

Exploratory Research Report (Final)

Deliverable ID:	D5.8
Project acronym:	E-CONTRAIL
Grant:	101114795
Call:	HORIZON-SESAR-2022-DES-ER-01
Topic:	HORIZON-SESAR-2022-DES-ER-01-WA1-6
Consortium coordinator:	Universidad Carlos III de Madrid
Edition date:	11 Junio 2025
Edition:	01.00
Status:	Draft
Classification:	PU

Abstract

The present deliverable constitutes the Experimental Research Report (ERR) - final version - for Project E-CONTRAIL. It presents the findings and outcomes of the experiments conducted within the project, tracking the timely and efficient execution as outlined in our initial plan. The document provides the context in which the research was carried out, including a summary of the experimental plan (validation objectives, exercises, activities, the expected performance contributions, primarily in the Key Performance Area (KPA) of environment, and the involvement of stakeholders (if any). Comprehensive details of the experimental procedures are provided, including the general approach adopted, and an analysis of the exercises performed.

Authoring & approval

Author(s) of the document

Organisation name	Date
Universidad Carlos III de Madrid	01/06/2025

Reviewed by

Organisation name	Date
UC3M	05/06/2025
BIRA	05/06/2025
KTH	05/06/2025
RMI	05/06/2025

Approved for submission to the SESAR 3 JU by

Organisation name	Date
UC3M	09/06/2025
BIRA	09/06/2025
KTH	09/06/2025
RMI	09/06/2025

Rejected by

Organisation name	Date

Document history

Edition	Date	Status	Organisation author	Justification
01.00	11/06/2025	Submission	UC3M	

Copyright statement

© (2025) – (UC3M, BIRA, KTH and RMI). All rights reserved. Licensed to SESAR 3 Joint Undertaking under conditions.

Disclaimer

The opinions expressed herein reflect the author's view only. Under no circumstances shall the SESAR 3 Joint Undertaking be responsible for any use that may be made of the information contained herein

E-CONTRAIL

ARTIFICIAL NEURAL NETWORKS FOR THE PREDICTION OF CONTRAILS AND AVIATION INDUCED CLOUDINESS

E-CONTRAIL

This document is part of a project that has received funding from the SESAR 3 Joint Undertaking under grant agreement No 101114795 under European Union's Horizon Europe research and innovation programme.



We provide a high-level summary of the project E-CONTRAIL:

Contrails and aviation-induced cloudiness effects on climate change show large uncertainties since they are subject to meteorological, regional, and seasonal variations. Indeed, under some specific circumstances, aircraft can generate anthropogenic cirrus with cooling. Thus, the need for research into contrails and aviation-induced cloudiness and its associated uncertainties to be considered in aviation climate mitigation actions becomes unquestionable.

We will blend cutting-edge AI techniques (deep learning) and climate science with application to the aviation domain, aiming at closing (at least partially) the existing gap in terms of understanding aviation-induced climate impact.

The overall purpose of E-CONTRAIL project is to develop artificial neural networks (leveraging remote sensing detection methods) for the prediction of the climate impact derived from contrails and aviation-induced cloudiness, contributing, thus, to a better understanding of the non-CO2 impact of aviation on global warming and reducing their associated uncertainties as essential steps towards green aviation.

Specifically, the objectives of E-CONTRAIL are:

- O-1 to develop remote sensing algorithms for the detection of contrails and aviation-induced cloudiness.
- O-2 to quantify the radiative forcing of ice clouds based on remote sensing and radiative transfer methods.
- O-3 to use of deep learning architectures to generate AI models capable of predicting the radiative forcing of contrails based on data- archive numerical weather forecasts and historical traffic.
- O-4 to assess the climate impact and develop a visualization tool in a dashboard.

Table of contents

Abstract	1
1 Executive summary	8
2 Introduction.....	9
2.1 Purpose of the document.....	9
2.2 Intended readership	10
2.3 Background	10
2.4 Structure of the document.....	10
2.5 Glossary of terms.....	11
2.6 List of acronyms	11
3 Context of the experimental research report	13
3.1 SESAR solution E-CONTRAIL “Climate Quantification and Hotspot Prediction Service”: a summary	13
3.2 Summary of the exploratory research plan.....	14
3.3 Deviations	20
4 Validation results.....	22
4.1 Summary of project E-CONTRAIL validation results.....	22
4.2 Preliminary Experiment: Satellite-Based Quantification of Contrail Radiative Forcing over Europe: a two-week analysis of Aviation-Induced Climate Effects.	23
4.3 Experiment 1: Satellite-Based Quantification of Contrail Radiative Forcing over Europe and ISSR forecasting: a Full Year 2023 analysis.....	33
4.4 Confidence in validation results	40
5 Conclusions and recommendations.....	59
5.1 Conclusions	59
5.2 Recommendations.....	60
6 References	64
6.1 Applicable documents	64
6.2 Reference documents	64

List of figures

Figure 1: Radiative Forcing (RF), measured in Watts per square meter, in the Short Wave (SW) (first image) and in the Long Wave (LW) (second image) for the field of view of Meteosat satellites on January 24, 2023, at 08:00 UTC. The final image illustrates the net RF obtained by aggregating the SW and the LW RF.....	25
Figure 2: Comparison of the cumulative Radiative Forcing (RF) curves (top) and the cumulative contrail RF curves (bottom) between the week of the 24th-30th of January of 2023 and the corresponding week in 2024.....	28
Figure 3: Comparison of the number of contrails detected across the entire field of view of Meteosat satellites in the week of the 24th-30th of January of 2023, and the same week in 2024, encompassing all times of day.	29
Figure 4: Correlation Matrix between contrail cloud parameters (COD, CTH, CER), zenith angles, surface temperatures (ST), RF values, and contrail sizes of all contrails detected over the two-week period. 30	
Figure 5: The top row displays the Radiative Forcing (RF) of contrail features by size, measured in terawatts, with separate panels for daytime (left) and nighttime (right). The bottom row shows the Short Wave (SW) and Long Wave (LW) RF components for daytime (left) and nighttime (right), which were combined to derive the RF values presented in the top row.	31
Figure 6: Contrail Outbreak over the Atlantic Ocean spanning a 9.5-hour period on January 30, 2023. The red colors indicate positive Radiative Forcing values (warming) while the blues indicate negative values (cooling).....	32
Figure 7: E-CONTRAIL visualization Dashboard.....	37
Figure 8: Average number of contrails detected across different times of the day (preliminary results).	37
Figure 9: RF estimation over 2023 (preliminary results).....	38
Figure 10: ISSR prediction Vs Ground truth.	39
Figure 11: ISSR prediction metrics.....	39
Figure 12: ISSR & RF prediction metrics	39
Figure 13: Contrail detection examples	41
Figure 14: Lagrangian Contrail Model (COCIP).....	42
Figure 15: the E-CONTRAIL contrail detection and RF estimation algorithm	42
Figure 16: Contrail detection above water clouds vs clear sky conditions-	43
Figure 17: Contrail CoCIP mask over clouds.....	43
Figure 18: E-CONTRAIL estimation of RF during the week of Feb 1st - 8th 2025	44
Figure 19: Comparison of E-CONTRAIL estimation (MSG) of RF w.r.t.....	44

Figure 20: Example of a contrail outbreak	45
Figure 21: Scatter plot between radiative forcing (RF) values estimated by using the Look-Up Tables (LUTs) (RFLUTs) and radiative transfer calculations using the actual atmospheric temperature vertical profiles (RFactual) for randomly selected pixels containing contrails above land surfaces in the (a) SW, and (b) LW, underlying water clouds (i.e., multi-layered) in the (c) SW and (d) LW, and ocean surfaces in the (e) SW and (f) LW.	47
Figure 22: Vertical (a) temperature and (b) humidity profiles of US Standard atmosphere, median profiles of the ECMWF vertical profiles corresponding to the largest discrepancies (i.e., large RMS error percentage between radiative forcing (RF) values estimated by using the Look-Up Tables (LUTs) (RF_{LUTs}) and radiative transfer calculations using the actual atmospheric temperature vertical profiles) for the six selected days.	48
Figure 23: Simulated radiative forcing values in the shortwave (i.e., solar) wavelength range (RF_{sol}) are shown as a function of various ice crystal habits and their degrees of roughness based on the parameterization of Yang et al. (2013) and Fu et al. (1998) for three different solar zenith angle (SZA) scenarios. The horizontal line (i.e., grey dashed line) represents the $RF_{sol(reference)}$ value for the selected ice crystal shape and roughness used in this study.	49
Figure 24: Comparison of Predicted vs. Actual Radiative Forcing from Contrails.....	53
Figure 25: ConvLSTM performance across different input configurations	55
Figure 26: CNN+LSTM performance across different input configurations	56
Figure 27: ViT performance across different input configurations	56
Figure 28: GeoSTANet performance across different input configurations	57
Figure 29: Main challenges associated with contrail detection models trained on OpenContrails. Dataset (Ortiz et al., 2025).	61

List of tables

Table 1: glossary of terms	11
Table 2: list of acronyms.....	12
Table 3: maturity levels table	13
Table 4. validation objectives and success criteria	15
Table 5: validation assumptions overview	16
Table 6: validation exercise layout	17
Table 7: Initial exercise #01 time planning.....	17
Table 8: stakeholders' expectations and involvement.....	19
Table 9: exercise #01 risks and mitigation actions.....	20
Table 10: review chronogram for the validation exercise #01 time planning	21
Table 11: summary of validation exercises results	23
Table 12: Average daily RF and CRF values for the 24-30th Jan, 2023 and the 24-30th Jan, 2024	29
Table 13: Monthly median of RF values (W/m ²) for Ice Clouds and Contrails in year 2023.	38
Table 14: The formulas for the averaged and global metrics are computed in terms of True Positives (TP), False Positives (FP), and False Negatives (FN) across a dataset consisting of N samples.....	40
Table 15: The 'GMax' columns utilize a global threshold that maximizes Dice Score (DS) across masks, while the 'APT' columns apply the adaptive probability thresholding technique presented in this work. The error margin values utilized for the $\gamma\beta$ -BDS metric are $\gamma = \beta = 2$	41
Table 16: Mean radiative forcing (RF) values over all the randomly selected pixels for the six selected days, bias, RMS error, RMS error percentage, and mean percent errors between RF values estimated by using the Look-Up Tables (LUTs) and by using the ERA5 atmospheric profile and the OCA cloud conditions for the SW, and LW estimated RFs.	47
Table 17: Mean radiative forcing (RF) values over all the randomly selected pixels for the 25th of September 2023, bias, RMS error between RF values when using an ice crystal habit of column with 8 elements and a plate with 10 elements for the SW, LW, and net estimated RFs.....	50
Table 18: Performance Metrics for Different Models and Feature Configurations.....	54

1 Executive summary

Specifically, the objectives of E-CONTRAIL are:

- O-1 to develop remote sensing algorithms for the detection of contrails and aviation-induced cloudiness.
- O-2 to quantify the radiative forcing of ice clouds based on remote sensing and radiative transfer methods.
- O-3 to use of deep learning architectures to generate AI models capable of predicting the radiative forcing of contrails based on data- archive numerical weather forecasts and historical traffic.
- O-4 to assess the climate impact and develop a visualization tool in a dashboard.

In the experimental research plan we defined 2 experiments (Experiment 1 focusing on a small dataset of two weeks of data; Experiment 2 focusing on a full year of data). Each experiment has been divided into 5 activities, one devoted to assessing each of the validation objectives:

- Activity 1.1 linked to Val-O1.1: to develop remote sensing algorithms for the detection of linear contrails.
- Activity 1.2 Val-O1.2: to develop remote sensing algorithms for the detection of aviation induced cloudiness.
- Activity 2 Val-O2: quantify the radiative forcing of ice clouds based on remote sensing and radiative transfer methods.
- Activity 3 Val-O3: To develop deep learning architectures to generate AI models capable of predicting the radiative forcing of contrails.
- Activity 4 Val-O4: is to assess the climate impact and develop a visualization tool in a dashboard.

The document provides the context in which the research was carried out, including a summary of the experimental plan (validation objectives, exercises, activities, the expected performance contributions, primarily in the Key Performance Area (KPA) of environment, and the involvement of stakeholders (if any)). Comprehensive details of the experimental procedures are provided, including the general approach adopted, and an analysis of the validation activities concluded so far (Activity 1.1).

2 Introduction

2.1 Purpose of the document

In this Experimental Research Report (ERR) for the E-CONTRAIL project, we present the outcomes of executing the Exploratory Research Plan for the “E-CONTRAIL Climate Hotspot Prediction Service.” As a SESAR exploratory research initiative, E-CONTRAIL aimed to develop innovative solutions for predicting climate hotspots influenced by aviation activities.

Purpose and Execution of the Experimental Plan

- The primary objective of the experimental plan was to ensure the application of scientific best practices in assessing the results of the E-CONTRAIL project. To achieve this, the following steps were undertaken:
- Identification of Reference Guidance Documents: We meticulously selected and reviewed relevant guidance documents to align our methodologies with established scientific standards. This ensured that our experimental design and assessment criteria were robust and credible.
- Definition of Research Questions and Hypotheses: Clear research questions were formulated to guide the investigation, and corresponding hypotheses were established to be tested through empirical data. This structured approach facilitated focused and meaningful experimentation.
- Design and Implementation of Experiments: A series of experiments were meticulously designed to address the research questions and test the hypotheses. These experiments incorporated both qualitative and quantitative methods to provide a comprehensive analysis of the Climate Hotspot Prediction Service.
- Development of Metrics and Assessment Methods: Specific metrics were defined to evaluate the performance and accuracy of the prediction service. Advanced statistical and computational methods were employed to assess the results, ensuring thorough and objective evaluation.

Results and Findings

The partial execution of the experimental plan so far has yielded significant insights and validated one of the key enablers of the E-CONTRAIL Climate Hotspot Prediction Service: the contrail detection algorithm. In particular:

- Alignment with Scientific Best Practices: By adhering to the identified reference guidance documents, the experimental procedures were conducted with high standards of scientific rigor. This alignment enhanced the reliability and validity of the findings.
- Successful Validation of Hypotheses: The experiments confirmed the initial hypotheses, demonstrating that the contrail detection algorithm of the Climate Hotspot Prediction Service accurately identifies the contrails in satellite images.
- Performance Metrics Achievement: The defined metrics indicated that the prediction service met and, in some cases, exceeded the expected performance criteria.
- Comprehensive Assessment Methods: The application of advanced assessment methods provided a detailed understanding of the service's capabilities and limitations. This

comprehensive evaluation facilitated the identification of areas for further improvement and optimization.

- **Stakeholder Engagement and Feedback:** Involving relevant stakeholders throughout the experimental process ensured that the prediction service aligns with practical needs and real-world applications. Feedback from stakeholders was instrumental in refining and enhancing its usability.

2.2 Intended readership

The intended readership for this Deliverable, focusing on the Experimental Research Report, comprises aviation researchers, industry innovators, regulatory authorities, and SESAR program stakeholders. This document is tailored to guide and inform those actively engaged in experimental research endeavours aimed at enhancing air traffic management and advancing the SESAR program's objectives.

2.3 Background

There is no previous project or activity in which E-CONTRAIL is building up.

The reference document to prepare this Report is the Experimental Research Plan (D5.2), which was based on the Experimental Approach guidance ER [AD1]. In addition, we rely on E-CONTRAIL's Grant Agreement [AD2], where the research questions and hypotheses were established.

2.4 Structure of the document

The Experimental Research Report is systematically structured to provide a comprehensive overview of the E-CONTRAIL project's objectives, methodologies, and findings. It begins with an Abstract that succinctly summarizes the key aspects of the research. The Executive Summary offers a high-level overview of the project's scope, main outcomes, and significance. The Introduction section details the purpose of the document, identifies the intended readership, provides essential background information, outlines the structure of the report, and includes a glossary of terms and a list of acronyms to ensure clarity. Any deviations from the SESAR 3 JU project handbook [AD3] are addressed.

Following the introduction, the Context of the Experimental Research Report elaborates on the SESAR solution "E-CONTRAIL Climate Quantification and Hotspot Prediction Service," summarizes the exploratory research plan, and discusses any deviations encountered during the project. The Validation Results section presents a summary of the validation outcomes so far in the project, including a detailed satellite-based quantification of contrail radiative forcing over Europe, an in-depth analysis of validation results aligned with each research objective, and an assessment of the confidence in these results.

The report concludes with Conclusions and Recommendations, which synthesize the findings and provide actionable suggestions for future research and implementation. Finally, the References section lists all applicable and reference documents that underpin the research, ensuring transparency and allowing for further exploration of the topics discussed.

2.5 Glossary of terms

Term	Definition	Source of the definition
Non-CO ₂	Effect of aviation due to nitrogen oxides (NO _x), vapour trails and cloud formation triggered by the altitude at which aircraft operate (i.e., not related to CO ₂ emissions).	www.transportenvironment.org
Contrails	When water vapour is released from jet engines at altitude under certain high humidity conditions (ice supersaturated regions) it can condense into exhaust carbon particles as well as into atmospheric aerosols. If the air is sufficiently humid, the water vapour can condense further into crystals and a cloud can be formed. Such clouds, formed from the condensation of exhaust aircraft water vapour, are called condensation trails or contrails.	www.iata.org
Aviation Induced Cloudiness	Aviation-induced cloudiness (AIC) is defined to be the sum of all changes in cloudiness associated with aviation operations.	https://archive.ipcc.ch
Radiative Forcing	Radiative forcing is what happens when the amount of energy that enters the Earth's atmosphere is different from the amount of energy that leaves it. Energy travels in the form of radiation: solar radiation entering the atmosphere from the sun, and infrared radiation exiting as heat. If more radiation is entering Earth than leaving—as is happening today—then the atmosphere will warm up. This is called radiative forcing because the difference in energy can force changes in the Earth's climate.	https://climate.mit.edu

Table 1: glossary of terms

2.6 List of acronyms

Term	Definition
<i>aCCF</i>	<i>algorithmic Climate Change Functions</i>
<i>AIC</i>	<i>Aviation-induced Cloudiness</i>
<i>ATM</i>	<i>Air traffic management</i>
<i>CoaT</i>	<i>Co-Scale Conv-Attentional Image Transformers)</i>
<i>CoCiP</i>	<i>Contrail Cirrus Prediction Model</i>
<i>CRF</i>	<i>Cumulative Radiative Forcing</i>
<i>CNN</i>	<i>Convolutional Neural Network</i>
<i>CTH</i>	<i>Cloud Top Height</i>
<i>DES</i>	<i>Digital European Sky</i>

ECMWF	<i>European Center of Medium Range Weather Forecast</i>
ERF	<i>Earth Radiative forcing</i>
ERP	<i>Experimental Research Plan</i>
EXE	<i>Exercise</i>
FPN	<i>Feature Pyramid Network</i>
GA	<i>Grant agreement</i>
GAN	<i>Generative Adversarial Network</i>
GHG	<i>Greenhouse Gas</i>
HE	<i>Horizon Europe</i>
ID	<i>Identifier</i>
KPA	<i>Key performance area</i>
KPI	<i>Key performance indicator</i>
LSTM	<i>Long-Short Terms Memory</i>
LUT	<i>Look Up Table</i>
MSG	<i>Meteosat Second Generation</i>
MTG	<i>Meteosat Third Generation</i>
OCA	<i>Optical Cloud Analysis</i>
OSD	<i>Operational service and environment description</i>
RF	<i>Radiative Forcing</i>
RMS	<i>Root Mean Square</i>
SESAR	<i>Single European sky ATM research</i>
SESAR 3 JU	<i>SESAR 3 Joint Undertaking</i>
SZA	<i>Solar Zenit Angle</i>
VAE	<i>Variational AutoEncoder</i>
ViT	<i>Vision Transformer</i>

Table 2: list of acronyms

3 Context of the experimental research report

3.1 SESAR solution E-CONTRAIL “Climate Quantification and Hotspot Prediction Service”: a summary

This project can be classified as "Pre-TRL1 Scientific Research". E-CONTRAIL has the overall objective of the future integration of the project's outcome into the ATM processes, thus a strategic goal is to be able to show readiness for TRL2. At the end of the project, we ambition to showcase the maturity level of the E-CONTRAIL solution and, thus, readiness for TRL1.

E-CONTRAIL Solution (which we have coined at this project stage “**E-CONTRAIL Climate Quantification and Hotspot Prediction Service**”) will consist of an AI-driven model (already trained using historical data) capable of predicting the volumes of airspace with the conditions for large global warming impact due to contrails and aviation-induced cloudiness. A user-friendly visualization tool tailored for stakeholders' needs will be also implemented.

The foreseen activities are:

- Scientific studies on remote sensing (of contrails, aviation-induced cloudiness) and deep learning.
- Algorithms for remote sensing (of contrails, aviation-induced cloudiness) and implementation of deep learning architectures.
- Concept analysis, via visualization tool, oriented towards aviation stakeholders.

The expected outcomes are:

- We will state the basic principles about the studies and algorithms related to remote sensing and deep learning architectures.
- We will identify the potential application and the end users. They will be invited to participate in the conceptual design of the visualization tool.
- We aim at formulating the technological concept and/or application as a met service.

Project/ Proposed SESAR solution(s) ID	Proposed SESAR solution(s) title	Initial level	maturity	Exit level	maturity	Reused validation material from past R&I Initiatives
E-CONTRAIL	E-CONTRAIL Climate Quantification and Hotspot Prediction Service	TRL0		TRL1		-

Table 3: maturity levels table

3.2 Summary of the exploratory research plan

The Experimental Research Plan was presented in D5.2. Interested readers are referred therein for a detailed and insightful description. Here we provide a summary

3.2.1 Exploratory research plan purpose

The experimental plan shall ensure that the specific objectives of the project are achieved, yet measured and quantified. The experimental approach will be based on: The quantitative assessment of the goodness of fit of the deep learning methods employed in the project. This requires the usage of state-of-the-art artificial intelligence metrics (see Section 4.1.1 of D5.2).

3.2.2 Summary of validation objectives and success criteria

The following table only lists the validation objectives for the models developed for the E-CONTRAIL Solution 1; the operational objectives will be added in the final VALP, where the stakeholder benefits analysis will be performed.

Object ive ID	Objective title	Objective description	Success Criteria	Research Questions
Val-O1.1	<u>to develop remote sensing algorithms for the detection of linear contrails</u>	Apply artificial intelligence algorithms for the detection of contrails from remote sensing devices.	Quantitative indicators in supervised learning, including Precision, Recall, Dies Score, F1 Score. We expect to obtain contrail detection accuracies greater than 80% (in F1 Score)	Linked to E-Contrail O-1 and RQ#01;
Val-O1.2	<u>to develop remote sensing algorithms for the detection of aviation induced cloudiness</u>	Combine optical flow techniques for time-interpolation and Computational Fluid Dynamics methods to understand the temporal evolution of the contrails	Qualitative analysis: 1) Compare with State-of-art methods in the literature: COCIP and NASA models. 2) Compare with contrails captures by visual camaras (all-sky camaras) We expect to qualitative improve the current state-of-art models	Linked to E-Contrail O-1 and RQ#01;
Val-O2	<u>quantify the radiative forcing of ice clouds based</u>	Starting from remote-sensing data and radiative transfer-based lookup tables, compute the	Comparison with short-wave and long-wave fluxes from the Cloud and Earth Radiant Energy System	Linked to E-Contrail O-

	<u>on remote sensing and radiative transfer methods</u>	effective radiative forcing for ice clouds	(CERES) Single Satellite Footprint (SSF) product from NASA. The CERES SSF product provides all-sky fluxes, which are also available in the lookup tables. We will average those fluxes for pixels identified as ice clouds, over the E-CONTRAIL region. We aim for 10% accuracy when comparing our results with the CERES data	2 and RQ#02
Val-O3	<u>To develop deep learning architectures to generate AI models capable of predicting the radiative forcing of contrails</u>	<u>based on data-archive numerical weather forecasts and historical traffic, we will make use of Convolutional Neural Networks (CNNs), together with transfer learning from already-existing models, as well as recurrent networks such as Long-Short Terms Memory (LSTM), and generative models such as Generative Adversarial Networks (GANs) and variational autoencoders (VAEs).</u>	Quantitative indicators in supervised learning, including Precision, Recall, Dies Score, F1 Score. We expect to obtain contrail detection accuracies greater than 80% (in precision)	Linked to E-Contrail O-3 and RQ#03
Val-O4	<u>is to assess the climate impact and develop a visualization tool in a dashboard.</u>	The use of Geospatial data (GeoTIFF, shapefiles, GeoJSON) associate with GeoServer and request from users, is a good way to introduce the time dimension in the visualization. RF from AIC (associated to flight trajectory and contrails labelling) will be shown on a world map is to operate. Forecasts of area with possible RF and climate impact will also be shown.	The assessment of RF from AIC will be quantitative.	Linked to E-Contrail O-4 and RQ#04

Table 4. validation objectives and success criteria

3.2.3 Validation assumptions

Assumption ID	Assumption title	Assumption description	Justification	Impact Assessment
Assu #1	Europe	Geographical scope limited to EUROPE and half of the North Atlantic. Due to the lack of a labelled dataset in Europe, we will train our models with data in US and then extrapolate the resulting model to US.	It is obvious, due to the funding, that we must focus on Europe.	We will label a small dataset in Europe to compare the validation scores.
Assu #2	MSG data	We will use the Meteosat Second Generation (MSG) data from years 2022-2024.	We wanted to use the Meteosat Third Generation (MTG) data, but they did not become available timely for the project	MSG has lower spatio-temporal resolution than MTG.

Table 5: validation assumptions overview

3.2.4 Validation exercises list

[EXE-01]

Identifier	<i>TVAL.01.0-[ProjectAcronym]-TRL1</i>
Title	<i>Contrail and aviation-induced cloudiness radiative forcing prediction using deep learning.</i>
Description	<i>Numerical simulations using a year (2023) of data over Europe.</i>
KPA/TA addressed	<i>Environment</i>
Addressed expected performance contribution(s)	<i>Reduce uncertainties in contrail and aviation-cloudiness radiative forcing estimates and predictions.</i>
Maturity level	<i>TRL1</i>
Use cases	<i><UC1> Europe-2023 using MSG data</i>
Validation technique	<i>AI performance metrics</i>
Validation platform	<i>N/A</i>
Validation location	<i>N/A</i>
Start date	<i>N/A</i>
End date	<i>N/A</i>
Validation coordinator	<i>UC3M</i>
Status	<i>Not Started</i>
Dependencies	<i>N/A</i>

[EXE 1-Trace]

Linked Element Type	N/A
<SESAR Solution>	N/A
<Project>	E-CONTRAIL
<Sub-Operating Environment>	N/A
<Validation Objective>	N/A

Table 6: validation exercise layout

3.2.4.1 Validation exercises planning

ID	Q3 2024			Q4 2024			Q1 2025		
	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
EXE01									

3.2.4.2 Exercise planning and management

Activities: The experiment will be divided into 5 activities, one devoted to assessing each of the validation objectives:

- Activity 1.1 linked to Val-O1.1: to develop remote sensing algorithms for the detection of linear contrails.
- Activity 1.2 Val-O1.2: to develop remote sensing algorithms for the detection of aviation induced cloudiness.
- Activity 2 Val-O2: quantify the radiative forcing of ice clouds based on remote sensing and radiative transfer methods.
- Activity 3 Val-O3: To develop deep learning architectures to generate AI models capable of predicting the radiative forcing of contrails.
- Activity 4 Val-O4: is to assess the climate impact and develop a visualization tool in a dashboard.

Time planning:

ID	Q3 2024			Q4 2024			Q1 2025		
	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
Activity 1.1									
Activity 1.2									
Activity 2									
Activity 3									
Activity 4									

Table 7: Initial exercise #01 time planning

3.2.5 Estimated performance contributions.

The main contribution of E-CONTRAIL in terms of Key Performance Areas (KPA) is on the environment, particularly on better understanding non-CO2 impacts.

The outcome related to the Environment specified in the call (as a high level ambition at European level) was: achievement of the objectives of a 55% reduction in greenhouse gas emissions by 2030 and net-zero greenhouse gas emissions by 2050, from a gate-to-gate perspective, by introducing new concepts enabling proper modelling of non-CO2 emissions and their impact on optimum green trajectories, taking into account the expected interoperability with new entrants (i.e., U-space flights) [RD2].

E-CONTRAIL's R&I goals will enable advanced AI-powered prediction of the radiative forcing of contrails and aviation-induced cloudiness, thereby enabling non-CO2 emission related climate mitigation actions by the aviation industry. Our unique contributions towards 55% reduction in climate impact of aviation by 2030 and climate neutrality by 2050 (a context indicator for reduction of 55% GHG emissions) will be:

1. Introducing AI driven models to predict, 24 hours in advance, the climate impact of contrails and aviation-induced cloudiness with 80-90% accuracy. We expect to achieve this accuracy based on the results obtained in our previous research activities related to predicting thunderstorms using AI. [RD2] [RD2].
2. Reducing the uncertainty in the climate impact (measured in terms of Radiative Forcing (RF) and/or Effective Radiative Forcing (ERF)) of contrails and aviation-induced cloudiness. We expect to contribute to this research on contrails and other non-CO2 effects by comparing E-CONTRAIL's AI-Driven approach with existing methods (e.g., the aCCFs) by better understanding the atmospheric conditions (related to ice-supersaturated regions) in which contrails form and persist.

By 2030, E-CONTRAIL will work towards the following, specific impact targets that will lead to the reduction of aviation-induced climate impact:

1. Identify contrail cirrus-forming in ice-supersaturated regions of the atmosphere and its radiative forcing.
2. Enable more efficient navigational avoidance & operations management and, therefore, the climate impact is reduced by 20-50%
3. Optimize airlines' operational costs and the climate impact. Trade-off solutions will be obtained, expecting an increase of the operational costs ranging from 0.5% to 3% to achieve climate mitigation reduction of 20-50%.
4. Developing indicators that enable the concept of green trajectories for the first time, and the quantified indicators lead evidence-based policy making (fees and incentives for the airlines to compensate the extra costs).

All in all, our Climate Quantification and Hotspot Prediction Service can be used as a meteorological enabler for the airlines and flight dispatchers towards reducing in 20-50% the aviation-induced climate-impact [RD2] by 2030 via climate-optimized trajectories [RD2], at an increased operational cost ranging from 0.5% to 3%, (FlyATM4E D4.4) [RD2].

3.2.6 Stakeholders' expectations and involvement

The involvement of Stakeholders to qualitatively assess the presentation of results in our dashboard will be also key.

Aviation Industry/ Airline operators: Contrails which heighten the effect of global warming may account for more than half (57%) of the entire climate impact of aviation [RD2]. However, more research is needed to bring knowledge about contrails and chemical interactions in the atmosphere to a level at which the aviation industry can be more confident about the route forward. The uncertainty distributions show that non-CO₂ forcing contributes about 8 times more than the CO₂ to the overall uncertainty in the aviation net forcing (EASA, 2018) [RD2]. **With our accurate prediction of persistent contrails and radiative forcing, we will enable the aviation industry to carry out operational changes and prompt mitigatory actions (such as navigation avoidance and others).**

Policy makers: Decarbonisation of aviation sector will continue (e.g., alternative fuels, electrical aircraft), however, requires high investments and results in stranded assets (aircraft, engines, etc.). Policy makers are increasingly looking for modernization of the Air Traffic Management (ATM) not only to consider both CO₂ and non-CO₂ effects in the long term, but also to enable short-term actions to mitigate aviation-induced climate change. The quantification of the cost linked to flying green and the development of indicators is mandatory to pave the road towards establishing fees and incentives. **Therefore, the results of E-CONTRAIL are highly relevant, and will result in actionable policy insights, to fast track the modernization of ATM.**

The ANSPs and the network manager can also benefit from E-CONTRAIL solution: **they can define volumes of airspace that are highly sensitive to climate change, thereby issues restrictions if they deem them necessary.**

The Met Offices can also benefit from E-CONTRAIL solution: **they can integrate the climate models into their meteorological services.**

Stakeholder	Involvement	Why it matters to the stakeholder
Airline Operators	We will involve them in the final workshop. We are also in touch with them within the Green Deal Flagship	With our accurate prediction of persistent contrails and radiative forcing, we will enable the aviation industry to carry out operational changes and prompt mitigatory actions (such as navigation avoidance and others)
Policy Makers	We will involve them in the final workshop. We are also in touch with them within the Green Deal Flagship	the results of E-CONTRAIL are highly relevant, and will result in actionable policy insights, to fast track the modernization of ATM
ANSPs and the NM	We will involve them in the final workshop. We are also in touch with them within the Green Deal Flagship	can also benefit from E-CONTRAIL solution: they can define volumes of airspace that are highly sensitive to climate change, thereby issues restrictions if they deem them necessary.
The Met. Offices	We will involve them in the final workshop. We are also in touch with them within the Green Deal Flagship	can also benefit from E-CONTRAIL solution: they can integrate the climate models into their meteorological services

Table 8: stakeholders' expectations and involvement

3.3 Deviations

3.3.1 Deviations with respect to the S3JU project handbook

There are no deviations from the SESAR 3 JU project handbook.

3.3.2 Deviations with respect to the exploratory research plan (ERP)

We initially carried out an intermediate integrated experiment including 2 weeks of data. The results were published in the SESAR innovation days 2024 [RD18]. The experiment with the full dataset of 2023 has just concluded with some delay due the cascade of delays that we have been accumulating.

The only risk was associated to the lack of MTG data, which has been materialized already. The mitigation action, which has been already put in place, was to use MSG.

Risks	Impact (1-low, 2-medium, 3-high)	Likelihood (1-low, 2-medium, 3-high)	Criticality (calculated based on likelihood and impact)	Mitigation actions
Risk 1: Lack of MTG data	Medium	Medium	Medium	Usage of MSG data

Table 9: exercise #01 risks and mitigation actions

This delay in the access to MTG data has had several implications:

- 1) MTG was initially supposed to be available at the end of 2023. A delay was announced in Early 2024 and a new date for delivery was established in April 2024. We then decided to request a 2-months delay in the WP1 Deliverables (D1.1, D1.2) to wait until MTG could be available in April. However, a problem was identified in the data and new delays on MTG availability were announced (without committing any new data for the data release).
- 2) As a consequence, we decided to use MSG data for the activities in WP1, which has a lower spatio-temporal resolution than MTG. The WP has been closed (including the approval of D1.1 and D1.2) using MSG data, though with a 2-months delay.
- 3) The 2-months delay has been also affecting WP3 activities (because the output of WP1 is the input of WP3), so we had to ask for a 2-months delay in D3.2, which was accepted by SJU.
- 4) On Sept. 24th 2024, MTG was finally made available in pre-operational mode. It comes, however, with limited information (in particular, information that we need for WP2 activities and that was available in MSG and was supposed to be available in MTG) and a line artifact in the image. We decided to pause the validation activities for one month and dedicate efforts in exploring to which extent this version of MTG is useful for the algorithms developed in WP1 and WP2 and, if needed, find workarounds. After careful analysis, we concluded that MTG could be used in the framework of contrail detection (WP1) but not for RF estimation (WP2).

Consequently, the initial validation planning for the different activities has been re-adjusted, adding a delay of 3 months in the validation of Activity 1.1, 1.2, and 2. Activity 3 will accumulate a Delay of 2 months. Activity 4 has finalized in May 2025 with 2 months of delay. **Table 10** includes a reviewed chronogram for the validation activities.

ID	Q3 2024			Q4 2024			Q1 2025		
	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
Activity 1.1									
Activity 1.2									
Activity 2									
Activity 3									
Activity 4									

Table 10: review chronogram for the validation exercise #01 time planning

4 Validation results

4.1 Summary of project E-CONTRAIL validation results

Project validation objective ID	Project validation objective title	Project success criterion ID	Project criterion success	Project validation results	Project validation objective status
Val-O1.1:	<u>to develop remote sensing algorithms for the detection of linear contrails</u>	<i>E-CONTRAIL Success criterion #1.1</i>	Quantitative indicators in supervised learning, including Precision, Recall, Dies Score, F1 Score. We expect to obtain contrail detection accuracies greater than 80% (in F1 Score)	<i>Scores based on $\gamma\beta$-Boundary Soft Dice (BSD) over 80% (See Section 4.3.1). See Ortiz et al. 2025 [RD19]</i>	OK
Val-O1.2:	<u>to develop remote sensing algorithms for the detection of aviation induced cloudiness.</u>	<i>E-CONTRAIL Success criterion #1.2</i>	Qualitative analysis: 3) Compare with State-of-art methods in the literature: COCIP and NASA models. 4) Compare with contrails captures by visual camaras (all-sky camaras) We expect to qualitative improve the current state-of-art models	<i>3) Comparison with COCIP performed. 4) Comparison with Ground Camaras not possible due to difference in resolution. See Ortiz et al., 2025.b [RD22]</i>	OK
Val-O2:	<u>quantify the radiative forcing of ice clouds based on remote sensing and radiative transfer methods.</u>	<i>E-CONTRAIL Success criterion #2</i>	Comparison with short-wave and long-wave fluxes from the Cloud and Earth Radiant Energy System (CERES) Single Satellite Footprint (SSF) product from NASA. The CERES SSF product provides all-sky fluxes, which are also available in the lookup tables. We will average those fluxes for pixels identified as ice clouds, over the ECONTRAIL region. We aim for 10% accuracy	<i>Completed. The RF comparison and the fluxes comparison provided errors around 10-15% when compared with the Look Up Tables. See Dimitropoulou et al, 2025 [RD23]</i>	OK

Project validation objective ID	Project validation objective title	Project success criterion ID	Project criterion	success	Project validation results	Project validation objective status
			when comparing our results with the CERES data			
Val-O3:	<u>To develop deep learning architectures to generate AI models capable of predicting the radiative forcing of contrails.</u>	<i>E-CONTRAIL Success criterion #3</i>	Quantitative indicators in supervised learning, including Precision, Recall, Dies Score, F1 Score. We expect to obtain contrail detection accuracies greater than 80% (in precision)		<i>Completed.</i>	<i>OK</i>
Val-O4:	<u>is to assess the climate impact and develop a visualization tool in a dashboard.</u>	<i>E-CONTRAIL Success criterion #4</i>	The assessment of RF from AIC will be qualitative.		<i>Completed</i>	<i>OK</i>

Table 11: summary of validation exercises results

4.2 Preliminary Experiment: Satellite-Based Quantification of Contrail Radiative Forcing over Europe: a two-week analysis of Aviation-Induced Climate Effects.

We present in this section the integrated experiment for two weeks of data, which was presented in the SIDs conference 2024 [RD18].

The experiment was conducted over two complete weeks, encompassing 24-hour periods each day. The selected periods were from January 24th to January 30th in both 2023 and 2024. Notably, exceptionally warm temperatures were reported in January 2024. While multiple factors could contribute to the observed temperature increase, visual inspection in the satellite imagery revealed a significant rise in both the number of contrails observed and those detected by models during this period. This increase in contrail coverage could be a potential contributor to the warming reported.

4.2.1 DATA

The data used in this work was obtained from the SEVIRI onboard the MSG satellites of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). In particular, the data from the MSG-3 (Meteosat-10) and MSG-4 (Meteosat-11) satellites have been used for this work. Positioned in geostationary orbit at 0° longitude, about 36,000 kilometres above Earth, these satellites provide spectral information across 11 channels, including visible and infrared regions. Observations are

captured every 15 minutes with a spatial resolution of approximately 3x3 km² at the subsatellite point, with a pixel size that grows as one looks further from the equator. The SEVIRI level 1b data serves two main purposes: (a) generating the false-color RGB images for contrail detection models, and (b) providing input to the Optical Cloud Analysis (OCA) system, which we use for the physical characterization of the clouds. These physical parameters serve as the input to the radiative forcing estimation. The specifics of this retrieved information are as follows:

(a) Ash Composite: This product generates false-color RGB images to enhance contrail visibility by combining several MSG thermal infrared (IR) bands. It is composed of the red Brightness Temperature (BT) difference IR12 μ m-IR10.8 μ m to highlight contrails by their higher transmissibility compared to natural cirrus, the green BT difference IR10.8 μ m-IR8.7 μ m to differentiate cloud phases, and the blue BT IR10.8 μ m to accentuate contrails by leveraging their colder temperatures relative to surrounding features. See [2] for the definition. The Ash RGB composite was also used in for detecting contrails.

(b) Cloud parameters: Cloud state parameters are characterized by cloud phase (CP), cloud top pressure (CTP), cloud optical thickness (COT), and cloud effective radius (CER). These parameters are obtained from the OCA of EUMETSAT, which employs the Optimal Estimation (OE) method along with SEVIRI spectral measurements simultaneously. The cloud information obtained can be separated into upper layer and lower layer clouds. The upper layer consists of ice clouds, and includes values both for cases where the ice clouds are alone and where they coexist with underlying water clouds. The lower layer shows data only when a water cloud is present beneath the upper ice layer.

(c) Forecast Data: We use the skin temperature from Numerical Weather Prediction (NWP) data obtained from the European Centre for MediumRange Weather Forecasts (ECMWF).

(d) Land Cover Data: We use the MODIS L3 500m Land Cover dataset MCD12Q1 v061.

We resample all input to a regularly spaced grid in latitude and longitude with a grid spacing of 0.04 degrees.

4.2.2 Methods

4.2.2.1 Radiative Forcing Calculations

The methodology employed to quantify the RF of contrail cirrus detectable by a geostationary satellite involves three key steps: First, a contrail detection model identifies the locations of all visible contrails within the scene. Second, the RF of all ice clouds present in the image is calculated. Third, the results from these two steps are intersected to determine the RF specifically attributable to the detected contrails. See Dimitropoulou et al, 2025 [RD23]

The radiative forcing estimation was derived using multidimensional interpolation on pre-built Lookup Tables (LUTs). These LUTs were constructed using the libRadtran radiative transfer library to simulate both shortwave and longwave radiative forcing for various combinations of thin to semi-transparent ice cloud parameters, along with other relevant factors like solar zenith angles, sea or land surface temperature, surface type and the presence of an underlying water cloud. By interpolating the simulated shortwave and longwave RFs from the LUTs as a function of the surface type, geometry, skin temperature, and OCA parameters for each pixel, a detailed and location-specific assessment of radiative forcing is provided. This approach streamlines the radiative forcing estimation process, removing the need to repeatedly run time-consuming radiative transfer simulations in future large-

scale analyses. Figure 1 provides an example of the shortwave, longwave, and net radiative forcing estimates obtained through this process. We use the sign convention of downward flux, so that a positive value of the RF represents a warming effect.

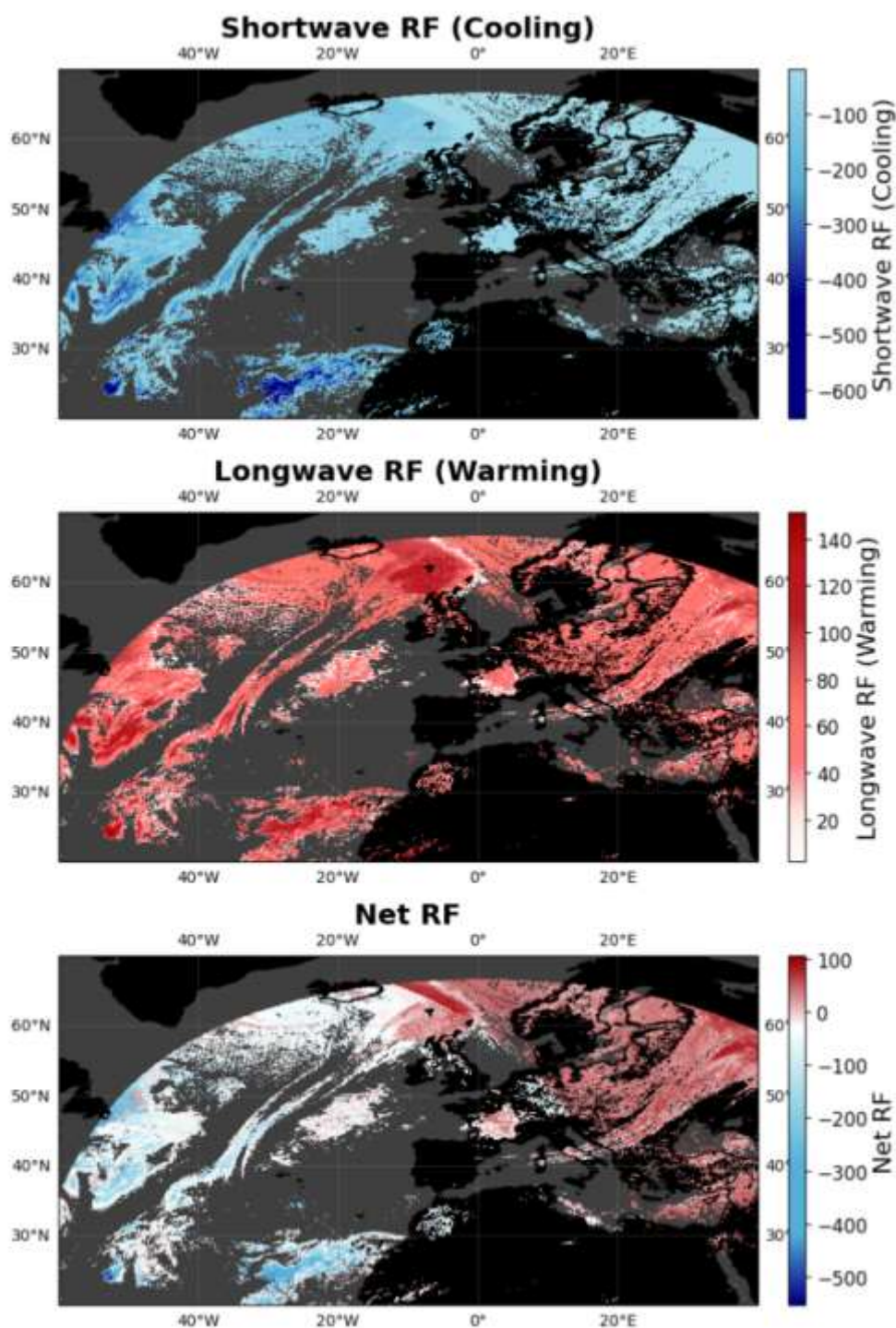


Figure 1: Radiative Forcing (RF), measured in Watts per square meter, in the Short Wave (SW) (first image) and in the Long Wave (LW) (second image) for the field of view of Meteosat satellites on January 24, 2023, at 08:00 UTC. The final image illustrates the net RF obtained by aggregating the SW and the LW RF.

4.2.2.2 Contrail Detection

For this experiment, we utilized a single-frame U-Netbased network previously trained on the OpenContrails Dataset, which comprises 22,000 images captured by the Geostationary Operational Environmental Satellites 16 Series (GOES-16) between April 2019 and April 2020 in different locations of North and South America. This dataset has been selected for training due to its substantial number of labelled scenes (55%), necessary for capturing all the variability of contrail features.

Network Architectural Details

The architecture employed is a hybrid neural network that combines a transformer-based encoder with a convolutional decoder. Because of the limitations in the computational resources, the encoder is a lightweight variant of the CoaT (Co-Scale Conv-Attentional Image Transformers) model, called CoaT-Lite Mini. This variant is optimized for efficient image processing, avoiding parallel blocks and incorporating a reduced channel depth in each layer. The model processes images through four sequential blocks, where feature maps are downsampled and converted into image tokens. These tokens are analyzed using convolutional operations for local pattern extraction and self-attention for capturing image interpart relationships. The output from each block is reshaped into a 2D feature map and forwarded to the next block and decoder via skip connections. The decoder employs sequential convolutional blocks with upsampling, producing feature maps at three different resolutions. The three feature maps are combined using a Feature Pyramid Network (FPN). The output is then regularized with a dropout layer, applying a 0.5 probability to deactivate weights to prevent overfitting to the training data. Finally, the features are upsampled to the original image size, with the number of channels reduced to one. Each pixel in this single-channel output mask represents the probability of being part of a contrail. The optimization of the weights of the network was performed using the AdamW optimizer over 30 epochs, minimizing a convex surrogate of the Dice loss function. Transfer learning was used to initialize the encoder's weights with those from a CoaT network pretrained on ImageNet.

Domain Adaptation

The trained network was used to detect contrails in MSG Ash RGB images, which have different characteristics from the original training data, especially in terms of geographical coverage and image resolution. MSG images cover Europe, Africa, and portions of the Atlantic Ocean, with a maximum thermal infrared resolution of 3x3 km². In contrast, the Ash RGB training images from GOES-16 focus on the United States, the Atlantic Ocean, South America, and the Caribbean, offering a finer resolution of 2x2 km² at nadir. Although the difference in geographical coverage is not expected to have an impact, the disparity in resolution could affect detection accuracy. To address this, the resolution of MSG images was adjusted using bilinear interpolation, simulating a 2x2 km² resolution at nadir. Given the large size of an MSG scene, a sliding window of 256x256 pixels was applied to divide the images into smaller overlapping sections. The detector was applied to each section, with results from overlapping areas combined to ensure accurate identification of objects partially visible across sections. Finally, after aggregating all the detections produced by the sliding window, the contrail mask for the entire scene was transformed back to the original resolution, preserving the true sizes of the segmented contrails.

4.2.3 Results

This section presents our results, starting with cumulative RF and CRF calculations every 15 minutes over two weeks to assess yearly variations. We then analyze cloud parameters, lightning, and other factors for over 700,000 contrails to identify key influences on net CRF.

4.2.3.1 Jan 2023- Jan 2024 Comparison

The analysis focuses on examining the changes observed during the two selected weeks. To draw more definitive conclusions, we plan to extend the study to cover a longer time frame in a future experiment. This extension will help account for seasonal variations and other factors that may influence the results.

Changes in Contrail Coverage

We compare the number of features detected by the model for the same week in 2023 and 2024 to evaluate changes in contrail coverage. Figure 2 presents the number of contrails detected at different times of day for each day of the weeks in 2023 (left) and 2024 (right). This comparison demonstrates a notable increase in contrail detections in 2024. Specifically, when aggregating the total number of contrails detected throughout the week, we observe a 41.03% increase in 2024 compared to the same week in 2023.

Changes in Contrail Warming

We now analyze the changes in RF to determine if the observed increase in contrail coverage is associated with a rise in total warming. The cumulative RF and CR values shown in Figure 3 were derived by aggregating the net RF of all clouds (top) and the net CRF of all contrails detected (bottom) across the entire field of view of the MSG satellites. The values for each time of day are averaged over the seven-day period.

The analysis indicates that both total cloud RF and CRF exhibited more extreme values in 2024 compared to 2023, with increased cooling (more negative RF/CRF values) during the day and heightened warming (more positive RF/CRF values) at night. Given the observed cooling effect during the day and the warming effect at night, we aggregate the total forcing across all times of day to estimate the overall warming effect during one complete day (see Table 11).

The data reveals two key findings: (1) the combined effect of nighttime warming and daytime cooling results in a net warming effect in both years, and (2) there is an increase in both cloud and contrail warming during the week of 2024. Specifically, total cloud RF rose by 19.51%, while total CRF surged by 128.7%, indicating that the increased warming in 2024 is primarily driven by a rise in contrail coverage.

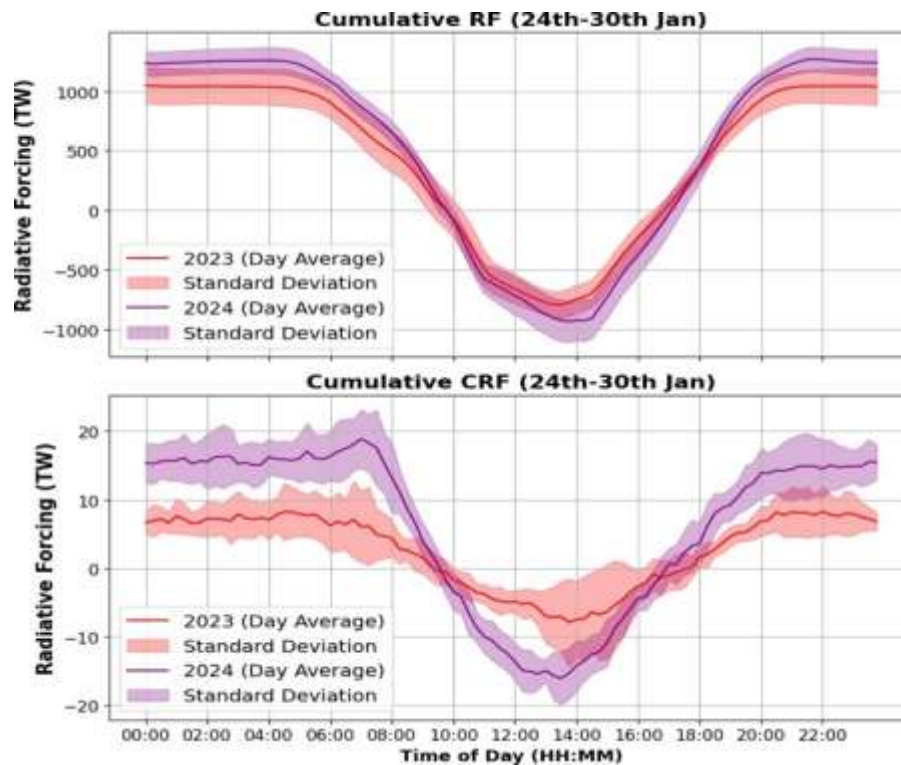


Figure 2: Comparison of the cumulative Radiative Forcing (RF) curves (top) and the cumulative contrail RF curves (bottom) between the week of the 24th-30th of January of 2023 and the corresponding week in 2024.

We now analyze the changes in RF to determine if the observed increase in contrail coverage is associated with a rise in total warming. The cumulative RF and CRF values shown in Figure 3 were derived by aggregating the net RF of all clouds (top) and the net CRF of all contrails detected (bottom) across the entire field of view of the MSG satellites. The values for each time of day are averaged over the seven-day period.

The analysis indicates that both total cloud RF and CRF exhibited more extreme values in 2024 compared to 2023, with increased cooling (more negative RF/CRF values) during the day and heightened warming (more positive RF/CRF values) at night. Given the observed cooling effect during the day and the warming effect at night, we aggregate the total forcing across all times of day to estimate the overall warming effect during one complete day (see Table 11).

The data reveals two key findings: (1) the combined effect of nighttime warming and daytime cooling results in a net warming effect in both years, and (2) there is an increase in both cloud and contrail warming during the week of 2024. Specifically, total cloud RF rose by 19.51%, while total CRF surged by 128.7%, indicating that the increased warming in 2024 is primarily driven by a rise in contrail coverage.

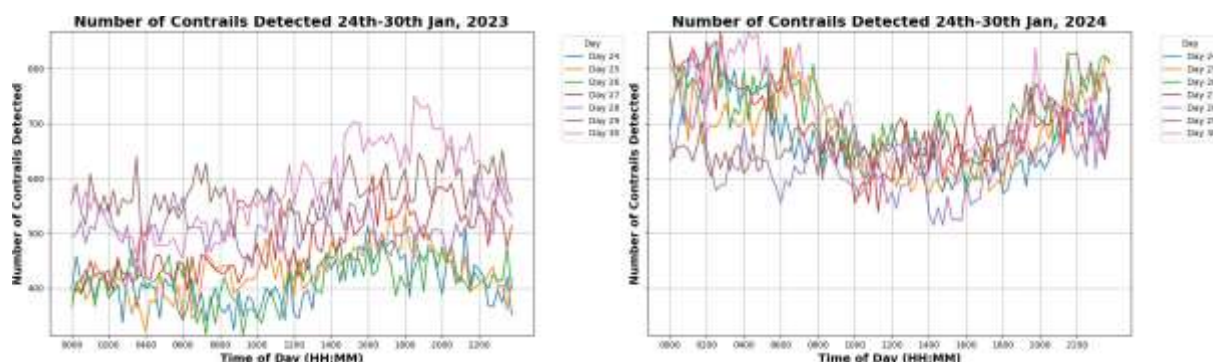


Figure 3: Comparison of the number of contrails detected across the entire field of view of Meteosat satellites in the week of the 24th-30th of January of 2023, and the same week in 2024, encompassing all times of day.

Table 12: Average daily RF and CRF values for the 24-30th Jan, 2023 and the 24-30th Jan, 2024

Year	Cumulative RF	Cumulative CRF
2023	41, 000 TW	258 TW
2024	49, 000 TW	590 TW
$\Delta(2024 - 2023)$	19.51%	128.7%

4.2.3.2 Analysis of Individual Contrails

For each contrail detected over the two-week period, we assessed the correlation between RF values and factors such as contrail parameters, size, surface temperatures, and lighting conditions (see Figure 4). The analysis excludes underlying water clouds, as they exhibited negligible correlation with the CRF values.

The key insights of this analysis include:

1. The Net RF (W), which sums reflected and emitted radiation across all contrail points, appears independent of contrail size. Larger contrails exhibit both stronger warming and cooling (under shortwave radiation), balancing each other out. This strong correlation with shortwave forcing suggests that nighttime contrails, without cooling, have the greatest warming impact, regardless of size. In other words, this means that a long, thick contrail formed during the day has a smaller warming impact compared to a small, thin contrail formed at night. Figure 6 illustrates a large outbreak with a net cooling effect during the day, which, after sunset, transitions to a few small warming contrails.
2. The Average RF (W/m²), which measures radiation at a single point of a contrail, is strongly correlated with the zenith angle, indicating that lighting conditions primarily determine whether a contrail warms the Earth. This is significant, as it suggests that most contrails appear to have a cooling effect during daytime. Figure 5 illustrates the evolution of the contrail warming effect as daylight decreases. Average RF is also influenced, though to a lesser extent, by cloud altitude (CTH), with higher altitude contrails generally trapping more heat.



Figure 4: Correlation Matrix between contrail cloud parameters (COD, CTH, CER), zenith angles, surface temperatures (ST), RF values, and contrail sizes of all contrails detected over the two-week period.

Warming Impact of Individual Contrails

Given that lighting conditions significantly influence a contrail's overall effect, we analyze each contrail's warming contribution by size, distinguishing between daytime and nighttime contrails. The top row bar plots in Figure 4 show that daytime contrails generally produce a cooling effect. Even the largest daytime contrails typically don't warm as much as an average-sized nighttime contrail. Given that most daytime contrails exhibit a cooling effect, a significant factor contributing to the overall warming effect throughout the day is the higher proportion of nighttime contrails. Only 38 % of all contrails detected over the 14-day period occurred during daylight, with the remaining 62% were observed at night.

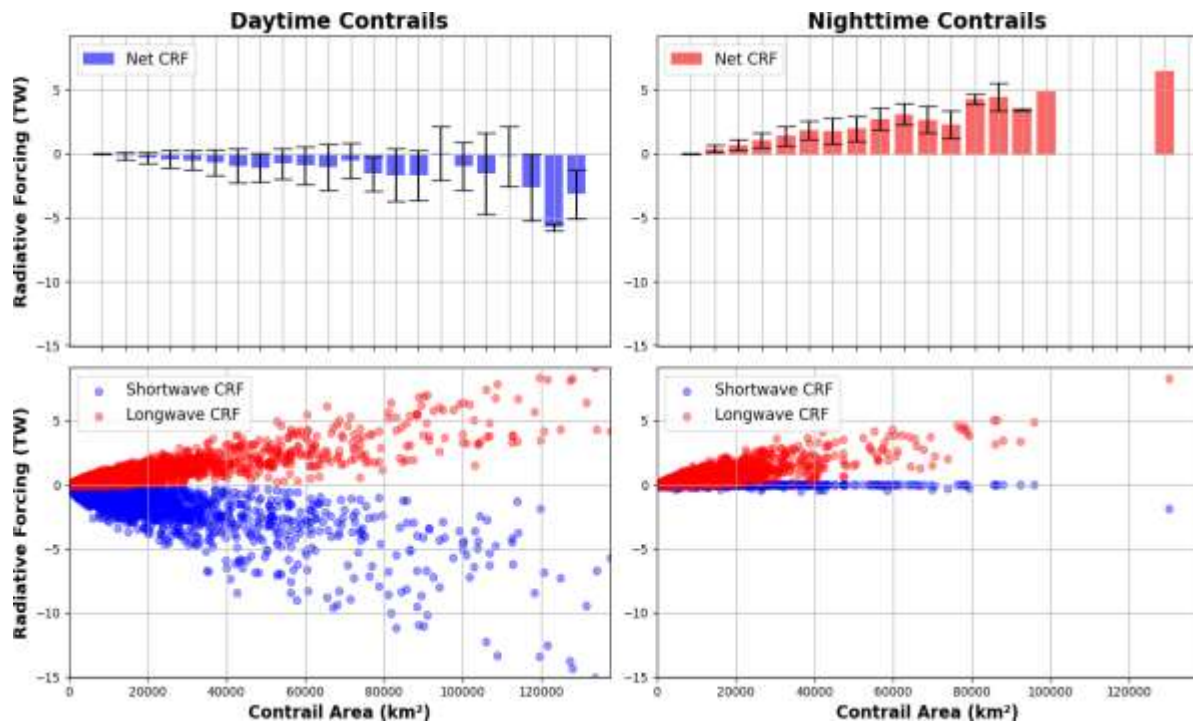


Figure 5: The top row displays the Radiative Forcing (RF) of contrail features by size, measured in terawatts, with separate panels for daytime (left) and nighttime (right). The bottom row shows the Short Wave (SW) and Long Wave (LW) RF components for daytime (left) and nighttime (right), which were combined to derive the RF values presented in the top row.

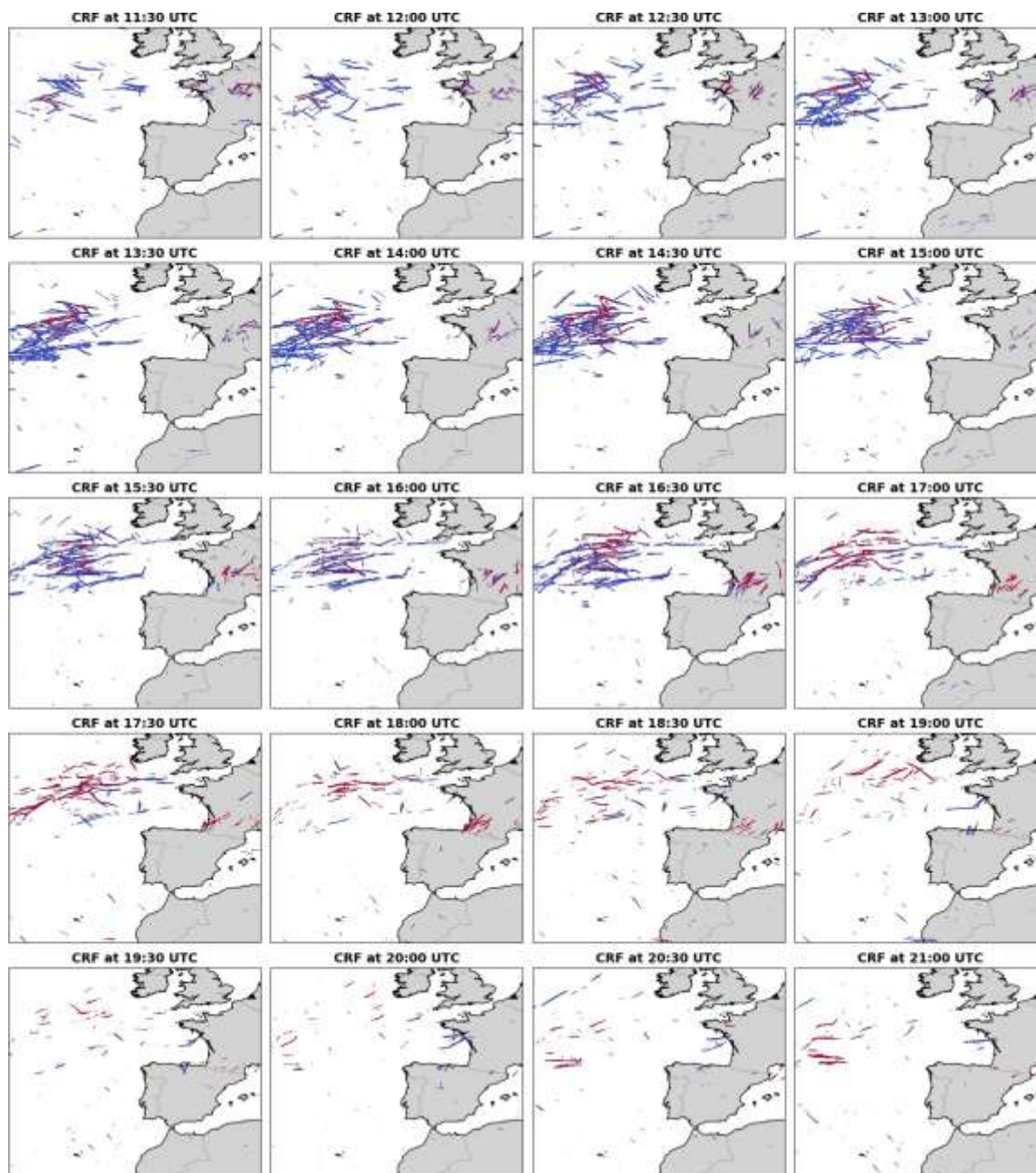


Figure 6: Contrail Outbreak over the Atlantic Ocean spanning a 9.5-hour period on January 30, 2023. The red colors indicate positive Radiative Forcing values (warming) while the blues indicate negative values (cooling).

4.3 Experiment 1: Satellite-Based Quantification of Contrail Radiative Forcing over Europe and ISSR forecasting: a Full Year 2023 analysis.

We present in this section the integrated experiment for one year of data. The experiment was conducted over the entire year 2023 (including also January 2024), encompassing 24-hour periods each day.

4.3.1 DATA

The data used in this work was obtained from the SEVIRI onboard the MSG satellites of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). In particular, the data from the MSG-3 (Meteosat-10) and MSG-4 (Meteosat-11) satellites have been used for this work. Positioned in geostationary orbit at 0° longitude, about 36,000 kilometres above Earth, these satellites provide spectral information across 11 channels, including visible and infrared regions. Observations are captured every 15 minutes with a spatial resolution of approximately 3x3 km² at the subsatellite point, with a pixel size that grows as one looks further from the equator. The SEVIRI level 1b data serves two main purposes: (a) generating the false-color RGB images for contrail detection models, and (b) providing input to the Optical Cloud Analysis (OCA) system, which we use for the physical characterization of the clouds. These physical parameters serve as the input to the radiative forcing estimation. The specifics of this retrieved information are as follows:

(a) Ash Composite: This product generates false-color RGB images to enhance contrail visibility by combining several MSG thermal infrared (IR) bands. It is composed of the red Brightness Temperature (BT) difference IR12μm-IR10.8μm to highlight contrails by their higher transmissibility compared to natural cirrus, the green BT difference IR10.8μm-IR8.7μm to differentiate cloud phases, and the blue BT IR10.8μm to accentuate contrails by leveraging their colder temperatures relative to surrounding features. See [2] for the definition. The Ash RGB composite was also used in for detecting contrails.

(b) Cloud parameters: Cloud state parameters are characterized by cloud phase (CP), cloud top pressure (CTP), cloud optical thickness (COT), and cloud effective radius (CER). These parameters are obtained from the OCA of EUMETSAT, which employs the Optimal Estimation (OE) method along with SEVIRI spectral measurements simultaneously. The cloud information obtained can be separated into upper layer and lower layer clouds. The upper layer consists of ice clouds, and includes values both for cases where the ice clouds are alone and where they coexist with underlying water clouds. The lower layer shows data only when a water cloud is present beneath the upper ice layer.

(c) Forecast Data: We use the skin temperature from Numerical Weather Prediction (NWP) data obtained from the European Centre for Medium Range Weather Forecasts (ECMWF).

(d) Land Cover Data: We use the MODIS L3 500m Land Cover dataset MCD12Q1 v061.

4.3.2 Methods

4.3.2.1 Radiative Forcing Calculations

The methodology employed to quantify the RF of contrail cirrus detectable by a geostationary satellite involves three key steps: First, a contrail detection model identifies the locations of all visible contrails within the scene. Second, the RF of all clouds present in the image is calculated. Third, the results from these two steps are intersected to determine the RF specifically attributable to the detected contrails. See Dimitropoulou et al, 2025 [RD23]

The radiative forcing estimation was derived using multidimensional interpolation on pre-built Lookup Tables (LUTs). These LUTs were constructed using the libRadtran radiative transfer library to simulate both shortwave and longwave radiative forcing for various combinations of thin to semi-transparent ice cloud parameters, along with other relevant factors like solar zenith angles, sea or land surface temperature, surface type and the presence of an underlying water cloud. By interpolating the simulated shortwave and longwave RFs from the LUTs as a function of the surface type, geometry, skin temperature, and OCA parameters for each pixel, a detailed and location-specific assessment of radiative forcing is provided. This approach streamlines the radiative forcing estimation process, removing the need to repeatedly run time-consuming radiative transfer simulations in future large-scale analyses. Figure 3.1 provides an example of the shortwave, longwave, and net radiative forcing estimates obtained through this process. We use the sign convention of downward flux, so that a positive value of the RF represents a warming effect.

4.3.2.2 Contrail Detection

For this experiment, we utilized a single-frame U-Netbased network previously trained on the OpenContrails Dataset, which comprises 22,000 images captured by the Geostationary Operational Environmental Satellites 16 Series (GOES-16) between April 2019 and April 2020 in different locations of North and South America. This dataset has been selected for training due to its substantial number of labelled scenes (55%), necessary for capturing all the variability of contrail features.

Network Architectural Details

The architecture employed is a hybrid neural network that combines a transformer-based encoder with a convolutional decoder. Because of the limitations in the computational resources, the encoder is a lightweight variant of the CoaT (Co-Scale Conv-Attentional Image Transformers) model, called CoaT-Lite Mini. This variant is optimized for efficient image processing, avoiding parallel blocks and incorporating a reduced channel depth in each layer. The model processes images through four sequential blocks, where feature maps are downsampled and converted into image tokens. These tokens are analyzed using convolutional operations for local pattern extraction and self-attention for capturing image interpart relationships. The output from each block is reshaped into a 2D feature map and forwarded to the next block and decoder via skip connections. The decoder employs sequential convolutional blocks with upsampling, producing feature maps at three different resolutions. The three feature maps are combined using a Feature Pyramid Network (FPN). The output is then regularized with a dropout layer, applying a 0.5 probability to deactivate weights to prevent overfitting to the training data. Finally, the features are upsampled to the original image size, with the number of channels reduced to one. Each pixel in this single-channel output mask represents the probability of being part of a contrail. The optimization of the weights of the network was performed using the AdamW optimizer over 30 epochs, minimizing a convex surrogate of the Dice loss function. Transfer

learning was used to initialize the encoder's weights with those from a CoaT network pretrained on ImageNet.

Domain Adaptation

The trained network was used to detect contrails in MSG Ash RGB images, which have different characteristics from the original training data, especially in terms of geographical coverage and image resolution. MSG images cover Europe, Africa, and portions of the Atlantic Ocean, with a maximum thermal infrared resolution of 3x3 km². In contrast, the Ash RGB training images from GOES-16 focus on the United States, the Atlantic Ocean, South America, and the Caribbean, offering a finer resolution of 2x2 km² at nadir. Although the difference in geographical coverage is not expected to have an impact, the disparity in resolution could affect detection accuracy. To address this, the resolution of MSG images was adjusted using bilinear interpolation, simulating a 2x2 km² resolution at nadir. Given the large size of an MSG scene, a sliding window of 256x256 pixels was applied to divide the images into smaller overlapping sections. The detector was applied to each section, with results from overlapping areas combined to ensure accurate identification of objects partially visible across sections. Finally, after aggregating all the detections produced by the sliding window, the contrail mask for the entire scene was transformed back to the original resolution, preserving the true sizes of the segmented contrails.

4.3.2.3 ISSR Forecasting

We explored architectures that combine convolutional layers (for spatial feature extraction) with recurrent or attention-based mechanisms (for temporal sequencing).

- CNN+LSTM and ConvLSTM:** One of the first spatio-temporal models we explored was a CNN+LSTM, where a convolutional neural network (CNN) extracts spatial features at each time step, and an LSTM models their temporal evolution. This setup allowed the model to process spatial structures dynamically but still treated each spatial frame separately, limiting its ability to capture motion continuity across frames. To address this, we implemented a Convolutional LSTM (ConvLSTM), which applies convolutions inside the LSTM cell. This enabled the model to learn how spatial structures evolve over time. The ConvLSTM architecture was particularly effective at tracking contrail movement. For example, it learned that a contrail cloud patch at time t might drift eastward at $t+1$ if the wind is from the west. The ConvLSTM model output at each time step was a predicted contrail mask or RF field for the next step. This model served as a baseline for spatial prediction performance, capturing local spatial dependencies and temporal continuity. However, it struggled with long-range dependencies, prompting the introduction of transformer-based models.
- Vision Transformer (ViT) for Spatial Feature Learning:** To further enhance spatial pattern recognition, we introduced Vision Transformer (ViT). Unlike CNNs, which rely on local convolutional filters, ViT treats an image (here, a contrail + weather map) as a sequence of patches and applies self-attention across all patches, enabling it to learn global spatial relationships. ViT was particularly effective at:
 1. Detecting contrail structures and clusters
 2. Identifying meteorological features influencing contrail formation
 3. Recognizing spatial dependencies over large distances

However, ViT lacks explicit temporal modelling, meaning it processes each time frame independently without learning temporal evolution. To address this, we explored two strategies: Time-Distributed ViT – Extracting ViT-based embeddings for each frame, then

applying a temporal model (LSTM or Transformer) to capture sequence dependencies. Geo-Spatiotemporal Attention Network (GeoSTANet) – A hybrid model integrating attention across both space and time in a single architecture.

- **Geo-Spatiotemporal Attention Network (GeoSTANet):** To overcome the limitations of CNN+LSTM, ConvLSTM, and ViT, we developed GeoSTANet, a custom spatio-temporal transformer that captures both:
 1. Local spatial features (using convolutional layers for feature extraction)
 2. Global spatial dependencies (using transformer-based attention)
 3. Temporal evolution (using a cross-frame attention mechanism)

GeoSTANet learns motion dynamics by attending to contrail features at t_0 and tracking their evolution at $t_1, t_2, \dots$. This architecture significantly outperformed ConvLSTM in predicting persistent contrails and long-term RF effects. Key advantages of GeoSTANet:

1. Captures both spatial and temporal relationships
 2. Learns contrail persistence and movement patterns
 3. Handles irregular motion (contrail drift, spread, and dissipation)
 4. Higher recall than ConvLSTM and ViT alone
- **Incorporating U-Net and Segmentation Models:** We explored U-Net-inspired architectures for purely spatial predictions, treating contrail forecasting as an image segmentation problem. U-Net processes meteorological and flight data to generate contrail mask predictions for a given time step. However, since our task also requires predicting contrail evolution, we explicitly incorporate the temporal dimension into our spatio-temporal models. In our approach, U-Net processes spatial features, while a time-distributed framework ensures that each frame is processed sequentially before applying a final temporal aggregation step. This integration captures temporal dependencies beyond static segmentation, enabling a more accurate representation of contrail formation and evolution.
 - **Handling Imbalanced Outputs:** A notable challenge is that contrail occurrence is a rare event relative to no-contrail, leading to a class imbalance in the data (most pixels are clear sky/no contrail). This can bias models to trivially predict “no contrail everywhere” and achieve high overall accuracy but zero usefulness. To counter this, we employed strategies such as focal loss and class-weighting in the loss function. Focal loss down-weights easy negatives and focuses the training on the hard examples (i.e., the minority contrail pixels) by dynamically scaling the cross-entropy loss for contrail vs non-contrail. The cross-entropy loss is scaled using focal loss with a gamma parameter of 2, ensuring higher weight is given to hard-to-classify contrail pixels while reducing the influence of easy negative samples. We also monitored metrics like Recall and Balanced Accuracy (the average of recall for contrail and no-contrail) to ensure the model is detecting contrails, not just maximizing overall accuracy. For the RF regression, the imbalance is less direct, but the magnitude of RF values can vary widely (many near zero). We normalized the RF outputs and, in some cases, trained the model to predict contrail occurrence and RF jointly (multi-task learning), so that the detection of contrails and estimation of their RF inform each other.

4.3.2.4 Dashboard Visualization

We have incorporated the results in a visualization dashboard (Figure 7).



Figure 7: E-CONTRAIL visualization Dashboard.

4.3.3 Results

Figure 8 shows the average number of contrails detected across different times of the day and months. Key findings include:

- **Months with the Highest Number of Contrail Detections:** April and May exhibited the highest frequency of contrail occurrences
- **Peak Detection Times:** 07:30 UTC and 15:30 UTC.

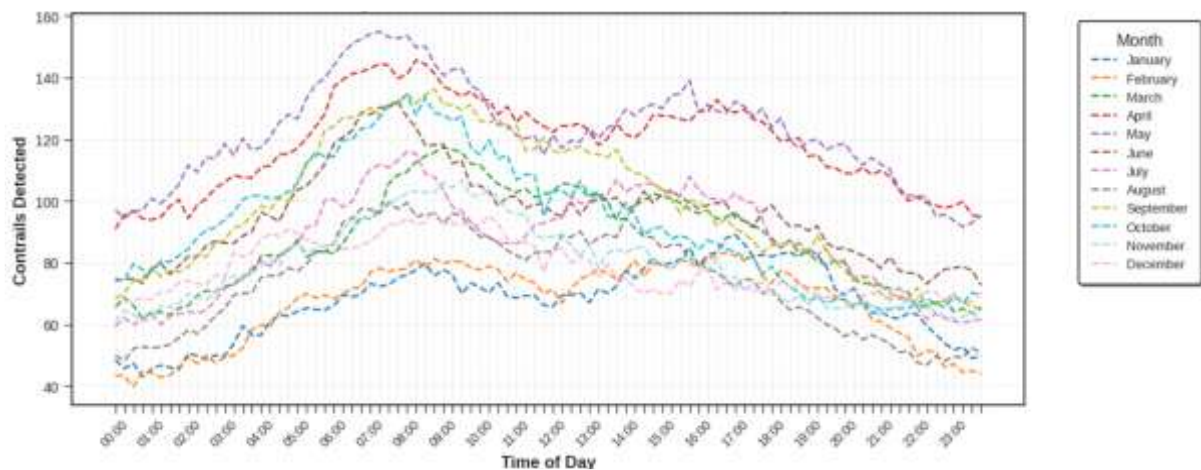


Figure 8: Average number of contrails detected across different times of the day (preliminary results).

Figure 9 presents the estimation of RF across different hours of the day and different months. Table 13 presents the monthly median of RF values (W/m²) for Ice Clouds and Contrails in year 2023. Key findings include:

- **There is an overall net cooling effect during the central hours of the day.**

- There is an overall net warming effect during the months of August to April, with more intense values in November to February.

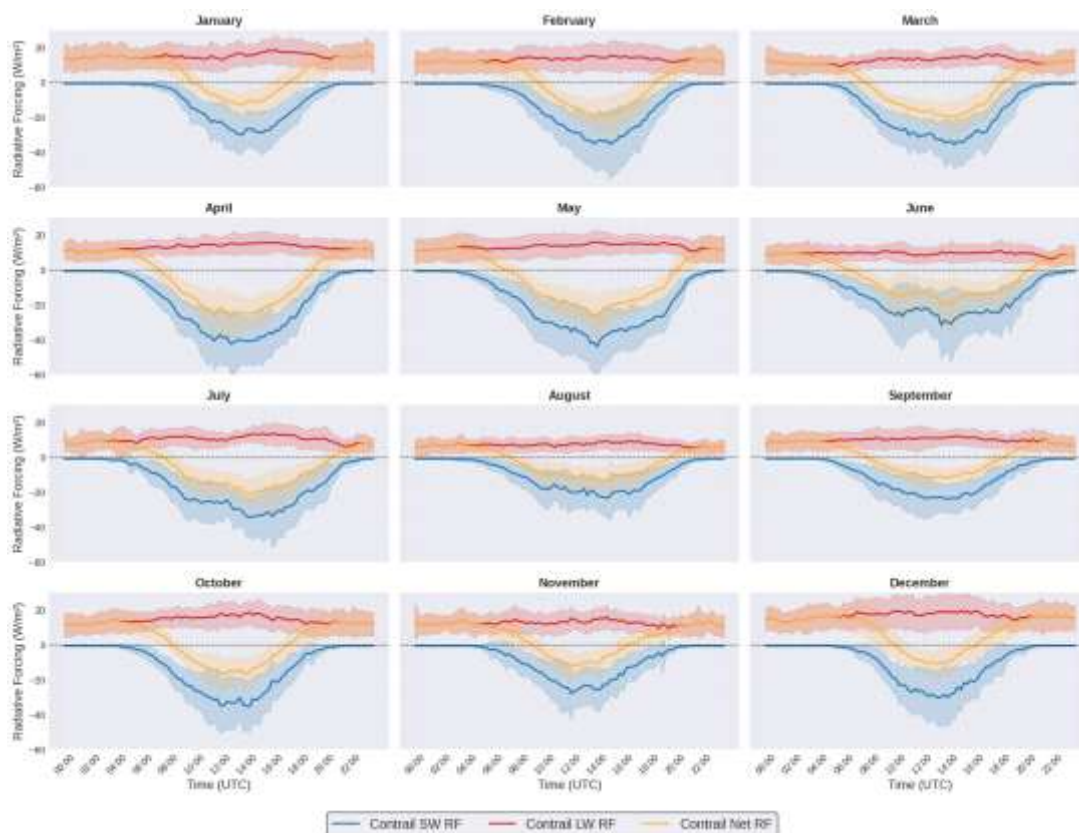


Figure 9: RF estimation over 2023 (preliminary results)

2023												
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Ice Clouds	5.21	3.46	1.38	0.27	-0.65	-1.04	-0.44	0.47	2.83	4.63	5.1	7.18
Contrails	7.31	4.2	2.8	0.6	-3.02	-3.02	-2.03	-0.63	1.78	4.8	5.6	9.04

Table 2: Monthly median RF values (W/m^2) for Ice Clouds and Contrails (Jan - Dec 2023). Ice Clouds refer to high-altitude clouds (above 300 hPa), including both multi-layered clouds (with underlying water clouds) and those above clear skies, as classified by OCA. The Contrails row includes all contrails detected by the Machine Learning model.

Table 13: Monthly median of RF values (W/m^2) for Ice Clouds and Contrails in year 2023.

Figure 10 presents an example of the prediction of ISSR regions, including SW and LW Radiative Forcing (left) executed 12 hours ahead of time and the comparison with a ground truth based on ERA5 data (right). Figure 11 and Figure 12 show the metrics for the prediction of ISSR region (binary classification) and RF prediction (Regression problem).

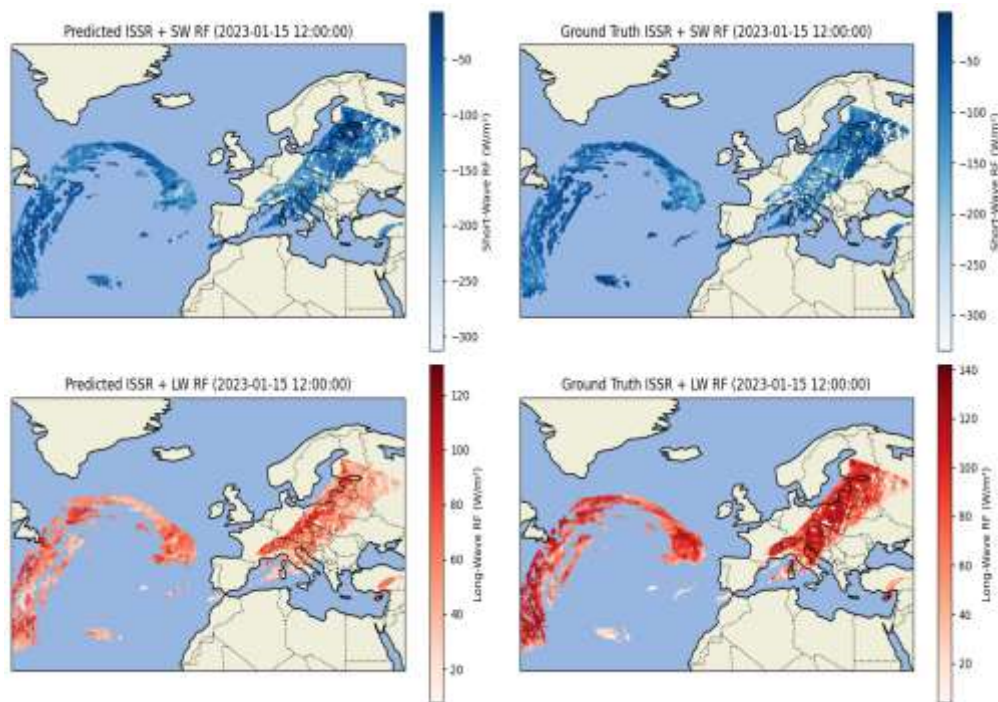


Figure 10: ISSR prediction Vs Ground truth.

Metric	Value
MAE	29.36
R Score	80.1%
Dice Score	80.1%
Variance Score	81.45%

Figure 11: ISSR prediction metrics

Long Wave RF		Short Wave RF	
Predicted Range	Ground truth Range	Predicted Range	Ground truth Range
4.01 to 142.64	3.71 to 149.01	-442.35 to 0.0	-486.61 to 0.0

Figure 12: ISSR & RF prediction metrics

4.4 Confidence in validation results

We provide insights on the limitations, quality, and significance of the validation results on an objective-wise manner

4.4.1 OBJ-01.1-TRL1-ERP-001 results

Six neural network architectures (including transformer-based U-Net variants and convolutional models like Mask-RCNN and YOLO11) were trained and evaluated on the OpenContrails dataset to segment young linear contrails (approximately 30 minutes to 4 hours old) in infrared imagery captured by the GOES-16 Advanced Baseline Imager (ABI) [RD19]. The performance of individual models is compared, with additional metrics derived from ensemble predictions to highlight improvements in accuracy and robustness. A qualitative analysis of individual images identifies specific scenarios where model performance is lower and general segmentation challenges. Furthermore, the temporal consistency of model predictions is assessed across time-series data.

4.4.1.1 Metrics

Table 14: The formulas for the averaged and global metrics are computed in terms of True Positives (TP), False Positives (FP), and False Negatives (FN) across a dataset consisting of N samples.

The $\gamma\beta$ -BSD Metric

In this work, we introduce a novel metric, the $\gamma\beta$ -Boundary Soft Dice (BSD), designed to more accurately evaluate the performance of segmentation models in identifying target instances. The metric is robust to slight misalignments between predicted and ground-truth features, which may arise from inherent errors or inaccuracies in the ground truths. By incorporating a soft constraint on boundary segmentation, the proposed metric provides a more reliable assessment that accounts for these discrepancies.

The approach we employ to impose soft constraints on the boundaries involves redefining the sets of pixels classified as false positives and false negatives for each prediction as follows:

$$FP_k^\gamma(y, \hat{y}_\alpha) = \sum_{i=1}^H \sum_{j=1}^W R_{T_k}^\gamma(\hat{y}_\alpha(i, j)^k) \cdot (1 - y(i, j)^k)$$

$$FN_k^\beta(y, \hat{y}_\alpha) = \sum_{i=1}^H \sum_{j=1}^W (1 - \hat{y}_\alpha(i, j)^k) \cdot R_{T_k}^\beta(y(i, j)^k)$$

The global metric formula is the defined as

$$\gamma\beta - BSD = \frac{2 \cdot TP}{2 \cdot TP + FP^\gamma + FN^\beta}$$

while the image metric formula is

$$\gamma\beta - BSD = \frac{1}{N} \sum_{k=1}^N \frac{2 \cdot TP_k}{2 \cdot TP_k + FP_k^\gamma + FN_k^\beta}$$

4.4.1.2 Results

Table 15: The 'GMax' columns utilize a global threshold that maximizes Dice Score (DS) across masks, while the 'APT' columns apply the adaptive probability thresholding technique presented in this work. The error margin values utilized for the $\gamma\beta$ -BDS metric are $\gamma = \beta = 2$.

Global Metrics across the OpenContrails validation set											
Single Models	Precision (%)		Recall (%)		PR-AuC (%)	DS (%)		IoU (%)		$\gamma\beta$ -BSD (%)	
	GMax	APT	GMax	APT	-	GMax	APT	GMax	APT	GMax	APT
Mask-RCNN	16.93	25.61	52.53	52.53	20.00	25.62	26.01	14.69	14.94	26.75	29.34
YOLO11	23.40	24.15	44.05	47.52	35.89	30.15	32.02	17.75	19.04	50.88	42.51
CoaT U-Net	59.11	61.64	70.90	65.27	63.49	62.31	64.52	45.25	47.62	75.72	75.25
NeXtViT U-Net	66.89	68.17	66.55	67.66	67.95	66.72	67.92	50.03	51.42	82.71	80.79
CoaT U-LSTM	67.92	68.40	66.35	67.35	67.91	67.14	68.08	50.53	51.60	83.57	80.64
NeXtViT U-LSTM	65.70	66.94	68.37	69.57	68.28	67.0	68.23	50.38	51.78	84.61	82.23

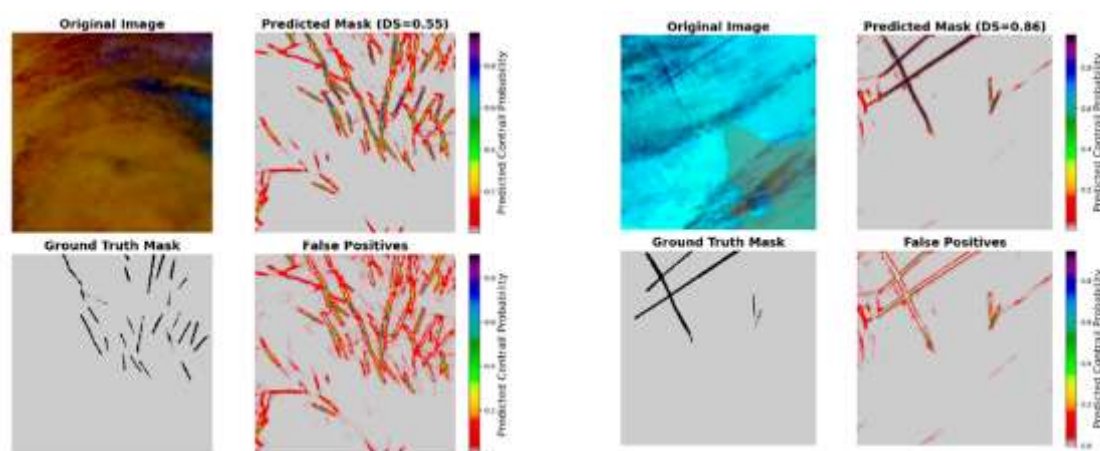


Figure 13: Contrail detection examples

4.4.2 OBJ-01.2-TRL1-ERP-001 results

We compare the contrails detected using the E-CONTRAIL detection algorithm with the ones obtained by the state-of-the-art physical model COCIP [RD8]. The analysis is conducted in one week of data, ranging 1st-8th February 2025 on a geographical coverage of the full field of view of the Meteosat [RD22].

Figure 14 shows the Lagrangian Contrail Model (COCIP) estimation of contrails and its associated RF. It is computed using Numerical Weather Forecasts and ADSB data. Figure 15 shows the E-CONTRAIL contrail detection algorithm combined with the RF estimation algorithm.

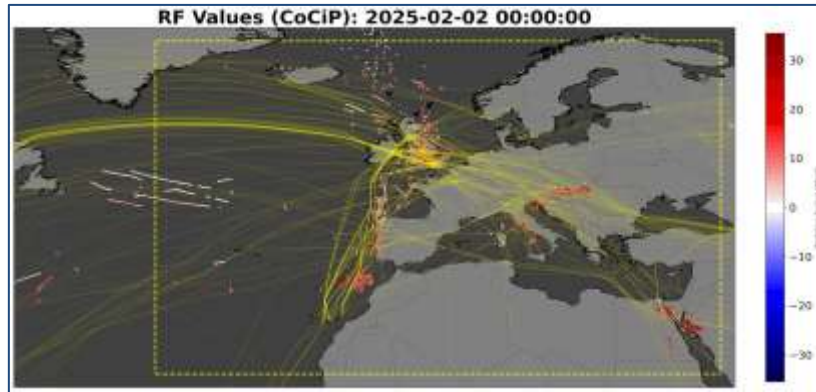


Figure 14: Lagrangian Contrail Model (COCIP).

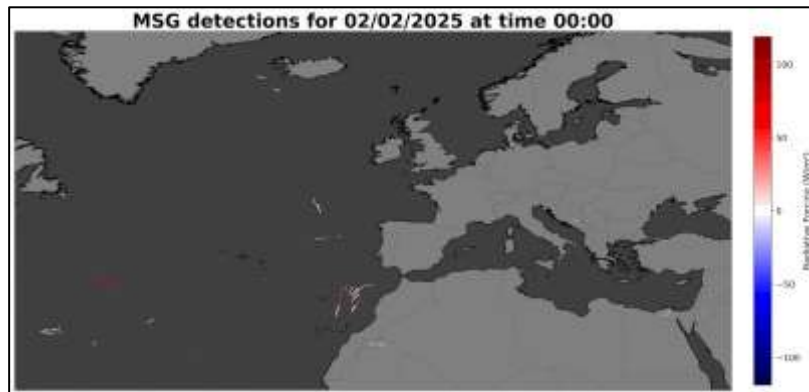


Figure 15: the E-CONTRAIL contrail detection and RF estimation algorithm.

Qualitative analysis: Figure 16 shows the Contrail detection above water clouds vs clear sky conditions. It can be observed that:

- The spatial location of contrail outbreaks in satellite observations generally matches regions containing CoCiP detections.
- Individual CoCiP contrails differ from observations in terms of frequency, orientation, and persistence.
- With CoCiP 66 % of simulated contrails intersect low-level water clouds (see also Figure 17)

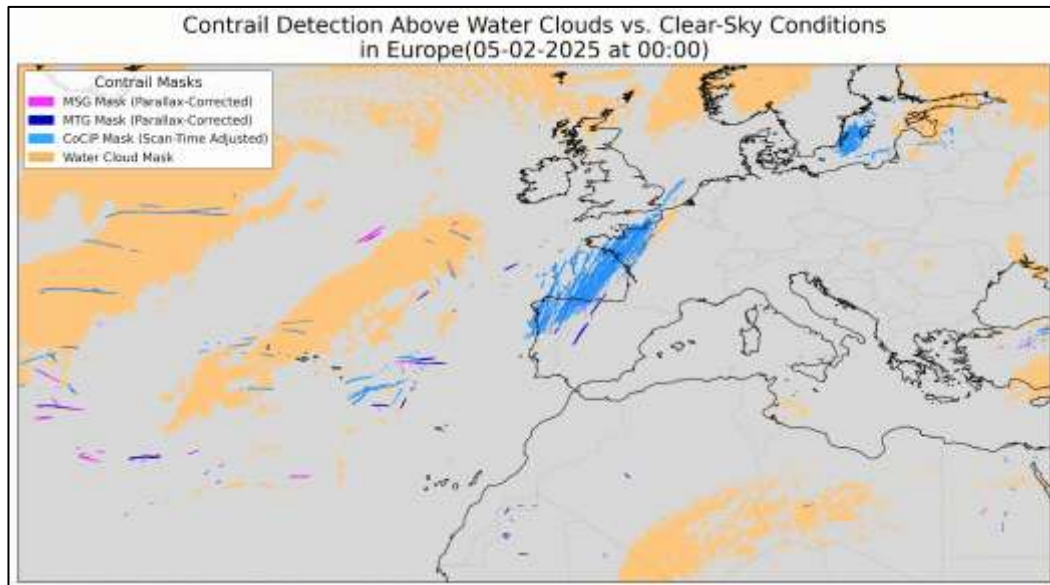


Figure 16: Contrail detection above water clouds vs clear sky conditions-

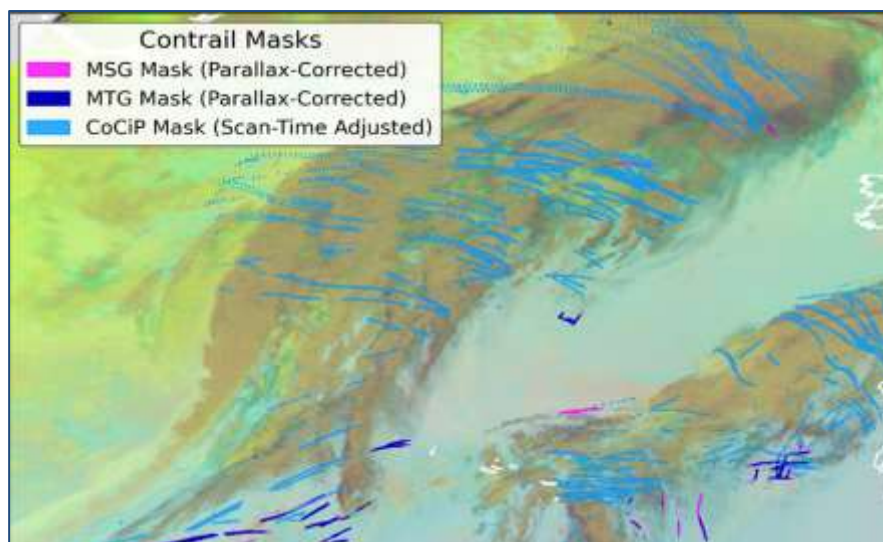


Figure 17: Contrail CoCiP mask over clouds

Quantitative analysis: Figure 18, Figure 19, and Figure 20 show the quantitative estimations of RF attributed to contrails using the E-CONTRAIL algorithm on MSG. It also compares the results with estimations on MTG and CoCiP. The following conclusions can be drawn:

- The net mean radiative forcing indicates an overall contrail warming effect.
- Contrail–cloud systems reduce both cooling and warming effects compared to clear-sky contrails [RD20]
- Overall cooling occurred for ~6 hours contrail-cloud systems, and for ~9 hours in clear-sky conditions.
- The reduced SW and LW values in CoCiP may be attributed to a rise in contrail formation above water clouds and cirrus cloud coverage [RD21]

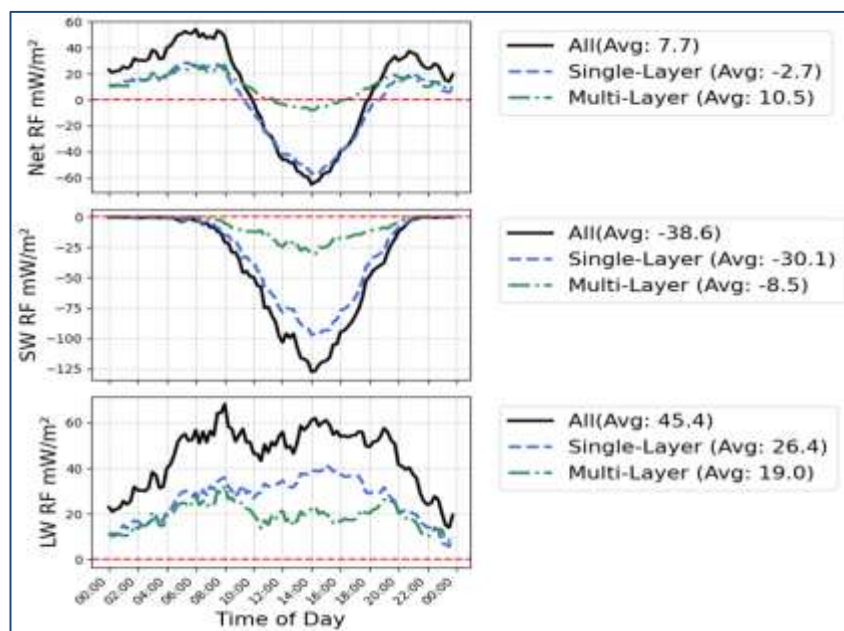


Figure 18: E-CONTRAIL estimation of RF during the week of Feb 1st - 8th 2025

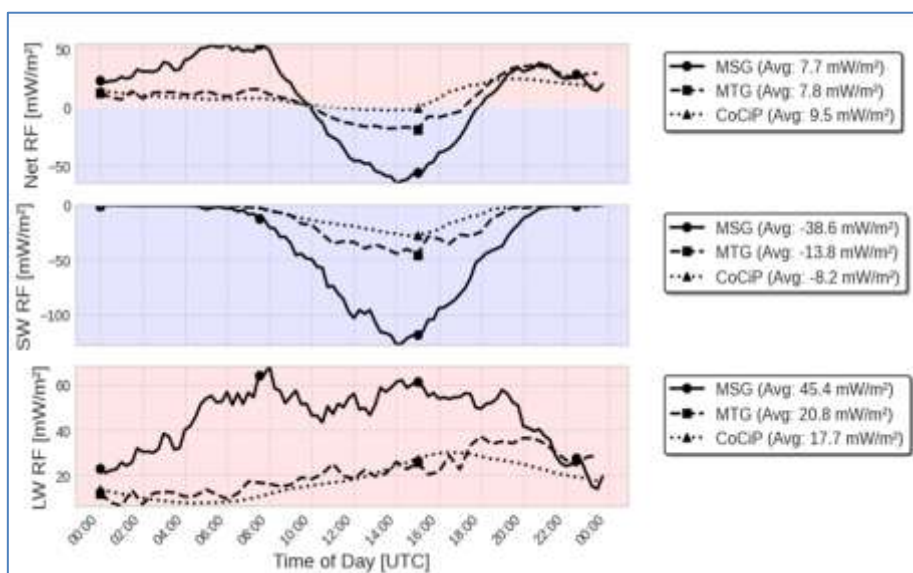


Figure 19: Comparison of E-CONTRAIL estimation (MSG) of RF w.r.t.

MTG and CoCiP during the week of Feb 1st - 8th 2025

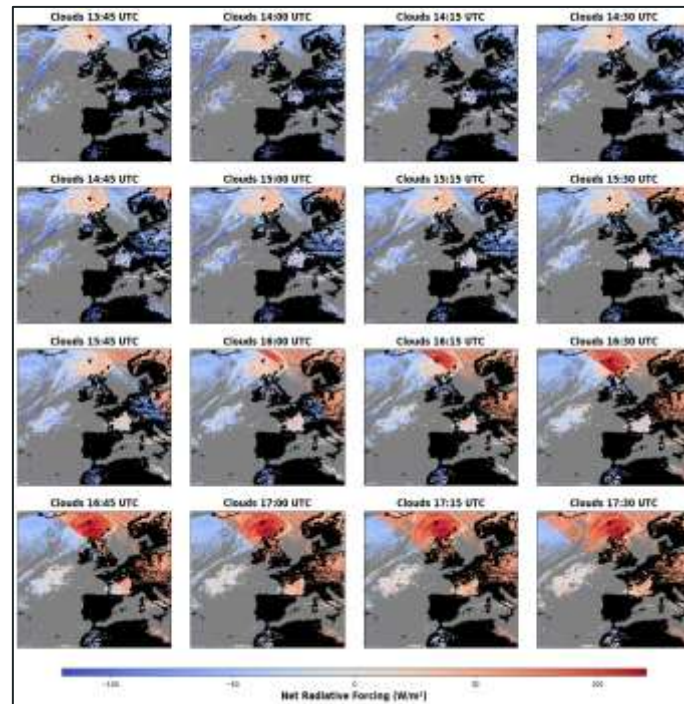


Figure 20: Example of a contrail outbreak

4.4.3 OBJ-02-TRL1-ERP-001 results

4.4.3.1 Validation of the Rapid Contrail-RF Estimation Approach

The accuracy and reliability of the Rapid Contrail-RF Estimation Approach in constructing RF maps for contrails have been investigated through four different validation exercises, presented in the following subsections [RD23]. These exercises focus on different aspects of the methodology. First, we evaluate the choice of using a single atmospheric vertical profile in the Radiative Transfer (RT) simulations. Next, by performing a small subset of RT simulations, we investigate the impact of selecting a certain ice cloud parameterization scheme. Additionally, we evaluate the impact of using Cloud Top Height (CTH) values estimated by a single atmospheric vertical profile on the RF estimations. Finally, we perform a comparison between the flux maps for contrails and polar-orbiting satellite observations.

Impact of vertical temperature profile on radiative transfer calculations

The core component of the Rapid Contrail-RF Estimation Approach is the construction of the ice cloud RF LUTs and their 370 merging with the re-gridded geostationary maps. As presented in previous WP2 deliverables, the atmospheric temperature vertical profile used in the RT simulations remains constant and corresponds to the U.S. Standard Atmosphere. To assess the validity of this choice and estimate the uncertainty associated with using a single constant temperature vertical profile, randomly selected pixels from the zoomed geographic regions of each day-containing contrails above land, ocean, and water clouds (i.e., multi-layered)-covering day- and night-time conditions were chosen as the sample of this investigation.

For these selected pixels, RT simulations were performed using the ERA5 vertical temperature profile from ECMWF as the input atmospheric profile. These profiles were also used to estimate CTH and wCTH (only in the presence of a water cloud). Additionally, for each pixel, the actual CER and COT

values from the OCA product were used, along with the real SZA. In the presence of a water cloud, we use the wCOT value from the OCA product.

In Figure 21, for each scene scenario, we present the comparison results between the RF values coming from the LUTs (RFUSstandard) and the RF values calculated by using the actual atmospheric and cloud conditions (RFERA5) per selected pixel in the SW and LW wavelength ranges, separately. As it can be seen, for all the scene scenarios in the SW wavelength range, overall good agreement is found with the correlation coefficient and slope values being close to unity, except for a few comparison points.

Table 16 provides some statistics for the two different methodologies followed in this Section per wavelength and scene scenario. In the SW wavelength range, the use of LUTs instead of real-time RT simulations per pixel can lead to RMS error equal to 6.13 W/m², 10.76 W/m², and 11.99 W/m² above land, ocean, and water cloud, respectively. The comparisons in the LW wavelength range (see Figure 21) reveal an overall good agreement with correlation coefficient values being around 1.00 and slope values in the range of 0.95 - 0.97. In contrast to the comparison in the SW, in the LW, we observe that a larger number of points appears to be scattered around the 1:1 line. This finding means that the RT simulations in the LW wavelength range are more sensitive in the choice of the atmospheric temperature vertical profile. The use of LUTs in the LW wavelength range leads to RMS error values of the same order of magnitude for the three scene scenarios. When focusing on the SW and LW RMS error percentage, we find that the largest values for both wavelength ranges are observed for the scene scenario of an ice cloud above a water cloud (multi-layered).

To explain the scattered points around the 1:1 line in the subplots of Figure 21,, we focus on the points with an RMS error value larger than the mean RMS error value plus two times the standard deviation of the RMS error. For these points, we first investigated whether there is a correlation between the large discrepancies in the two RF datasets and the differences between the values of each actual cloud parameter and the closest values used during the multi-dimensional interpolations in the LUTs. The comparison results showed no correlation.

Additionally, for these points, we examine the corresponding ECMWF vertical profiles used in the RTM simulations. Figure 22 illustrates the temperature and humidity of the US Standard profile, along with the median profile of the ECMWF vertical profiles, as well as the coverage. We observe that the coverage of the ECMWF vertical profiles shows different values for surface temperatures, but their median profile agrees very well with the US Standard atmospheric profile. In contrast, the humidity ECMWF vertical profiles show a large difference at the surface compared to the US Standard profile.

	Mean RF value (USstandard) (W/m ²)	Mean RF value (ERA5) (W/m ²)	Bias (W/m ²)	RMS Error (W/m ²)	RMS Error percentage (%)	Mean bias percentage (%)
Land/ SW	-95.28	-97.27	1.99	6.13	6.30	2.05
Multi-layered/ SW	-71.21	-68.28	-2.93	11.99	17.56	4.29
Ocean/ SW	-145.49	-146.97	1.48	10.76	7.32	1.01
Land/ LW	84.46	86.74	-2.29	7.53	8.68	2.64
Multi-layered LW	61.35	65.23	-3.88	7.01	10.75	5.95
Ocean/ LW	95.32	98.19	-2.88	7.24	7.37	2.93

Table 16: Mean radiative forcing (RF) values over all the randomly selected pixels for the six selected days, bias, RMS error, RMS error percentage, and mean percent errors between RF values estimated by using the Look-Up Tables (LUTs) and by using the ERA5 atmospheric profile and the OCA cloud conditions for the SW, and LW estimated RFs.

Overall, in the SW wavelength range, the use of a standard profile in the construction of the LUTs lead to mean bias percentage of about 2.05%, 1.01%, and 4.29% for a contrail above land, ocean, and water cloud, respectively. In the LW wavelength range, the mean percent errors equal to 2.64%, 2.93%, and 5.95% for a contrail above land, ocean, and water cloud, respectively.

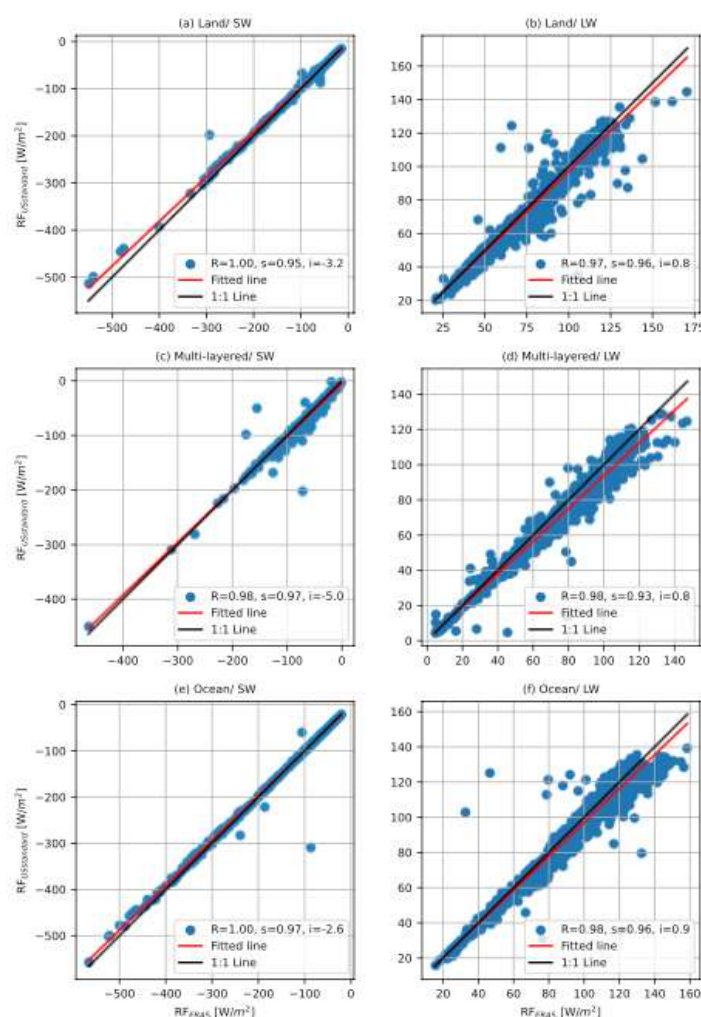


Figure 21: Scatter plot between radiative forcing (RF) values estimated by using the Look-Up Tables (LUTs) (RFLUTs) and radiative transfer calculations using the actual atmospheric temperature vertical profiles (RF_{actual}) for randomly selected pixels containing contrails above land surfaces in the (a) SW, and (b) LW, underlying water clouds (i.e., multi-layered) in the (c) SW and (d) LW, and ocean surfaces in the (e) SW and (f) LW.

Impact of ice cloud parameterization on radiative transfer calculations

The micro-physical properties of the ice crystals, which are part of the cirrus clouds and contrails, play a crucial role in their single scattering properties and, consequently, the RF of these clouds. Here, we assess the impact related to the choice of ice cloud parameterization in the RT simulations. The parameterization determines how the ice water content and CER are translated into optical properties. Since the ice crystal shape is an unknown parameter, we have selected the parameterization by Yang et al. 2013 [RD24], assuming the ice crystal habit to be a column composed of 8 elements with a moderate degree of roughness, as this is the habit most frequently observed for thin ice clouds (Forster Mayer, 2022 [RD25]). According to the same study, 60 % of cirrus clouds are a mixture of ice crystals with severe roughness, while 40 % a mixture of smoothed ones. Similarly to Wolf et al. 2023 [RD26], we have chosen a moderate degree of roughness for the simulations included in the LUTs.

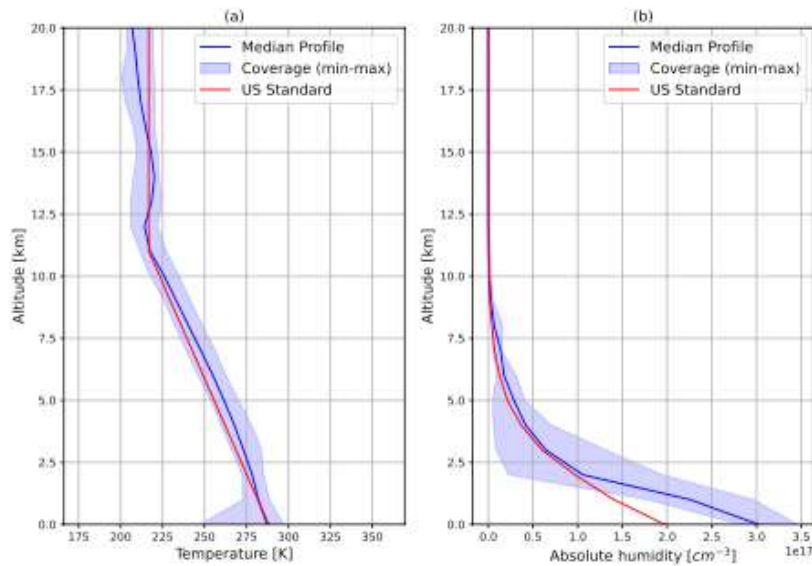


Figure 22: Vertical (a) temperature and (b) humidity profiles of US Standard atmosphere, median profiles of the ECMWF vertical profiles corresponding to the largest discrepancies (i.e., large RMS error percentage between radiative forcing (RF) values estimated by using the Look-Up Tables (LUTs) (RFLUTs) and radiative transfer calculations using the actual atmospheric temperature vertical profiles) for the six selected days.

For this sensitivity study, we performed a small subset of RT simulations in the SW and LW wavelength ranges, varying the choice of ice cloud parameterization. We selected all the available ice crystal shapes from the parameterization by Yang et al. (2013) [RD24]. In addition, we included the parameterization by Fu (1996) [RD27]; Fu et al. (1998) [RD28], which is operationally applied in the ECMWF Integrated Forecasting System (IFS) and assumes ice crystals as pristine hexagonal columns. The simulations are always performed for an ice cloud with a COT equal to 0.5 to maximize its semi-transparency and, subsequently, the effect of cloud microphysics. We have chosen three different SZA values (10° , 40° , and 70°), a CER of $20 \mu\text{m}$, and a CTH of 10 km. For these simulations, the ice cloud is located above an ocean surface characterized by three different SST values (273 K, 293 K, and 303 K).

Figure 23 shows RF_{sol} as a function of various ice crystal habits based on the parameterization of Yang et al. (2013) [RD24] (i.e., column with 8 elements, droxtal, hollow bullet rosette, hollow column, plate, plate with 10 elements, plate with 5 elements, solid bullet rosette, and solid column) and their degrees of roughness (smooth, moderate, and severe) for three different SZAs. The ice crystal habit of an hexagonal column by Fu (1996) [RD27] is included as well. Additionally, the figure presents the relative differences in RF_{sol} compared to the selected ice crystal shape and degree of roughness for the construction of the LUTs. As observed, the choice of ice crystal habit and roughness degree can result in large differences, which can be up to 60% (e.g., the case for $\text{SZA} = 10^\circ$ for smooth plates of 10

elements) in the SW wavelength range. In addition, the parameterization of Fu (1996) [RD27], which assumes a pristine hexagonal column results in differences up to approximately 20 % for the case of a small SZA. For the three SZA scenarios, RFsol of the selected ice crystal shape and roughness appears to have the lowest values compared to other ice crystal shapes and degrees of roughness.

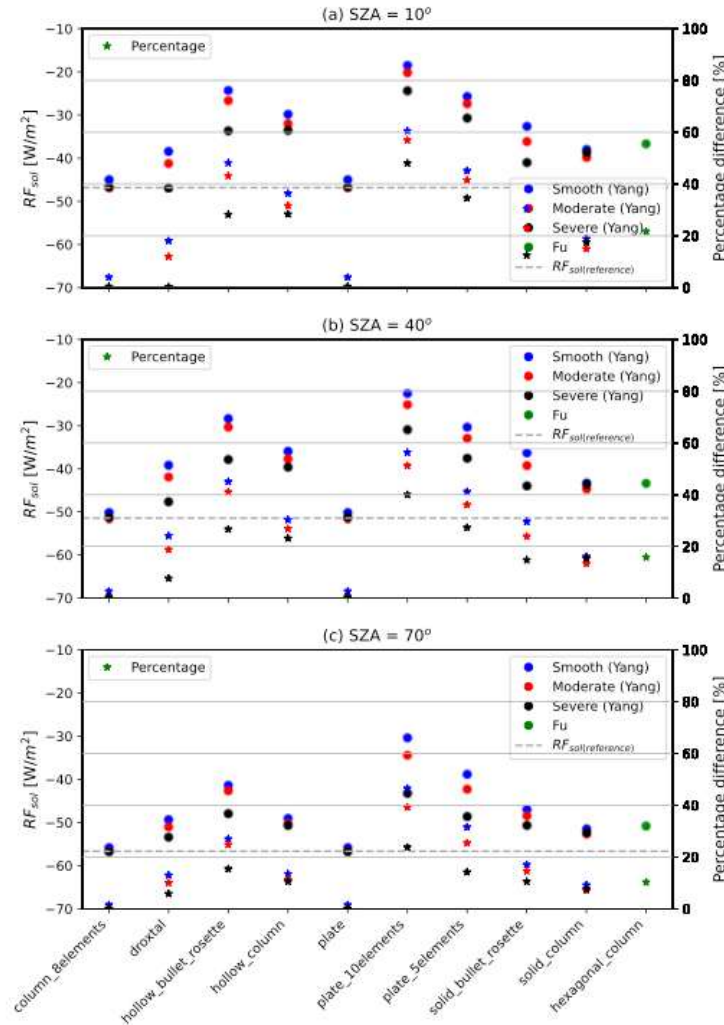


Figure 23: Simulated radiative forcing values in the shortwave (i.e., solar) wavelength range (RF_{sol}) are shown as a function of various ice crystal habits and their degrees of roughness based on the parameterization of Yang et al. (2013) and Fu et al. (1998) for three different solar zenith angle (SZA) scenarios. The horizontal line (i.e., grey dashed line) represents the $RF_{sol}(reference)$ value for the selected ice crystal shape and roughness used in this study.

Figure 23 shows RF_{tir} as a function of the same ice crystal habits and roughness degrees for three different SST scenarios, along with their relative differences. In contrast to the shortwave range, the differences in the LW (RF_{tir}) are much smaller, not exceeding 12%.

From the sensitivity tests, we conclude that ice crystal habit and roughness can lead to significant differences in RT simulations in the SW wavelength region, while these factors play a less significant role in the LW wavelength region. When investigating the simulated upward and downward irradiance at TOA in the SW wavelength region, we find that the largest differences between the selected ice crystal shape and roughness (i.e., column of 8 elements with moderate roughness) and a plate of 10

elements with a smooth degree of roughness (i.e., largest differences in RF_{sol}) occur in the following wavelength ranges: 1122 - 1135 nm, 1346 - 1471 nm, 1800 - 1954 nm, and 2486 - 2752 nm. Similarly, in the LW wavelength, the simulated spectrum is affected the most by the choice of the ice crystal shape and roughness in the following wavelength ranges: 3487 - 4171 nm, 4645 - 5502 nm, and 8113 - 9153 nm.

To estimate the uncertainty associated with the selection of a specific ice cloud parameterization in the RF_{sol} and RF_{tir} maps, we have re-performed the RT simulations for the randomly selected pixels. Consequently, the comparison is made between RF_{sol} and RF_{tir} , where the default ice cloud parameterization was applied, and those generated by employing the same input values for the RT simulations but differing the choice of ice cloud parameterization. For the comparison, we have used the ice crystal habit and roughness, which exhibits the largest difference with our default settings: plate of 10 elements with a smooth degree of roughness (Yang et al., 2013) [RD24].

Table 17 summarizes the findings of the above-mentioned comparison. As expected by the sensitivity study, the use of another ice crystal habit and roughness can lead to large differences in the SW and slightly affects the LW wavelength range. For the SW wavelength range and for all the scene scenarios, the mean RF values for columnar and plate ice crystals differ by a negative bias, with the largest bias found for contrails above ocean surfaces (-49.33 W/m^2).

For the LW wavelength range, the bias values are smaller, with the largest bias being equal to 5.89 W/m^2 for ice clouds above water clouds (i.e., multi-layered).

We should keep in mind that actual measurements of the micro-physical properties of ice crystals in contrail clouds are rare and difficult to obtain. There have been in-situ measurements, such as those in Järvinen et al. (2018) [RD29], which found that the primary ice crystal habit is aggregates (i.e., the one used in this study), though the presence of other crystal shapes has been reported. Consequently, we used the most common one to optimize the representation of ice crystals. However, applying a single ice crystal shape and roughness for the overall number of detected contrails during different seasons, and above various scenes may not be fully representative.

	Mean RF value (column 8elements) (W/m^2)	Mean RF value (plate 10elements) (W/m^2)	Bias (W/m^2)	RMS Error (W/m^2)
Land/ SW	-77.14	-42.70	-34.44	37.36
Multi-layered/ SW	-66.65	-34.39	-32.26	41.57
Ocean/ SW	-133.52	-84.19	-49.33	60.08
Land/ LW	63.20	58.37	4.83	5.24
Multi-layered LW	47.08	41.19	5.89	7.13
Ocean/ LW	65.55	67.29	-1.75	15.26

Table 17: Mean radiative forcing (RF) values over all the randomly selected pixels for the 25th of September 2023, bias, RMS error between RF values when using an ice crystal habit of column with 8 elements and a plate with 10 elements for the SW, LW, and net estimated RFs.

4.4.4 OBJ-03-TRL1-ERP-001 results

To achieve OBJ-03, we leverage temporal and spatial deep learning modelling for contrail evolution and RF estimation. The goal is to integrate multi-modal datasets—including satellite-derived contrail masks, ADS-B flight trajectories, radiative forcing calculations, and meteorological reanalyses into a unified spatio-temporal deep learning framework.

Methodology: To capture both spatial structure and temporal dynamics, we harmonized data onto a $0.1^\circ \times 0.1^\circ$ resolution grid with 15-minute intervals, ensuring pixel-wise learning. The following deep learning architectures were implemented:

- ConvLSTM – A baseline model capturing spatio-temporal dependencies.
- Vision Transformer – Extracts rich spatial features but lacks explicit temporal encoding.
- GeoSTANet – A novel geospatial spatio-temporal network combining transformer-based attention with convolutional encodings to track contrail movement and persistence.

Models were trained and evaluated on two high-activity periods (Jan 24–30, 2023 & 2024) to assess contrail formation and RF prediction accuracy. To address extreme class imbalance, we applied focal loss, while the Schmidt-Appleman Criterion was used to filter training data based on atmospheric conditions conducive to contrail formation.

4.4.4.1 Contrail Occurrence Prediction Performance

One of the primary goals was to accurately predict where and when contrails will form (or persist). This is essentially a binary classification problem on each grid cell at each future time, under highly imbalanced conditions (very few positive contrail pixels). Our best models achieved a balanced accuracy of slightly above 50% on the test dataset. While this may seem modest, note that a naive guess (no contrails everywhere) would score 50% by default in balanced accuracy (0% recall for contrails, 100% for clear sky). Thus, any improvement above 50% indicates the model is indeed detecting some contrails correctly. The GeoSTANet model in the “combined features” scenario (using both ADS-B and weather inputs) attained ~52.1% balanced accuracy, with a contrail Recall of ~40% and precision (Accuracy on contrail class) around 60%. In contrast, the ViT-based model had a higher overall accuracy (~69% vs 61%) but lower recall (~32%) in the combined input case. This suggests that GeoSTANet is better at not missing contrails (higher recall), whereas ViT is more conservative (fewer false positives, hence higher precision). GeoSTANet’s strength likely comes from its ability to leverage sequence information it “remembers” contrail presence over time, whereas the ViT (with less temporal context) might default to predicting contrails only when conditions are very evidently favourable (hence missing some that were subtle).

We also evaluated models with different input feature sets to understand the contribution of each data source:

- **Using Meteorological Data (MET) only:** The models in this setup try to infer contrail formation purely from weather conditions. We found that they can learn the SAC-like criteria to some extent (e.g., picking out cold, humid areas), but performance was limited. For instance, the ViT with MET-only had higher accuracy (~80%) but very low recall (~22%). It predicts contrails only in the most obvious situations (high confidence from weather), missing many contrails that occurred perhaps due to slight violations of ideal conditions.

- **Using ADS-B Data only:** These models know where planes are, but not the environment. An ADS-B-only model can identify where contrails could potentially be (high flight density regions), but it cannot tell if the atmosphere allowed a contrail to form. This yielded many false positives (predicting contrails whenever a plane flew by, even though contrails might not always form). We observed low precision in this case. (Exact metrics for ADS-B only were not highlighted, as they were inferior; the combination of ADS-B+MET is far more informative).
- **Using Combined ADS-B + MET:** This was the most successful input configuration. The model knows both where planes were and if the air was conducive. The results above for GeoSTANet and ViT with “Combined” inputs show the synergy neither flights nor weather alone suffices, but together they allow the prediction of contrails with improved reliability. Notably, adding ADS-B data helped the recall (catch more contrails) because the model is aware that without a plane, a contrail cannot exist, so it focuses on those areas to decide if conditions tipped it into actual contrail formation.

Figure 24 illustrates a sample prediction from the GeoSTANet model compared to the true radiative forcing induced by contrails, across three-time steps from the test set. The model accurately captured the spatial extent and general intensity of contrail-induced RF over key air traffic corridors, such as the North Atlantic and Central Europe. While the predicted RF fields closely follow the observed patterns, some discrepancies are visible such as over-smoothing of fine-scale features and occasional false positives, particularly in areas with high flight density but borderline atmospheric conditions. These visualizations provide insight into the model's behaviour; in this instance, the model tends to prioritize sensitivity, occasionally predicting RF where conditions are marginal, consistent with a bias toward capturing all possible contrail events.

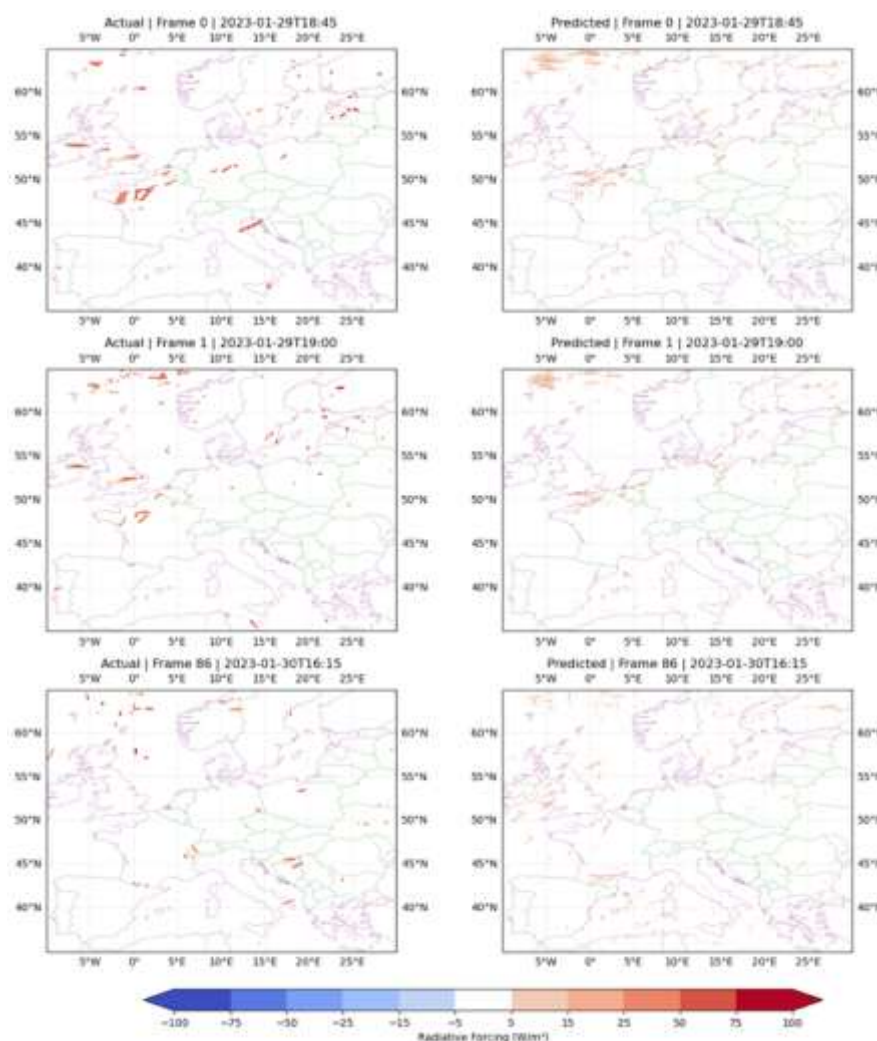


Figure 24: Comparison of Predicted vs. Actual Radiative Forcing from Contrails

Note: Left panels show the actual radiative forcing (RF) induced by contrails observed from satellite data, while the right panels show the predicted RF generated by the GeoSTANet model using input features from earlier time steps. The predictions were made at 15-minute intervals over Europe, covering multiple frames between 29 January 2023, 18:45 UTC and 30 January 2023, 16:15 UTC. The color scale represents RF values in W/m^2 , where red tones indicate positive forcing (warming effect) and blue tones indicate negative forcing (cooling effect).

Overall, the ability to predict contrail occurrence, while not perfect, is significantly better than chance and provides a basis for targeted mitigation (e.g., if a model indicates with moderate confidence a contrail will form on a given route, action can be taken). It's worth noting that even state-of-the-art contrail forecasts in research are an emerging capability, so these results are a valuable step forward. The balanced accuracy in the low 50s indicates there is room for improvement likely through more data (to cover more scenarios) and further model tuning (discussed in Section 5, Future Directions). Encouragingly, the models rarely predicted contrails in completely wrong places/times; most false positives occurred in areas very close to actual contrails or under marginal conditions, which suggests the models learned physically sensible decision boundaries.

4.4.4.2 Radiative Forcing Prediction Performance

Beyond predicting contrail occurrence, our models estimate the radiative forcing (RF) associated with contrails. This can be viewed as a regression problem, often conditioned on first predicting contrails. The best models achieved an RMSE of around 2.0 to 2.5 W/m² on the test set, compared to real-world contrail RF variations from near 0 up to 10 W/m² locally. An RMSE of 2 W/m² suggests that the model's RF estimates are reasonably accurate, though with some underprediction of high-RF cases.

Interestingly, the ViT model had the lowest RF error (RMSE ~1.7 W/m²) in the MET-only setup, indicating that it excels at general RF estimation when contrail locations are assumed known. However, when contrail location uncertainty is involved (combined input), its RMSE increased to ~2.3 W/m². Meanwhile, GeoSTANet had RMSE ~3.3 W/m² in the combined input case, likely due to its higher recall (catching more contrails but sometimes overestimating their RF contribution).

A noteworthy finding was that when trained on 2023 data and tested on 2024 data, the models correctly predicted a higher total contrail RF in 2024, though it underestimated the increase compared to observations (~80% predicted vs. ~128% actual increase). This suggests the model generalizes but could benefit from additional training data.

RF Prediction Accuracy: The models showed a moderate correlation between predicted and actual RF across the test dataset (Pearson correlation ~0.6 to 0.7), meaning they captured general trends but struggled with outliers. The ViT model was best at predicting low-RF cases, while GeoSTANet had higher recall on extreme RF values.

4.4.4.3 Model Interpretability and Insights

To provide a comprehensive comparison, we evaluated four deep learning architectures: CNN+LSTM, ConvLSTM, GeoSTANet, and ViT. The models were tested across different data configurations (ADS-B, Meteorological, and Combined), and their performance was summarized using heatmaps (Figure 25, Figure 26, Figure 27, and Figure 28) and detailed in Table 18.

<i>Model</i>	<i>Accuracy</i>	<i>Recall</i>	<i>Balanced Acc</i>	<i>AUC</i>	<i>Specificity</i>	<i>MAE</i>	<i>RMSE</i>	<i>MSE</i>
<i>CNN+LSTM</i>	85.42%	22.29%	54.27%	54.27%	86.26%	2.96	6.88	47.87
<i>ConvLSTM</i>	47.05%	62.38%	54.64%	54.64%	46.89%	6.00	10.12	102.48
<i>GeoSTANet</i>	62.63%	41.95%	52.43%	52.43%	62.91%	2.80	4.81	24.63
<i>ViT</i>	80.06%	21.43%	51.62%	51.62%	81.81%	0.64	1.35	2.02

Table 18: Performance Metrics for Different Models and Feature Configurations

Experimental Design

- Data: The dataset included contrail masks, radiative forcing, ADS-B aircraft data, and meteorological fields from January 2023 and 2024.
- Resolution: Data was harmonized to 0.1° x 0.1° spatial resolution with a 15-min. temporal interval.
- Models: CNN+LSTM, ConvLSTM, ViT, and GeoSTANet were tested with different input feature combinations (ADS-B only, MET only, and combined).
- Evaluation Metrics: Accuracy, Recall, Balanced Accuracy, AUC, Specificity, MAE, RMSE, and MSE were analysed.

Figure 25, Figure 26, Figure 27, and Figure 28 show performance heatmaps for the tested models:

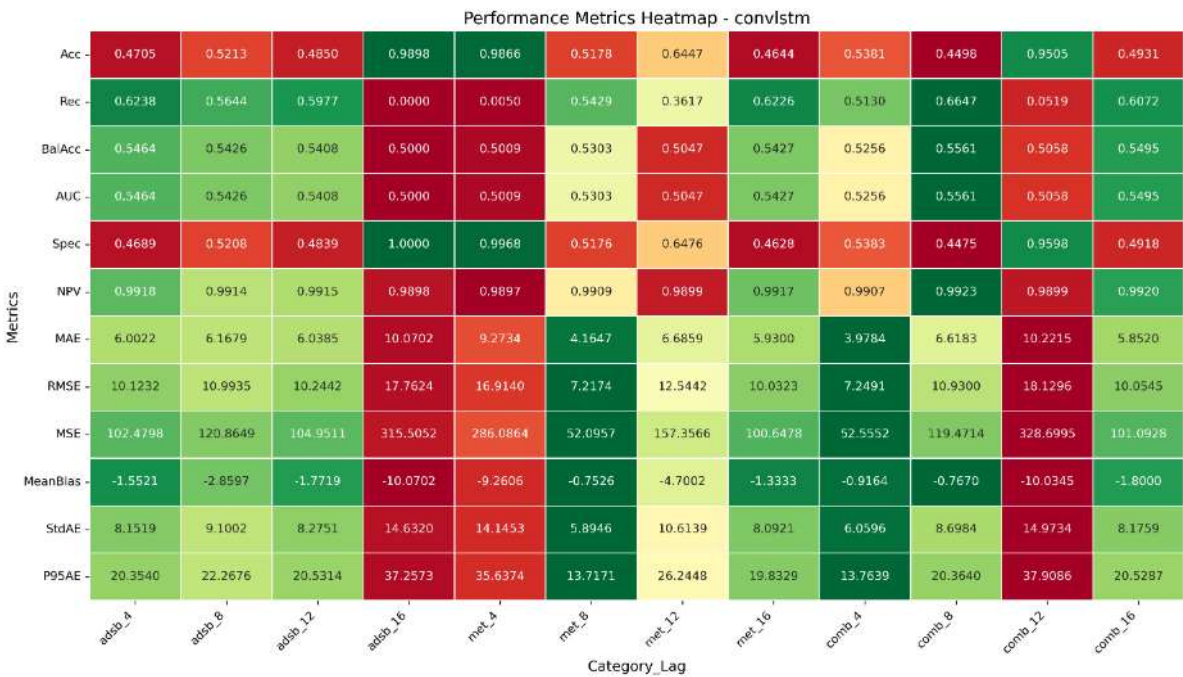


Figure 25: ConvLSTM performance across different input configurations



Figure 26: CNN+LSTM performance across different input configurations



Figure 27: ViT performance across different input configurations



Figure 28: GeoSTANet performance across different input configurations

These heatmaps provide insights into the strengths and weaknesses of each model, helping to guide further improvements in contrail prediction modelling.

Analysing the models' behaviours reveals that:

- The models latched onto humidity and temperature at cruise altitudes as primary indicators (replicating Schmidt-Appleman Criterion internally).
- The GeoSTANet model's higher recall stems from its ability to track contrails over time, ensuring persistence is considered.
- The ViT model excelled in identifying contrail shape patterns, though it lacked explicit time encoding.

Several key findings emerged:

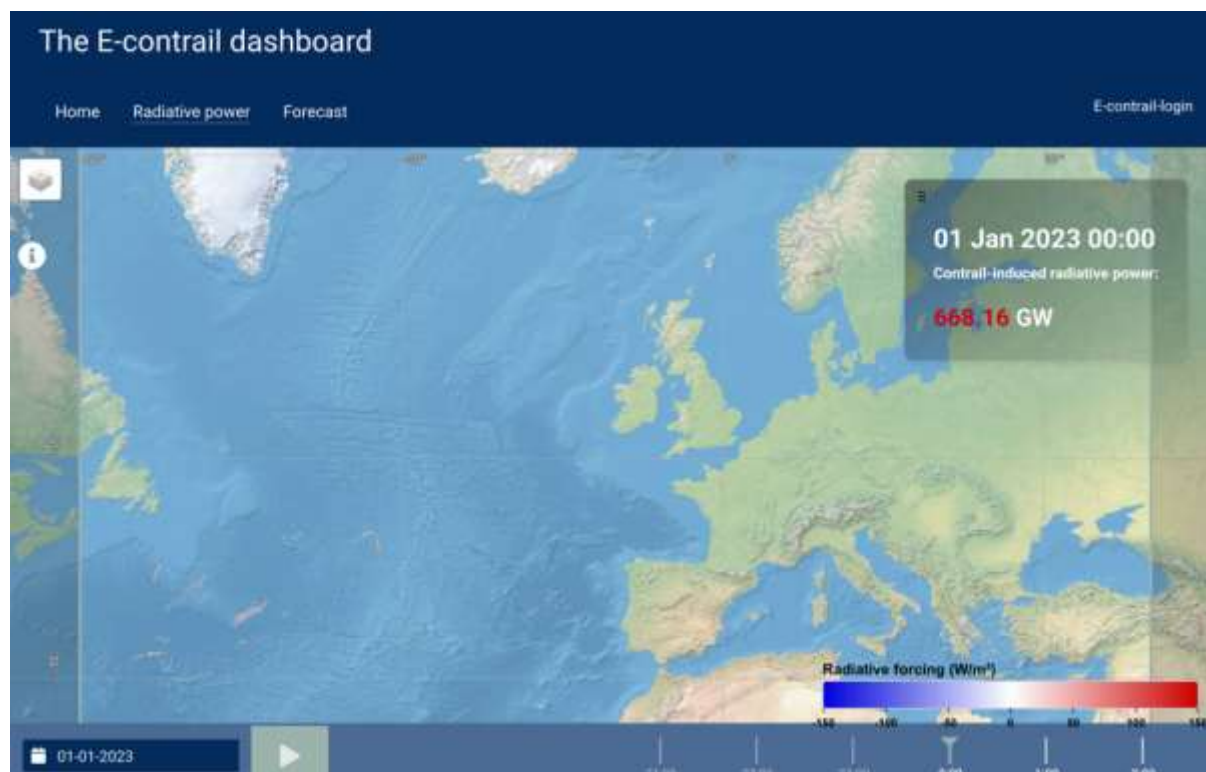
- Longer temporal context improves recall: Increasing input sequence length from 8 to 12 lag steps improved accuracy (~2% gain in balanced accuracy).
- Handling class imbalance is crucial: Focal loss significantly improved contrail detection recall (~40% vs. <10% without focal loss).

Flight and meteorology data are both required: Models trained on only one dataset failed to generalize.

4.4.5 Obj-04-TRL1-ERP-001 results

A visualization tool has been developed to inform stakeholders. The visualization tool can be divided into two visualization services: climate impact quantification and ISSR forecasting.

It can be accessed via this link <https://econtrail-test.aeronomie.be/test>



The validation has been performed on a qualitative way, gathering insights from different stakeholders. The E-CONTRAIL team will continue to show and improve the dashboard in different forums and to different stakeholders with the aim at continuously improving it.

5 Conclusions and recommendations

5.1 Conclusions

5.1.1 Conclusions of the preliminary experiment: “Satellite-Based Quantification of Contrail Radiative Forcing over Europe: A Two-Week Analysis of Aviation-Induced Climate Effects.”

The preliminary experiment [18] evaluates the RF of contrails detected by our model using MSG satellite data during two complete weeks: January 24-30 in 2023 and 2024. The main conclusions are as follows:

- (1) Daytime contrails typically produce a cooling effect, with maximum net CRF values reaching approximately -8 TW, whereas nighttime contrails contribute to warming with CRF values up to 6 TW.
- (2) Despite the cooling effect of daytime contrails, the overall daily impact is warming due to the higher number of nighttime contrails, which account for 62 % of the total detected.
- (3) A comparison of data from 2023 and 2024 shows a significant increase in contrail coverage, with detections rising by 41.03 % and CRF values increasing by 128.7 % in 2024, suggesting an increased warming effect from the additional number of contrails.
- (4) Analysis of individual contrail features reveals that larger daytime contrails have a smaller warming impact compared to smaller nighttime contrails, reinforcing the overall warming trend observed.

Overall, the study highlights the significant impact of contrail timing on their net warming effect, emphasizing the need to consider both daytime and nighttime contrails in evaluations of RF. Understanding these temporal differences is crucial for accurately assessing the influence of contrails on climate change and for developing effective mitigation strategies to address aviation-induced warming.

In a future study, we will examine a full year of data to capture seasonal variations and account for fluctuations in flight patterns, offering a thorough assessment of contrail impacts across different times of the year. Additionally, we will implement comprehensive validation of the detected contrails or introduce uncertainty metrics to address potential errors and improve the accuracy of our statements.

5.1.2 Conclusions of the experiment 2: Satellite-Based Quantification of Contrail Radiative Forcing over Europe and ISSR forecasting: a Full Year 2023 analysis

These conclusions should be considered as preliminary:

1. We have been able to run the E-CONTRAIL detection algorithm, applying it for the first time in EUROPE (MSG data) for a full year of data:

- a. The E-CONTRAIL detection algorithm has shown to have very powerful metrics in terms of true positives and false negatives (with Dice Scores over 80%) when assessed against the training dataset.
 - b. Results show that the **Months with the Highest Number of Contrail Detections** are April and May, which exhibited the highest frequency of contrail occurrences in 2023.
 - c. Results show that the **Peak Detection Times** are 07:30 UTC and 15:30 UTC.
2. We have been able to run the E-CONTRAIL radiative forcing estimation algorithm, applying it for the first time in EUROPE (MSG data) for a full year of data:
 - a. The E-CONTRAIL radiative forcing algorithm has shown to benchmark well when comparing the fluxes of the radiometers on board CERES LEO satellite (with accuracies above 80%).
 - b. There is an overall net cooling effect during the central hours of the day.**
 - c. There is an overall net warming effect during the months of August to April, with more intense values in November to February.**
3. We have been able to run the E-CONTRAIL ISSR prediction algorithm, applying it for the first time in EUROPE (MSG data) for a full year of data:
 - a. The E-CONTRAIL ISSR & RF prediction algorithm has shown good metrics in terms of success rate (with Dice scores above 80%) and prediction of RF values, with Mean Average Errors of 10-20%.
4. We have been able to build a visualization dashboard and show all the results.

5.1.3 Conclusions on project/ SESAR solution maturity

The conclusions of the project coincide with those already listed in Section 5.1.3

5.2 Recommendations

The most important recommendation that can be drawn from the project is that the scientific understanding of contrails and its associated impact is still low (though is growing fast).

Results indicate that:

- Any mitigation action to be taken should be prioritizing flights occurring at dark (or mostly at dark) and preferably in winter months.
- Any mitigation action to be taken in flights occurring during daylight (or mostly exposed to daylight) and/or in spring, fall, and summer season still require further scientific understanding about the cooling effects and the intensity of the RF when compared to CO₂.

5.2.1 Recommendations for next R&I phase

The following open problems remain open:

- Contrail detection:

Current satellite-based contrail detection models are challenged by contrail age, morphology, and cloud interference, leading to frequent underdetection and underestimation of their climate impact (Figure 29). The absence of labeled individual contrails limits model applicability, while the lack of labeled imagery outside the U.S., particularly in Europe, prevents proper validation and generalization for these regions.

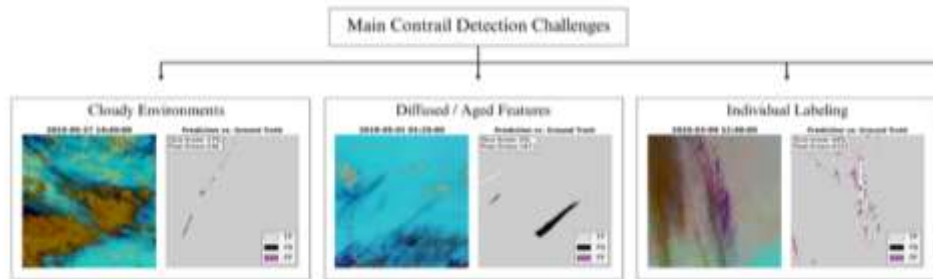


Figure 29: Main challenges associated with contrail detection models trained on OpenContrails. Dataset (Ortiz et al., 2025).

- Contrail physical modelling:

Existing contrail models such as CoCiP (Schumann et al., 2010 and Schumann et al., 2012), APCEMM (Fritz et al., 2020), and ACSM (Li et al., 2023) leverage meteorological data to identify regions conducive to contrail formation, particularly those that satisfy the Schmidt-Appleman criterion (Schumann et al., 2012, Schumann et al., 1995, Gierens et al., 2003). Then, with the use of air traffic data, they simulate contrail plume development and estimate the RF associated with these features, thereby facilitating more accurate assessments of the global environmental impact of contrails.

However, existing contrail physics models still lack reliability due to the following reasons:

- **Inadequate Microphysical Representation of Ice Crystals:** Current models often lack accurate representations of ice crystal habits. While assuming spherical particles is reasonable for crystals smaller than ~5 micrometers in radius, larger ice crystals subjected to varying supersaturation and temperature conditions tend to develop faceted, complex shapes. These morphological changes significantly influence plume dynamics and radiative forcing.
- **Insufficient Macroscale Modeling of Ice Crystal Propagation:** A major limitation is the absence of proper coupling between the ice crystal (particle) phase and the surrounding fluid phase at larger scales. While such multi-phase interactions are typically included in Phase 1 models (short-term plume behavior), they are largely missing in Phase 2 models, which are essential for assessing contrails' long-term atmospheric impacts. This lack of coupling affects predictions of plume vertical extent (which is linked to a turbulent self-diffusion effect) and related optical depth calculations critical to radiative forcing estimates.
- **Complexity of Physics-Driven Models:** Physics-based models rely on nonlinear differential equations (either partial or ordinary) to simulate the physical processes underlying contrail formation and evolution. Solving these equations is inherently challenging, compounded by significant uncertainties in the numerous physical and rheological parameters involved.
- **Uncertainty in Ice Crystal Properties at Later Stages:** There exists large uncertainty on the concentration, ice crystal distribution, morphology and number density of the crystals at the dissipation and diffusion regimes and its dependency on fuel composition and engine operating conditions for their use in well-established contrail models (aCCFs and CoCiP).
- **Validation** of physics-driven models is a challenging task due to lack of scientific experimental setup able to capture detailed properties of contrail plumes at a microphysical level.

In addition, physical models permit modelling the contrails of future fuels/aircraft technologies (contrary to data-driven methods, which can only learn from existing data). This is the case of H2 aircraft and sustainable aviation fuels (SAF), where non-CO2 impact assessment is under debate.

Scientific Gaps: To accurately assess the climate benefits for sustainable aviation including new propulsion technologies operated with next generation of synthetic fuels, further research is needed to deepen our understanding of their emission characteristics and associated climate impacts. This research should ultimately support the development of models that enable climate studies encompassing both current aviation systems and emerging technologies and fuels.

- **Physic-Informed contrail detection:**

To overcome some of the limitations and open problems related to contrail detection, the blending of physical models and machine learning for contrail detection remains as an interesting line of research.

- **Radiative Forcing Estimation:**

Although broadband radiometers, like CERES or BBR, can measure the total reflected solar and emitted thermal fluxes directly, their coarse spatial resolution (on the order of 20 km at nadir) often precludes accurate detection and flux quantification of narrow, linear contrails. Consequently, many studies rely on radiative-transfer (RT) simulations or simplified parameterizations to derive contrail radiative forcing, but co-located broadband measurements rarely validate these methods, and they lack supporting lidar data to improve contrail-top height estimates. Furthermore, partial pixel coverage of contrails introduces additional uncertainties, compromising the accuracy of flux calculations. This underscores the need for a refined methodology, one that converts narrowband signals to broadband fluxes at higher resolution and that integrates lidar or other auxiliary data, where available, to better constrain contrail properties and improve RF estimates.

Scientific Gap: There remain significant gaps due to the coarse resolution of broadband instruments (e.g., CERES), the lack of co-located broadband observations to validate narrowband-to-broadband conversions, and the absence of lidar data for precise contrail-top height estimates. Uncertainties increase when contrails fill only part of a pixel, making flux calculations less reliable. As a result, current radiative forcing estimates rely on partially validated parameterizations or radiative-transfer simulations that do not fully capture the true impact of contrails.

-

- **ISSR and Radiative Forcing Estimation:**

Scientific Gap: Significant scientific gaps remain despite growing interest in predicting ISSRs and their role in persistent contrail formation. One core challenge is the limited accuracy and resolution of current ISSR forecasts, especially regarding their dependence on altitude, pressure levels, and regional meteorological conditions. Traditional numerical weather prediction systems, such as ERA5 and meteorological services, provide coarse, static outputs that are insufficient for real-time operational decision-making in aviation contexts. Furthermore, attributing aviation-induced radiative forcing to specific contrail-producing regions remains poorly constrained. This limits the effectiveness of climate mitigation strategies, including trajectory optimization to avoid high-risk ISSR zones. While recent advances in machine learning have shown promise in atmospheric modeling, most existing efforts either focus on surface-level variables or lack the spatiotemporal resolution required for accurate upper-tropospheric forecasting. Current models often fall short in several key areas: (i) *Forecast Accuracy*: Existing approaches struggle to match the temporal and spatial resolution needed for accurate ISSR predictions. (ii) *Operational Readiness*: Many models are computationally intensive and

not designed for real-time or near-real-time deployment. (iii) *Integration with Mitigation Tools*: There is a lack of ISSR prediction systems that directly support contrail avoidance and RF reduction strategies. (iv) *Spatiotemporal Dynamics*: Modeling the evolution of atmospheric variables at the relevant scales remains a complex challenge, especially in three-dimensional domains. (v) *Model Generalizability*: Few studies have demonstrated that deep learning models for ISSR forecasting can generalize well across seasons, geographic regions, or varying atmospheric profiles.

6 References

6.1 Applicable documents

This ERP complies with the requirements set out in the following documents:

Project and programme management

- [AD1] SESAR 2020 Experimental Approach guidance ER, 11/12/2020, edition 1.
- [AD2] 101114795 E-CONTRAIL Grant Agreement, [02/06/2023]
- [AD3] SESAR 3 JU Project Handbook – Programme Execution Framework, 11/04/2022, edition 1.

6.2 Reference documents

- [RD1] Kulik, L. (2019). Satellite-based detection of contrails using deep learning (Doctoral dissertation, Massachusetts Institute of Technology).
- [RD2] Young, Alisa H., et al. "The international satellite cloud climatology project H-Series climate data record product." *Earth System Science Data* 10.1 (2018): 583-593.
- [RD3] Poulsen, Caroline A., et al. "Cloud_cci ATSR-2 and AATSR data set version 3: a 17-year climatology of global cloud and radiation properties." *Earth System Science Data* 12.3 (2020): 2121-2135.
- [RD4] Illingworth, Anthony J., et al. "The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation." *Bulletin of the American Meteorological Society* 96.8 (2015): 1311-1332.
- [RD5] Jardines, A., et al. "Convection Indicator for Pre-Tactical Air Traffic Flow Management using Neural Network. *Machine Learning with Applications*". 2021-06 | DOI:
- [RD6] Jardines, A., et al. "Thunderstorm prediction during pre-tactical air-traffic-flow management using convolutional neural networks. *Expert Systems with Applications*." Jardines, A., 2024, Vol 241, 10.1016/j.mlwa.2021.100053
- [RD7] Dietmüller, S., et al. (2022): A python library for computing individual and merged non-CO2 algorithmic climate change functions: CLIMaCCF V1.0, Geoscientific Model Development (GMD). Under Review.
- [RD8] Schumann, U. (2012). A contrail cirrus prediction model. *Geoscientific Model Development*, 5(3), 543-580.
- [RD9] Fundamental science and outreach for aviation green deal (HORIZON-SESAR-2022-DES-ER-01-WA1-6)

- [RD10] Jardines, A., Soler, M., et al. Convection Indicator for Pre-Tactical Air Traffic Flow Management using Neural Network. Machine Learning with Applications. 2021-06 | DOI: 10.1016/j.mlwa.2021.100053
- [RD11] Jardines, A., Eivazi, H., Zea, E., Otero, E., Soler, M., et al. (2023) Thunderstorm prediction during pre-tactical air-traffic-flow management using convolutional neural networks Expert Systems with Applications.
- [RD12] EASA report on aviation's non-CO2 climate impacts highlights need for policy decisions and more research – GreenAir News
- [RD13] Abolfazl Simorgh, Manuel Soler, et al. Robust 4D Climate Optimal Flight Planning in Structured Airspace using Parallelized Simulation on GPUs. Geoscientific Model Development (GMD). Under Review. Egusphere-2022-1010
- [RD14] Soler et al., FlyATM4E D4.4. “Report on robust and climate impact reducing ATM operations including an overall environmental evaluation and implementation analysis from a hindcast analysis”
- [RD15] The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018 – ScienceDirect
- [RD16] 2020.Updated analysis of the non-CO2 climate impacts of aviation and potential policy measures pursuant to EU Emissions Trading System Directive Article 30(4). European Commission. <https://www.easa.europa.eu/en/downloads/120860/en>
- [RD17] EUROCAE ED-78A Guidelines for Approval of the Provision and Use of Air Traffic Services supported by Data Communications, December 2020.
- [RD18] Irene Ortiz, Ermioni Dimitropoulou, Pierre de Buyl, Nicolas Clerbaux, Javier García-Heras, Amin Jafarimoghaddam, Hugues Brenot, Jeroen van Gent, Klaus Sievers, Evelyn Otero, Parthiban Loganathan, Manuel Soler. **Satellite-Based Quantification of Contrail Radiative Forcing over Europe: A Two-Week Analysis of Aviation-Induced Climate Effects**. SESAR Innovation Days 2024. Pre-Print @ <https://arxiv.org/abs/2409.10166>
- [RD19] Ortiz, I., García-Heras, J., Jafarimoghaddam, A., & Soler, M. (2025, January). Enhancing GOES-16 contrail segmentation through ensemble neural network modeling and optical flow corrections. Institute of Electrical and Electronics Engineers (IEEE). <https://doi.org/10.36227/techrxiv.173749955.56653418/v1>
- [RD20] Sanz-Morère, I., Eastham, S. D., Allroggen, F., Speth, R. L., and Barrett, S. R. H.: Impacts of multi-layer overlap on contrail radiative forcing, Atmos. Chem. Phys., 21, 1649–1681, <https://doi.org/10.5194/acp-21-1649-2021>, 2021.
- [RD21] Spangenberg, D. A., Minnis, P., Bedka, S. T., Palikonda, R., Duda, D. P., & Rose, F. G. (2013). Contrail radiative forcing over the Northern Hemisphere from 2006 Aqua MODIS data. Geophysical Research Letters, 40(2), 276–281. <https://doi.org/10.1002/qlr.50168>
- [RD22] Ortiz, I., Simorgh, A., García-Heras, J., Dimitropoulou, E., De Buyl, P., Clerbaux, N., and Soler, M.: Assessing Contrail Radiative Effects: A Comparison of MTG Satellite Detections and Physics-Based Simulations, EGU General Assembly 2025, Vienna, Austria, 27 Apr–2 May 2025, EGU25-19879, <https://doi.org/10.5194/egusphere-egu25-19879>, 2025.

- [RD23] Dimitropoulou, E., de Buyl, P., and Clerbaux, N.: Satellite-based estimation of contrail cirrus cloud radiative forcing derived through a Rapid Contrail-RF Estimation Approach, EGU sphere [preprint], <https://doi.org/10.5194/egusphere-2025-697>, 2025.
- [RD24] Yang, P., Bi, L., Baum, B. A., Liou, K.-N., Kattawar, G. W., Mishchenko, M. I., and Cole, B.: Spectrally consistent scattering, absorption, and polarization properties of atmospheric ice crystals at wavelengths from 0.2 to 100 μm , *Journal of the atmospheric sciences*, 70, 330–347, 2013.
- [RD25] Forster, L. and Mayer, B.: Ice crystal characterization in cirrus clouds III: Retrieval of ice crystal shape and roughness from observations of halo displays, *Atmospheric Chemistry and Physics*, 22, 15 179–15 205, 2022.
- [RD26] Wolf, K., Bellouin, N., and Boucher, O.: Sensitivity of cirrus and contrail radiative effect on cloud microphysical and environmental parameters, *Atmospheric Chemistry and Physics*, 23, 14 003–14 037, 2023
- [RD27] Fu, Q.: An accurate parameterization of the solar radiative properties of cirrus clouds for climate models, *Journal of climate*, 9, 2058–2082, 1996.
- [RD28] Fu, Q., Yang, P., and Sun, W.: An accurate parameterization of the infrared radiative properties of cirrus clouds for climate models, *Journal of climate*, 11, 2223–2237, 1998.
- [RD29] Järvinen, E., Jourdan, O., Neubauer, D., Yao, B., Liu, C., Andreae, M. O., Lohmann, U., Wendisch, M., McFarquhar, G. M., Leisner, T., et al.: Additional global climate cooling by clouds due to ice crystal complexity, *Atmospheric Chemistry and Physics*, 18, 15 767–15 781, 2018.

