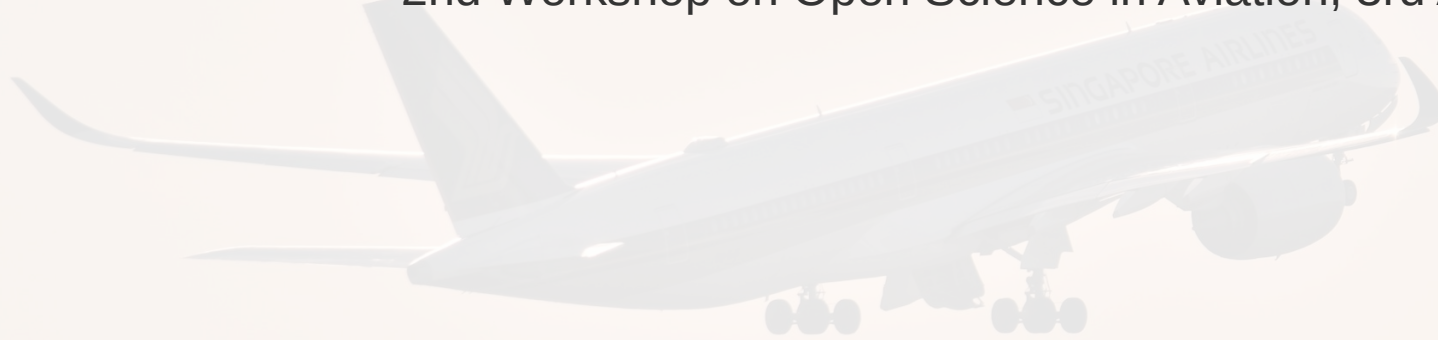


A large-scale and open dataset of aircraft interactions

RAÚL LÓPEZ-MARTÍN & MASSIMILIANO ZANIN

Institute for Cross-Disciplinary Physics and Complex Systems (IFISC)

2nd Workshop on Open Science in Aviation, 3rd April 2025, Delft



EXCELENCIA
MARÍA
DE MAEZTU
2023 - 2027



Universitat
de les Illes Balears



@ifisc_mallorca



Facebook.com/ifisc

<http://ifisc.uib-csic.es> - Mallorca - Spain

What is this dataset?

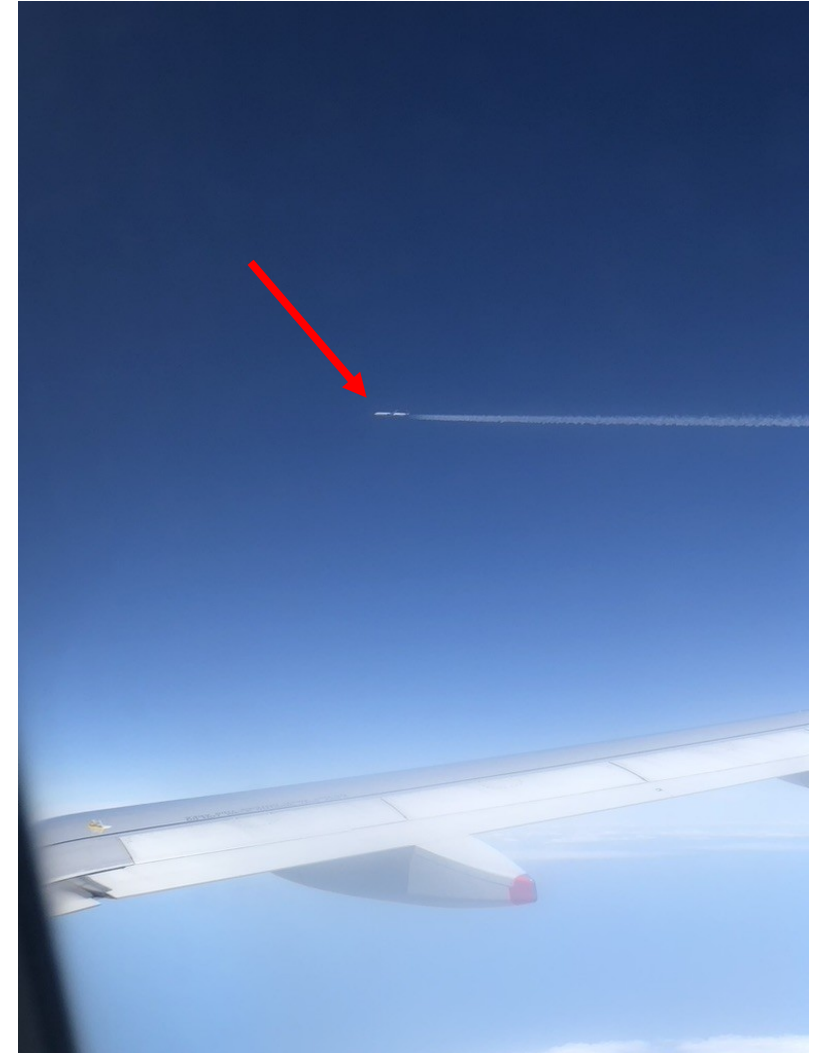
- A collection of daily pairwise aircraft interactions occurred over Europe.
- For each day, macro-scale information about the airspace's traffic is additionally given.

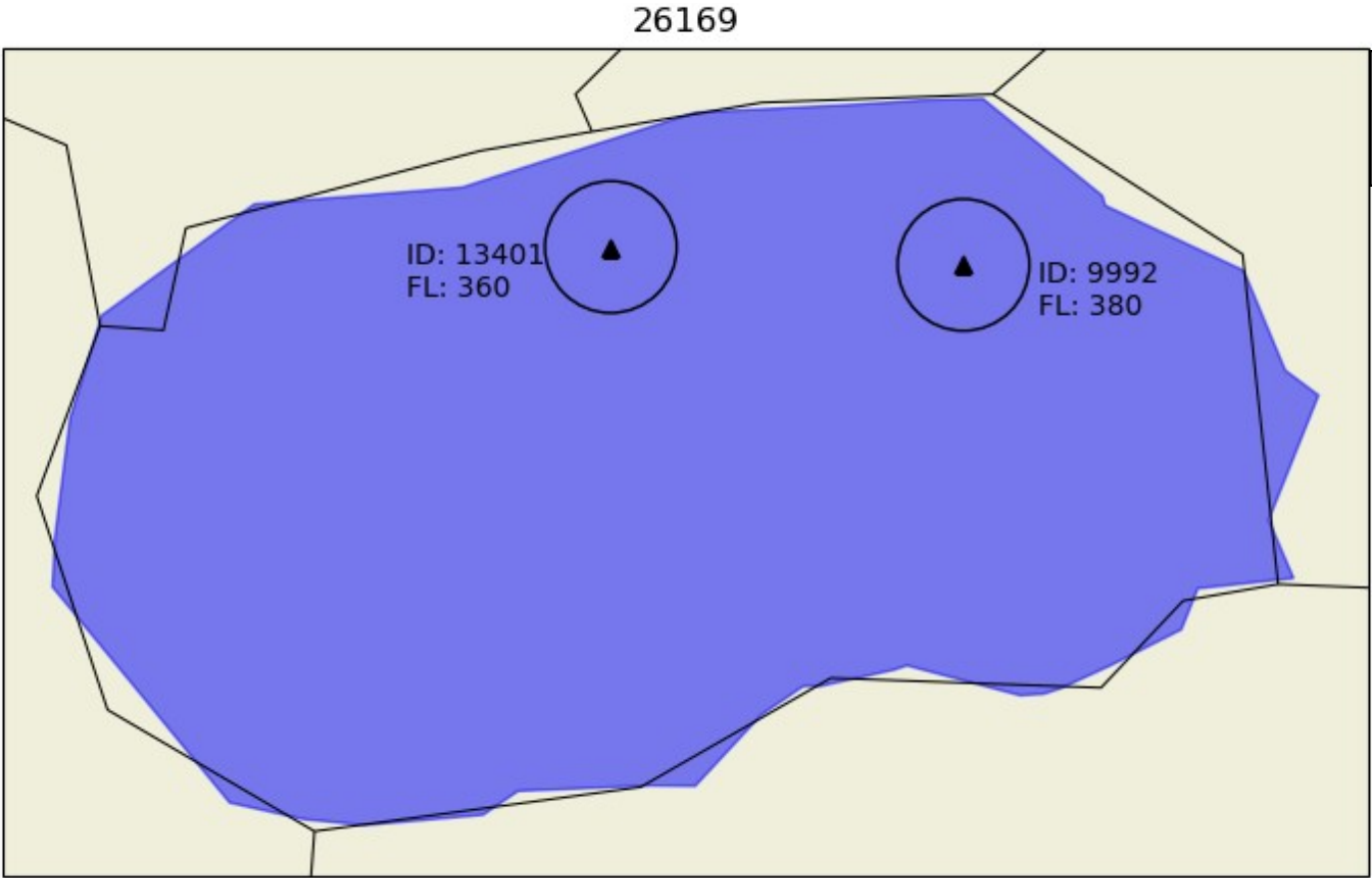
What are aircraft interactions?

- En-route (over 100FL) pairwise distance reductions (not LoS).

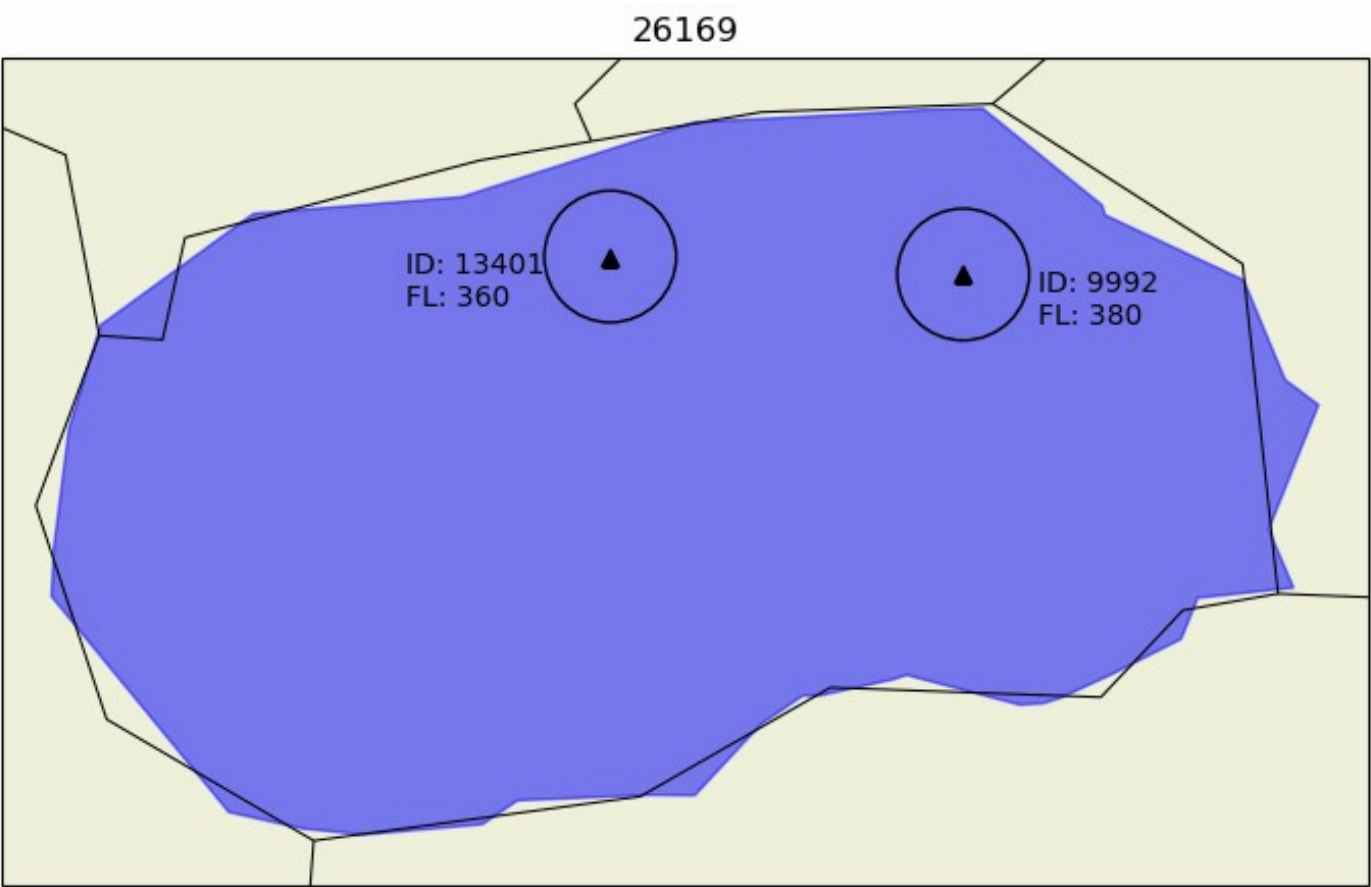
Why aircraft interactions?

- Relevant in aviation:
 1. Safety
 2. Efficiency

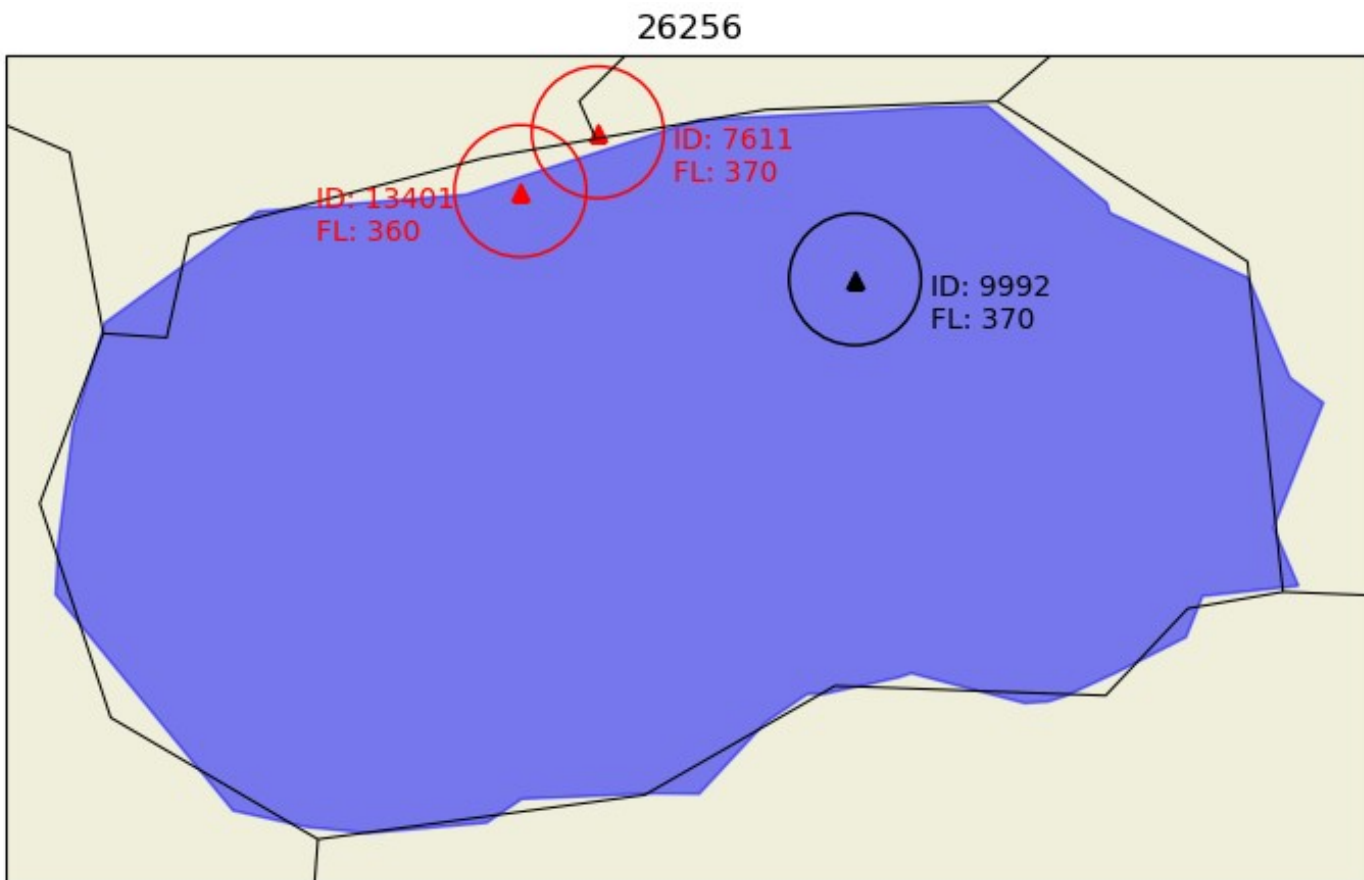




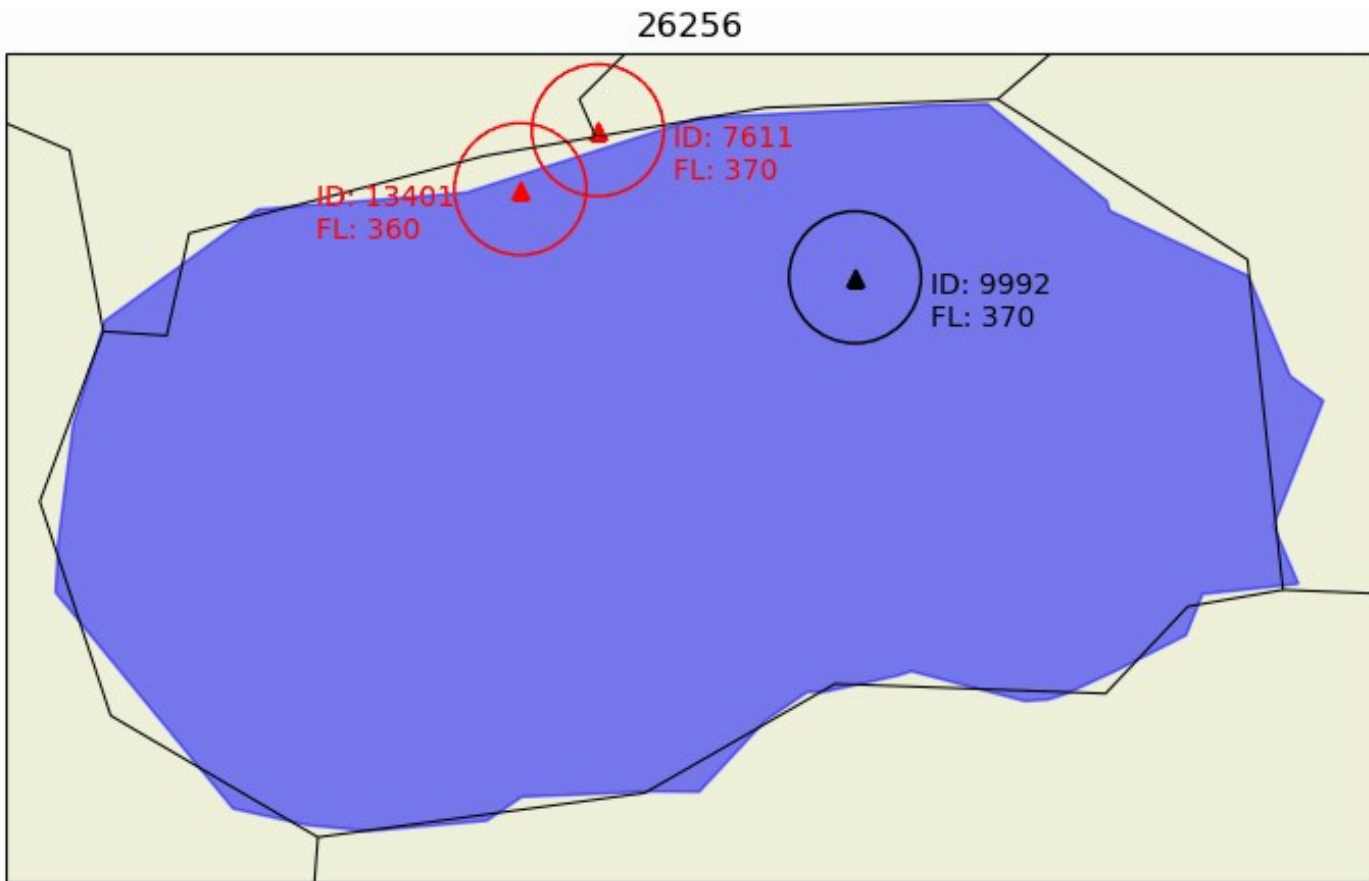
Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
-------------	-------------	----------	---------------	---------------	--------



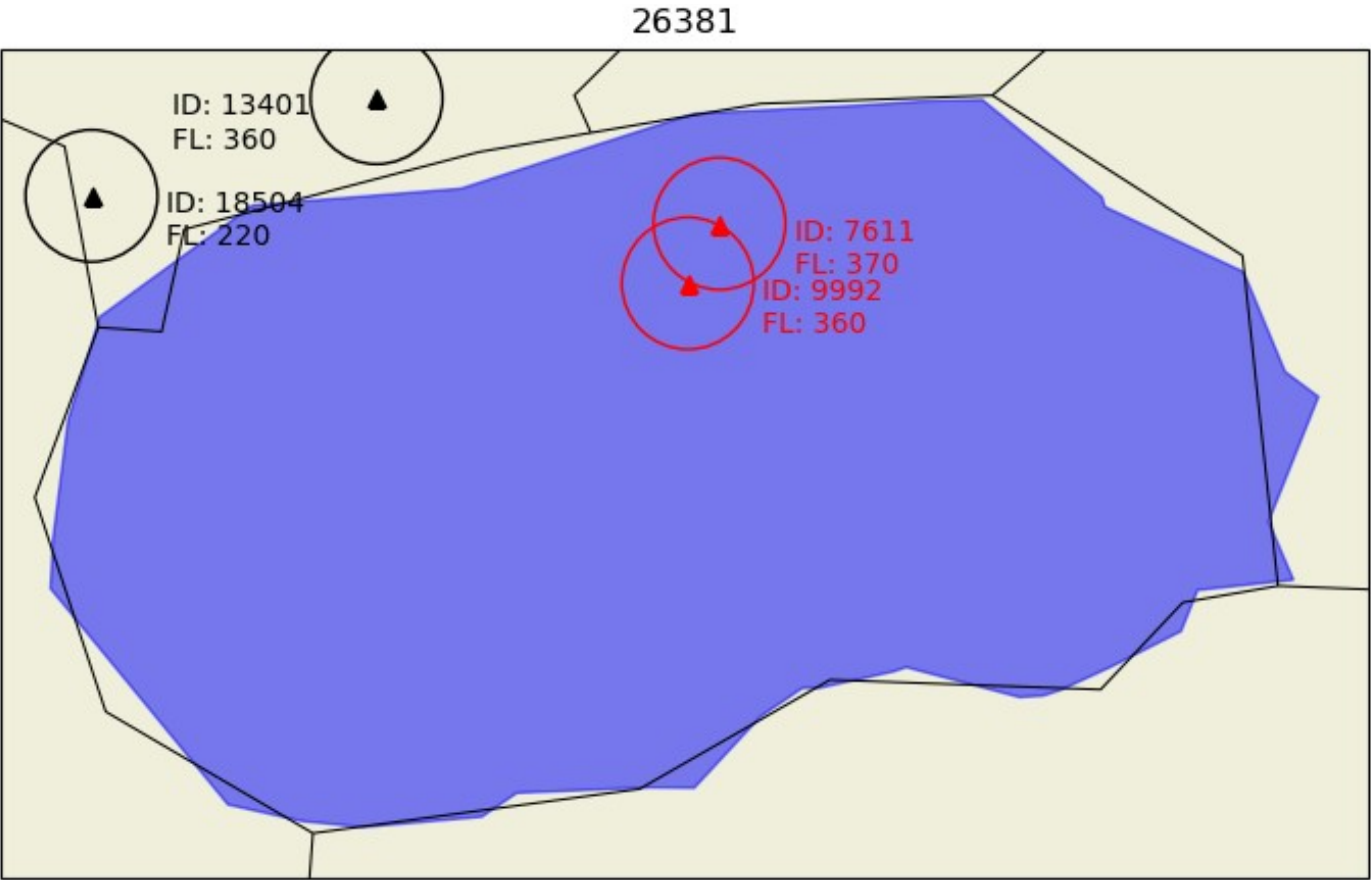
Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
-------------	-------------	----------	---------------	---------------	--------



Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
7611	13401	26256	9.868	360	65



Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
7611	13401	26256	9.868	360	65



Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
7611	13401	26256	9.868	360	65
7611	9992	26381	7.714	370	65

Total dataset size

- **854 days analysed**
Spanning from 2015-2021, for all days available at EUROCONTROL's R&D Data Archive [1]
- **Both executed and planned trajectories analysed**
Given in separated files
- **3,416 files in total**

[1] <https://www.eurocontrol.int/dashboard/rnd-data-archive>

Daily size

For each day, interactions are given for:

- ~24 hours of flights...
- ...flying over a vast area of the European airspace
- Classification of events in +80 FIRs

On average:

- 22,000 flights per day
- 49,000 interactions per day



Interaction files:

- **Flights involved**

Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
7611	13401	26256	9.868	360	65
7611	9992	26381	7.714	370	65

Interaction files:

- Flights involved
- Time

Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
7611	13401	26256	9.868	360	65
7611	9992	26381	7.714	370	65

Interaction files:

- Flights involved
- Time
- **Minimum horizontal distance**

Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
7611	13401	26256	9.868	360	65
7611	9992	26381	7.714	370	65

Interaction files:

- Flights involved
- Time
- Minimum horizontal distance
- **Altitude**

Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
7611	13401	26256	9.868	360	65
7611	9992	26381	7.714	370	65

Interaction files:

- Flights involved
- Time
- Minimum horizontal distance
- Altitude
- **FIR**

Flight ID 1	Flight ID 2	Time (s)	Distance (NM)	Altitude (FL)	FIR ID
7611	13401	26256	9.868	360	65
7611	9992	26381	7.714	370	65

Metadata files:

- **Number of flights**

FIR ID	Num. flights	Distance (NM)	Time (s)	Entropy	Entropy (norm. area)
-1	18958	6.3×10^6	8.6×10^7	13.987	0.844
65	230	1.6×10^4	1.4×10^5	8.040	0.790

For all the airspace
&
Each individual FIR

Metadata files:

- Number of flights
- Flown distance

FIR ID	Num. flights	Distance (NM)	Time (s)	Entropy	Entropy (norm. area)
-1	18958	6.3×10^6	8.6×10^7	13.987	0.844
65	230	1.6×10^4	1.4×10^5	8.040	0.790

For all the airspace
&
Each individual FIR

Metadata files:

- Number of flights
- Flown distance
- **Flown time**

FIR ID	Num. flights	Distance (NM)	Time (s)	Entropy	Entropy (norm. area)
-1	18958	6.3×10^6	8.6×10^7	13.987	0.844
65	230	1.6×10^4	1.4×10^5	8.040	0.790

For all the airspace
&
Each individual FIR

Metadata files:

- Number of flights
- Flown distance
- Flown time
- Entropy
- Area-normalised entropy

FIR ID	Num. flights	Distance (NM)	Time (s)	Entropy	Entropy (norm. area)
-1	18958	6.3×10^6	8.6×10^7	13.987	0.844
65	230	1.6×10^4	1.4×10^5	8.040	0.790

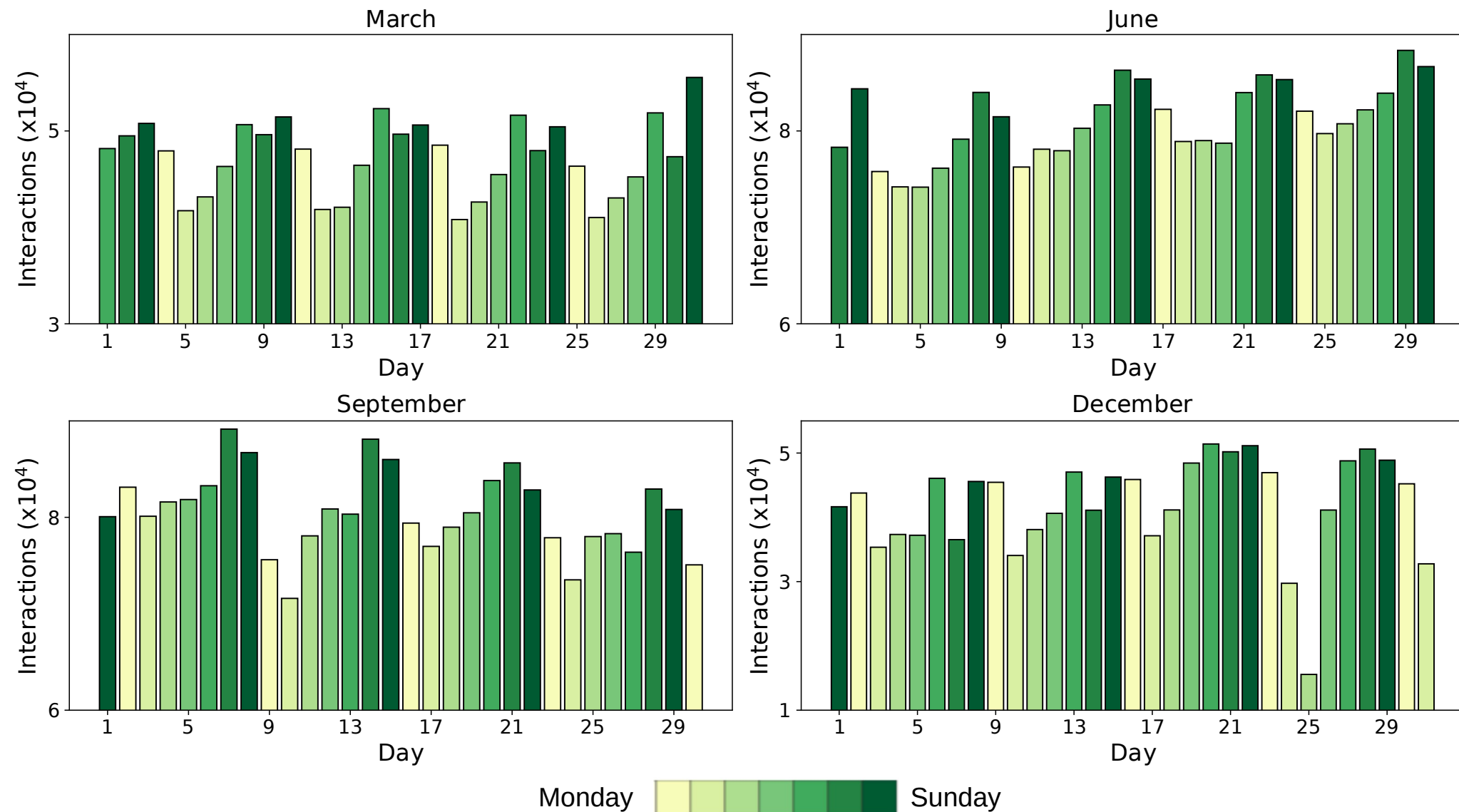
For all the airspace
&
Each individual FIR

Extended information*:

- Exact longitude and latitude of the events
- Flights extra information
- UTC time
- FIR names

Validation:

- Unit testing
- Study of discarded data
- Correlation found between reported LoS and interactions



Dataset is able to capture...

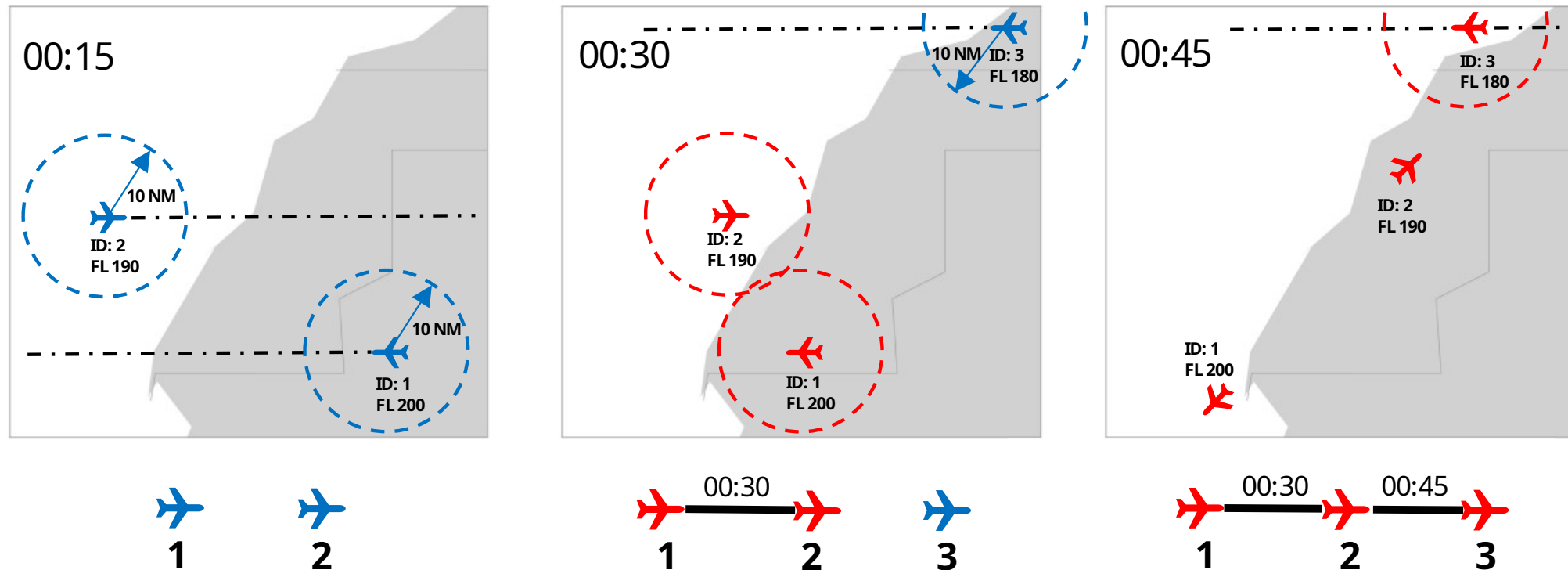
- Changes in months
- Weekly trends
- Outliers (i.e. Christmas day)

How can this dataset be useful?

1. Extend existing methodologies

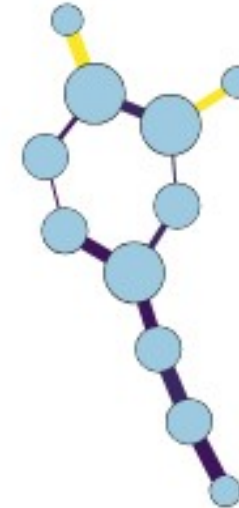
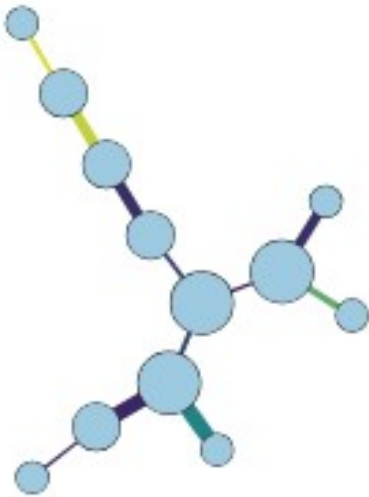
➤ **Modeling and Feature Analysis of Air Traffic Complexity Propagation by H. Wang et. al. [2]**

- Reconstructs interaction temporal networks to find aircraft's complexity.
- Range under analysis: 2 days over Shanghai region.



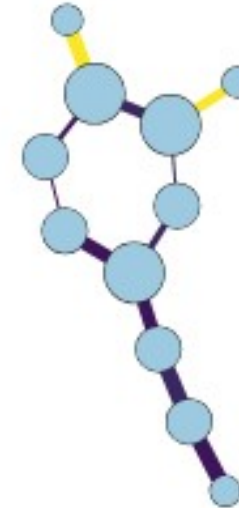
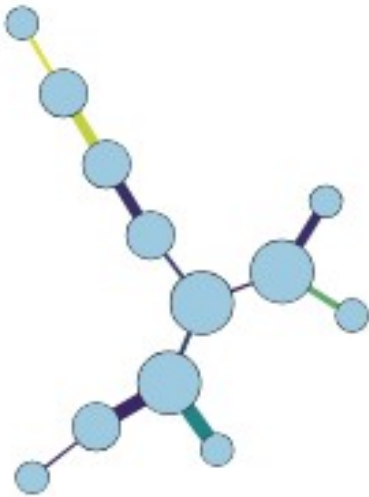
[2] H. Wang, P. Xu, and F. Zhong, "Modeling and feature analysis of air traffic complexity propagation," Sustainability, vol. 14, no. 18, 2022.

2. Identify relevant air traffic configurations



2. Identify relevant air traffic configurations

For better understanding ATCs workload

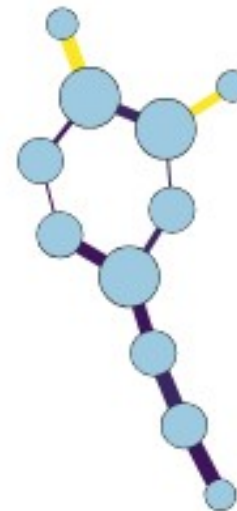
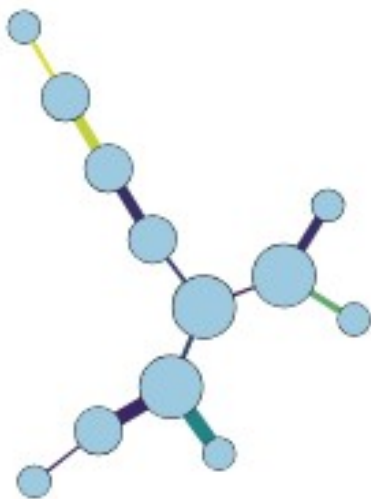


[3] J. Y. Jingyu Zhang and C. Wu, "From trees to forest: relational complexity network and workload of air traffic controllers," *Ergonomics*, vol. 58, no. 8, pp. 1320–1336, 2015.

[4] J. Zhang, X. E, F. Du, J. Yang, and S. Loft, "The difficulty to break a relational complexity network can predict air traffic controllers' mental workload and performance in conflict resolution," *Human Factors*, vol. 63, no. 2, pp. 240–253, 2021.

2. Identify relevant air traffic configurations

For evaluating CD&R algorithms



3. Study air traffic dynamics

- Interactions can be understood as an indicator of how close aircraft fly on average

3. Study air traffic dynamics

- Interactions can be understood as an indicator of how close aircraft fly on average
- Through interactions we can understand the structure of air traffic dynamics
 1. Analysis of interaction networks structure as a function of air traffic variables
 2. Observation of changes in its structure

[6] R. López-Martín and M. Zanin, "Large-scale analysis of trajectory interaction networks in Europe," in SESAR Innovation Days 2024, 2024

[7] R. López-Martín and M. Zanin, "Propagation of interactions among aircraft trajectories: A complex network approach," Aerospace, vol. 10, no. 3, 2023.

How can this dataset be useful?

1. Extend existing methodologies
2. Identify relevant air traffic configurations
3. Study air traffic dynamics

How can this dataset be useful?

1. Extend existing methodologies
2. Identify relevant air traffic configurations
3. Study air traffic dynamics
4. Your next work 😊

Access here this dataset:

<https://doi.org/10.5281/zenodo.15017762>



Past works that use this dataset:

[6] R. López-Martín and M. Zanin, “Large-scale analysis of trajectory interaction networks in Europe,” in SESAR Innovation Days 2024, 2024

[7] R. López-Martín and M. Zanin, “Propagation of interactions among aircraft trajectories: A complex network approach,” Aerospace, vol. 10, no. 3, 2023.

Get in contact with us!

raullopez@ifisc.uib-csic.es

mzanin@ifisc.uib-csic.es

Acknowledgements:



ARCTIC
No 851255



FPU22/03765



CEX2021-
001164-M

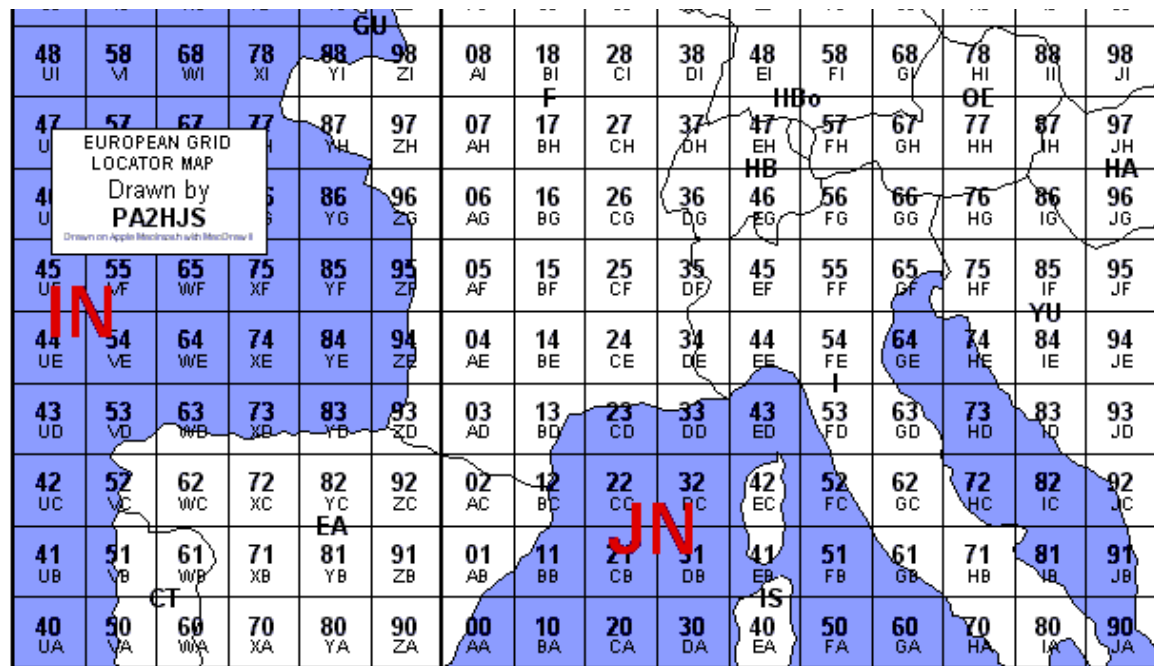
Supplementary Slides

Computational complexity

To obtain this dataset a total of **8.88×10^{12} pairs of flights** were analysed.

Entropy

Matrix with the number of times
each cell has been flown over



Source: <http://www.telesupport.nl/wave/locator.htm>

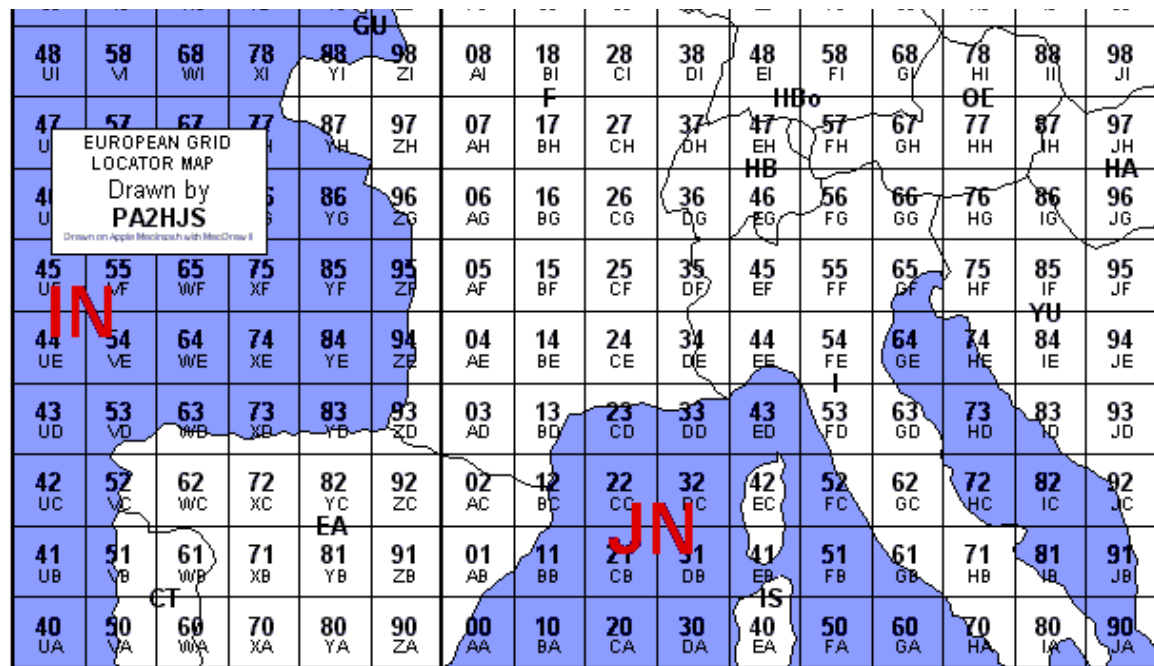
$$N = \begin{pmatrix} 10 & 34 & \dots \\ 15 & 7 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$\downarrow / \sum_{i,j} N_{ij}$$

$$p = \begin{pmatrix} 0.002 & 0.0044 & \dots \\ 0.003 & 0.0016 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$S = - \sum_{i,j} p_{ij} \cdot \ln p_{ij}$$

Area-normalised Entropy



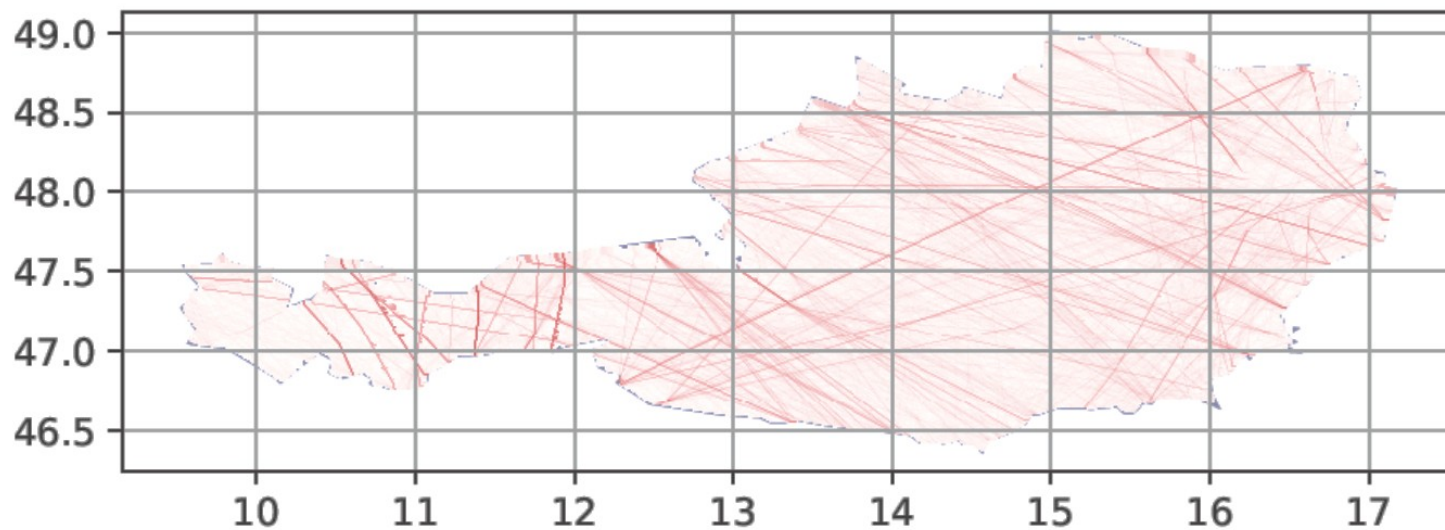
Source: <http://www.telesupport.nl/wave/locator.htm>

$$\bar{S} = \frac{S}{\ln M}$$

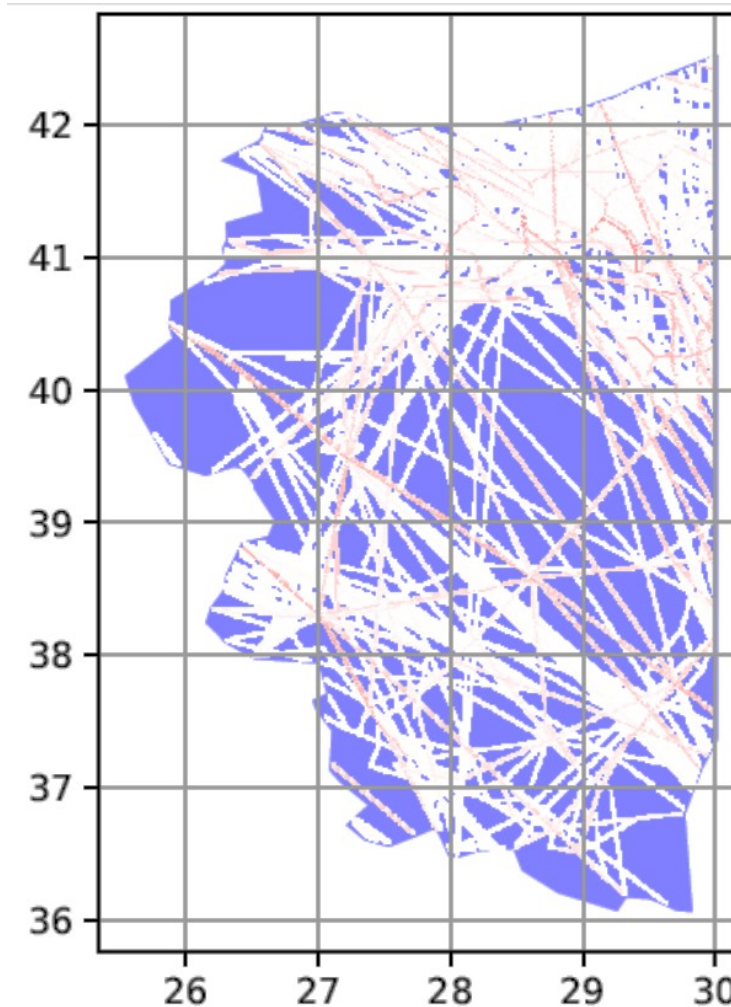
M : Number of cells in the grid

Entropy

LOVVFIR – High entropy



LTBBFIR – Low entropy



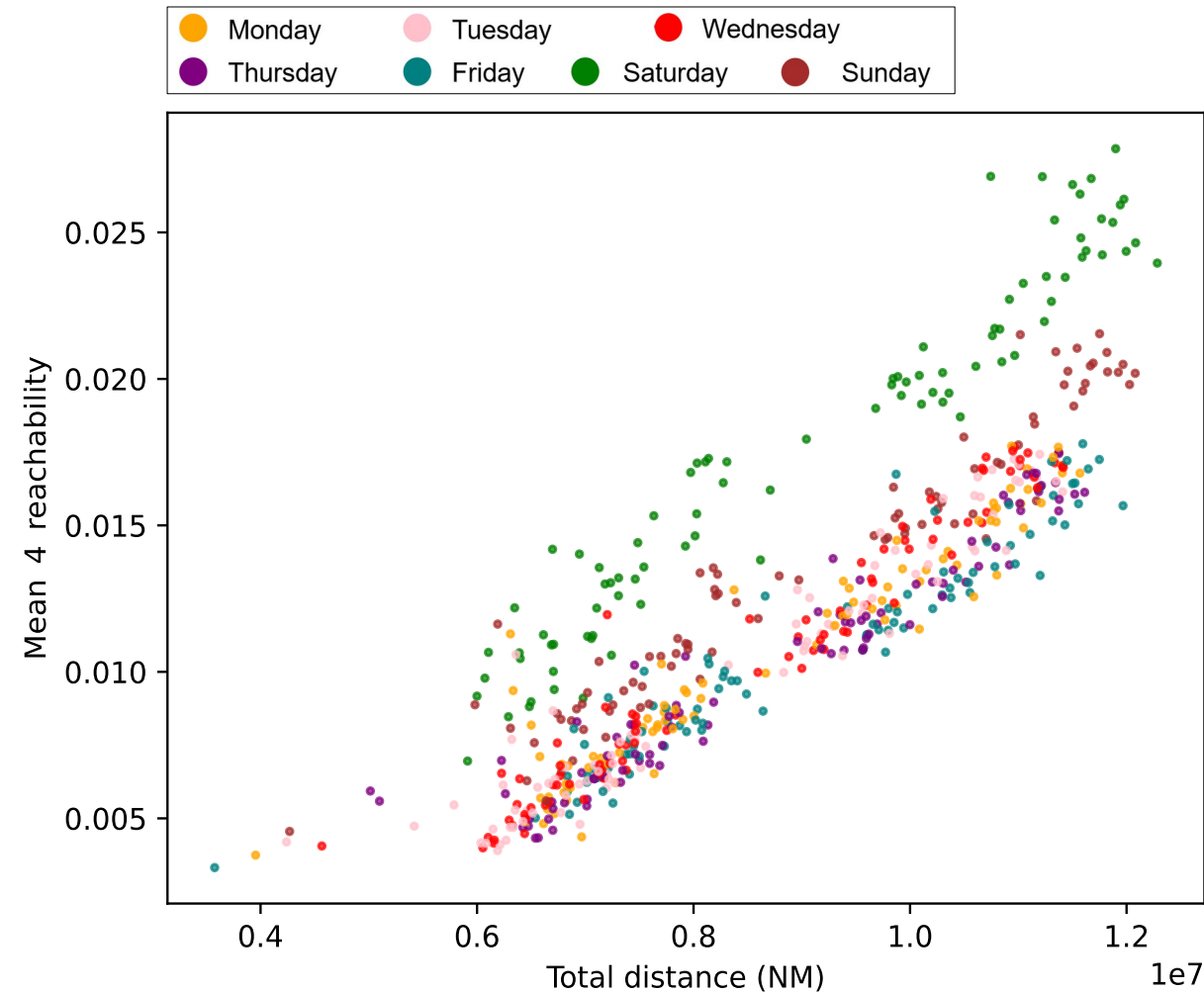
Validation

Trajectory type	# flights	# pre-interpolation deleted flights	# post-interpolation deleted flights	# sub FL100 flights
Executed	34421	98	1618	499
Planned	34421	117	1610	443

Information about number of flights deleted during the pre-processing of the trajectory data, for June 28th, 2019.

3. Study air traffic dynamics

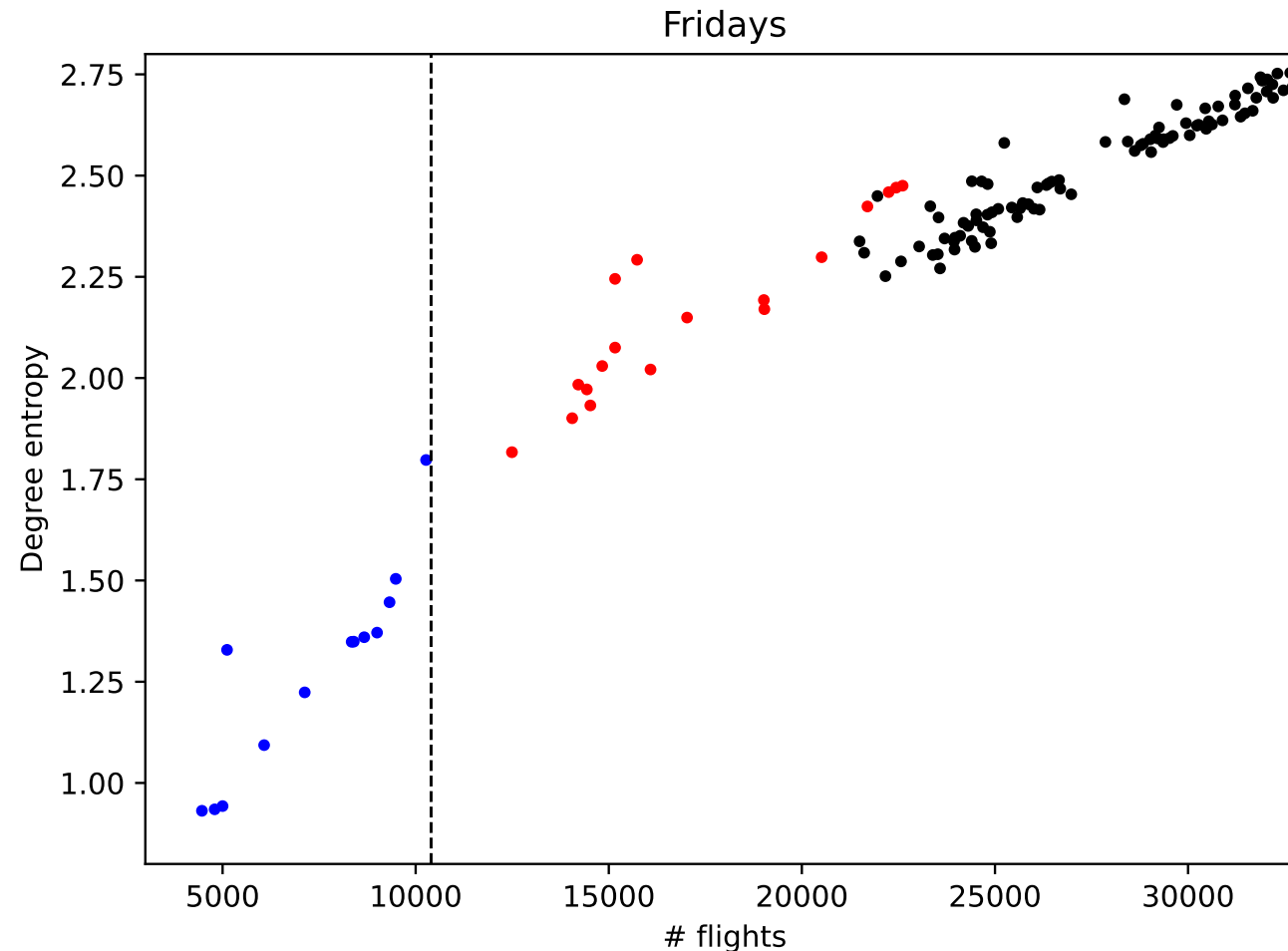
- Analysis of interaction networks structure as a function of air traffic variables



[6] R. López-Martín and M. Zanin, "Large-scale analysis of trajectory interaction networks in Europe," in SESAR Innovation Days 2024, 2024

3. Study air traffic dynamics

- Observation of changes in interaction networks' structure



[6] R. López-Martín and M. Zanin, "Large-scale analysis of trajectory interaction networks in Europe," in SESAR Innovation Days 2024, 2024



Article

Modeling and Feature Analysis of Air Traffic Complexity Propagation

Hongyong Wang ^{1,*}, Ping Xu ¹ and Fengwei Zhong ²

¹ College of Air Traffic Management, Civil Aviation University of China, Tianjin 300300, China
² Air Traffic Management Bureau of China Civil Aviation Administration, Beijing 100015, China
* Correspondence: hy_wang@cauc.edu.cn; Tel.: +86-139-2081-9396

5. Case Validation

With the real radar data from two days (1–2 November 2019) in the Shanghai region, the validation process is shown below.

modeling is proposed. First, a temporal network is built with aircraft as nodes and between-aircraft proximity relations as edges. Second, the disease propagation model is introduced to simulate the evolution course of between-aircraft proximity relations, and the propagation model is solved using Runge–Kutta algorithm and particle swarm optimization. Third, based on the solved results of the propagation model, the aircraft are divided into three groups with high, medium, and low propagation capability, respectively. Finally, the effects of different factors on the propagation course

Table 1. Results of K-means algorithm.

	Q1		Q2		Q3	
	DAY1	DAY2	DAY1	DAY2	DAY1	DAY2
Clustering center of infection rate	0.6891	0.7016	0.4312	0.428	0.4622	0.4867
Clustering center of recovery rate	0.2933	0.3042	0.1687	0.1571	0.6901	0.7092
Aircraft number	1231	1126	2317	1926	278	242
Proportion of aircraft number	32%	34%	61%	58%	7%	8%