. Indeed, allocation decisions not only influence climate impact but can also significantly affect traffic manageability. For example, instead of rerouting a kerosene-powered aircraft and potentially causing congestion, it may be preferable to utilize a hydrogen-powered aircraft that can maintain a direct, shortest-path trajectory. Therefore, a coupled optimization framework is required to simultaneously allocate the available technologies and fuels and to plan operationally feasible trajectories, ultimately aiming for the most climate-optimal performance. This remains an open problem for future research.

- **Using constrained MARL for end-to-end network-scale climate-optimal flight planning without relying on predefined trajectories:** The constrained MARL framework presented in Chapter 6 relies on predefined flight plans, which are modified to meet climate and traffic manageability objectives. A potential direction for future work is to extend this approach to plan complete flight trajectories directly from origin-destination pairs, eliminating the need for initial flight plans. In this setting, the framework would optimize trajectories by minimizing operational costs while enforcing upper bounds on climate hotspot crossings and traffic complexity. This extension would enable a fully integrated, end-to-end flight planning solution capable of delivering operationally feasible, climate-aware trajectories from the outset.

- **Extending MARL frameworks to broader aviation applications:** The MARL-based frameworks presented in this thesis were designed for aircraft trajectory optimization at the planning stage. However, the proposed strategies hold potential for real-time air traffic applications, where they could serve as advisory tools to support air traffic controllers. Another promising application is flow and capacity management in en-route airspace, where MARL could be used to dynamically balance demand across sectors and mitigate airspace complexity under evolving traffic and weather conditions.

  Beyond en-route operations, the underlying MARL frameworks, especially the constrained MARL approach, are transferable to other trajectory planning domains, such as unmanned aerial vehicles, where different objectives and constraints must be handled in autonomous and coordinated environments. Additionally, future research could explore applying these methods to airport operations, including taxi routing

and gate assignment, where coordinated decision-making can reduce delays, fuel consumption, and emissions, thereby minimizing environmental impact.

Bibliography

[1] U. FCCC, "Framework convention on climate change. adoption of the paris agreement," 2015.

[2] International Air Transport Association (IATA), "Global outlook for air transport: Deep change," June 2024. Accessed: December 4, 2024.

[3] D. S. Lee, D. Fahey, A. Skowron, M. Allen, U. Burkhardt, Q. Chen, S. Doherty, S. Freeman, P. Forster, J. Fuglestvedt, *et al.*, "The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018," *Atmospheric Environment*, vol. 244, p. 117834, 2021.

[4] G. P. Brasseur, M. Gupta, B. E. Anderson, S. Balasubramanian, S. Barrett, D. Duda, G. Fleming, P. M. Forster, J. Fuglestvedt, A. Gettelman, *et al.*, "Impact of aviation on climate: FAA's aviation climate change research initiative (ACCRI) phase II," *Bulletin of the American Meteorological Society*, vol. 97, no. 4, pp. 561–583, 2016.

[5] D. S. Lee, D. W. Fahey, P. M. Forster, P. J. Newton, R. C. Wit, L. L. Lim, B. Owen, and R. Sausen, "Aviation and global climate change in the 21st century," *Atmospheric Environment*, vol. 43, no. 22-23, pp. 3520–3537, 2009.

[6] International Civil Aviation Organization, "Resolution A41-21: Consolidated statement of continuing ICAO policies and practices related to environmental protection—Climate change," tech. rep., International Civil Aviation Organization, Montréal, Canada, Oct. 2022. Adopted at the 41st ICAO Assembly, 7 October 2022.

[7] J. Larsson, A. Elofsson, T. Sterner, and J. Åkerman, "International and national climate policies for aviation: a review," *Climate Policy*, vol. 19, no. 6, pp. 787–799, 2019.

[8] V. Grewe, A. G. Rao, T. Grönstedt, C. Xisto, F. Linke, J. Melkert, J. Middel, B. Ohlenforst, S. Blakey, S. Christie, *et al.*, "Evaluating the climate impact of aviation emission scenarios towards the paris agreement including COVID-19 effects," *Nature Communications*, vol. 12, no. 1, pp. 1–10, 2021.

[9] L. Dray, A. W. Schäfer, C. Grobler, C. Falter, F. Allroggen, M. E. Stettler, and S. R. Barrett, "Cost and emissions pathways towards net-zero climate impacts in aviation," *Nature Climate Change*, vol. 12, no. 10, pp. 956–962, 2022.

[10] A. R. Gnadt, R. L. Speth, J. S. Sabnis, and S. R. Barrett, "Technical and environmental assessment of all-electric 180-passenger commercial aircraft," *Progress in Aerospace Sciences*, vol. 105, pp. 1–30, 2019.

[11] J. Ribeiro, F. Afonso, I. Ribeiro, B. Ferreira, H. Policarpo, P. Peças, and F. Lau, "Environmental assessment of hybrid-electric propulsion in conceptual aircraft design," *Journal of Cleaner Production*, vol. 247, p. 119477, 2020.

[12] F. Stöckl, W.-P. Schill, and A. Zerrahn, "Optimal supply chains and power sector benefits of green hydrogen," *Scientific reports*, vol. 11, no. 1, p. 14191, 2021.

[13] E. J. Adler and J. R. Martins, "Hydrogen-powered aircraft: Fundamental concepts, key technologies, and environmental impacts," *Progress in Aerospace Sciences*, vol. 141, p. 100922, 2023.

[14] M. D. Staples, R. Malina, P. Suresh, J. I. Hileman, and S. R. Barrett, "Aviation CO_2 emissions reductions from the use of alternative jet fuels," *Energy Policy*, vol. 114, pp. 342–354, 2018.

[15] R. Teoh, U. Schumann, C. Voigt, T. Schripp, M. Shapiro, Z. Engberg, J. Molloy, G. Koudis, and M. E. Stettler, "Targeted use of sustainable aviation fuel to maximize climate benefits," *Environmental Science & Technology*, vol. 56, no. 23, pp. 17246–17255, 2022.

[16] C. Bergero, G. Gosnell, D. Gielen, S. Kang, M. Bazilian, and S. J. Davis, "Pathways to net-zero emissions from aviation," *Nature Sustainability*, vol. 6, no. 4, pp. 404–414, 2023.

[17] R. Sacchi, V. Becattini, P. Gabrielli, B. Cox, A. Dirnaichner, C. Bauer, and M. Mazzotti, "How to make climate-neutral aviation fly," *Nature Communications*, vol. 14, no. 1, p. 3989, 2023.

[18] A. Simorgh, M. Soler, González-Arribas, *et al.*, "A comprehensive survey on climate optimal aircraft trajectory planning," *Aerospace*, vol. 9, no. 3, p. 146, 2022.

[19] J. van Manen and V. Grewe, "Algorithmic climate change functions for the use in eco-efficient flight planning," *Transportation Research Part D: Transport and Environment*, vol. 67, pp. 388–405, 2019.

[20] A. Simorgh and M. Soler, "Climate-optimized flight planning can effectively reduce the environmental footprint of aviation in europe at low operational costs," *Communications Earth & Environment*, vol. 6, no. 1, p. 66, 2025.

[21] A. Martin Frias, M. Shapiro, Z. Engberg, R. Zopp, M. Soler, and M. E. J. Stettler, "Feasibility of contrail avoidance in a commercial flight planning system: an operational analysis," *Environmental Research: Infrastructure and Sustainability*, 2024.

[22] A. Simorgh, M. Soler, D. González-Arribas, F. Linke, B. Lührs, M. M. Meuser, S. Dietmüller, S. Matthes, H. Yamashita, F. Yin, *et al.*, "Robust 4D climate-optimal flight planning in structured airspace using parallelized simulation on gpus: ROOST V1. 0," *Geoscientific model development*, vol. 16, no. 13, pp. 3723–3748, 2023.

[23] EUROCONTROL, "European aviation overview: Traffic trends and performance review," 2025. Accessed: March 17, 2025.

[24] Z. Lea Zengerling and A. Lau, "Reducing the climate impact of flight trajectories considering network effects," *SESAR Innovation Days*, 2024.

[25] E. Roosenbrand, J. Sun, and J. Hoekstra, "Contrail minimization through altitude diversions: A feasibility study leveraging global data," *Transportation Research Interdisciplinary Perspectives*, vol. 22, p. 100953, 2023.

[26] C. Demouge, M. Mongeau, N. Couellan, and D. Delahaye, "Climate-aware air traffic flow management optimization via column generation," *Transportation Research Part C: Emerging Technologies*, vol. 166, p. 104792, 2024.

[27] D. González-Arribas, F. Baneshi, E. Andrés, M. Soler, A. Jardines, and J. García-Heras, "Fast 4D flight planning under uncertainty through parallel stochastic path simulation," *Transportation Research Part C: Emerging Technologies*, vol. 148, p. 104018, 2023.

[28] F. Baneshi, M. Cerezo-Magaña, and M. Soler, "Integrating Non-CO2 climate impact considerations in air traffic management: Opportunities and challenges," *Transport Policy*, 2024.

[29] F. Baneshi, M. Soler, and A. Simorgh, "Conflict assessment and resolution of climate-optimal aircraft trajectories at network scale," *Transportation Research Part D: Transport and Environment*, vol. 115, p. 103592, 2023.

[30] F. Baneshi, M. Cerezo-Magaña, and M. Soler, "Network-level aircraft trajectory planning via multi-agent deep reinforcement learning: Balancing climate considerations and operational manageability," *Expert Systems with Applications*, p. 126604, 2025.

[31] F. Baneshi, M. Soler, and M. Cerezo-Magaña, "Climate-conscious aircraft trajectory planning considering air traffic manageability using deep multi-agent reinforcement learning," *Knowledge-Based Systems (Under revision)*, 2025.

[32] F. Baneshi, M. C. Magana, M. Soler, T. Ni, and M. Kamgarpour, "Aircraft trajectory planning for climate hotspot avoidance considering air traffic complexity: A constrained multi-agent reinforcement learning approach," *SESAR Innovation Days.*, 2024.

[33] C. Celis, V. Sethi, D. Zammit-Mangion, R. Singh, and P. Pilidis, "Theoretical optimal trajectories for reducing the environmental impact of commercial aircraft operations," *Journal of Aerospace Technology and Management*, vol. 6, no. 1, pp. 29–42, 2014.

[34] B. Sridhar, H. K. Ng, and N. Y. Chen, "Aircraft trajectory optimization and contrails avoidance in the presence of winds," *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 5, pp. 1577–1584, 2011.

[35] H. K. Ng, B. Sridhar, S. Grabbe, and N. Chen, "Cross-polar aircraft trajectory optimization and the potential climate impact," in *2011 IEEE/AIAA 30th Digital Avionics Systems Conference*, pp. 3D4–1, IEEE, 2011.

[36] Y. Lim, A. Gardi, and R. Sabatini, "Optimal aircraft trajectories to minimize the radiative impact of contrails and CO2," *Energy Procedia*, vol. 110, pp. 446–452, 2017.

[37] M. Soler, B. Zou, and M. Hansen, "Flight trajectory design in the presence of contrails: Application of a multiphase mixed-integer optimal control approach," *Transportation Research Part C: Emerging Technologies*, vol. 48, pp. 172–194, 2014.

[38] J. Rosenow, S. Förster, M. Lindner, and H. Fricke, "Multicriteria-optimized trajectories impacting today's air traffic density, efficiency, and environmental compatibility," *Journal of Air Transportation*, vol. 27, no. 1, pp. 8–15, 2019.

[39] F. Jelinek, "The advanced emission model (AEM3)-validation report," *Ratio*, vol. 306, no. 193, pp. 1–13, 2004.

[40] D. DuBois and G. C. Paynter, ""Fuel Flow Method 2" for estimating aircraft emissions," *SAE Transactions*, pp. 1–14, 2006.

[41] U. Schumann, "On conditions for contrail formation from aircraft exhausts," *Meteorologische Zeitschrift*, pp. 4–23, 1996.

[42] E. Schmidt, "Die entstehung von eisnebel aus den auspuffgasen von flugmotoren," *Schriften der Deutschen Akademie der Luftfahrtforschung, Verlag R. Oldenbourg, München, Heft 44*, vol. 5, no. 44, pp. 1–15, 1941.

[43] H. Appleman, "The formation of exhaust condensation trails by jet aircraft," *Bulletin of the American Meteorological Society*, vol. 34, no. 1, pp. 14–20, 1953.

[44] B. Zou, G. S. Buxi, and M. Hansen, "Optimal 4-D aircraft trajectories in a contrail-sensitive environment," *Networks and Spatial Economics*, vol. 16, no. 1, pp. 415–446, 2016.

[45] M. Niklaß, V. Grewe, V. Gollnick, and K. Dahlmann, "Concept of climate-charged airspaces: a potential policy instrument for internalizing aviation's climate impact of non-CO_2 effects," *Climate Policy*, vol. 21, no. 8, pp. 1066–1085, 2021.

[46] S. Matthes, U. Schumann, V. Grewe, C. Frömming, K. Dahlmann, A. Koch, and H. Mannstein, "Climate optimized air transport," in *Atmospheric physics*, pp. 727–746, Springer, 2012.

[47] C. Frömming, V. Grewe, P. Jöckel, S. Brinkop, S. Dietmüller, H. Garny, M. Ponater, E. Tsati, and S. Matthes, "Climate cost functions as a basis for climate optimized flight trajectories," *Air Traffic Semin*, vol. 239, pp. 1–9, 2013.

[48] V. Grewe, C. Frömming, S. Matthes, S. Brinkop, M. Ponater, S. Dietmüller, P. Jöckel, H. Garny, E. Tsati, K. Dahlmann, *et al.*, "Aircraft routing with minimal climate impact: The REACT4C climate cost function modelling approach (V1. 0)," *Geoscientific Model Development*, vol. 7, no. 1, pp. 175–201, 2014.

[49] S. Matthes, V. Grewe, K. Dahlmann, C. Frömming, E. Irvine, L. Lim, F. Linke, B. Lührs, B. Owen, K. Shine, *et al.*, "A concept for multi-criteria environmental assessment of aircraft trajectories," *Aerospace*, vol. 4, no. 3, p. 42, 2017.

[50] F. Yin, V. Grewe, F. Castino, P. Rao, S. Matthes, K. Dahlmann, S. Dietmüller, C. Frömming, H. Yamashita, P. Peter, *et al.*, "Predicting the climate impact of aviation for en-route emissions: the algorithmic climate change function submodel ACCF 1.0 of EMAC 2.53," *Geoscientific Model Development*, vol. 16, no. 11, pp. 3313–3334, 2023.

[51] S. Dietmüller, S. Matthes, K. Dahlmann, H. Yamashita, A. Simorgh, M. Soler, F. Linke, B. Lührs, M. Mendiguchia Meuser, C. M. Weder, *et al.*, "A python library for computing individual and merged non-CO_2 algorithmic climate change functions: CLIMaCCF V1. 0," *Geoscientific model development*, vol. 16, no. 15, pp. 4405–4425, 2023.

[52] S. Matthes, S. Dietmüller, K. Dahlmann, C. Frömming, H. Yamashita, V. Grewe, F. Yin, and F. Castino, "Algorithmic climate change functions (aCCFs) V1.0a: Consolidation of the approach and note for usage," *submitted to GMDD*, 2023.

[53] U. Schumann, "A contrail cirrus prediction model," *Geoscientific Model Development*, vol. 5, no. 3, pp. 543–580, 2012.

[54] M. Shapiro, Z. Engberg, B. Zugic, R. Teoh, M. Stettler, U. Schumann, and I. McKay, "Forecasting contrail climate forcing for flight planning and air traffic management applications," in *The 5th International Conference on Transport, Atmosphere and Climate (TAC)*, 2022.

[55] J. Li, J.-H. Kim, B. Sridhar, and H. K. Ng, "Ames contrail simulation model: Modeling aviation induced contrails and the computation of contrail radiative forcing using air traffic data," tech. rep., 2023.

[56] T. M. Fritz, S. D. Eastham, R. L. Speth, and S. R. Barrett, "The role of plume-scale processes in long-term impacts of aircraft emissions," *Atmospheric Chemistry and Physics*, vol. 20, no. 9, pp. 5697–5727, 2020.

[57] H. Yamashita, F. Yin, V. Grewe, P. Jöckel, S. Matthes, B. Kern, K. Dahlmann, and C. Frömming, "Newly developed aircraft routing options for air traffic simulation in the chemistry–climate model EMAC 2.53: AirTraf 2.0," *Geoscientific Model Development*, vol. 13, no. 10, pp. 4869–4890, 2020.

[58] H. Yamashita, F. Yin, V. Grewe, P. Jöckel, S. Matthes, B. Kern, K. Dahlmann, and C. Frömming, "Analysis of aircraft routing strategies for north atlantic flights by using AirTraf 2.0," *Aerospace*, vol. 8, no. 2, p. 33, 2021.

[59] F. Yin, V. Grewe, F. Castino, P. Rao, S. Matthes, K. Dahlmann, S. Dietmüller, C. Frömming, H. Yamashita, P. Peter, *et al.*, "Predicting the climate impact of aviation for en-route emissions: the algorithmic climate change function submodel ACCF 1.0 of EMAC 2.53," *Geoscientific Model Development*, vol. 16, no. 11, pp. 3313–3334, 2023.

[60] A. Simorgh, M. Soler, S. Dietmüller, S. Matthes, H. Yamashita, F. Castino, and F. Yin, "Robust 4D climate-optimal aircraft trajectory planning under weather-induced uncertainties: Free-routing airspace," *Transportation Research Part D: Transport and Environment*, vol. 131, p. 104196, 2024.

[61] D. González Arribas and A. Simorgh, "Robust optimization of structured trajecoties (ROOST) python library," Dec. 2022.

[62] A. Simorgh and M. Soler, "Climate-optimized flight planning can effectively reduce the environmental footprint of aviation in europe at low operational costs," *Communications Earth & Environment*, vol. 6, no. 1, p. 66, 2025.

[63] R. Teoh, U. Schumann, E. Gryspeerdt, M. Shapiro, J. Molloy, G. Koudis, C. Voigt, and M. E. J. Stettler, "Aviation contrail climate effects in the north atlantic from 2016 to 2021," *Atmospheric Chemistry and Physics*, vol. 22, no. 16, pp. 10919–10935, 2022.

[64] B. Lührs, F. Linke, S. Matthes, V. Grewe, and F. Yin, "Climate impact mitigation potential of european air traffic in a weather situation with strong contrail formation," *Aerospace*, vol. 8, no. 2, p. 50, 2021.

[65] F. Castino, F. Yin, V. Grewe, H. Yamashita, S. Matthes, S. Dietmüller, S. Baumann, M. Soler, A. Simorgh, M. Mendiguchia Meuser, *et al.*, "Decision-making strategies implemented in SolFinder 1.0 to identify eco-efficient aircraft trajectories: application study in AirTraf 3.0," *Geoscientific Model Development*, vol. 17, no. 9, pp. 4031–4052, 2024.

[66] R. Sausen, S. M. Hofer, K. M. Gierens, L. Bugliaro Goggia, R. Ehrmanntraut, I. Sitova, K. Walczak, A. Burridge-Diesing, M. Bowman, and N. Miller, "Can we successfully avoid persistent contrails by small altitude adjustments of flights in the real world?," *Meteorologische Zeitschrift*, 2023.

[67] WMO, "Guidelines on ensemble prediction systems and forecasting," *World Meteorological Organization Weather Climate and Water*, vol. 1091, 2012.

[68] B. Hilburn, "Cognitive complexity in air traffic control: A literature review," *EEC note*, vol. 4, no. 04, pp. 1–80, 2004.

[69] T. S. Clair, "On-demand assessment of air traffic impact of blocking airspace," *The Aeronautical Journal*, vol. 122, no. 1258, pp. 1985–2009, 2018.

[70] N. M. Smith, C. Brasil, P. U. Lee, N. Buckley, C. Gabriel, C. P. Mohlenbrink, F. Omar, B. Parke, C. Speridakos, and H.-S. Yoo, "Integrated demand management: Coordinating strategic and tactical flow scheduling operations," in *16th AIAA Aviation Technology, Integration, and Operations Conference*, p. 4221, 2016.

[71] M. E. Miller and W. D. Hall, "Collaborative trajectory option program demonstration," in *2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC)*, pp. 1C1–1, IEEE, 2015.

[72] E. C. Fernández, J. M. Cordero, G. Vouros, N. Pelekis, T. Kravaris, H. Georgiou, G. Fuchs, N. Andrienko, G. Andrienko, E. Casado, *et al.*, "Dart: A machine-learning approach to trajectory prediction and demand-capacity balancing," *SESAR Innovation Days, Belgrade*, pp. 28–30, 2017.

[73] T. Kravaris, G. A. Vouros, C. Spatharis, K. Blekas, G. Chalkiadakis, and J. M. C. Garcia, "Learning policies for resolving demand-capacity imbalances during pre-tactical air traffic management," in *German Conference on Multiagent System Technologies*, pp. 238–255, Springer, 2017.

[74] I. Roychoudhury, L. Spirkovska, M. Oconnor, and C. Kulkarni, "Survey of methods to predict controller workload for real-time monitoring of airspace safety," tech. rep., 2018.

[75] V. Socha, L. Hanáková, V. Valenta, L. Socha, R. Ábela, S. Kušmírek, T. Pilmannová, and J. Tecl, "Workload assessment of air traffic controllers," *Transportation research procedia*, vol. 51, pp. 243–251, 2020.

[76] S. M. Casner and B. F. Gore, "Measuring and evaluating workload: A primer," *NASA Technical Memorandum*, vol. 216395, p. 2010, 2010.

[77] J. D. Welch, J. Y. Cho, N. K. Underhill, and R. A. DeLaura, "Sector workload model for benefits analysis and convective weather capacity prediction," in *Tenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2013), Chicago, IL*, 2013.

[78] M. Prandini, J. Lygeros, A. Nilim, and S. Sastry, "A probabilistic framework for aircraft conflict detection," in *Guidance, Navigation, and Control Conference and Exhibit*, p. 4144, 1999.

[79] M. Prandini and O. J. Watkins, "Probabilistic aircraft conflict detection," *HYBRIDGE, IST-2001*, vol. 32460, no. 2, 2005.

[80] U. Kjellén, "The safety measurement problem revisited," *Safety Science*, vol. 4, no. 47, pp. 486–489, 2009.

[81] W. Kelly III and M. Eby, "Advances in force field conflict resolution algorithms," in *AIAA guidance, navigation, and control conference and exhibit*, p. 4360, 2000.

[82] Y. Hong, Y. Kim, and K. Lee, "Conflict management in air traffic control using complexity map," *Journal of Aircraft*, vol. 52, no. 5, pp. 1524–1534, 2015.

[83] L. F. Vismari and J. B. C. Junior, "A safety assessment methodology applied to cns/atm-based air traffic control system," *Reliability Engineering & System Safety*, vol. 96, no. 7, pp. 727–738, 2011.

[84] M. Jardin, "Analytical conflict detection and resolution for air traffic management," *Journal of Guidance, Control, and Dynamics*, vol. 28, no. 4, pp. 682–690, 2005.

[85] H. Kharoufah, J. Murray, G. Baxter, and G. Wild, "A review of human factors causations in commercial air transport accidents and incidents: From to 2000–2016," *Progress in Aerospace Sciences*, vol. 99, pp. 1–13, 2018.

[86] A. Pasquini and S. Pozzi, "Evaluation of air traffic management procedures—safety assessment in an experimental environment," *Reliability Engineering & System Safety*, vol. 89, no. 1, pp. 105–117, 2005.

[87] T. Schouwenaars, J. How, and E. Feron, "Decentralized cooperative trajectory planning of multiple aircraft with hard safety guarantees," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, p. 5141, 2004.

[88] P. Brooker, "Air traffic safety: continued evolution or a new paradigm.," 2007.

[89] A. Bicchi and L. Pallottino, "On optimal cooperative conflict resolution for air traffic management systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 4, pp. 221–231, 2000.

[90] S. Loft, S. Bolland, M. S. Humphreys, and A. Neal, "A theory and model of conflict detection in air traffic control: Incorporating environmental constraints.," *Journal of Experimental Psychology: Applied*, vol. 15, no. 2, p. 106, 2009.

[91] F. Netjasov and M. Janic, "A review of research on risk and safety modelling in civil aviation," *Journal of Air Transport Management*, vol. 14, no. 4, pp. 213–220, 2008.

[92] P. Flener, J. Pearson, M. Ågren, C. Garcia-Avello, M. Celiktin, and S. Dissing, "Air-traffic complexity resolution in multi-sector planning," *Journal of Air Transport Management*, vol. 13, no. 6, pp. 323–328, 2007.

[93] M. Prandini, V. Putta, and J. Hu, "Air traffic complexity in future air traffic management systems," *Journal of Aerospace Operations*, vol. 1, no. 3, pp. 281–299, 2012.

[94] E. A. Group *et al.*, "Complexity metrics for ANSP benchmarking analysis," *EUROCONTROL, April*, 2006.

[95] R. H. Mogford, J. Guttman, S. Morrow, and P. Kopardekar, "The complexity construct in air traffic control: A review and synthesis of the literature.," 1995.

[96] S. Mondoloni and D. Liang, "Airspace fractal dimensions and applications," 12 2001.

[97] A. Perera and P. Kamalaruban, "Applications of reinforcement learning in energy systems," *Renewable and Sustainable Energy Reviews*, vol. 137, p. 110618, 2021.

[98] T. Pejovic, F. Netjasov, and D. Crnogorac, "Relationship between air traffic demand, safety and complexity in high-density airspace in europe," in *Risk Assessment in Air Traffic Management*, IntechOpen, 2020.

[99] A. J. Masalonis, M. B. Callaham, and C. R. Wanke, "Dynamic density and complexity metrics for realtime traffic flow management," in *Proceedings of the 5th USA/Europe Air Traffic Management R & D Seminar*, p. 139, Budapest, Hungary, 2003.

[100] P. H. Kopardekar, A. Schwartz, S. Magyarits, and J. Rhodes, "Airspace complexity measurement: An air traffic control simulation analysis," *International Journal of Industrial Engineering: Theory, Applications and Practice*, vol. 16, no. 1, pp. 61–70, 2009.

[101] P. Kopardekar and S. Magyarits, "Measurement and prediction of dynamic density," in *Proceedings of the 5th usa/europe air traffic management R & R seminar*, vol. 139, 2003.

[102] G. Chatterji and B. Sridhar, "Measures for air traffic controller workload prediction," in *1st AIAA, aircraft, technology Integration, and operations Forum*, p. 5242, 2001.

[103] D. Kudumija, B. Antulov-Fantulin, P. Andraši, and T. Rogošić, "The effect of the croatian free route airspace implementation on the air traffic complexity," *Transportation Research Procedia*, vol. 64, pp. 356–363, 2022.

[104] P. Kopardekar, "Dynamic density: A review of proposed variables," *FAA WJHTC internal document. overall conclusions and recommendations, Federal Aviation Administration*, 2000.

[105] M. Xiao, J. Zhang, K. Cai, and X. Cao, "Atcem: a synthetic model for evaluating air traffic complexity," *journal of Advanced Transportation*, vol. 50, no. 3, pp. 315–325, 2016.

[106] D. Delahaye, S. Puechmorel, J. Hansman, and J. Histon, "Air traffic complexity based on non linear dynamical systems," in *ATM 2003, 5th USA/Europe Air Traffic Management Research and Development Seminar*, pp. pp–xxxx, 2003.

[107] D. Delahaye and S. Puechmorel, "Air traffic complexity based on dynamical systems," in *49th IEEE Conference on Decision and Control (CDC)*, pp. 2069–2074, IEEE, 2010.

[108] D. Delahaye, P. Paimblanc, S. Puechmorel, J. Histon, and R. Hansman, "A new air traffic complexity metric based on dynamical system modelization," in *Proceedings. The 21st Digital Avionics Systems Conference*, vol. 1, pp. 4A2–4A2, IEEE, 2002.

[109] X. Jiang, X. Wen, M. Wu, M. Song, and C. Tu, "A complex network analysis approach for identifying air traffic congestion based on independent component analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 523, pp. 364–381, 2019.

[110] M. S. Eby and W. E. Kelly, "Free flight separation assurance using distributed algorithms," in *1999 IEEE Aerospace Conference. Proceedings (Cat. No. 99TH8403)*, vol. 2, pp. 429–441, IEEE, 1999.

[111] J. Krozel, T. Mueller, and G. Hunter, "Free flight conflict detection and resolution analysis," in *Guidance, Navigation, and Control Conference*, p. 3763, 1996.

[112] I. Annex, "Aeronautical telecommunications," *Volume III Part I Digital Data Communication Systems*, vol. 10, 10.

[113] B. Bakker, M. Griffin, A. Hendriks, and D. McMillan, "Eurocontrol specification for short term conflict alert," tech. rep., Technical report, EUROCONTROL, 2007.

[114] T. Prevot and P. U. Lee, "Trajectory-based complexity (TBX): A modified aircraft count to predict sector complexity during trajectory-based operations," in *2011 IEEE/AIAA 30th Digital Avionics Systems Conference*, pp. 3A3–1, IEEE, 2011.

[115] I. V. Laudeman, S. G. Shelden, R. Branstrom, and C. Brasil, "Dynamic density: An air traffic management metric," tech. rep., 1998.

[116] E. Hernández Romero, "Probabilistic aircraft conflict detection and resolution under the effects of weather uncertainty," 2020.

[117] M. Ribeiro, "Conflict resolution at high traffic densities with reinforcement learning," 2023.

[118] V. Courchelle, M. Soler, D. González-Arribas, and D. Delahaye, "A simulated annealing approach to 3D strategic aircraft deconfliction based on en-route speed changes under wind and temperature uncertainties," *Transportation research part C: emerging technologies*, vol. 103, pp. 194–210, 2019.

[119] M. Ribeiro, J. Ellerbroek, and J. Hoekstra, "Improvement of conflict detection and resolution at high densities through reinforcement learning," *Proceedings of the ICRAT*, 2020.

[120] M. Brittain and P. Wei, "Autonomous separation assurance in an high-density en route sector: A deep multi-agent reinforcement learning approach," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 3256–3262, IEEE, 2019.

[121] M. W. Brittain and P. Wei, "One to any: Distributed conflict resolution with deep multi-agent reinforcement learning and long short-term memory," in *AIAA Scitech 2021 Forum*, p. 1952, 2021.

[122] E. Hernández Romero, A. Valenzuela Romero, and D. Rivas Rivas, "Probabilistic aircraft conflict detection considering ensemble weather forecast," 2016.

[123] J. Mollinga and H. van Hoof, "An autonomous free airspace en-route controller using deep reinforcement learning techniques," *arXiv preprint arXiv:2007.01599*, 2020.

[124] M. Pelegrín and C. d'Ambrosio, "Aircraft deconfliction via mathematical programming: Review and insights," *Transportation science*, vol. 56, no. 1, pp. 118–140, 2022.

[125] N. P. Tran, D.-T. Pham, S. K. Goh, S. Alam, and V. Duong, "An intelligent interactive conflict solver incorporating air traffic controllers' preferences using reinforcement learning," in *2019 Integrated Communications, Navigation and Surveillance Conference (ICNS)*, pp. 1–8, IEEE, 2019.

[126] M. Gariel and E. Feron, "3D conflict avoidance under uncertainties," in *2009 IEEE/AIAA 28th Digital Avionics Systems Conference*, pp. 4–E, IEEE, 2009.

[127] M. A. Christodoulou and S. G. Kodaxakis, "Automatic commercial aircraft-collision avoidance in free flight: the three-dimensional problem," *IEEE transactions on intelligent transportation systems*, vol. 7, no. 2, pp. 242–249, 2006.

[128] A. E. Vela, S. Solak, J.-P. B. Clarke, W. E. Singhose, E. R. Barnes, and E. L. Johnson, "Near real-time fuel-optimal en route conflict resolution," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 4, pp. 826–837, 2010.

[129] W. P. Niedringhaus, "Stream option manager (SOM): Automated integration of aircraft separation, merging, stream management, and other air traffic control functions," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 25, no. 9, pp. 1269–1280, 1995.

[130] K. Bilimoria, "A geometric optimization approach to aircraft conflict resolution," in *18th Applied aerodynamics conference*, p. 4265, 2000.

[131] E. Frazzoli, Z.-H. Mao, J.-H. Oh, and E. Feron, "Resolution of conflicts involving many aircraft via semidefinite programming," *Journal of Guidance, Control, and Dynamics*, vol. 24, no. 1, pp. 79–86, 2001.

[132] J. Omer, "A space-discretized mixed-integer linear model for air-conflict resolution with speed and heading maneuvers," *Computers & Operations Research*, vol. 58, pp. 75–86, 2015.

[133] J. Cai and N. Zhang, "Mixed integer nonlinear programming for aircraft conflict avoidance by applying velocity and altitude changes," *Arabian Journal for Science and Engineering*, vol. 44, pp. 8893–8903, 2019.

[134] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 1999.

[135] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.

[136] P. E. Gill, W. Murray, and M. A. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization," *SIAM review*, vol. 47, no. 1, pp. 99–131, 2005.

[137] R. H. Byrd, J. Nocedal, and R. A. Waltz, "Knitro: An integrated package for nonlinear optimization," 2006. `https://www.artelys.com/solvers/knitro/`.

[138] S. Mondoloni and S. Conway, "An airborne conflict resolution approach using a genetic algorithm," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, no. AIAA Paper 2001-4054, 2001.

[139] H. Liu, F. Liu, X. Zhang, X. Guan, J. Chen, and P. Savinaud, "Aircraft conflict resolution method based on hybrid ant colony optimization and artificial potential field," *Science China. Information Sciences*, vol. 61, no. 12, p. 129103, 2018.

[140] E. Hernández-Romero, A. Valenzuela, and D. Rivas, "Probabilistic multi-aircraft conflict detection and resolution considering wind forecast uncertainty," *Aerospace Science and Technology*, vol. 105, p. 105973, Oct. 2020.

[141] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[142] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[143] S. Ravichandiran, *Deep Reinforcement Learning with Python: Master classic RL, deep RL, distributional RL, inverse RL, and more with OpenAI Gym and TensorFlow*. Packt Publishing Ltd, 2020.

[144] Z. Wang, W. Pan, H. Li, X. Wang, and Q. Zuo, "Review of deep reinforcement learning approaches for conflict resolution in air traffic control," *Aerospace*, vol. 9, no. 6, p. 294, 2022.

[145] P. Razzaghi, A. Tabrizian, W. Guo, S. Chen, A. Taye, E. Thompson, A. Bregeon, A. Baheri, and P. Wei, "A survey on reinforcement learning in aviation applications," *Engineering Applications of Artificial Intelligence*, vol. 136, p. 108911, 2024.

[146] D.-T. Pham, P. N. Tran, S. Alam, V. Duong, and D. Delahaye, "Deep reinforcement learning based path stretch vector resolution in dense traffic with uncertainties," *Transportation research part C: emerging technologies*, vol. 135, p. 103463, 2022.

[147] S. Li, M. Egorov, and M. Kochenderfer, "Optimizing collision avoidance in dense airspace using deep reinforcement learning," *arXiv preprint arXiv:1912.10146*, 2019.

[148] M. Ribeiro, J. Ellerbroek, and J. Hoekstra, "Distributed conflict resolution at high traffic densities with reinforcement learning," *Aerospace*, vol. 9, no. 9, p. 472, 2022.

[149] H. Wen, H. Li, Z. Wang, X. Hou, and K. He, "Application of DDPG-based collision avoidance algorithm in air traffic control," in *2019 12th International Symposium on Computational Intelligence and Design (ISCID)*, vol. 1, pp. 130–133, IEEE, 2019.

[150] D.-T. Pham, N. P. Tran, S. K. Goh, S. Alam, and V. Duong, "Reinforcement learning for two-aircraft conflict resolution in the presence of uncertainty," in *2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*, pp. 1–6, IEEE, 2019.

[151] S. Dong, X. Weiping, and K. Zhang, "Study on the resolution of multi-aircraft flight conflicts based on an IDQN," *Chinese Journal of Aeronautics*, vol. 35, no. 2, pp. 195–213, 2022.

[152] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, "Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications," *IEEE transactions on cybernetics*, vol. 50, no. 9, pp. 3826–3839, 2020.

[153] M. Brittain, X. Yang, and P. Wei, "A deep multi-agent reinforcement learning approach to autonomous separation assurance," *arXiv preprint arXiv:2003.08353*, 2020.

[154] R. Dalmau and E. Allard, "Air traffic control using message passing neural networks and multi-agent reinforcement learning," *Proceedings of the 10th SESAR Innovation Days, Virtual Event*, pp. 7–10, 2020.

[155] P. Zhao and Y. Liu, "Physics informed deep reinforcement learning for aircraft conflict resolution," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8288–8301, 2021.

[156] Y. Chen, M. Hu, L. Yang, Y. Xu, and H. Xie, "General multi-agent reinforcement learning integrating adaptive manoeuvre strategy for real-time multi-aircraft conflict resolution," *Transportation Research Part C: Emerging Technologies*, vol. 151, p. 104125, 2023.

[157] G. Papadopoulos, A. Bastas, G. A. Vouros, I. Crook, N. Andrienko, G. Andrienko, and J. M. Cordero, "Deep reinforcement learning in service of air traffic controllers to resolve tactical conflicts," *Expert Systems with Applications*, vol. 236, p. 121234, 2024.

[158] S. Ghosh, S. Laguna, S. H. Lim, L. Wynter, and H. Poonawala, "A deep ensemble multi-agent reinforcement learning approach for air traffic control," *arXiv preprint arXiv:2004.01387*, 2020.

[159] W. Guo, M. Brittain, and P. Wei, "Safety enhancement for deep reinforcement learning in autonomous separation assurance," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 348–354, IEEE, 2021.

[160] T. Kravaris, K. Lentzos, G. Santipantakis, G. A. Vouros, G. Andrienko, N. Andrienko, I. Crook, J. M. C. Garcia, and E. I. Martinez, "Explaining deep reinforcement learning decisions in complex multiagent settings: towards enabling automation in air traffic flow management," *Applied Intelligence*, vol. 53, no. 4, pp. 4063–4098, 2023.

[161] C. Spatharis, A. Bastas, T. Kravaris, K. Blekas, G. A. Vouros, and J. M. Cordero, "Hierarchical multiagent reinforcement learning schemes for air traffic management," *Neural Computing and Applications*, pp. 1–13, 2023.

[162] Z. Wang, H. Li, J. Wang, and F. Shen, "Deep reinforcement learning based conflict detection and resolution in air traffic control," *IET Intelligent Transport Systems*, vol. 13, no. 6, pp. 1041–1047, 2019.

[163] M. Hermans, "Towards explainable automation for air traffic control using deep q-learning from demonstrations and reward decomposition," *Aerosp Eng*, 2021.

[164] F. Baneshi, M. Soler, and A. Simorgh, "Conflict assessment and resolution of climate-optimal aircraft trajectories at network scale," *Transportation Research Part D: Transport and Environment*, vol. 115, p. 103592, 2023.

[165] D. González Arribas, *Robust aircraft trajectory optimization under meteorological uncertainty*. PhD thesis, 2019.

[166] E. Gallo, F. Navarro, A. Nuic, and M. Iagaru, "Advanced aircraft performance modeling for ATM: BADA 4.0 results," in *2006 IEEE/AIAA 25TH Digital Avionics Systems Conference*, pp. 1–12, IEEE, Oct. 2006.

[167] S. Matthes, S. Dietmüller, K. Dahlmann, C. Frömming, G. V. Yamashita, Hiroshi, F. Yin, and F. Castino, "Algorithmic climate change functions (aCCFs) V1.0a: Consolidation of the approach and note for usage," *submitted to GMDD*, 2023.

[168] S. Matthes, B. Lührs, K. Dahlmann, V. Grewe, F. Linke, F. Yin, E. Klingaman, and K. P. Shine, "Climate-optimized trajectories and robust mitigation potential: Flying ATM4E," *Aerospace*, vol. 7, no. 11, p. 156, 2020.

[169] D. González Arribas, E. Andrés-Enderiz, M. Soler, A. Jardines, and J. García-Heras, "Probabilistic 4D flight planning in structured airspaces through parallelized simulation on gpus," *9th International Conference for Research in Air Transportation (ICRAT)*.

[170] P. Bauer, A. Thorpe, and G. Brunet, "The quiet revolution of numerical weather prediction," *Nature*, vol. 525, pp. 47–55, Sept. 2015.

[171] A. Simorgh, M. Soler, D. González-Arribas, S. Matthes, V. Grewe, S. Dietmüller, S. Baumann, H. Yamashita, F. Yin, F. Castino, *et al.*, "A comprehensive survey on climate optimal aircraft trajectory planning," *Aerospace*, vol. 9, no. 3, p. 146, 2022.

[172] J. T. Betts, *Practical methods for optimal control and estimation using nonlinear programming*. Advances in Design and Control, SIAM, 2ed ed., 2010.

[173] L. T. Biegler and V. M. Zavala, "Large-scale nonlinear programming using ipopt: An integrating framework for enterprise-wide dynamic optimization," *Computers & Chemical Engineering*, vol. 33, no. 3, pp. 575–582, 2009.

[174] N. Andrei, "A SQP algorithm for large-scale constrained optimization: SNOPT," in *Continuous nonlinear optimization for engineering applications in GAMS technology*, pp. 317–330, Springer, 2017.

[175] M. Jardin, "Grid-based strategic air traffic conflict detection," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, p. 5826, 2005.

[176] E. J. Hastings, J. Mesit, and R. K. Guha, "Optimization of large-scale, real-time simulations by spatial hashing," in *Proc. 2005 Summer Computer Simulation Conference*, vol. 37, pp. 9–17, 2005.

[177] K. A. Dowsland and J. Thompson, "Simulated annealing," *Handbook of natural computing*, pp. 1623–1655, 2012.

[178] A. Oroojlooy and D. Hajinezhad, "A review of cooperative multi-agent deep reinforcement learning," *Applied Intelligence*, vol. 53, no. 11, pp. 13677–13722, 2023.

[179] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[180] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *International conference on machine learning*, pp. 1587–1596, PMLR, 2018.

[181] S. Fujimoto, D. Meger, and D. Precup, "Off-policy deep reinforcement learning without exploration," in *International conference on machine learning*, pp. 2052–2062, PMLR, 2019.

[182] L. Canese, G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, D. Giardino, M. Re, and S. Spanò, "Multi-agent reinforcement learning: A review of challenges and applications," *Applied Sciences*, vol. 11, no. 11, p. 4948, 2021.

[183] A. Simorgh, M. Soler, D. González-Arribas, F. Linke, B. Lührs, M. M. Meuser, S. Dietmüller, S. Matthes, H. Yamashita, F. Yin, *et al.*, "Robust 4D climate-optimal flight planning in structured airspace using parallelized simulation on GPUs: ROOST V1. 0," *Geoscientific Model Development*, vol. 16, no. 13, pp. 3723–3748, 2023.

[184] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *International conference on machine learning*, pp. 1889–1897, PMLR, 2015.

[185] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[186] C. Yu, A. Velu, E. Vinitsky, J. Gao, Y. Wang, A. Bayen, and Y. Wu, "The surprising effectiveness of PPO in cooperative multi-agent games," *Advances in Neural Information Processing Systems*, vol. 35, pp. 24611–24624, 2022.

[187] J. Foerster, I. A. Assael, N. De Freitas, and S. Whiteson, "Learning to communicate with deep multi-agent reinforcement learning," *Advances in neural information processing systems*, vol. 29, 2016.

[188] J. K. Gupta, M. Egorov, and M. Kochenderfer, "Cooperative multi-agent control using deep reinforcement learning," in *Autonomous Agents and Multiagent Systems: AAMAS 2017 Workshops, Best Papers, São Paulo, Brazil, May 8-12, 2017, Revised Selected Papers 16*, pp. 66–83, Springer, 2017.

[189] R. Lowe, Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," *Advances in neural information processing systems*, vol. 30, 2017.

[190] S. Dietmüller, S. Matthes, K. Dahlmann, H. Yamashita, A. Simorgh, M. Soler, F. Linke, B. Lührs, M. M. Meuser, C. Weder, *et al.*, "A python library for computing individual and merged non-CO2 algorithmic climate change functions: CLIMaCCF V1. 0," *Geoscientific Model Development*, vol. 16, no. 15, pp. 4405–4425, 2023.

[191] K. Zhang, Z. Yang, and T. Başar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms," *Handbook of reinforcement learning and control*, pp. 321–384, 2021.

[192] F. Christianos, G. Papoudakis, M. A. Rahman, and S. V. Albrecht, "Scaling multi-agent reinforcement learning with selective parameter sharing," in *International Conference on Machine Learning*, pp. 1989–1998, PMLR, 2021.

[193] A. Ray, J. Achiam, and D. Amodei, "Benchmarking safe exploration in deep reinforcement learning," *arXiv preprint arXiv:1910.01708*, vol. 7, no. 1, p. 2, 2019.