DART Data Pool

D1.3

DART

Grant:	699299
Call:	H2020-SESAR-2015-1
Topic:	SESAR-02-2015 Data Science in ATM
Consortium coordinator:	University of Piraeus Research Center (UPRC)
Edition date:	15/01/2018
Edition:	00.01.01

Authoring & Approval

Authors of the document

Name/Beneficiary	Position/Title	Date
MIGUEL GARCÍA MARTÍNEZ (CRIDA)	Project Member	10/01/2018

Reviewers internal to the project

Name/Beneficiary	Position/Title	Date
George Vouros /UPRC	Project Coordinator	13/02/2018
David Scarlatti /BR&T-E	Project Member	13/02/2018
George Fuchs/FHFR	Project Member	13/02/2018
Jose Manuel Cordero/CRIDA	Project Member	13/02/2018

Approved for submission to the SJU By — Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date
George Vouros /UPRC	Project Coordinator	13/02/2018
David Scarlatti /BR&T-E	Project Member	13/02/2018
George Fuchs/FHFR	Project Member	13/02/2018
Jose Manuel Cordero/CRIDA	Project Member	13/02/2018

Rejected By - Representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date

Document History

Edition	Date	Status	Author	Justification
0.1	15/01/2018	Draft	Miguel García Martínez	New Document
0.2	31/01/2018	Draft	Miguel García Martínez	Release 1st draft, internal CRIDA review
0.3	09/02/2018	Draft	Enrique Iglesias Martínez / Elena López Ballesteros	Added Comments after Internal review. Version for project member's review.
1.0	13/02/2018	Release	Miguel García Martínez	Added modifications after project member's review. Final Version for SJU review.
1.1	09/04/2018	Release	Miguel García Martínez	Added modifications after SJU assessment report.

DART

DATA-DRIVEN AIRCRAFT TRAJECTORY PREDICTION RESEARCH

This document is part of a project that has received funding from the SESAR Joint Undertaking under grant agreement No 699299 under European Union's Horizon 2020 research and innovation programme.

Abstract

This document describes the DART Data Pool internal updates providing an historical report. The DART Data Pool host all the datasets required for the DART project success. This datasets contains the information from all the sources that WPs required since the beginning of the project but also some sources that were created to solve needs after continuous feedback between WP2/WP3 and WP1. This procedure has last from M0 to M20 where this document is delivered, nevertheless WP1 will be listening for possible feedback from WP2/WP3 (including possible errors on datasets) till the end of the project M24 just in case any update is required. [1]

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

Table of Contents

Executive Summary

This document describes the contents and the evolution of the DART Data Pool. The Data Pool is the container for all the data sources that WP1 has gather for DART project. All the data sources have the same scope, they describe the years 2016 and 2017 for the Spanish Airspace.

Most of the sources are generated by computers and collected by different ETL (Extract Transform Load) processes. Nevertheless there are a couple of sources that have been generated manually by WP1 to satisfy the needs of WP2 and WP3 during the process of loading and merging data.

This way, the document firstly contains the description of the DART Data Pool from the point of view of the data. No hardware or software is described; see DART D1.2 Data Transaction Pipeline for further information about it. This section consists of a brief description of each released data set and of each data source collected. Then the evolution of the DART Data Pool will be detailed in the way of a change log listing each of the updates during the project, including its contents and an approximate size of the dataset.

Finally, the results at M20 will be exposed and summarized to demonstrate that the DART Data Pool has been completed for M20 and then the milestone of the project achieved.

As an annex of this document a detailed description of each data source is included. This description has been used all along the project since the very beginning as an analysis of potential data sources. It is from the point of view of the user of the data and it focus on the description of the content of each data source (fields meaning) and the way that different data sources can be merged or linked.

1 Introduction

1.1 Purpose of the document

All along the DART project internal updates of the Data Pool have been released. The designated period for updates from T0+3 to T0+20 could be overlapped by special updates. During first months of the project, WP1 provides samples of candidates for DART data sources and, during last months of the project responding to WPs feedback. This final deliverable covers the whole process including description, historic and results.

This document lists the changes in the datasets since the creation dataset until the current date. The document pretends to collect all the information related to the update date and the reason of the update with regard to the original dataset.

1.2 Intended readership

This document is intended to be used by DART members.

1.3 Acronyms and Terminology

Term	Definition
ANS	Air Navigation Service
ANSP	Air Navigation Service Provider
ATC	Air Traffic Control
ATM	Air Traffic Management
AU	Airspace User
CFMU	Central Flow Management Unit
CFS	Certificate on the Financial Statements
CSV	comma-separated values
DCB	Demand and Capacity Balancing
DMP	Data Management Plan
FP	Flight Plan(s)
GA	General Assembly
GIPV	Integrated Flight Plans Management (Gestión Integrada de Planes de Vuelo)
GRIB	GRIdded Binary or General Regularly-distributed Information in Binary form
H	Humans
Horizon 2020	EU Research and Innovation programme implementing the Innovation Union, a Europe 2020 flagship initiative aimed at securing Europe's global competitiveness.

IFS	InForme de Seguimiento
IPR	Intellectual Property Rights
KPI	Key Performance Indicator
METAR	METeorological Aerodrome Report
NMOC	Network Manager Operations Centre (previously called CFMU)
NOAA	National Oceanic and Atmospheric Administration
PMP	Project Management Plan
POPD	Protection of Personal Data
PRC	Performance Review Commission
SACTA	Automatic System of Air Traffic Control (Sistema Automatizado de Control de Tránsito Aéreo)
SC	Sector Configuration(s)
SESAR	Single European Sky ATM Research Programme
SESAR Programme	The programme which defines the Research and Development activities and Projects for the SJU.
SIGMET	Significant Meteorological Information
SJU	SESAR Joint Undertaking (Agency of the European Commission)
SJU Work Programme	The programme which addresses all activities of the SESAR Joint Undertaking Agency.
TMA	Terminal control area
TRL	Technology Readiness Level
UTC	Coordinated Universal Time
WBS	Work Breakdown Structure
WP	Work Package

Table 1: Acronyms and Terminology

2 Description

2.1 Data Sets

The information was divided in seven big releases of datasets designated as:

- First Dataset (January 2017). Data from January 2016 to November 2016
- Second Dataset (June 2017). Data from December 2016 and beginning of 2017 if available.
- Third Dataset (started February 2017, provided in May 2017). Data from January to December 2016 of processed NOAA (csv files with Spain region filtered instead of binary grid files with everything).
- Fourth Dataset (June 2017). Data from January to December 2016. Link between Sectors (SACTA-Operational).
- Fifth Dataset (September 2017). Updating of Link between Sectors.
- Sixth Dataset (December 2017). Sample data of new data sources (CFMU – Regulations).
- Seventh Dataset (February 2018). All sources are completed with the remaining datasets for 2017. This is the biggest release since the first one and it is expected to be the last one. In case of errors on the datasets that could arise during the next steps of the project, further releases could be done.

2.2 Data Sources

Data sources are structured in four different categories: Weather, Radar, Airspace and Flight Plan. All of them are essential "bricks" for the different use cases.

A scheme of data sources organization in this document is as follows:

Weather	NOAA
	METAR
	SIGMET
Radar	IFS
Airspace	Sector configuration
	Link Sectors SC <->GIPV
Flight Plan	GIPV
Network Management	CFMU

Table 1: Sources

2.2.1 Weather

This data category comprises of three sources of weather information: NOAA, METAR and SIGMET. The first are weather predictions at world level, every 6 hours with information 7 days in advance, from NOAA (National Oceanic and Atmospheric Administration). The second one is airports' weather information, but limited to Spanish airports. The third one, also limited to Spanish airspace, is information of actual or anticipated adverse weather conditions (en-route or at the airports approach).

2.2.1.1 NOAA

This data source is used mainly to obtain the weather conditions at the position an aircraft is at any given time of the flight. So, for each 4D position (latitude, longitude, altitude and time) we can obtain the value of different variables. The most relevant variables are the Temperature, the Pressure, and the two horizontal components of the Wind Speed, u and v (a positive u component represents wind blowing to the East and a positive v component is wind to the North), since they affect the performance of the aircraft.

These variables can be forecasted, so we have to add another variable that reflects when the forecast was done. Typically, variables are available at a number of isobaric levels (pressure altitude) (26 levels is common), which works at the same time as a proxy for the altitude since the height is known for each isobaric level using this equation:

$$Height\ (Feet) = (1 - (millibars/1013.25)\hat{}.190284)) * 145366.45$$

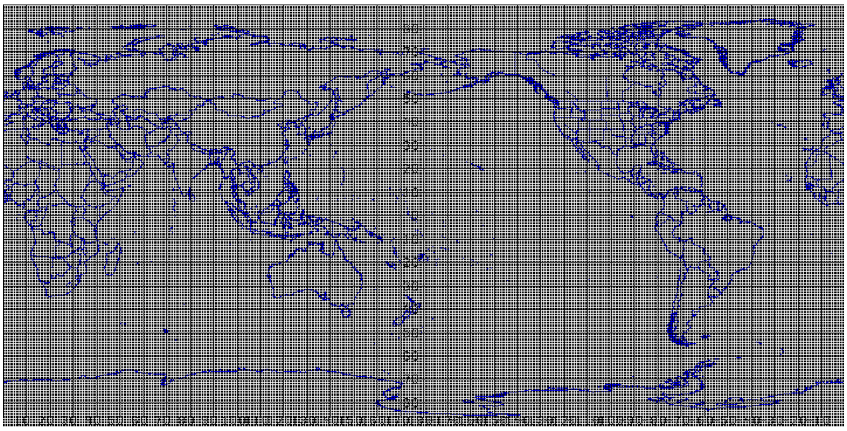
It is important to notice we are talking about weather models; there is no way to know the real value of these variables at all given positions in the atmosphere. The models can represent a prediction of the weather (a forecast) or represent the "best guess" of reality (a reanalysis).

Weather models use a Grid which may have more or less resolution. For DART we'll work with NCEP Grid 4 which has a resolution of 0.5° [2].

(see http://www.nco.ncep.noaa.gov/pmb/docs/on388/tableb.html)

Forecast models can be run several times a day, for DART we'll typically use the latest forecast available previous to the time we are interested in. Forecast models has too a time resolution, or "forecast step", which we expect to be 1 hour.

For Dart we are collecting the first forecast of each day, the one with the forecastHour field set to 0.



NCEP Grid 4

Figure 1: Illustration of NECP Grid 4 information

2.2.1.2 METAR

METAR is a format for reporting weather information. A METAR weather report is predominantly used by pilots in fulfilment of a part of a pre-flight weather briefing, and by meteorologists, who use aggregated METAR information to assist in weather forecasting.

METARs typically come from airports or permanent weather observation stations. Reports are generated once an hour or half-hour, but if conditions change significantly, a report known as a special (SPECI) may be issued. Some METARs are encoded by automated airport weather stations located at airports, military bases, and other sites. Some locations still use augmented observations, which are recorded by digital sensors, encoded via software, and then reviewed by certified weather observers or forecasters prior to being transmitted. Observations may also be taken by trained observers or forecasters who manually observe and encode their observations prior to transmission.

A typical METAR contains data for the temperature, dew point, wind speed and direction, precipitation, cloud cover and heights, visibility, and barometric pressure. A METAR may also contain information on precipitation amounts, lightning, and other information that would be of interest to pilots or meteorologists such as a pilot report or PIREP and runway visual range (RVR).

Raw METAR is the most common format in the world for the transmission of observational weather data. It is highly standardized through the International Civil Aviation Organization (ICAO) [3], which allows it to be understood throughout most of the world.

2.2.1.3 SIGMET

SIGMET data are weather advisory that contains meteorological information concerning the safety of all aircraft. This information is usually broadcast on the ATIS at ATC facilities, as well as over VOLMET stations. They are assigned analphabetic designator from N through Y (excluding S and T). SIGMETs are issued as needed, and are valid up to four hours. SIGMETS for hurricanes and volcanic ash outside the CONUS are valid up to six hours.

2.2.2 Radar IFS (Surveillance)

A single data source (IFS), will be considered for this category. This data source contains information from Spanish secondary radars. It includes the flight call sign, altitude, speed, position, direction and time. The information is updated every 5 seconds.

This data source provides radar tracks of the Spanish airspace controlled by the Spanish ATC provider ENAIRE. A radar track file consists on tabular data rows with a timestamp key and several columns of geospatial information for each one of these timestamps. The update interval is 5 seconds. The area provided is separated into 5 different regions delivered each one on a different plain text file (ifs files). The regions (M, B, C, P and S) are approximately shown on the next figure.

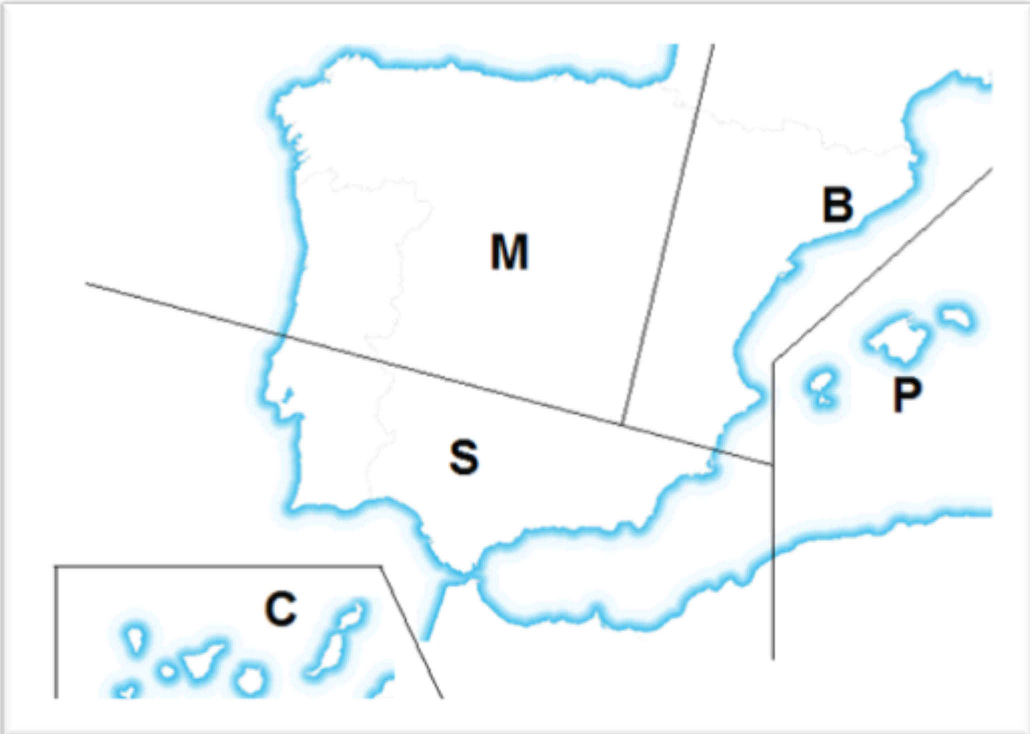


Figure 2: IFS Regions

Each region is completely independent from the others so according regions files could have repeated information while flights can be at the same time under the coverage zone of 2 radars (never more than 2).

Also there is not an unique id per flight on the files since ids of the radar tracks are reused after a certain time interval once they are released (close to the half hour but never a fixed value) and call signs of flights are used several times per day for common routes (the same plane/flight goes and returns the same route on the day). Besides system errors, a flight should be on the same radar track for the whole time under the coverage of a specific radar but this track id is different on the rest of radar regions.

All this means that the way to provide meaning to these files is using mainly timestamp and call sign information with some help of the geospatial information. The files are named with the first letter according to the region and the date on format yyMMdd.

All this Information has been processed and structured in a single table with only the meaningful data to provide a whole flight unique track identified by a unique flight key.

2.2.3 Airspace

Airspace data source can be seen from a dual perspective. On the one hand, describes the existing airspace organization, with no gaps or overlaps, and all the possible ways of combining volumes to generate different operational sector configurations, also with the associated sector capacities, or flights that a sector can manage in a period of time (this static data generically called "adaptation data"), and on the other hand, described the schedule of effective sector configurations that have been effectively put in place in Spanish airspace.

2.2.3.1 Sector configuration (SC)

Air traffic control (ATC) is a service provided by ground-based controllers who direct aircraft on the ground and through controlled airspace. The primary purpose of ATC worldwide is to prevent collisions, organize and expedite the flow of air traffic, and provide information and other support for pilots.

Airspace can be divided in a set of ways, with a different numbers of pieces (sectors). A sector configuration 9A means that a particular airspace (a region in Spain) is divided in 9 sectors, in a particular way. 9B also mean 9 sectors, but divided in a different way. Typically, due to low traffic at nights, the configuration set at those times is a 1A, meaning that a single sector (thus, a single controller) is in place.

This leads to the fact that configurations available are fixed, but configuration “in place” varies during day, adapting capacity resources (Air Traffic Controllers, mainly, as more sectors open mean more capacity, but also more controllers) to the expected demand.

2.2.3.2 Link Sectors SC <-> GIPV

This data source appears after feedback from WPs, there were some difficulties to link sectors between SC (2.2.3.1 - Sector configuration (SC)) and sectors from GIPV (2.2.4 - Flight Plans GIPV). Metrics for ANSP always refer to “official” sectors which are included in SC source but flight plans refer to SACTA-System sectors which are not exactly the same. So this link is required to allow calculating metrics such as the demand of traffic for the officials sectors.

These differences are not majors (in geographic terms) but they make impossible to establish a 1:1 relationship between sectors of both sources. In addition, a relationship of 1:N is not applicable neither because sectors are just different in more or less percentage between sources.

So this manual in-house source describes the relationships for the smallest piece of the Airspace that we have available, Volumes. Sectors are compound with Volumes and then a relationship map between Sectors SC and Volumes GIPV can be defined by 1:N deterministic links. This allows data users to know for each flight plan point related to a sector-volume SACTA the corresponding official sector.

To complete the operation WPs have to use the help of Sectorizations to know what sectors where operative at each instant and then being able to figure the correct sector out.

2.2.4 Flight Plans GIPV

Flight Plan is an essential category as contains the information that triggers a lot of operational decision, both in planning and execution phase, and both on the Air Navigation Service Provision side, and in the Airline one.

A single data source (GIPV, Flight Plan Information Management System) will be considered for this category, a subsystem of the Spanish ATC platform. This contains information on every flight plan currently in flight or scheduled to fly. All the changes and cancellations that affect flight plans are constantly updated and registered in the system.

The GIPV is a Flight Plan Report Manager Subsystem that contains information of flight plans are flying or going to fly in the near future (to 15 hours) in the responsibility airspace of the Flight Plan Central Treatment.

This data source provides 4 different data tables always linked by unique keys that refer to the flight:

- Flights: data table with the identification of flights and information of the flight plan.
- Route points: data table with overflown waypoints of the route for a specific flight plan.
- Segments: data table with overflown segments for a specific flight plan.
- Vol SACTA: data table with overflown volumes for a specific flight plan.

2.2.5 Network Management CFMU

This category covers the data sources that contain Network Management (also known as Flow Management) information, meaning this the regulations put in place to ensure a proper Demand Capacity balance in a tactical way. There is a single source considered: CFMU, coming from the Network Management organization (Eurocontrol), thus covering European airspace. CFMU is the former name of the current Network Manager, and this source will be referred as CFMU.

This data source provides list of flights and the regulations that the network manager can apply to them. The source is one table for flights. When a flight has a regulation, the code of the regulation applied is provided on the row.

3 Historic

3.1 First Dataset – First major update with 2016 data

This was the first and one of the two big updates of data (with the last one). This first dataset (released during the beginning of 2017) was intend to make available for WPs a whole year of all data sources for the stages of study and test of the different machine learning techniques that Dart Project could use.

CONTENT:

Airspace Structures:

- Sectorizations (2016-01-01 to 2016-11-30)
- Configurations (2016-01-01 to 2016-11-30)

Flight Plans:

- Flights (2016-01-01 to 2016-11-30)
- Route points (2016-01-01 to 2016-11-30)
- Segments (2016-01-01 to 2016-11-30)
- Vol Sacta (2016-01-01 to 2016-11-30)

Radar Tracks:

- IFS (2016-01-01 to 2016-11-30)

Weather:

- METAR (2016-01-01 to 2016-11-30)
- SIGMET (2016-01-01 to 2016-07-31)
- NOAA (2016-01-01 to 2016-12-01)

SIZE:

7 TB

CHANGELOG:

DATASET 1	Update	Modification
	15/12/2016	Dataset creation
	20/01/2017	RADAR IFS (including date format fixed 0-24H)(January to November 2016) GIPV (January to November 2016) AIRSPACE SC (January to November 2016) METAR (January to November 2016) SIGMET (January to November 2016) NOAA (raw data) (January 2016)
	03/02/2017	NOAA (processed files) (January to May 2016)
	20/05/2017	NOAA (processed files) (June to November 2016)

Table 2 - Changelog Dataset 1

3.2 Second Dataset – 2016 completion

This second dataset purpose is to complete the 2016 year.

CONTENT:

Airspace Structures:

- Sectorizations (2016-12-01 to 2017-01-01)
- Configurations (2016-12-01 to 2017-01-01)

Flight Plans:

- Flights (2016-12-01 to 2017-02-01)
- Route points (2016-12-01 to 2017-02-01)
- Segments (2016-12-01 to 2017-02-01)
- Vol Sacta (2016-12-01 to 2017-02-01)

Radar Tracks:

- IFS (2016-12-01 to 2017-04-01)

Weather:

- METAR (2016-12-01 to 2017-01-01)
- SIGMET (2016-12-01 to 2017-01-01)

- NOAA (2016-12-01 to 2017-03-01)

SIZE:

1 TB

CHANGELOG:

	Update	Modification
DATASET 2	01/06/2017	Dataset creation
	07/06/2017	RADAR IFS (December 2016) METAR (December 2016) SIGMET (December 2016) AIRSPACE SC (December 2016) NOAA (December 2016)

Table 3: Changelog Dataset 2

3.3 Third Dataset – NOAA Processed

This dataset comes to solve a need of the WPs that were not able at this point to process NOAA files. NOAA is the biggest data source of Dart Project and requires third party libraries to read the provided binary data files. Then WP1 created this data set "NOAA – Processed" with structured and plain text readable CSV files.

CONTENT:

Weather:

- NOAA Processed (2016-01-01 to 2016-12-31)

SIZE:

0.5 TB

CHANGELOG:

	Update	Modification
DATASET 3	24/01/2017	Dataset creation
	15/02/2017	NOAA PROCESSED (test version January to June 2016)
	22/05/2017	NOAA PROCESSED (final version January to December 2016)

Table 4: Changelog Dataset 3

3.4 Fourth Dataset – SC vs FP

As the previous case, this dataset comes to solve another need that appears during the work of WPs. There were some problems to link two sources (GIPV – SC) on the field Sector (Airspace spatial structure) and WP1 creates a new manual source that links all sectors from one source to the other.

CONTENT:

Linkage assistance Sectors/Volumes GIPV (FP) <-> AIRSPACE (SC):

- Link volumes (2016-01-01 to 2016-12-31)

SIZE:

< 100 MB

CHANGELOG:

DATASET 4	Update	Modification
	01/06/2017	Dataset creation
	13/06/2017	• Remove non-operational volumes • Correction: replace “AXX” by “APP”
	14/06/2017	Inclusion of “LSA” and “LSB” volumes as part of “LSX” volume
	07/07/2017	Replace “SMNUL” SACTA volume by “SMNU” volume
	01/08/2017	Inclusion of volumes "RCE”, RCW, RNE, RNW, RST and RWW". Sometimes these volumes are named with the first “S” and others without it.
	04/09/2017	• LEMDEXS: Replace “DWN” by “DWS” • LEMDENS: "DEN" and "EN" addedd no this volume • LEBLT4E: remove one "S" in "SSGON" volume
	25/09/2017	Fix of some detected errors regarding Sectors relationships

Table 5: Changelog Dataset 4

3.5 Fifth Dataset – CFMU Regulations

In order to compare results in WP3 to delays imposed by Network Managers (CFMU) WP1 includes the data source that contain Network Management information.

This dataset gives to the WPs the historic of the regulations put in place to ensure a proper Demand Capacity balance in a tactical way.

CONTENT:

Network Management

- CFMU Regulations (2016-01-01 to 2016-12-31)

SIZE:

< 100 MB

CHANGELOG:

DATASET 5	Update	Modification
	01/12/2017	Dataset creation CFMU Regulations Small sample data
	22/01/2018	CFMU Regulations Final version 2017-2018

Table 6: Changelog Dataset 5

3.6 Sixth Dataset

This is the last expected update of the Dart Data Pool. It includes the remaining data to complete 2017 year for all original sources of the data pool (CFMU and SC_vs_FP were created with the whole time scope available).

CONTENT:

Airspace Structures:

- Sectorizations (2017-01-01 to 2018-01-01)
- Configurations (2017-01-01 to 2018-01-01)

Flight Plans:

- Flights (2017-02-01 to 2018-01-01)
- Route points (2017-02-01 to 2018-01-01)
- Segments (2017-02-01 to 2018-01-01)
- Vol Sacta (2017-02-01 to 2018-01-01)

Radar Tracks:

- IFS (2017-04-01 to 2018-01-01)

Weather:

- METAR (2017-01-01 to 2018-01-01)
- SIGMET (2017-01-01 to 2018-01-01)
- NOAA (2017-01-01 to 2018-01-01)

SIZE:

7 TB

CHANGELOG:

DATASET 6	Update	Modification
	30/01/2018	Dataset creation RADAR IFS (Remaining 2017) METAR (Remaining 2017) SIGMET (Remaining 2017) AIRSPACE SC (Remaining 2017) GIPV-FP (Remaining 2017) NOAA (Remaining 2017)

Table 7: Changelog Dataset 6

4 Results

4.1 Current Situation (M20 February 2018)

This deliverable follows the expected last upload of datasets for Dart project to complete them all along the time (2016-1017) and spatial (Spanish Airspace) scopes. The picture at this stage is that the Dart Data Pool is completed.

The next figure describes it with a timeline evolution:

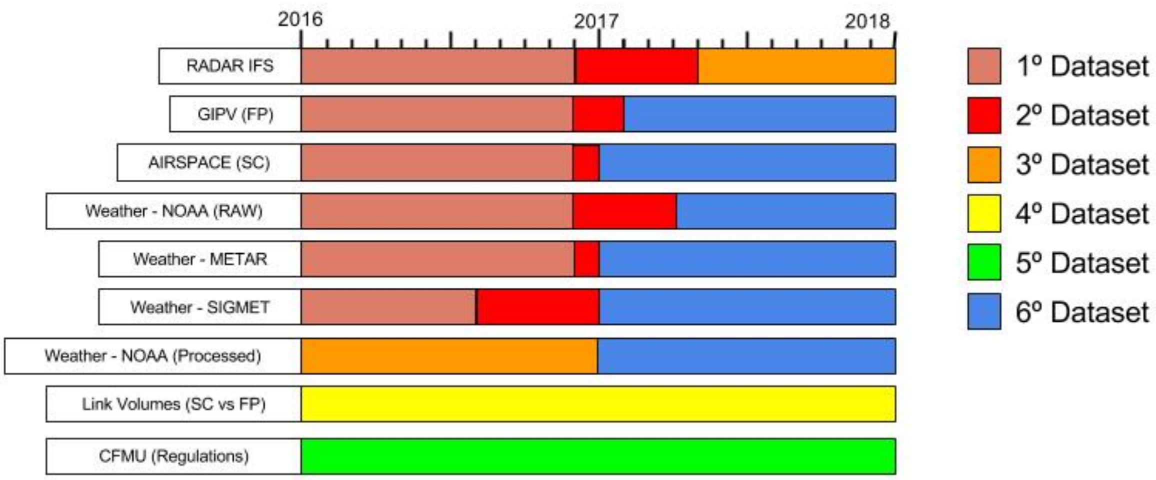


Figure 3: DART Data Pool Evolution

SIZE SUMMARY:

AirspaceStructures-SC	354.6 MB
FlightPlans-GIPV	52.81 GB
NetworkManagement-CFMU	908 MB
RadarTracks-IFS	95 GB
SC_vs_FP	1 MB
Weather-METAR	106 MB
Weather-NOAA	13 TB
Weather-SIGMET	15 MB
TOTAL	14 TB

SOME EXTRA INFORMATION ABOUT DART DATA POOL:

4 Millions of Flights (4106320 flights)

2.5 Billions of Radar points (2714042496 points)

1 Billion of Flight Plan Route points (1003734563 points)

89 Millions of Flight Plan messages (89903772 messages)

5 Annex – Data Sources Detailed Description

5.1 NOAA

5.1.1 Metadata

Data for weather models is typically distributed in "GRIB" format files. GRIB (GRIdded Binary or General Regularly-distributed Information in Binary form) format allows compressing a lot the weather data and includes metadata about the content of the file, so it is very convenient for transferring the data. The data can be extracted with many available tools (I.e. GRIB API from ECMWF available at https://software.ecmwf.int/wiki/display/GRIB/Home).

5.1.2 Fields detailed description (structured data)

Once the "dataset" is extracted from the GRIB file, these fields are expected:

- Reftime - "GRIB reference time" - Is the time when the forecast was produced
- Time - "GRIB forecast or observation time" - Is the time for the forecast, typically as hours since the Reftime.
- Lat – Latitude typically in "degrees_north"
- Lon – Longitude typically in "degrees_east"
- Temperature_isobaric – Temperature in K
- u-component_of_wind_isobaric – Wind speed vector component u in m/s.
- v-component_of_wind_isobaric– Wind speed vector component v in m/s.
- Geopotential_height_isobaric – Geopotential Height in gpm.
- Isobaric surface – The pressure in Pa for each isobaric level
- Relative_humidity_isobaric
- Absolute_vorticity_isobaric in 1/s
- Cloud_mixing_ratio_isobaric in kg/kg
- Wave_Geopotential_Height_isobaric in gpm
- Per_cent_frozen_precipitation_surface
- Pressure_height_above_ground in Pa
- Pressure_maximum_wind in Pa
- Pressure_reduced_to_MSL_msl in Pa
- Specific_humidity_height_above_ground in kg/kg
- Temperature_altitude_above_msl in K
- Temperature_height_above_ground in K
- u-component_of_wind_altitude_above_msl – Wind speed vector component u in m/s.
- v-component_of_wind_altitude_above_msl – Wind speed vector component v in m/s.
- u-component_of_wind_height_above_ground – Wind speed vector component u in m/s.
- v-component_of_wind_height_above ground – Wind speed vector component v in m/s.
- Wind_speed_gust_surface in m/s

5.1.3 Scale

The data stored are 745 files per day (amounting 650 Gb of raw data) from NOAA source, for one month (April 2016) inside the entire world [4].

NOAA data is contained in the *.grb2 files. In the example below there are two file with meteorological information, the first one (gfs_4_20160417_0000_000.gbr2) is performed at 00:00 with a forecast for the same instant, and the second one (gfs_4_20160417_0000_003.gbr2) is performed also at 00:00 but it is a forecast for inside 3 hours.

gfs_4_20160417_0000_000.grb2	07/08/2016 4:09	Archivo GRB2	56.177 KB
gfs_4_20160417_0000_000.inv	07/08/2016 3:55	Archivo INV	15 KB
gfs_4_20160417_0000_003.grb2	07/08/2016 4:26	Archivo GRB2	61.877 KB
gfs_4_20160417_0000_003.inv	07/08/2016 3:55	Archivo INV	20 KB

Figure 4: NOAA data stored

5.1.4 Range of available dates and areas

NOAA is available for the whole European airspace, from 2014 to present..

5.1.5 Join procedure with other sources

- METAR: METAR initial date and time (columns B and C) crosses with the nearest prediction available for NOAA (both point and time). [NOAA information is to be crossed with METAR information in order to validate NOAA information at surface, especially in the vicinities of airports]
- SIGMET: As SIGMET is defined in 4D (specific volume in the airspace plus timespan), several data coming from NOAA will fall in that region at that specific time. [NOAA information is to be crossed with SIGMET messages in order to predict different phenomena and define dangerous meteorology even at the lack of SIGMET messages]
- IFS (Radar): For every point in the radar track together with its time, its nearest NOAA prediction (again for both point and time) provides the meteorological required information. [With the different points along the track of a flight, NOAA information can provide different meteorological information to each of these points.]
- Airspace: Definition by time and location of the different points from NOAA that are inside specific sectors in the airspace. [Expected to be at the maximum possible sectorisation, the definition of which NOAA nodes belong to each sector may affect the sectorisation in the future]
- Flight Plans: Similarly to what happens with IFS, but using predictions instead of real data. [Crossing information coming from NOAA (specifically the predictions) and the flight plan to follow by an aircraft, the different expected meteorological aspects of the flight can be defined].

5.2 METAR

5.2.1 Metadata

N/A

5.2.2 Fields detailed description (structured data)

The structured METAR data is organized in the next six tables. First table (metar_Facts) contains those meteorological variables which contain only one item in same METAR message.

The rest of the tables (metarCloudInformation_Facts, metarRunwayState_Facts, metarRvr_Facts, metarSignificantMeteo_Facts and metarWindShear_Facts) contain other meteorological variables which can have more than one item in the same METAR message, so it was decided to structure in different tables. All tables are linked through the fields: “initialValidDateKey”, “initialValidDateKey” and “aerodromeCode” (“column B, C and D” of each table).

Table 2: metar_Facts table

<table>
<tr><th>Column</th><th>Ítem – Description</th></tr>
<tr><td>A</td><td>metarKey</td></tr>
<tr><td>B</td><td>initialValidDateKey (yyyymmdd)
UTC Date corresponds to the hour of the last meteorological observation. (Valid from)</td></tr>
<tr><td>C</td><td>initialValidTimeKey (9HHMMSS)
UTC Time corresponds to the hour of the last meteorological observation. (Valid from)</td></tr>
<tr><td>D</td><td>aerodromeCode
ICAO AERODROME INDICATOR. It corresponds to the ICAO code of the Aerodrome. E.g. LEMD</td></tr>
<tr><td>E</td><td>isSpeci: (0=METAR, 1=SPECI)
Type of message, it can be:
METAR: Aerodrome Routine Meteorological Report. Sent every hour.
SPECI: Similar to the METAR message sent not evenly but punctually.</td></tr>
<tr><td>F</td><td>WindDirection
WIND DIRECTION (degrees): Average direction from which the wind blows 10 minutes before the observation rounded to the nearest ten.
VARIABLE DIRECTION: the direction is considered as variable and this field is filled out with “VRB”, if:
the velocity is less than 3KT and the variation of the direction is equal or higher than 60 degrees
the velocity is equal or higher than 3KT and the variation of the direction is equal or higher than 180 degrees or and an indeterminate direction</td></tr>
<tr><td>G</td><td>windIntensity
WIND INTENSITY (knots)</td></tr>
<tr><td>H</td><td>windGustIntensity
WIND GUST INTENSITY (knots): The maximum velocity during the previous 10 minutes to the observation (only if this velocity is equal or higher than the intensity indicated + 10).</td></tr>
<tr><td>I</td><td>windVariationMinor
TOTAL VARIATION OF THE WIND DIRECTION (MINOR DIRECTON): Used when there are extreme directions, it means, the velocity is equal or higher than 3 kt and the variation of the direction is between 60 and 180 degrees. There are two fields to show the interval.</td></tr>
<tr><td>J</td><td>windVariationMajor
TOTAL VARIATION OF THE WIND DIRECTION (LARGEST DIRECTION)</td></tr>
<tr><td>K</td><td>Visibility
PREDOMINANT VISIBILITY (m): Is the one observed from the aerodrome (360 degrees).
PARTICULAR CASES: If:
The visibility is higher than 10km the field is filled out with “9999”.</td></tr>
</table>

Column	Ítem – Description
	○ The visibility is lower than 50m the field is filled out with "0000".
L	minVisibility • **MINIMUN VISIBILITY (m):** Is the minimum visibility if and when this one is less than 1500 m or 50% than the predominant visibility.
M	minVisibilityDir • **DIRECTION OF MINIMUN VISIBILITY:** Is the direction of the minimum visibility with respect to one of the eight cardinal points (N, NE, E, SE, S, SW, W, and NW).than 1500 m or 50% than the predominant visibility.
N	isCavok **CAVOK (Ceiling and visibility OK):** This term substitutes the groups of visibility, RVR, significant time and cloudiness or vertical visibility when simultaneously: • Visibility is higher than 10km • There is lack of clouds under 5000 feet or under the highest minimum altitude of the sector when this one is higher than 5000 feet and without Cumulonimbus.
O	QNH **QNH (hPa):** It is defined as, "barometric pressure adjusted to sea level." It is a pressure setting used by pilots, air traffic control (ATC), and low frequency weather beacons to refer to the barometric setting which, when set on an aircraft's altimeter, will cause the altimeter to read altitude above mean sea level within a certain defined region.
P	temperature **TEMPERATURE (Celsius):** An M in front of the number indicates that the temperature is below zero (0) Celsius.
Q	DewPoint **DEW POINT (Celsius):** An M in front of the number indicates that the dew point is below zero (0) Celsius.
R	referenceLoadFile

Table 3: metarCloudInformation_Facts table

Column	Ítem – Description
A	metarKey
B	initialValidDateKey (yyyymmdd) UTC Date corresponds to the hour of the last meteorological observation. (Valid from)
C	initialValidTimeKey (9HHMMSS) UTC Time corresponds to the hour of the last meteorological observation. (Valid from)
D	aerodromeCode ICAO AERODROME INDICATOR. It corresponds to the ICAO code of the Aerodrome. E.g. LEMD
E	isSpeci: (0=METAR, 1=SPECI) Type of message, it can be: • **METAR:** Aerodrome Routine Meteorological Report. Sent every hour. • **SPECI:** Similar to the METAR message sent not evenly but punctually.
F	cloudQuantity **CLOUD QUANTITY**. Dividing the quantity of clouds into eighth

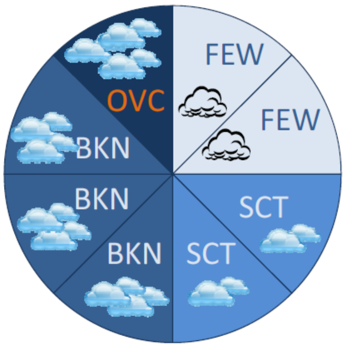
Column	Ítem – Description
	○ FEW: 1/8 a 2/8 ○ SCT (SCATTERED) 3/8 a 4/8 ○ BKN: BROKEN 5/8 a 7/8 ○ OVC: OVERCAST 8/8 
G	cloudHeight **CLOUD ALTITUDE**: It is measured in hundreds of feet. There are some methods to calculate the cloud ceiling but the tool the most used is based on the measurement of time that one lighted thrust takes in go back to the ground after being reflected in the cloud.
H	cloudType **CLOUD KIND/TYPE**: It is fill out with CB: Cumulonimbus and TCU: towering cumulus.
I	referenceLoadFile

Table 4: metarRunwayState_Facts table

Column	Ítem – Description
A	metarKey
B	initialValidDateKey (yyyymmdd) UTC Date corresponds to the hour of the last meteorological observation. (Valid from)
C	initialValidTimeKey (9HHMMSS) UTC Time corresponds to the hour of the last meteorological observation. (Valid from)
D	aerodromeCode ICAO AERODROME INDICATOR. It corresponds to the ICAO code of the Aerodrome. E.g. LEMD
E	isSpeci: (0=METAR, 1=SPECI) Type of message, it can be: • **METAR:** Aerodrome Routine Meteorological Report. Sent every hour. • **SPECI:** Similar to the METAR message sent not evenly but punctually.
F	runway **RUNWAY STATE: R+ INDICATOR/ RUNWAY or ALLRWY – XyzzWW.** i.e. R15/550493 • **RUNWAY or ALLRWY:** ▪ 88 to designate all the runways and the 99 99 one to repeat the last message.
G	state **RUNWAY STATE: R+ INDICATOR/ RUNWAY or ALLRWY – XyzzWW.** i.e. R15/550493 • **State: XyzzWW** ○ **X: Runway deposit** ▪ **0:** clear and dry runway ▪ **1:** wet ▪ **2:** wet with puddle

Column	Ítem – Description
	▪ **3:** covered by frost (thickness less than 1mm) ▪ **4:** dry snow ▪ **5:** wet snow ▪ **6:** melting snow ▪ **7:** ice ▪ **8:** compact snow ▪ **9:** ice bank ▪ **/:** kind of deposit not notified (out service runway) ○ **Y: Pollution State of runway** ▪ **1:** less than 10% of the covered runway ▪ **2:** covered runway between 11 to 25% ▪ **5:** covered runway between 26 to 50% ▪ **9:** covered runway between 51 to 100% ▪ **//:** kind of deposit not notified (e.g. the runway is being cleared) ○ **Zz: Deposit thickness** ▪ **00 <1mm** ▪ **01: 1mm** ▪ **02: 2mm** ▪ **03: 3mm** ▪ **…** ▪ **90: 90mm** ▪ **92: 10cm** ▪ **93: 15cm** ▪ **94: 20cm** ▪ **95: 25cm** ▪ **96: 30cm** ▪ **97: 35cm** ▪ **98 >=40cm** ▪ **99: runway out of service due to the deposit or cleaning** ▪ **//:** Deposit thickness no measurable or not important from an operational point of view. ○ **WW: Friction coefficient and stop efficiency** ▪ **Friction coefficient.** It is used the stop values without zero and comma. • **28:** 0,28 coefficient • **35:** 0,35 coefficient • **…** ▪ **Stop efficiency** • **95:** good • **94:** good/medium • **93** medium • **92:** medium/deficient

Column	Ítem – Description
	• **91:** deficient • **99:** unreliable **//:** Stop conditions not notified, runway out of service.
H	referenceLoadFile

Table 5: metarRvr_Facts table

Column	Ítem – Description
A	metarKey
B	initialValidDateKey (yyyymmdd) UTC Date corresponds to the hour of the last meteorological observation. (Valid from)
C	initialValidTimeKey (9HHMMSS) UTC Time corresponds to the hour of the last meteorological observation. (Valid from)
D	aerodromeCode ICAO AERODROME INDICATOR. It corresponds to the ICAO code of the Aerodrome. E.g. LEMD
E	isSpeci: (0=METAR, 1=SPECI) Type of message, it can be: • **METAR:** Aerodrome Routine Meteorological Report. Sent every hour. • **SPECI:** Similar to the METAR message sent not evenly but punctually.
F	runway **Runway (Runway Visual Range):** e.g. 32R
G	minVisibility **Visual Range, RVR (m)**
H	referenceLoadFile

Table 6: metarSignificantMeteo_Facts table

Column	Ítem – Description
A	metarKey
B	initialValidDateKey (yyyymmdd) UTC Date corresponds to the hour of the last meteorological observation. (Valid from)
C	initialValidTimeKey (9HHMMSS) UTC Time corresponds to the hour of the last meteorological observation. (Valid from)
D	aerodromeCode ICAO AERODROME INDICATOR. It corresponds to the ICAO code of the Aerodrome. E.g. LEMD
E	isSpeci: (0=METAR, 1=SPECI) Type of message, it can be: • **METAR:** Aerodrome Routine Meteorological Report. Sent every hour. • **SPECI:** Similar to the METAR message sent not evenly but punctually.
F	intensity **SIGNIFICANT METEOROLOGY, INTENSITY OR PROXIMITY:** ○ **–:** Minor ○ **Empty:** Moderate ○ **+:** Strong or well developed in case of whirld sand and clouds with shape of bottleneck **VC:** In the vicinity
G	descriptor **SIGNIFICANT METEOROLOGY, DESCRIPTOR:**

Founding Members
EUROPEAN UNION EUROCONTROL

Column	Ítem – Description
	○ **MI:** Low ○ **BC:** Bank ○ **PR:** Partial (Aerodrome partially covered) ○ **DR:** lift by the wind (altitude lower than 2m) ○ **BL:** lift by the wind (altitude higher than 2m) ○ **SH:** Squall ○ **TS:** Storm ○ **FZ:** Frozen
H	phenomena **SIGNIFICANT METEOROLOGY, PHENOMENON:** ○ **DZ:** Drizzle ○ **RA:** Rain ○ **SN:** Snow ○ **SG:** Snow grain ○ **IC:** Ice Crystal ○ **PL:** Granulated Ice ○ **GR:** Hail ○ **GS:** Granulated snow ○ **UP:** Unknown precipitation ○ **BR:** Mist ○ **FG:** Fog ○ **FU:** Fume ○ **VA:** Volcanic Ash ○ **DU:** Dust ○ **SA:** Sand ○ **HZ:** Haze ○ **PO:** Dust whirlwind ○ **SQ:** Squall ○ **FC:** clouds with shape of bottleneck ○ **SS:** Sand storm ○ **DS:** Dust storm e.g. +SHRA-> Strong rain squalls
I	referenceLoadFile

Table 7: metarWindShear_Facts table

Column	Ítem – Description
A	metarKey
B	initialValidDateKey (yyyymmdd) UTC Date corresponds to the hour of the last meteorological observation. (Valid from)
C	initialValidTimeKey (9HHMMSS) UTC Time corresponds to the hour of the last meteorological observation. (Valid from)

Column	Ítem – Description
D	aerodromeCode ICAO AERODROME INDICATOR. It corresponds to the ICAO code of the Aerodrome. E.g. LEMD
E	isSpeci: (0=METAR, 1=SPECI) Type of message, it can be: • **METAR:** Aerodrome Routine Meteorological Report. Sent every hour. • **SPECI:** Similar to the METAR message sent not evenly but punctually.
F	runway **WS+INDICATOR/RUNWAY or ALLRWY.** i.e. WS R17 Windshear in lower layers in the takeoff or landing trajectories or both of them, over the head of the runway specified.
G	minVisibility **Visual Range, RVR (m)**
H	referenceLoadFile

5.2.3 Scale

The data stored are 6 CSV files (amounting 10.7 Mb of structured data) from METAR source, for one month (April 2016) inside the entire Spanish airspace [4].

5.2.4 Range of available dates and areas

METAR is available for Spanish aerodromes, from 2012 to present.

Table 8: Spanish aerodromes

GCFV	LEAS	LEGA	LEMG	LETL
GCGM	LEBA	LEGE	LEMH	LETO
GCHI	LEBB	LEGR	LEMO	LEVC
GCLA	LEBG	LEGT	LEPA	LEVD
GCLP	LEBL	LEHC	LEPP	LEVS
GCRR	LEBT	LEIB	LERI	LEVT
GCTS	LEBZ	LEJR	LERS	LEVX
GCXO	LECH	LELC	LERT	LEXJ
LEAB	LECO	LELL	LESA	LEZG
LEAL	LECV	LELN	LESO	LEZL
LEAM	LEDA	LELO	LEST	
LEAO	LEEC	LEMD	LESU	

5.2.5 Join procedure with other sources

- NOAA: METAR initial date and time (columns B and C) crosses with the nearest prediction available for NOAA (both point and time). [METAR information is to be crossed with NOAA information in order to validate NOAA information at surface, especially in the vicinities of airports]

5.3 SIGMET

5.3.1 Metadata

N/A

5.3.2 Fields detailed description (structured data)

Table 9: SIGMET table

Column	Ítem - Description
A	sigmetKe
B	receptionDate (yyyymmdd) UTC Date corresponds to the hour of message reception.
C	receptionTime (9HHMMSS) UTC Time corresponds to the hour of message reception.
D	startDateKey (yyyymmdd) UTC Date corresponds to the hour of validity beginning. (Valid from)
E	startTimeKey (9HHMMSS) UTC Time corresponds to the hour of validity beginning. (Valid from)
F	endDateKey (yyyymmdd) UTC Date corresponds to the hour of validity end. (Valid to)
G	endTimeKey (9HHMMSS) UTC Date corresponds to the hour of validity end. (Valid to)
H	airblockName The smallest division of the airspace is the Airblock. These Airblocks then form Sectors. This field mark which are the Airblocks affected by the message.
I	offifcialSectorName Name of the sector affected by the message. This name corresponds to the official sectorization at the moment of validity of the message.
J	meteoArea Volume definition of the area where the message holds. Written as a polygon for SQL Server format
K	lowerBound Lower limit in FL
L	upperNound Upper limit in FL
M	movement Movement of the phenomena defined at the message: • En Movimiento: Moving • Estacionario: Stationary
N	direction If the phenomenon is moving, indicates the direction where it is going. Defined by cardinal points
O	speed If the phenomenon is moving (and if specified in the message), indicates the speed at which it is moving. Defined in knots
P	intensityChange Defines the changes in intensity of the phenomena: • Sin Cambios: No Changes • En debilitación: Weakening • En intensificación: Intensifying
Q	meteoPhenomena Defines the phenomena itself (in the list only appears the cases in the provided first set of

Column	Ítem - Description
	data, this list is to be amplified in a later version of the document): • Turbulencia Fuerte: Strong Turbulence • Tormentas Inmersas: Embedded Thunderstorms
R	fir FIR where the message was issued
S	isPrediction Defines whether the message comes after the phenomenon was either forecasted or obversed: • 0: Observation • 1: Forecasted
T	sourceFilename

5.3.3 Scale

The data stored is a CSV files (amounting 400 Kb of structured data) from SIGMET source, for one month (April 2016) inside the entire Spanish airspace [4].

5.3.4 Range of available dates and areas

SIGMET is available for the Spanish airspace, from 2012 to present.

5.3.5 Join procedure with other sources

- NOAA: As SIGMET is defined in 4D (specific volume in the airspace plus timespan), several data coming from NOAA will fall in that region at that specific time. [NOAA information is to be crossed with SIGMET messages in order to predict different phenomena and define dangerous meteorology even at the lack of SIGMET messages]
- IFS (Radar) and Flight Plans: Similarly to what is used for NOAA, consider if each point falls inside the volume defined in the SIGMET message (column J) and then compare the deviations of the radar track with respect to the flight plan. [With information coming from both sources and having defined an area where SIGMET messages affect, it can be determined the horizontal and vertical efficiency of the flight]
- Airspace: Already crossed. The airspace affected by SIGMET is provided at column J, and more specifically the airblocks affected (column H) and the sector affected (column I), which holds for the available and official sectorization at the SIGMET initial validity.

5.4 IFS

5.4.1 Metadata

On each IFS file some metadata is listed at the beginning and the end of the files. First lines describe some system information without relevance for the rows, indicating the period contained on the file. In fact, this information may be useless too because each day is processed on different files so all files contains the info of a single day that can be discovered by the name of the file.

The raw data files are named with the first letter according to the region and the date on format yyMMdd. On the structured version of the data, each row contains a column with the source file providing traceability to the files.

5.4.2 Fields detailed description (structured data)

Table 10: IFS table

Field	Ítem - Description
factId	This field identifies uniquely the row of data.
flightKey	This Key identifies uniquely the flight
callsign	This string is the call sign of the flight. These call signs are reused permanently by flights so it would be repeated over the data.
Adep	Aerodrome of departure (If is indicated)
ades	Aerodrome of destination (If is indicated)
flightRule	Rules of the flight (I – Instrumental; V – Visual)
wake	Wake of the plane (J = Super, H = High, M = Medium, L = Low). This means how close a plane can be following to another one. The bigger is the plane the farer has to be.
Aircraft	Aircraft type
Date	Point's date [yyyy-MM-dd]
Time	Point's time of the day [HH :mm:ss]
lat	Latitude [DDD]
lng	Longitude [DDD]
modo_c	Altitude (Hundreds of feet]
vel_mod	Module of the horizontal speed [Knots]
hdg	Heading [degrees from 0 to 360]
Vel X	Speed on X axis [Knots]
Vel Y	Speed on Y axis [Knots]
Vel Z	Speed on Z axis [Feet/min]

5.4.3 Scale

The data stored are csv files for each week (amounting 1GB of structured compress data) from IFS source, for one month (**April 2016**) inside the entire **Spanish airspace**.

5.4.4 Range of available dates and areas

Available for the Spanish airspace, from 2013 to present.

5.4.5 Join procedure with other sources

- NOAA: For every point in the radar track together with its time, its nearest NOAA prediction (again for both point and time) provides the meteorological required information. [With the different points along the track of a flight, NOAA information can provide different meteorological information to each of these points.]
- SIGMET: : Similarly to what is used for NOAA, consider if each point falls inside the volume defined in the SIGMET message (column J) and then compare the deviations of the radar track with respect to the flight plan. [With information coming from both sources and having defined an area where SIGMET messages affect, it can be determined the horizontal and vertical efficiency of the flight]

5.5 Sector configuration

5.5.1 Metadata

This data source has no metadata information.

5.5.2 Fields detailed description (structured data)

The structured Sector Configuration data is organized in two tables. SC_CONFIGURATIONS table contains actual ATC configurations used for each Spanish ATC, and SC_SECTORISATION table provides the division of Spanish airspace in sectors.

These tables are linked, by the columns:

- ATC Units Name: ‘SC_CONFIGURATIONS table Column E’ with ‘SC_SECTORISATION table Column D’
- Configuration: ‘SC_CONFIGURATIONS table Column D’ with ‘SC_SECTORISATION table Column F’

Also respect to ATC Units Name, the relation between ‘SC_CONFIGURATIONS Column E’ and ‘SC_SECTORISATION Column D’ is:

SC_CONFIGURATIONS table	SC_SECTORISATION table
CANARIAS	Canarias
BARCELONA_APROXIMACION	Barcelona TMA
BARCELONA_RUTA	Barcelona Ruta
LEVANTE	Valencia APP
MADRID_NORTE	Madrid Ruta 1
MADRID_SUR	Madrid Ruta 2
PALMA	Palma APP

SEVILLA	Sevilla
MADRID_APROXIMACION	Madrid TMA
-	Málaga APP

- *SC_CONFIGURATIONS table*

This table contains information of the actual configuration (column D) used for each Spanish ATC units (column E), according to the period (column B and column C) of day (column A).

Table 11: SC_CONFIGURATIONS table

Column	Ítem	Description
A	day	yyyy-mm-dd
B	Hour from	HH:MM:ss
C	Hour to	HH:MM:ss
D	Configuration	
E	ATC units Name	

- *SC_SECTORISATION table*

The airspace is divided into a number of sectors (column G), each one of them is assigned to an active Control Working Position, into tasks with manageable workload (capacity, column H). This table contains the division of Spanish airspace in sectors, as well as the configurations (Column F) and ATC unit (column D) to which each sector belongs, according to the period of validity (column A and column B), since the design of airspace is variable.

Table 12: SC_SECTORISATION table

Column	Ítem	Description
A	Date From	Beginning of the validity period
B	Date to	End of the validity period
C	ATC	
D	ATC Units Name	
E	ATC Configuration	
F	Configuration	
G	Sector	Each sector is made up of volumes
H	Capacity	
I	Volume	Each volume is made up of airblocks
J	Airblock ID	Each airblock is define by a polygon and it lower and upper
K	Lower Bound	
L	Uper Bound	

Column	Ítem	Description
M	Airblock Polygon	
N	Airblock Polygon String	
O	Airblock Polygon WKB	

5.5.3 Scale

The data stored are two CSV files (amounting 81.3 MB of structured data) for one month (April 2016) inside the entire Spanish airspace:

- SC_CONFIGURATIONS_20160: Structured data (CSV): 216 KB (1 month - April 2016)
- SC_SECTORISATION_201604: Structured data (CSV): 79.8 MB (1 month: for April 2016 are 2 periods of validity)

5.5.4 Range of available dates and areas

Available for Spanish airspace.

- SC_CONFIGURATIONS table: More than 3 years of structured data (from 02/03/2007 to present)
- SC_SECTORISATION table: More than 9 years of structured data (from 01/01/2013 to present)

5.5.5 Join procedure with other sources

The key field for linking this data with other sources is 'Table 2. Column I' Volume with 'Table 3. Column E' Volume name from GIPV source.

- NOAA: Definition by time and location of the different points from NOAA that are inside specific sectors in the airspace. [Expected to be at the maximum possible sectorization, the definition of which NOAA nodes belong to each sector may affect the sectorization in the future]
- SIGMET: Already crossed. The airspace affected by SIGMET is provided at SIGMET's column J, and more specifically the airblocks affected (column H) and the sector affected (column I), which holds for the available and official sectorization at the SIGMET initial validity.

5.6 GIPV

5.6.1 Metadata

N/A

5.6.2 Fields detailed description (structured data)

The structured GIPV data is organized in the next four tables. FP_FLIGHTS table contains main information (Flight plan message reception time, Type of message, Call sign, Departure, Destination…) which contain only one item in same METAR message.

The rest of the tables (FP_ROUTE_POINTS, FP_VOL_SACTA and FP_SEGMENTS) contain information (route points, sectors and segments) which can have more than one item in the same message, so it was decided to structure in different tables. These tables are linked with the main table (FP_FLIGHTS table) through the keys field “idMsgPV” and “fileName”, (“column A” and “column C” of each table).

All times in the GIPV data source are in UTC, except the “instant” item (FP_FLIGHTS table_Column F) which is in Spanish local time.

Each row of this table contains a flight plan message with its corresponding information. Each flight has several flight plan messages.

The usual behaviour of flight plan messages by each flight includes a creation message and a cancellation message. The intermediate messages are modifications/updates, activations (i.e., when the flight enters the FIR) or terminations (when the flight has either landed or left the FIR).

For univocally identify each particular flight and all its flight plan messages, next items are used: Callsign (“column G”), Departure airport (“column R”), Destination airport (“column AC”), and IOBT – initial off block time- (“column I”).

Table 13: FP_FLIGHTS table

Column	Item	Description
A	idMsgPV	This Key identifies uniquely a flight plan message (for those messages with the same filename (“column C”)
B	dateReference	File Date reference
C	fileName	File Name
D	nBytes	Bytes number
E	mType	Type of message CREAP: Creation ACTP: Activation TERP: Termination CANCP: Cancellation
F	instant	Reception Date of the flight plan message (Spanish local time))Spain is GMT/UTC + 1h during Standard Time and GMT/UTC + 2h during Daylight Saving Time)
G	Indicativo	Flight Identification
H	ETOT	Estimated Take-Off Time
I	IOBT	Initial Estimated Off-Block Time
J	SSRE	Surveillance equipment
K	DCFL	CFL in Departure Aerodrome
L	SLCL	SELCAL code
M	TDR	Crossing point is a radar note (S/N)

N	TRFL	Transition Flight Level (100 ft)
O	RUTA	Route ponts in plain text (This information is structured **FP_ROUTE_POINTS table**)
P	CONF	Confirmed flight plan Indicator F:Finished C:Controlled E:Inbound D:Departure Confirmed N:No confirmed
Q	MOBT	DMAN off block date and time
R	DPT	ICAO location indicator of the aerodrome of departure
S	ELDT	Estimated Landing date and time
T	EOBT	Estimated Off block date and time
U	ACGR	Aircraft group
V	DEFL	Estimated Flight Level in Departure Aerodrome (100 ft)
W	RMK	Remark
X	ABI	ABI or PAC reception flag (S/N)
Y	SDFL	Second flight Level **(100 ft)**
Z	VOL_SACTA_SECTORS	Overflown sectors in plain text (This information is structured in **FP_VOL_SACTA table**)
AA	DPRW	Departure runway
AB	AOBT	Actual off block date and time
AC	DST	ICAO location indicator of the aerodrome of Destination
AD	CTOT	Calculated Take-Off Time
AE	FLTY	Type of Flight
AF	CSPD	Cruise speed
AG	ORUT	AFTN original route
AH	OBS	Observations
AI	SID	Standard Instrument Departure
AJ	REG	Identifier of a Regulation concerning a flight
AK	SEGM	Segments in plain text (This information is structured in **FP_SEGMENTS table**)
AL	ARRW	Arrival runway
AM	MTOT	DMAN take-off date and time
AN	AADR	The ICAO location indicator of the first alternate aerodrome of destination
AO	HPAC	Pre-activation Hour
AP	CREG	Code of a Regulation Concerning a Flight
AQ	FLRL	Flight rules
AR	RFL	Requested Flight Level (100 ft)
AS	SRUT	Calculated route provided by flight data processing system
AT	IFP	Indication of known errors within a FPL
AU	STAR	Standard Arrival

AV	SRC	Indication of the data source
AW	ATOT	Actual take-off date and time
AX	PDST	Previous destination aerodrome
AY	MLDT	AMAN Landing date and time
AZ	HFIR	Estimated date and time to FIR
BA	HSLC	SLC Time
BB	ACNB	Aircraft number and Aircraft Group
BC	NCAE	Radio communication, navigation and approach aid equipment
BD	DPFL	Departure Planned Flight Level (100 ft)
BE	QUBC	Replacement Flight Plan (RFP) indicator
BF	MCFP	Manually coordinated flight plan flag (S/N)
BH	TRWP	Previous departure aerodrome
BG	PDPT	Flight Plan WayPoint
BI	UOBT	Last received EOBT (date and time)
BJ	HOBT	Manual EOBT
BK	ALDT	Actual Landing date and time
BL	ACWT	Wake Turbulence Category
BM	HTOT	Hand Take-Off Time

- *FP_ROUTE_POINTS table*

Each row of this table contains a route point information for an specific flight plan message. This table is linked with FP_FLIGHTS table through the keys field “idMsgPV” and “fileName”, (“column A” and “column C” of each table).

Table 14: FP_ROUTE_POINTS table

Column	Item	Description
A	idMsgPV	This Key identifies uniquely a flight plan message (for those messages with the same filename (“column C”)
B	dateReference	File Date reference
C	fileName	File Name
D	POINT_SEQUENCE_NUMBER	POINT_SEQUENCE_NUMBER
E	POINT_NAME	POINT_NAME
F	X_Y_COORDINATE	X_Y_COORDINATE
G	GAT_OAT_INDICATOR	GAT/OAT Indicator
H	TYPE_OF_STIMAGE	Type of estimate A: automatic. M: manual. T: CORR-CMON.

Column	Item	Description
I	ETO	Estimated Time Over
J	CALCULATED_CROSSING_LEVEL	Calculated Crossing level (100 ft)
K	CFL	Cleared Flight Level (100 ft)
L	ECL	En-route Cruising Level (100 ft)
M	VFR_IFR_INDICATOR	VFR-IFR Indicator
N	SPEED	Speed
O	SECTION_INDICATION	Section Indicator
P	TYPE_OF_POINT	Type of point
Q	LATITUDE	Latitude (degree)
R	LONGITUDE	Longitude (degree)

- *FP_VOL_SACTA table*

Each row of this table contains overflown volume information for a specific flight plan message. This table is linked with FP_FLIGHTS table through the keys field "idMsgPV" and "fileName", ("column A" and "column C" of each table).

Table 15: FP_VOL_SACTA table

Column	Item	Description
A	idMsgPV	This Key identifies uniquely a flight plan message (for those messages with the same filename ("column C")
B	dateReference	File Date reference
C	fileName	File Name
D	VOLUME_SEQUENCE_NUMBER	Volume sequence number
E	VOLUME_NAME	Volume name
F	ENTRY_LEVEL	Entry level (100 ft)
G	EXIT_LEVEL	Exit level (100 ft)
H	ENTRY_TIME	Entry Time

- *FP_SEGMENTS table*

Each row of this table contains segment information for a specific flight plan message. This table is linked with FP_FLIGHTS table through the keys field "idMsgPV" and "fileName", ("column A" and "column C" of each table).

Table 16: FP_SEGMENTS table

Column	Item	Description
A	idMsgPV	This Key identifies uniquely a flight plan message (for those messages with the same filename ("column C")
B	dateReference	File Date reference
C	fileName	File Name
D	SEGMENT_SEQUENCE_NUMBER	Segment sequence number

Column	Item	Description
E	STATE_IN_SEGMENT	**State in Segment** PTE: Pending, COO: Coordinated, PRE: Preactive, ACT: Active, ETR: Hold, ETE: Terminated.
F	FIRST_VOLUME	First volume
G	LAST_VOLUME	Last volume
H	FIRST_FIX_POINT	First Fix Point
I	LAST_FIX_POINT	Last Fix Point
J	ASSIGEND_SSR_CODE	Assigned SSR code
K	PREVIOUS_SSR_CODE	Previous SSR code
L	SACTA_REGION	**SACTA Region** LECM: Madrid LECB: Barcelona LEPA: Palma LECS: Sevilla GCCC: Canarias

5.6.3 Scale

The data stored are 4 CSV files per 10-days period (amounting 12.9 Gb of Structured data), from GIPV source, for one month (April 2016) inside the entire Spanish airspace.

5.6.4 Range of available dates and areas

GIPV is available for Spanish airspace, from 2013 to present.

5.6.5 Join procedure with other sources

- NOAA: For every point in the flight plan together with its time, its nearest NOAA prediction (for both point and time) provides the forecasted meteorological required information. [Crossing information coming from NOAA (specifically the predictions) and the flight plan to follow by an aircraft, the different expected meteorological aspects of the flight can be defined]
- IFS (Radar) and SIGMET: Similarly to what is used for NOAA, consider if each point falls inside the volume defined in the SIGMET message (column J) and then compare the deviations of the radar track with respect to the flight plan. [With information coming from both sources and having defined an area where SIGMET messages affect, it can be determined the horizontal and vertical efficiency of the flight]

5.7 CFMU

5.7.1 Metadata

N/A

5.7.2 Fields detailed description (structured data)

Column	Item	Description
A	LOBT	Local on board time
B	FlightUID	Flight unique ID
C	ADEP	Aerodrome of departure
D	ADES	Aerodrome of destination
E	EFTimeOfEntry	Estimated First Time Of Entry of the flight in the FPM (Date)
F	AFTimeOfEntry	Actual First Time Of Entry of the flight in the FPM
G	AiracCycle	Id of the Airac cycle
H	AircraftId	Call sign
I	AircraftType	Aircraft type (i.e. A320)
J	ATOT	Actual Take Off Time
K	CTOT	Calculated Take Off Time
L	ETOT	Estimated Take Off Time
M	IOBT	Initial Off Block Time
N	TaxiTime	Time of taxi from the stand to the take-off position.
O	ATA	Actual Time of Arrival
P	CTA	Calculated Time of Arrival
Q	ETA	Estimated Time of Arrival
R	mostPenalizingRegulationId	Most penalizing regulation that the flight has suffered.
S	mostPenalReguATO	ATO of the most penalizing regulation that the flight has suffered.
T	mostPenalReguEtO	ETO of the most penalizing regulation that the flight has suffered.
U	RegulationsCount	Number of regulations that the flight has suffered
V	ATFDelay	Delay assigned to the flight.
W	RegulationReasonCode	Regulation Reason · C: ATC capacity · I: ATC industrial action · R: ATC routeing · S: ATC staffin · T: ATC equipment · A: Accident / incident · G: Aerodrome capacity · D: De-Icing · E: Equipment NON-ATC · N: Industrial action NON-ATC · M: Airspace management · P: Special event · W: Weather · V: Environmental issue like noise · O: Other

Column	Item	Description
X	ReferenceLocation	A code used in the Central Flow Management Unit (CFMU) Air Traffic Flow and Capacity Management (ATFCM) systems to describe one or more aerodromes, an Air Traffic Control (ATC) waypoint or an entire ATC sector.
Y	ReferenceLocationType	Type of the regulation location (AD - Aerodrome, AS - Airspace)
Z	RegulationStart	Start time of the regulation
AA	RegulationEnd	End time of the regulation

5.7.3 Scale

The data stored are 1 CSV files per 2-years period (amounting 840 Mb of Structured data) inside the entire Spanish airspace.

5.7.4 Range of available dates and areas

CFMU is available for Spanish airspace, from 2013 to present.

5.7.5 Join procedure with other sources

Just using time fields (i.e: RegulationStart) and ReferenceLocation as Sector.

References

[1] DART, DART 699299 Grant Agreement, v1.0

[2] Grid identification, master list of NCEP storage grids. Grib edition 1
http://www.nco.ncep.noaa.gov/pmb/docs/on388/tableb.html

[3] ICAO Guide – Manual of Aeronautical Meteorological Practice (Doc 8896)

[4] AEMET (Spanish State Meteorological Agency) Guide
http://www.aemet.es/documentos/es/conocermas/aeronautica/AU-GUI-0102.pdf

[5] DART, DART 699299 Consortium Agreement, v1.0

[6] European Comission, Guidelines on Data Management in Horizon 2020, Version 2.1, 15/02/2016

[7] DART, DART D 1.1, Data Management Plan

[8] DART, DART D 1.2, Data Transaction Pipeline

[9] DART, DART MS1, Internal Dataset Description

[10] DART, DART, Internal Dataset Updating