EXPLORATORY RESEARCH

# Data Management Plan

| | |
|---|---|
| **Deliverable ID:** | **D8.4** |
| **Dissemination Level:** | **PU** |
| **Project Acronym:** | **FMP-Met** |
| **Grant:** | **885919** |
| **Call:** | **H2020-SESAR-2019-2** |
| **Topic:** | **SESAR-ER4-05-2019 - Environment and Meteorology for ATM** |
| **Consortium Coordinator:** | **USE** |
| **Edition Date:** | **30/05/2022** |
| **Edition:** | **00.03.00** |
| **Template Edition:** | **02.00.03** |

EUROPEAN PARTNERSHIP

Co-funded by the European Union

## Authoring & Approval

### Authors of the document

| Name / Beneficiary | Position / Title | Date |
|---|---|---|
| USE | Project coordinator | 20/05/2022 |

### Reviewers internal to the project

| Name / Beneficiary | Position / Title | Date |
|---|---|---|
| AEMET | Project partner | 30/05/2022 |
| ACG | Project partner | 30/05/2022 |
| CCL | Project partner | 30/05/2022 |
| LiU | Project partner | 30/05/2022 |
| MetSol | Project partner | 30/05/2022 |
| PLUS | Project partner | 30/05/2022 |
| UC3M | Project partner | 30/05/2022 |
| ZFOT | Project partner | 30/05/2022 |

### Reviewers external to the project

| Name / Beneficiary | Position / Title | Date |
|---|---|---|
| | | |

### Approved for submission to the SJU By - Representatives of all beneficiaries involved in the project

| Name / Beneficiary | Position / Title | Date |
|---|---|---|
| USE | Project coordinator | 30/05/2022 |

### Rejected By - Representatives of beneficiaries involved in the project

| Name and/or Beneficiary | Position / Title | Date |
|---|---|---|
| | | |

### Document History

| Edition | Date | Status | Author | Justification |
|---|---|---|---|---|
| 00.00.01 | 10/10/2020 | Initial Draft | USE | New document |
| 00.01.00 | 25/10/2020 | First Issue | USE | Updated document |

| | | | | |
|---|---|---|---|---|
| 00.01.01 | 25/07/2021 | Updated Issue | USE | Update including data used in WP4 |
| 00.02.00 | 30/04/2022 | Second Issue | USE | Update including data used in WPs 5-7 |
| 00.03.00 | 30/05/2022 | Third Issue | USE | Updated document following S3JU assessment |

## Copyright Statement





EUROPEAN PARTNERSHIP

Co-funded by the European Union

# FMP-Met

METEOROLOGICAL UNCERTAINTY MANAGEMENT FOR FLOW MANAGEMENT POSITIONS

This deliverable is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 885919 under European Union's Horizon 2020 research and innovation programme.

## Abstract

In this document the Data Management Plan (DMP) of the FMP-Met project is presented. Its target audience is the SESAR Joint Undertaking and the consortium members. This plan is intended to describe the data management life cycle for all datasets to be collected, processed or generated by the project. The data management policy is based on making the data Findable, Accessible, Interoperable and Reusable (FAIR data management scheme). Since the FMP-Met project will not generate any research data, the focus of the DMP is on the use of pre-existing research data. In particular, the DMP addresses the data needed to validate the results presented in the scientific publications produced within the project.

# Table of Contents



# List of Figures



# List of Tables

# 1 Introduction[1]

In this document, the Data Management Plan (DMP) of the project entitled 'Meteorological Uncertainty Management for Flow Management Positions — FMP-Met' is presented. This plan elaborates further the initial version of the DMP presented in the Grant Agreement (GA) [1] (Section 2.2.3 of Annex 1 – Part B). The DMP is primarily intended for the SJU and the consortium members participating in the project (USE, AEMET, ACG, CCL, LiU, MetSol, PLUS, UC3M, ZFOT).

According to the Guidelines on FAIR Data Management in H2020 [2], the Data Management Plan (DMP) is intended to describe the data management life cycle for all datasets to be collected, processed or generated by the project. The DMP should address, at least, the data needed to validate the results presented in scientific publications, including information on:

- the handling of research data during and after the end of the project,
- what data will be collected, processed and/or generated,
- what methodology and standards will be applied,
- whether data will be shared /made open access, and
- how data will be curated and preserved (including after the end of the project).

The FMP-Met project does not participate in the extended Open Research Data Pilot; however, the delivery of a DMP was foreseen in the Grant Agreement on a voluntary basis, because a DMP is a key element of good data management, and implementing a good data management is considered to be a research best practice.

This DMP describes a data management policy based on making the research data Findable, Accessible, Interoperable and Reusable (FAIR data management scheme) in order to enhance knowledge discovery and innovation, and subsequent data and knowledge integration and reuse. Furthermore, the data management policy is consistent with exploitation and Intellectual Property Rights requirements, as stated in the Consortium Agreement [3].

The data handled in the FMP-Met project can be in principle classified into two different categories: research data generated within the project and research data used within the project. According to this classification, this document is organized as follows. A short explanation on generated research data is included in Section 2, whereas the used research data is described in Section 3. Some provisions for updating the DMP during the FMP-Met project life cycle are given in Section 4. The input data used in the project will be collected and identified in the Appendix, which will be updated accordingly.

[1] 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.

## 1.1 Acronyms and Terminology

| Acronym | Description |
|---|---|
| AIRAC | Aeronautical Information Regulation And Control |
| ANSP | Air Navigation Service Provider |
| ATM | Air Traffic Management |
| BADA | Base of Aircraft Data |
| DMP | Data Management Plan |
| DWD | Deutscher Wetterdienst |
| ECMWF | European Centre for Medium-Range Weather Forecasts |
| EPS | Ensemble Prediction System |
| FAIR | Findable, Accessible, Interoperable and Reusable |
| FMP | Flow Management Position |
| GA | Grant Agreement |
| ICAO | International Civil Aviation Organization |
| MARS | Meteorological Archive and Retrieval System |
| MET | Meteorology |
| SAF | Satellite Application Facility |
| SESAR | Single European Sky ATM Research Programme |
| SJU | SESAR Joint Undertaking |
| WP | Work Package |

## 1.2 FMP-Met Consortium

| Acronym | Description |
|---|---|
| USE | Universidad de Sevilla |
| AEMET | Agencia Estatal de Meteorología |
| ACG | Austro Control GmbH |
| CCL | Croatia Control Limited |
| LiU | Linköping University |
| MetSol | MeteoSolutions GmbH |
| PLUS | Paris-Lodron Universität Salzburg |
| UC3M | Universidad Carlos III de Madrid |
| ZFOT | University of Zagreb |

# 2 Generation of Research Data

The FMP-Met project will not generate any research data, but it will develop methodologies to provide probabilistic predictions of sector demand, sector complexity, and sector capacity reduction under convective weather for a forecasting horizon of 8 hours. In fact, the expected outcome of the project is the development of a methodology to provide Flow Management Positions (FMP) with an intuitive and interpretable probabilistic assessment of the impact of convective weather on the traffic, up to 8 hours in advance, to allow better-informed decision making, hence improving the decision-making process in traffic flow management under convective weather.

Additionally, a methodology to generate a probabilistic nowcast for convective weather will be developed in this project. This development is required in order to have the appropriate probabilistic MET input to obtain the traffic predictions mentioned above.

Accordingly, open access to research data is not required in FMP-Met (as stated in Article 29.3 of the GA [1]).

# 3 Use of Research Data

The FMP-Met project will make use of the following pre-existing research data:

a) Meteorological (MET) data (provided by MET Offices)
   - Ensemble Prediction Systems (EPS)
   - Nowcast data

b) Aircraft model data (provided by EUROCONTROL)
c) Air traffic data (provided by EUROCONTROL)
d) Airspace data (provided by Austro Control and Croatia Control)

## 3.1 Data Description

Data used (EPS, Nowcast, aircraft model, air traffic and airspace) constitute the input to the project, as they are needed to define the methodologies described in Section 2.

### EPS data

Ensemble Prediction Systems are short-range forecasts which provide probabilistic convective weather predictions. An EPS is a collection of typically 10 to 50 weather forecasts (referred to as members) with a common valid time, which can be obtained using different Numerical Weather Prediction models based on time-lagged, multi-model, multi-initial conditions and/or stochastic physics approaches. The EPS provides a set of deterministic forecasts and, from them, the probabilistic information is derived.

From EPS we will obtain the uncertainty of wind and temperature fields and also the uncertainty in the location of convective areas, which are areas where the necessary conditions for having convective activity are met. The wind and temperature fields are directly extracted from the EPS. In contrast, the convective weather intensity must be derived from a set of atmospheric variables provided by the members of the EPS (like instability indices, convective available potential energy, convective precipitation, and lightning flash density).

Two kinds of EPS will be used in FMP-Met, with a lead time of 1 to 8 hours:

- global EPS from the European Centre for Medium-Range Weather Forecasts [4], ECMWF-EPS, which has 50 members, for big areas,
- two different limited-area high-resolution EPSs for smaller areas:
  - COSMO-D2 (20 members) – from DWD (Deutscher Wetterdienst),
  - gSREPS (20 members) – from AEMET.

The ECMWF-EPS will be retrieved from the MARS (Meteorological Archive and Retrieval System) [5] database. The data will be downloaded as files in GRIB format which contain meteorological parameters on a regular latitude-longitude grid for the desired forecast times. The data will be extracted from the model grid to cover only the desired analysis region in time and space for the flights to be examined. The region to be extracted is defined by the minimum and maximum of latitude/longitude and the pressure level where the flights will take place.

The COSMO-D2 EPS will be provided by DWD or Fraunhofer IEE. Depending on the dates of the use cases there are two options: Historic data may be provided by DWD if not older than 15 months from the day of data base retrieval. This is free of charge. Older data can be provided by Fraunhofer IEE. Here a provision fee is due. Unfortunately, they cannot provide humidity which affects the calculation of the convection parameter (Totals-Totals-Index).

The data is provided in GRIB format. We will use all available model runs (00, 03, 06, 09, 12, 15, 18, 21 UTC), forecast times up to 8 hours (with outputs every hour), a horizontal resolution of 0.02º x 0.02º in a rotated grid, with the complete model coverage. Similarly, the data from gSREPS will be provided by AEMET, with the same format, resolution and forecast times, covering the area around the Iberian Peninsula.

### Nowcast data

Nowcast systems work on the regional scale and provide forecasts of convective weather with a lead time of up to 1 hour. The information consists mainly of forecasts of the area with embedded convective cells, and, for the individual cells, their positions, extents, strengths, and cloud heights.

The probabilistic nowcast to be developed in this project must provide a set of possible convective weather situations, which will be used to derive the uncertainty of the individual convective cells. This ensemble nowcast will have 15 members. The methodology used will be based on radar and satellite observations. For the European area the following data sources will be used:

- OPERA radar composites from EUMETNET (instantaneous surface rain rate),
- satellite data from EUMETSAT (convective rainfall intensity and cloud top height).

In addition to the ensemble of convective cells, we will use deterministic forecasts of top cloud heights derived from the cloud top height satellite product.

The OPERA radar composites will be retrieved from Odyssey [6], the OPERA Data Centre, that generates and archives composite products from raw single site radar data using common pre-processing and compositing algorithms. The data is obtained in HDF5 format, with Lambert azimuthal equal-area projection and a time resolution of 15 minutes.

The satellite products will be provided by AEMET. They are obtained after processing the raw Meteosat Second Generation data using the software and algorithms developed by the EUMETNET SAF (Satellite Application Facility) project [7]. The data is in HDF5 format, with Geostationary Satellite View projection and a time resolution of 15 minutes.

### Aircraft model data

The aircraft model data will be retrieved from the Base of Aircraft Data (BADA), from EUROCONTROL [8]. The main application of BADA is trajectory simulation and prediction. Although detailed data on aircraft performance is confidential and commercially sensitive, BADA transforms it so that aircraft models can be made available to a wide audience. BADA is made of two components: The model specifications, which provide the theoretical fundaments used to calculate aircraft performance parameters, and the datasets containing the aircraft-specific coefficients necessary to perform calculations, which includes:

- Aircraft operating parameters.
- Aerodynamic model.
- Fuel consumption model.
- Available thrust model.

The use of BADA 4 (the version we intend to use in FMP-Met) is regulated through a license agreement which stipulates the terms and conditions of use based on the contractual constraints EUROCONTROL is committed to with its data providers. The provision of access is subject to EUROCONTROL's approval. The agreement and its conditions are accessible through the BADA Users Interface. The agreement does not include any constraint regarding the publication of results from these data. UC3M (the only partner to use BADA 4 in FMP-Met) will request access to BADA 4 data.

### Air traffic data

Historical air traffic data will be retrieved from EUROCONTROL's service R&D Data Archive, available on the OneSky Online extranet. This service is offered by Statfor [9]. The traffic datasets cover historical commercial flights in four fixed, sample months (March, June, September and December) of specific years (from 2015 onwards). The data are released with a delay of two years. Flight data is filtered to include flights of ICAO flight types 'S' (scheduled) and 'N' (non-scheduled flight), excluding ICAO types General aviation, Military and Other.

For each flight, two trajectories are accessible: the trajectory computed from the last filed flight plan, and the actually flown trajectory. Each trajectory is given as a sequence of points including latitude, longitude, flight level, and time. The trajectory computed from the last filed flight plan is generated by Network Manager's Air Traffic Flow Management systems from the flight plans submitted by airlines and other aircraft operators to the Network Manager. The actual flown trajectory includes some updates from radar observation of the flight's path.

According to the Terms and Conditions of the R&D Data Archive service, the datasets are available only for research and development purposes. Access to these datasets has been granted to USE.

### Airspace data

The airspace data comprise a description of the airspace organization (geographical, operational and procedural), and basic civil aviation structure data (routes and significant points). In particular, it includes:

- Geometry of the traffic volumes (lateral and vertical boundaries).
- Nominal monitoring values of the traffic volumes.
- Entry and exit points to each traffic volume, and route structure inside the traffic volume.
- Available sector configurations and opening schemes.

This information will be provided by the ANSPs participating in the project, Austro Control and Croatia Control. Each ANSP will provide the data corresponding to its airspace (Austria and Croatia, respectively).

## 3.2 FAIR data management scheme

### Making data findable

In order to make the data used in the project identifiable and locatable, unique identifiers must be defined. In particular, the following is proposed:

- EPS and nowcast datasets will be identified by a string containing the following attributes: name, issuing office, date, delivery time, time step, coverage area, spatial grid resolution, barometric altitude, and variable name.
- Datasets from BADA will be identified by the aircraft code and the BADA version.
- Historical traffic datasets retrieved from R&D Data Archive will be identified by the common airspace entered by the flights (Austrian or Croatian flight information regions, LOVVFIR or LDZOFIR, respectively) and the time window they enter this airspace.
- Datasets for airspace data will be identified by the airspace to which they correspond (LOVVFIR or LDZOFIR) and the Aeronautical Information Regulation And Control (AIRAC) cycle for which they are effective.

### Making data accessible

Each dataset used in the FMP-Met project is obtained from a pre-existing data base, whose access is restricted to registered users. Some FMP-Met project partners have been granted access to those data bases, but they have accepted terms and conditions of use including non-disclosure clauses. Therefore, the datasets cannot be made openly available or shared. However, obtaining access for this databases is not difficult (at least for researchers in ATM community), and therefore there is good data accessibility.

### Making data interoperable

To facilitate interoperability of the data used, standard vocabulary from ATM and MET disciplines will be used throughout the project. No uncommon nor project specific ontologies or vocabularies will be generated. Moreover, interoperability is supported by the use of common data formats (HDF5, netCDF, GRIB, JSON).

### Making data reusable

There are no restrictions in data re-use by third parties, other than the fact that these third parties must have granted access to the aforementioned pre-existing data bases.

## 3.3 Data storage

As already indicated in Section 2, the FMP-Met project will not generate any research data.

The input data used in the project have been stored in two private repositories, managed by PLUS and UC3M, to be used by FMP-Met partners. Using two different repositories make the storage more secure, being one back-up for the other. These data are not subject to further updates.

All the input data, and the pre-existing data bases where they come from, are clearly identified in the Appendix, so that they could be retrieved/purchased by any researcher.

The used data will be kept in the repositories at least during two years after the end of the project. After that time, to obtain the data, anyone interested will have to retrieve/purchase them from the original data bases.

# 4 Updating the Data Management Plan

Following the Guidelines on FAIR Data Management in H2020 [2], updates in the DMP are foreseen as new input data are used in the project. This does not exclude the possibility of updating the DMP whenever unexpected significant changes arise. The DMP will be updated in time with the periodic reviews of the project (intermediate and final reviews). Hence, the timetable for DMP review can be summarized as follows (Table 1):

**Table 1. DMP Updating Timetable**

| Updating cause | Due date |
|---|---|
| Include data used in WP4 | T0+15 (31/07/2021) |
| Include data used in WP5, WP6 and WP7 | T0+24 (30/04/2022) |

### Updates

The input research data used in the project will be collected and identified in the Appendix, according to the timetable given above.

# 5 References

[1] Grant Agreement number: 885919 — FMPMet — H2020-SESAR-2019-2, 2020.

[2] European Commission, «Guidelines on FAIR Data Management in Horizon 2020,» July 2016. [On line]. Available: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

[3] FMP-Met Consortium Agreement, 15 April 2020.

[4] http://www.ecmwf.int/en/research/modelling-and-prediction/atmospheric-physics

[5] https://confluence.ecmwf.int/display/UDOC/MARS+user+documentation

[6] https://www.eumetnet.eu/activities/observations-programme/current-activities/opera/

[7] https://www.eumetsat.int/website/home/Satellites/GroundSegment/Safs/SupporttoNowcastingandVeryShortRangeForecasting/index.html

[8] EUROCONTROL, «Base of Aircraft Data (BADA),» [On line]. Available: http://www.eurocontrol.int/services/bada.

[9] EUROCONTROL /Statfor, "R&D Data Release – Metadata", version 0.4, February 2020.

# Appendix – Input data

As described in Section 3, the FMP-Met project will make use of the following types of input data:

a) Meteorological (MET) data
b) Aircraft (AC) model data
c) Air traffic data
d) Airspace data

In this Appendix the data used in the project (in WPs 4, 5, 6 and 7) are identified.

## 1. Input data used in WP4

A description of the data used in WP4 is provided in Table 2. Note that Airspace Data has not been used in WP4. In this table reference is made to the sections of Deliverable D4.1, where the different data are used.

**Table 2. Input data used in WP4**

| | |
|---|---|
| **MET data #1** | **Data type:** MET data (Rain and cloud data).<br>**Purpose:** Characterizing operational uncertainty (uncertainty in the avoidance strategy, Section 5.2) and performing two illustrative examples of trajectory prediction through the unified framework (Sections 7.2 and 7.4). The first one is a flight from LFPG airport to LORP airport, whereas the second one is a segment of a flight from KJFK airport to OMAA airport, beginning at (55.00 N, 16.60 W) and ending at (46.16 N, 20.40 E).<br>**Description of the data:** Radar and satellite observations from Opera RADAR and SAF EUMETNET Satellite information (including Cloud Top Height, CTH). The meteorological data corresponds to 27th July 2019.<br>**Source of the data:** Radar data retrieved from the Odyssey database; satellite data retrieved from AEMET database.<br>**Data format:** hdf5 file. |
| **MET data #2** | **Data type:** MET data (EPS data).<br>**Purpose:** Performing an illustrative example of EPS-based trajectory prediction for a flight from LEMD airport to EDDT airport (Section 4.2.6).<br>**Description of the data:** Wind speed and temperature fields (U, V, and T) at several pressure levels (200, 300, 400, 500, 700, 850, and 1000 hPa) from ECMWF-EPS. The meteorological data corresponds to 27th July 2019 at 12:00, with forecast steps from 3 to 6 hours. Temporal interpolation was applied to increase the time resolution of the data.<br>**Source of the data:** Retrieved from ECMWF's MARS database.<br>**Data format:** GRIB file. |
| **MET data #3** | **Data type:** MET data (EPS data).<br>**Purpose:** Performing three illustrative examples. The first one is an EPS-based trajectory prediction for a segment of a flight from KJFK airport to OMAA airport (Section 4.2.7); the piece of trajectory begins at (55.00 N, 15.00 W) and ends at |

| | |
|---|---|
| | (46.26 N, 21.10 E). The second one is a trajectory prediction through the unified framework for a flight from LFPG airport to LORP airport (Section 7.2). The third one is a trajectory prediction through the unified framework for a segment of a flight from KJFK airport to OMAA airport; this second piece of trajectory begins at (55.00 N, 16.60 W) and ends at (46.16 N, 20.40 E) (Section 7.4).<br>**Description of the data:** Wind speed, temperature, convective precipitation, and Total Totals fields (U, V, T, CP, and TT) at several pressure levels (200, 300, 400, 500, 700, 850, and 1000 hPa) from ECMWF-EPS. A convective indicator was calculated based on CP and TT. The meteorological data corresponds to 27th July 2019 at 00:00, with forecast steps from 15 to 21. Temporal interpolation was applied to increase the time resolution of the data.<br>**Source of the data:** Retrieved from ECMWF's MARS database.<br>**Data format:** GRIB file. |
| **MET data #4** | **Data type:** MET data (EPS data).<br>**Purpose:** Performing two illustrative examples of trajectory prediction through the unified framework. The first one is a flight from LFPG airport to LORP airport (Section 7.2), whereas the second one is a segment of a flight from KJFK airport to OMAA airport, beginning at (55.00 N, 16.60 W) and ending at (46.16 N, 20.40 E) (Section 7.4).<br>**Description of the data:** Wind speed, temperature, total precipitation, and Total Totals fields (U, V, T, TP, and TT) at several pressure levels (200, 300, 400, 500, 700, 850, 950, 975, and 1000 hPa) from COSMO-D2-EPS. A convective indicator was calculated based on TP and TT. The meteorological data corresponds to 27th July 2019 at 12:00, with forecast steps from 3 to 6 hours. Temporal interpolation was applied to increase the time resolution of the data.<br>**Source of the data:** Retrieved from DWD public server.<br>**Data format:** GRIB file. |
| **AC data #1** | **Data type:** Aircraft model data.<br>**Purpose:** Performing two illustrative examples. The first one is an EPS-based trajectory prediction for a flight from LEMD airport to EDDT airport (Section 4.2.6). The second one is a trajectory prediction through the unified framework for a flight from LFPG airport to LORP airport (Section 7.2).<br>**Description of the data:** Airbus A320 model parameters.<br>**Source of the data:** Retrieved from EUROCONTROL BADA 3.12.<br>**Data format:** APF, OPF, and PTD files. |
| **AC data #2** | **Data type:** Aircraft model data.<br>**Purpose:** Performing two illustrative examples. The first one is an EPS-based trajectory prediction for a segment of a flight from KJFK airport to OMAA airport (Section 4.2.7); the piece of trajectory begins at (55.00 N, 15.00 W) and ends at (46.26 N, 21.10 E). The second one is a trajectory prediction through the unified framework for a segment of a flight from KJFK airport to OMAA airport; this second piece of trajectory begins at (55.00 N, 16.60 W) and ends at (46.16 N, 20.40 E) (Section 7.4).<br>**Description of the data:** Airbus A380 model parameters.<br>**Source of the data:** Retrieved from EUROCONTROL BADA 3.12.<br>**Data format:** APF, OPF, and PTD files. |

| | |
|---|---|
| **Traffic data #1** | **Data type:** Air traffic data.<br>**Purpose:** Running an experiment to find a functional relationship between delay and exposure to convection (Section 4.2.5.1).<br>**Description of the data:** Flights in June 2018 that were regulated due to weather with every waypoint within longitudes (-20, 25) and latitudes (25, 60). This results in 52241 flights used for the experiment conducted.<br>**Source of the data:** Retrieved from EUROCONTROL Demand Data Repository V2.<br>**Data format:** NEST dataset file. |
| **Traffic data #2** | **Data type:** Air traffic data<br>**Purpose:** Characterizing uncertainty in the take-off time (Section 5.1).<br>**Description of the data:** Statistical characterization (namely, a number of percentiles) of the error committed by the ETFMS (Enhanced Traffic Flow Management System) when estimating the take-off time of flights crossing the airspace of Maastricht Upper Area Control Centre (MUAC). The data was collected by Dalmau et al.[2] in the context of a study on how to improve take-off time predictions by applying an explainable machine learning approach.<br>**Source of the data:** Provided by EUROCONTROL.<br>**Data format:** CSV file. |
| **Traffic data #3** | **Data type:** Air traffic data.<br>**Purpose:** Performing an illustrative example of EPS-based trajectory prediction for a flight from LEMD airport to EDDT airport (Section 4.2.6).<br>**Description of the data:** Airspace structure (i.e., waypoints and airways defining the route network) for AIRAC 1907. In particular, the waypoints defining the flight route considered in Section 4.2.6 were the following: '*MD47', 'RBO', 'PINAR', 'BRITO', 'LARDA', 'RONNY', 'TOPTU', 'TOU', 'GAI', 'MEN', 'MINDI', 'VEROT', 'MURRO', 'LSE', 'DEPUL', 'ARGIS', 'TOKDO', 'PAS', 'GVA', 'SPR', 'REVLI', 'FRI', 'WIL', 'ZUE', 'SONOM', 'NELLI', 'NOTGA', 'LAMPU', 'NIKUT', 'TOSTU', 'RASPU', 'GUDOM', 'GORKO', 'BOKNI', 'PILAM', 'VAGAB', 'BAMKI', 'NENAN', 'ERF', 'WEMAR', 'GALMA', 'OSKAT', 'TADUV', 'MILGU'.<br>**Source of the data:** Retrieved from EUROCONTROL Demand Data Repository V2.<br>**Data format:** NEST dataset file. |
| **Traffic data #4** | **Data type:** Air traffic data.<br>**Purpose:** Performing two illustrative examples. The first one is an EPS-based trajectory prediction for a segment of a flight from KJFK airport to OMAA airport (Section 4.2.7); the piece of trajectory begins at (55.00 N, 15.00 W) and ends at (46.26 N, 21.10 E). The second one is a trajectory prediction through the unified framework for a segment of a flight from KJFK airport to OMAA airport; this second piece of trajectory begins at (55.00 N, 16.60 W) and ends at (46.16 N, 20.40 E) (Section 7.4).<br>**Description of the data:** Waypoints coordinates and aircraft model of the flight with EUROCONTROL's flight ID: 203856023 (callsign ETD102, from KJFK to OMAA, planned to depart on 15$^{th}$ January 2017 at 19:50:00). The data was retrieved from |

[2] R. Dalmau, F. Ballerini, H., Naessens, S. Belkoura, and S. Wangnick, "An explainable machine learning approach to improve take-off time predictions", Journal of Air Transport Management, Vol 95, August 2021 (https://doi.org/10.1016/j.jairtraman.2021.102090).

| | |
|---|---|
| | its initial trajectory. A modification was introduced so that the trajectory prediction begins at a modified starting location and time; from there, the aircraft was supposed to continue with the initial route. A single cruise level was considered as well.<br>**Source of the data:** Retrieved from EUROCONTROL Demand Data Repository V2.<br>**Data format:** NEST dataset file. |
| **Traffic data #5** | **Data type:** Air traffic data.<br>**Purpose:** Performing an illustrative example of the unified framework for trajectory prediction for a flight from LFPG airport to LORP airport (Section 7.2).<br>**Description of the data:** Waypoints coordinates and aircraft model of the flight with EUROCONTROL's flight ID: 203867109 (callsign AFR1888, from LFPG to LORP, planned to depart on 16th January 2017 at 09:23:00). The data was retrieved from its initial trajectory. The departing time and date were modified, and a single cruise level was considered.<br>**Source of the data:** Retrieved from EUROCONTROL Demand Data Repository V2.<br>**Data format:** NEST dataset file. |

# 2. Input data used in WPs 5, 6 and 7

A description of the use case considered in WPs 5, 6 and 7 is provided next. This use case is developed within the Austrian airspace for **June 12th, 2018**. This scenario corresponds to a day with high convection intensity between 10:30 and 22:45 UTC (note that all times below are UTC). The prediction is performed at 12:00 for the next 8 hours.

## Meteorological data

The three probabilistic weather forecasts considered in this application, one ensemble nowcast and two Ensemble Prediction Systems (EPS), are described next. In all cases, the last available forecasts at 12:00 are used.

### *Rain and cloud data*

Radar and satellite observations from Opera RADAR and SAF EUMETNET Satellite information (including Cloud Top Height, CTH). Radar data have been retrieved from the Odyssey database, and satellite data from the AEMET database. Data format**:** hdf5 file.

These data are used to generate the ensemble nowcast used in the project following the Short-Term Ensemble Prediction System (STEPS) methodology[3]. This nowcast is produced every 15 minutes, interpolated every 5 minutes and processed to identify the convective cells (using a threshold of radar

[3] N. E. Bowler, C. E. Pierce, and A. W. Seed, "STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP", Quarterly Journal of the Royal Meteorological Society, 132(620): 2127–2155, 2006.

reflectivity of 38 dBZ). The actual weather cells are enlarged with a safety margin of 13.5 NM. A common cloud top height for all the nowcast coverage area is also provided (the flights can overfly them with a margin of 5000 ft). The last available nowcast at the moment of the prediction is the one generated at 11:45. The number of members is 15, and they are statistically independent among them.

***COSMO-D2-EPS***

This is a limited-area, high-resolution EPS. This EPS provides wind and air temperature data, and convection indicators. Specifically, two convection indicators have been considered: Lifted Index and Precipitation Intensity. Convective areas are identified when the Lifted Index is less than -4 and the Precipitation Intensity is above 5 mm/hour.

Note that a transition zone with unrealistic gradients has been identified in the contour of COSMO-D2-EPS coverage area, resulting from its boundary conditions during its generation. As a result, the outer 25 grid points on each side (about 50 km) have been discarded, so that the coverage area of COSMO has been slightly downsized.

In this application, the last available COSMO-D2-EPS is the one generated at 09:00. It has been interpolated every 15 minutes and processed to identify the convective areas. The number of members is 20, and they are statistically independent among them.

COSMO-D2-EPS has been purchased from the Fraunhofer Institute for Energy Economics and Energy System Technology (IEE). Data format: GRIB file.

***ECMWF-EPS***

This is a global EPS from the European Centre for Medium-Range Weather Forecasts (ECMWF). This EPS provides wind and air temperature data, and convection indicators. Specifically, two convection indicators have been considered: Total Totals and Convective Precipitation. Convective areas are identified when the Total Totals is above 44 K, and the Convective Precipitation is above 0.

In this application, the last available ECMWF-EPS is the one generated at 00:00. It has been interpolated every 15 minutes and processed to identify the convective areas. The number of members is 50, and they are statistically independent among them.

ECMWF-EPS has been retrieved from ECMWF's MARS database. Data format: GRIB file.

### Aircraft model data

The aircraft performance model used is BADA 3.12. Data format: APF, OPF, and PTD files.

### Air traffic data

In this application we consider the flight plan's data: positions of airborne aircraft at 12:00, nominal take-off times, and nominal routes to be followed.

The traffic consists of aircraft airborne at 12:00 or expected to take-off in the next 8 hours (including the uncertainty in the take-off time) which plan to cross the Austrian airspace plus a surrounding area of 50 NM. A total number of 2542 flights are considered: 393 flights are airborne at 12:00, and 2149 flights are expected to depart in the next 8 hours.

The historical traffic data has been retrieved from Eurocontrol's R&D Data Archive. Data format: CSV file.

### Airspace data

The Austrian airspace (LOVV) under the control of the Wien Area Control Centre (ACC WIEN) is shown in Figure 1 (AIRAC cycle 1806). It is divided into five geographical regions (B, E, N, S and W), and each region into 5 vertical layers:

- ACC WIEN B: B1, B2, B3, B4, and B5;
- ACC WIEN E: E1, E2, E3, E4, and E5;
- ACC WIEN N: N1, N2, N3, N4, and N5;
- ACC WIEN S: S1, S2, S3, S4, and S5; and
- ACC WIEN W: W1, W2, W3, W4, and W5.

In total, 38 elementary volumes are used to define this airspace, which lead to near 60 possible different ATC sectors (defined by their lateral and vertical boundaries) and 190 different sector configurations.

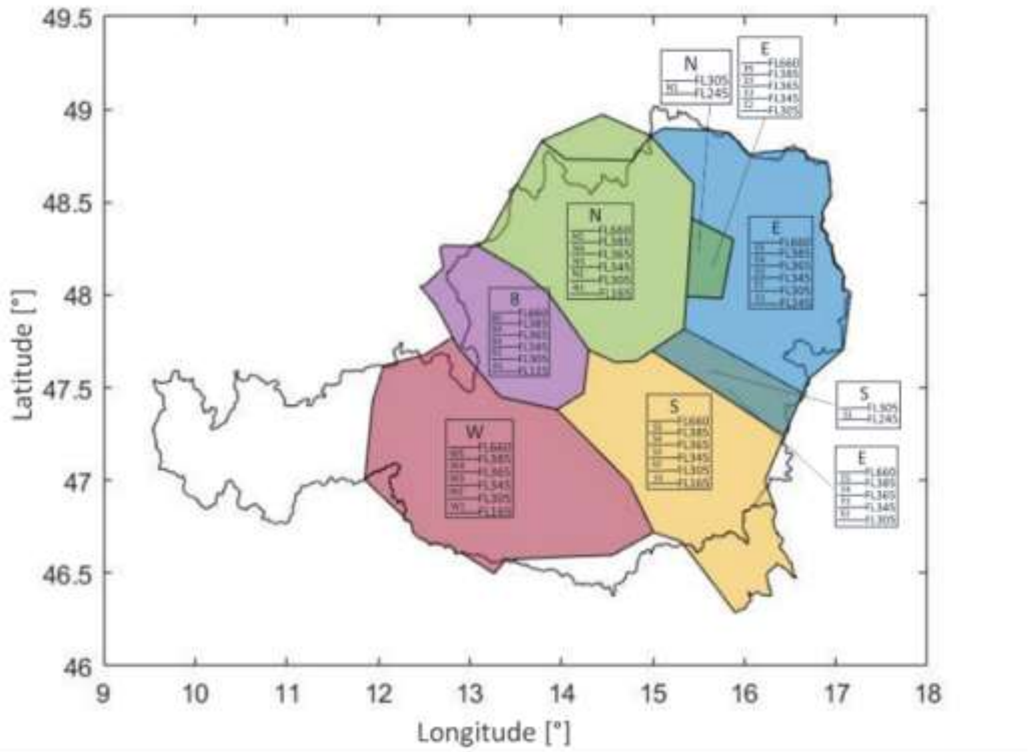


**Figure 1. Vertical span of LOVV sectors**

The sector configuration chosen is 10A1 (the one active at 12:00), composed of the following 10 sectors:

- B15, formed by all the elementary volumes from ACC WIEN B.
- N12, formed by the elementary volumes from N1 and N2.
- N35, formed by the elementary volumes from N3, N4 and N5.
- E13, formed by the elementary volumes from E1, E2 and E3.
- E45, formed by the elementary volumes from E4 and E5.
- S12, formed by the elementary volumes from S1 and S2.
- S35, formed by the elementary volumes from S3, S4 and S5.
- W12, formed by the elementary volumes from W1 and W2.
- W3, formed by W3.
- W45, formed by W4 and W5.

Airspace data (sector configurations and geometries) have been provided by Austro Control GmbH. Data format: xlsx, gpkg, pdf files.



EUROPEAN PARTNERSHIP

Co-funded by the European Union