H2020-ICT-2015

Grant Agreement No: 687591

Big Data Analytics for Time Critical Mobility Forecasting

datAcron

D6.3 Aviation Experiments Specification

Deliverable Form	
Project Reference No.	H2020-ICT-2015 687591
Deliverable No.	6.3
Relevant Work Package:	WP 6
Nature:	R
Dissemination Level:	PU
Document version:	1.0
Due Date:	31/12/2016
Date of latest revision:	10/01/2017
Completion Date:	10/01/2017
Lead partner:	CRIDA
Authors:	Iglesias Martínez, Enrique; Scarlatti, David; López Leones, Javier; López, Elena; García-Ovies Carro, Iciar
Reviewers:	Vouros, George; Santipantakis, Giorgos; Doulkediris, Christos; Fuchs, Georg; Andrienko, Natalia; Alevizos, Elias; Pelekis, Nikos; Cordero, Jose Manuel
Document description:	This deliverable details the experiments specification for the aviation domain.
Document location:	/WP6/Deliverables/D63

History of changes

Version	Date	Changes	Author	Remarks
0.1	14/12/2016	Creation	Iglesias Martínez, Enrique; Scarlatti, David; López Leones, Javier; López Elena	
0.2	23/12/2016	Inclusion of reviewers' comments.	Iglesias Martínez, Enrique; Scarlatti, David; López Leones, Javier; López, Elena; García-Ovies Carro, Iciar	
0.3	03/01/2017	Inclusion of reviewers' comments.	Iglesias Martínez, Enrique; Scarlatti, David; López Leones, Javier; López, Elena; García-Ovies Carro, Iciar	
1.0	10/01/2017	Final version	Iglesias Martínez, Enrique; Scarlatti, David; López Leones, Javier; López, Elena; García-Ovies Carro, Iciar	

TABLE OF CONTENTS

TERMS & ABBREVIATIONS

ACC	Area Control Center
ADS-B	Automatic Dependent Surveillance-Broadcast
AIC	Aeronautical Information Circulars
AIP	Aeronautical Information Publication
ALS	Alarm Service
AMDT	AIP Amendments
ANSP	Air Navigation Service Providers
ASM	Air Space Management
ATC	Air Traffic Control
ATCO	Air Traffic Controller
ATFM	Air Traffic Flow Management
ATM	Air Traffic Management
ATZ	Aerodrome Traffic Zones
CAS	Calibrated Air Speed
CFMU	Control Flow Management Unit
CM	Center of Mass
CNS	Communication, Navigation and Surveillance
CTA	Control Traffic Areas
CTOT	Calculated Take Off Time
CTR	Controlled Traffic Regions
D-ATIS	D-ATIS Digital Automatic Terminal Information Service
DDR	Demand Data Repository
DOC	Direct Operational Costs
DSTs	Decision Support Tools
EAS	Equivalent Air Speed
ECAC	European Civil Aviation Conference
ETA	Estimated Time of Arrival
ETD	Estimated Time of Departure
ETOT	Estimated Take Off Time

FDPs	Flight Data Processing Systems
FIR	Flight Information Region
FIS	Flight Information Service
FMS	Flight Management System
GS	Ground Speed
IAF	Initial Approach Fix
IAS	Indicated Air Speed
IATA	International Air Transport Association
ICAO	International Civil Air Organization
IFR	Instrumental Flight Rules
IMC	Instrument Meteorological Conditions
IOC	Indirect Operational Costs
MET	Meteorological Information Services
NARI	The Naval Academy Research Institute
NOAA	National Oceanic and Atmospheric Administration
NOTAM	Notice to Airmen
QNE	pressure when the altimeter is regulated with standard 1013.25 hPa isobar
QNH	barometric pressure adjusted to sea level – Query: Nautical Height
QoS	Quality of Service
RMSE	RMSE (Root Mean Square Error)
RNAV	Area Navigation
RNP	Required Navigation Performance
SID	Standard Instrument Departure
STAR	Standard Terminal Arrival Route
SUP	AIP Supplements
TAS	True Air Speed
TCAS	Traffic Alert and Collision Avoidance System
TMA	Terminal Manoeuvring Areas
ToC	Top of Climb
ToD	Top of Descent
TP	Trajectory Prediction

UIR	Upper Information Region
UOCs	User Operations Centers
VFR	Visual Flight Rules
VMC	Visual Meteorological Conditions

LIST OF TABLES

1. INTRODUCTION

1.1 Purpose and Scope

Experiments specification describes all the experiments to be conducted for the validation of datAcron developments. This document intended to be used by all datAcron partners, and not only for subject matter experts.

Once the basis of the different ATM scenarios has been understood by all partners, a comprehensive description of the experiments to be performed by the different work packages is needed. This deliverable specifies the experiments to be conducted in detail. The specifications include the datasets to use, the steps to perform, what to measure and the targeted values. The document will specify all the details needed to perform the experiments in a rigorous and repeatable way: What exactly will be measured and how, thresholds defining good or bad results will be included.

This document also aims at clarifying the tasks that need to be performed by each work package, as well as the expected results of each of them in every scenario.

1.2 Approach for the Work package and Relation to other Deliverables

The technological developments in datAcron will be validated and evaluated in user-defined challenges that aim at increasing the safety, efficiency and economy of operations concerning moving entities in the Air-Traffic Management and Maritime domains. The overall objective of work package 6 (WP6) is to validate the research results by means of experiments relevant to an Aviation Industry (ATM) use case. It relates directly to proposal objective 5: **"[O.5] Validation and evaluation of the datAcron system and individual components on the surveillance of moving entities in the ATM and marine domains."**

This document constitutes the definition of the experiments needed to carry out the validation of the research results and it is based on the description of the validation scenarios described in document D6.1 (Aviation use case detailed definition) and D6.2 (Aviation data preparation and curation). Future deliverables include the final version of the Aviation data preparation and curation (D6.4), datAcron prototype (D6.5) and final validation report (D6.6).

1.3 Methodology and Structure of the Deliverable

This document has been developed by industry experts and provides a comprehensive description of the experiments to be conducted in each aviation scenario.

For each scenario the same structure has been followed:

- Objectives: general description of the scenarios, as stated in D6.1.
- Experiments description: detailed explanation of the experiments to develop, providing a step-by-step methodology to carry them out.
- Data: specification of the data sources and datasets required to perform the experiments.
- Expected results: brief description of the outcomes expected in each work package.
- Evaluation metrics: Metrics to assess the quality of results and components' performance.
- Validation criteria: criteria to use in order to validate the results obtained from each of the experiments.

2. datAcron INITIATION PHASE EXPERIMENTS FOR FLOW MANAGEMENT SCENARIOS

2.1 Scenario FM01 – Flow Management – Regulations detection and prediction

Objectives

This scenario objective is to demonstrate how datAcron regulations detection and prediction capability is useful for reproducing Flow Management Behavior. This behavior is mainly represented by the applied regulations, it means, the system must be able to reproduce those regulations to anticipate a realistic scenario. These regulations are consequences of specific situations as those in which there is an excess of demand vs capacity in sectors, or those caused by different weather conditions, among others.

Regulation is a measure that a flow manager takes to solve a specific situation, in a punctual moment in a certain sector and it is applied over those flights that have not yet took off. The main consequence of a regulation is the re-schedule of an ETOT (Estimated take off time) by a CTOT (Calculated take off time) that is a new time to take off after the first one that causes a delay. As it has been said before, those ETOTs that are replaced by the CTOTs correspond to those flights that were going to fly in the affected sector during a punctual moment.

Hence, there are three objectives:

1. Investigate the available historical data in order to identify patterns in the emergence of the regulations. The patterns thus identified should suggest possible approaches to regulation prediction.
2. Develop a method or methods for regulation prediction based on the patterns identified.
3. Verify the method(s) by comparing predictions based on available historical data (without regulation data) with the real regulations.

Investigate historical data

It is not known at the beginning what kinds of patterns can exist. It is therefore necessary to look at the data from various perspectives using interactive visual displays as well as various filters and data transformations. The possible types of patterns that need to be searched for are:

- Temporal patterns, such as regularities with respect to the daily and weekly time cycles.
- Spatial patterns, e.g., regulations emerging in a certain area usually affect flights associated with certain origin and/or destination areas.
- Spatio-temporal patterns, e.g., different temporal patterns of regulations in different areas.
- Dependencies, e.g., some kinds of regulations with certain properties lead (after some time lag) to other regulations.

The visual analytics toolkit needs to provide the following functionality:

Visualization	Interactive tools
Visualize in summarized form information about the regulations: spatial positions (reference locations), times of issuing, start, end, and cancellation, numbers	Select subsets of regulations based on the spatial locations, times, reasons, severity in terms of the number affected flights and the

of affected flights, total and average delays.	produced delay.
Visualize the overall effects of the regulations: origin and destination airports or areas, numbers of affected flights and total or average delays by the airports or areas, affected origin-destination links and respective numeric information (flights and delays).	Same as above Select subsets of regulations based on the affected origin and/or destination airports or areas.
Visualize the distribution of the assigned delay durations among the delayed flights regarding their origins, destinations, distances, durations, and aircraft sizes.	Same as above.
(For a chosen regulation) Visualize relevant details of the affected flight plans; particularly, show where and when the flights were supposed to enter and exit the affected sector(s).	Select a regulation of interest.
(For a chosen regulated flight) Visualize the differences between the flight plan and the actual flight: (a) the temporal differences for the departure and arrivals; (b) the deviations from the planned route; (c) the differences between the times and locations of entering and exiting the affected sectors.	Select a flight of interest.
(For a subset of regulations, particularly, caused by bad weather conditions) Visualize the weather conditions relevant to the regulations.	Select subsets of regulations based on the reason (particularly, weather), spatial location (affected sectors), and time. Select weather attributes to be visualized.
(For a chosen sector) Find and visualize sequences of regulations that occurred in this sector shortly one after another (given a maximal allowed time gap between the consecutive regulations). Visualize the corresponding information concerning the regulated flights.	Select a sector to explore. Interactively set and modify the time gap threshold.
(For a chosen origin-destination pair or subset of pairs) Visualize the spatial and temporal distribution of the regulations that affected the corresponding flights. Show the reasons of these regulations and their severity with respect to the selected OD links, in particular, the proportions of the regulated flights among all flights and the average delays.	Select subsets of flights based on the origins and destinations.

Table 1: Visualization and interactive tools

Experiments description

The experiments for this scenario involve flight plan data, sector configuration data, weather information data and flow management information data.

datAcron needs to know the reason of the given regulations to know how each one affect the initial flight plans and how to predict them.

Data found in the applied regulations are:

- Day of the regulation
- Regulation ID
- Reference Location Type
- Traffic volume
- Period Start
- Period End
- Regulation Reason Code
- Regulated flights

Once regulations are known (and for which reason, e.g.: weather, ATC capacity, accident/incident…) flight plans have to be checked to know how these regulations affect them. Experiments will be divided in two parts because information and data used differ depending on the regulation reason code.Weather-related regulations are studied in a different way than those for other causes (under-capacities).

In the experiments it will be evaluated:

- Usability and responsiveness: How easy to use and flexible is the system while querying under different specific information (e.g: date of regulation, reason code, regulation traffic volumes, etc).
- Performance: Necessary time to link CFMU data with DDR and NOAA.
- Completeness: Information which is lost for several reasons: impossibility to link it with the rest of data sources, corrupted files, etc.
- Accuracy: Regulations which are correctly predicted in time and space.

datAcron prototype will be used by the user to retrieve and search information from the regulated sectors and filter them according to different types of queries (e.g., date, departure airport, arrival airport, callsign, airline or aircraft type) to check the regulated flight plans.
Regulated flight plans and the regulation reason codes could be studied to predict how future regulations affect to foreseen flight plans.

Experiment 1 (capacity regulations)

1. Structuring CFMU data in order to get general information about regulations (particularly those in which the regulation reason is specified as ATC Capacity (C)). CFMU data is divided in two tables: regulations and flights, and the latter is the one that contains the specific flights implied in a certain regulation.
2. CFMU data should be then linked with DDR Airspace data sources in order to obtain the sectorisation which was active at the time of the regulation.
3. Finally, DDR Flight Plan data should be linked with CFMU data and DDR Airspace data in order to include the trajectory of the flight and complete the information. In this way, it is possible to know where a certain flight has flown.

Experiment 2 (weather regulations)

1. Structuring CFMU data in order to get general information about regulations (particularly those in which the regulation reason is specified as Weather (W)).
2. CFMU data should be then linked with DDR Airspace data sources in order to obtain the sectorisation which was active at the time of the regulation.

3. In order to know which meteorological phenomena have caused a certain weather regulation, it is necessary to link the previous information with NOAA.
4. Finally, DDR Flight Plan data should be linked with CFMU data and DDR Airspace data in order to include the trajectory of the flight and complete the information. In this way, it is possible to know where a certain flight has flown.

Data

The data involved in this scenario is:

- Flight plans: DDR service
- Sector Configuration data: DDR service
- Weather Information data: NOAA
- Flow Management Information data: CFMU source

Expected results

The main goal of this scenario is to check the possibility to predict regulation events based on patterns existing in historical data.
For this goal, it is necessary to investigate historical data to establish a good performance based on integrated information. This information must be studied to identify patterns for regulations and predict how future regulations affect foreseen flight plans. Methods to predict regulations have to be developed, exploiting the identified patterns. Also, completeness must be assured before accuracy of the prediction methods is checked.
The expected result may be positive or negative. A positive result would mean that the historical data contain patterns that are suitable for creation of predictive models, and these models can sufficiently well reproduce the real regulations described in the historical data. A negative result would mean that the real regulations are unpredictable, particularly, due to the absence of clear regular patterns in the historical data. Both results should be treated as valid.
The aviation domain experts will review the dissimilarity measure(s) and the results of applying the measure(s) and give their judgments concerning the dissimilarity tolerance threshold(s).
In a case of a negative result, the validity of this result needs to be proven by visualizations clearly demonstrating the absence of regular patterns that could allow making reasonable predictions.

Evaluation Metrics

- Usability and responsiveness: Subjective evaluation is needed here. The user will need to select Flight Plan Data and weather information, when it is needed, according to several domain-specific parameters (e.g. number of hours to evaluate in the experiment, dates, number of flights) and be able to predict regulations.
- Performance: Average time to link CFMU data with DDR and NOAA.
- Completeness: Percentage of information which is lost (not linked with other data sources)
- Accuracy: Percentage of regulations correctly predicted in type, time and space.

datAcron WP contribution to the metrics:

	WP1	WP2	WP3	WP4
Usability and responsiveness	Provide historical data with an appropriate format to be			Determine patterns using WP1 information.

	used by other WP's.			
Performance	Information integration.		Methods for regulation prediction using WP4 information.	
Completeness	Assure the completeness of data.			
Accuracy			Assessment of the prediction accuracy.	Verify the methods for regulation prediction.

Table 2: datAcron WP contribution to FM01

Validation Criteria

The validation criteria will be applied for an experiment with a sufficient wide and representative set of regulations and flight plan data.

- Usability and responsiveness: If the information is easy to be integrated and be examined for patterns. Ok/NOT-Ok by user experience.
- Performance: Average time to integrate CFMU data with DDR and NOAA is below 10 seconds.
- Completeness: Percentage of lost information is below 1%.
- Accuracy: Percentage of regulations correctly predicted is over 99% or proven negative result.

2.2 Scenario FM02– Flow Management – Demand and capacity imbalance detection and prediction

Objectives

This scenario objective is to demonstrate how datAcron events' detection and prediction capability is useful for detecting demand and capacity imbalance by means of indicators monitoring.
Those indicators are based on real demand (Hourly Entry Count: for a given sector is defined as the number of flights entering in this sector during one hour) and declared capacity (Maximum number of flights allowed to enter in a sector during one hour) of the current configuration of airspace, by calculating them from the initial flight plan (deregulated traffic), denoted as M1 below, instead of the real flight plan.

- M1: flight plan that each flight has just before taking off. It is the last filed flight plan from the airline.
- M2: This message contains information only if the flight has been regulated.
- M3: the last updated flight plan enhanced with radar data. It is created when the flight has finished.

The main reason for that lays on the fact that if a flight has been regulated, its flight plan is modified to take off later and avoid the previously detected excess of demand. Thus if the real flight plan were used in that scenario, it would not be possible to detect the imbalance because this one would be avoided thanks to the regulation already applied.
Although in theory an imbalance could be produced by an excess of capacity compared with the demand, it should be an unusual situation that is out of our scope. datAcron will be focus on the excess of demand vs capacity.

Scenario FM02 focuses on detection and prediction of the events of excessive demand, i.e., when the demand for a sector (= the number of flights that need to enter the sector) exceeds the sector capacity. Based on detected excessive demands, it is supposed to find the relevant regulations in the historical database and simulate the application of these regulations: for the flights that were affected by the regulations, replace their original flight plans by the modified flight plans prescribed by the regulations. The modifications could lead to new events of excessive demand, which could require application of further regulations. Therefore, the sequence of steps "detect excessive demands – find and apply regulations – take modified flight plans" needs to be performed iteratively.
The final objective is to reconstruct the system's behavior in handling demand-capacity imbalances. This will allow us, in particular, to investigate propagation of the consequences of the regulations, that is, delaying some flights in a given time window may lead to increasing the demand in the next time window. It may also be useful to investigate the dependence of the behavior on the chosen time window length, e.g., what would happen if the currently adopted time window length of 1 hour is replaced by 30-minutes time window. Furthermore, it may be also reasonable to compare the use of fixed time windows with the use of a sliding time window. In the latter approach, the demand is calculated not from the beginning of an hour but from the time when each flight enters a zone.

Experiments description

The experiments for this scenario involve flight plan data, sector configuration data, weather information data and flow management information data, the same as in the first flow management scenario.

datAcron prototype will predict when demand will be higher than capacity studying the different flight plans.

In the experiments it will be evaluated:

- Usability and responsiveness: how easy is to detect an imbalance due to demand is higher than capacity.
- Performance: Time to calculate the demand in a certain sector based on DDR Flight Plans.
- Completeness: Information which is lost for several reasons: impossibility to link it with the rest of data sources, corrupted files, etc.
- Accuracy: Imbalances which are correctly predicted in time and space.

datAcron prototype will be used by the user to detect demand and capacity imbalances.

Experiment 1 (capacity-demand M1)

1. Extract information from DDR Flight Plan data in order to get M1 flight plans.
2. Obtain expected demand studying M1.
3. Obtain the declared capacity of each sector from 'Sector configuration data'.
4. Compare demand-capacity from M1 information.
5. With the information of previous steps, provide an accurate forecasting of imbalances using flight plan information (M1).

Experiment 2 (capacity-demand M3)

1. Extract information from DDR Flight Plan data in order to get M3 flight plans.
2. Obtain real demand studying M3.
3. Obtain the declared capacity of each sector from 'Sector configuration data'.
4. Compare demand-capacity from M3 information.
5. With the information of previous steps, provide an accurate forecasting of imbalances using flight plan information enhanced by radar data (M3)

Once both experiments are completed, it is necessary to assess the differences between them.

Data

The data involved in this scenario is:

- Flight Plan Data: DDR service
- Sector Configuration data: DDR service

Expected results:

In this scenario, datAcron develops a model capable to reconstruct the current behavior of the ATFM in terms of appearance of imbalances. To do so, it is necessary to carry out an investigation of the historical data available, in order to do the information linking with completeness. This will allow the construction of demand which will be compared with the existing capacity to detect imbalances.

After that, the validity of the model is assessed in terms of the correspondence between the real and reconstructed behaviours. Finally, the model will be used for testing "what-if" scenarios with the use of time windows of different lengths and sliding time windows. The model is expected to uncover how changes of the approach to assessing the demands may affect the number of the necessary regulations, the number of affected flights, and the overall delays. Ideally, the model will help to find an optimal approach to assessing the demands and balancing them against the capacities.

Evaluation Metrics

- Usability and responsiveness: Subjective evaluation is needed here. The user will need to evaluate in which window length it is reliable to detect an imbalance.
- Performance: Average time to calculate the demand in each sector based on DDR Flight Plans.
- Completeness: Percentage of information which is lost (not integrated with other data sources)
- Accuracy: Percentage of imbalances correctly predicted in time and space.

datAcron WP contribution to the metrics:

	WP1	WP2	WP3	WP4
Usability and responsiveness	Provision of historical data with an appropriate format to be used by other WP's.		Detect excessive demand events (hotspots).	
Performance	Information integration performance.	Trajectory reconstruction performance.	Event detection performance (entrance to sector and sector exit).	Explore consequences of excessive demand events.
Completeness	Assure completeness			Analyse detected events, regulations caused by these events and consequences of these regulations.
Performance			Accuracy of events detection.	Verify the methods. for imbalance prediction

Table 3: datAcron WP contribution to FM02

Validation Criteria

The validation criteria will be applied for an experiment with a sufficient wide and representative set of regulations and flight plan data.

- Usability and responsiveness: Reliable window length to predict an imbalance: 20-30min before the imbalance.
- Performance: Average time to calculate the demand in each sector based on DDR Flight Plans is below 10 seconds.
- Completeness: Percentage of information which is lost is below 1%.
- Accuracy: Percentage of imbalances correctly predicted in time and space is over 99%.

Type of regulation and regulated traffic volumes are specified in CFMU data. However, specific flight plans (M1, M2 and M3) must be checked in DDR files.

2.3 Scenario FM03– Flow Management – Resilience assessment

Objectives

This scenario objective is to demonstrate how datAcron resilience assessment and prediction capability is useful for evaluating and forecasting the system capacity to assume imbalance under nominal conditions. For that purpose, it will be necessary to compare each detected imbalance (from the scenario FM02) with the decision taken by the flow manager assessed by means of the scenario FM01 (a real imbalance with a regulation application to solve it may be found, but also it is possible to find an imbalance without any regulation because the flow manager considered that the system was able to absorb it). Once these situations are characterized, datAcron will establish patterns of those which get a better system behavior characterization.

The scenario aims at (1) detecting the cases when flow managers did not issue regulations despite of the demands exceeding the capacities, (2) investigating the properties of these cases, and (3) discovering the conditions when an excess of the demand over the capacity can be tolerated without issuing a regulation.

Experiments description

The experiments for this scenario involve flight management scenario FM01 "Regulations detection and prediction" and FM02 "Demand and capacity imbalance detection and prediction".

datAcron prototype will detected imbalances and how these imbalances affect flights. Also, it will have to detect when there is an imbalance but there is not a regulation.

Data used in these experiments are the same as in scenarios FM01 and FM02. In fact, this scenario is to compare differences between both of them.

In the experiments it will be evaluated:

- Usability and responsiveness: How easy to use, flexible and responsive is the system while querying under different specific information about experiments of flow management scenarios FM01 and FM02.
- Performance: Time to compare scenarios FM01 and FM02, that is, regulations and imbalances.
- Completeness: Information which is lost for several reasons: impossibility to link it with the rest of data sources, corrupted files, etc.
- Accuracy: Correctly detection of imbalances which have a regulation associated to them. Correct detection of imbalances with no regulation associated.

datAcron prototype will detect when an imbalance is going to take place in regulated and no regulated sectors to improve the decision taken by the flow manager.

Experiment 1 (imbalance with regulation)

1. Study FM02 scenario outputs in order to get information about demand and capacity imbalances.
2. Contrast imbalances (FM02) with regulations (FM01) to detect those imbalances with a regulation associated.
3. Evaluate which is the reason of the regulation.

Experiment 2 (imbalance without regulation)

1. Study FM02 scenario outputs in order to get information about demand and capacity imbalances.
2. Contrast imbalances (FM02) with regulations (FM01) to detect those imbalances with no regulation associated.
3. Evaluate which is the reason of the not applied regulation.

Data

The data involved in this scenario is:

- Scenario FM01: Regulations detection and prediction.
- Scenario FM02: Demand and capacity imbalance detection and prediction.

Expected results

The expected results are, first, to gain understanding (mental model) of the ATFM system resilience, i.e., capability to tolerate certain cases of excessive demand and, second, to build a classification model capable to recognize such cases. The model will be validated by testing against the historical data.

Evaluation Metrics

- Usability and responsiveness: How easy to use, flexible and responsive is the system while querying under different specific information about experiments of flow management scenarios 1 and 2. Detect imbalances in scenario FM02, and regulated flights in FM01.
- Performance: Average time to compare scenarios FM01 and FM02, that is, regulations and imbalances
- Completeness: Percentage of lost information (not linked with other data sources)
- Accuracy: Percentage of imbalances with/without a regulation associated which are correctly predicted.

datAcron WP contribution to the metrics:

	WP1	WP2	WP3	WP4
Usability and responsiveness	Provision of data and events detected in scenarios FM01 and FM02 with an appropriate format to be used by other WP's.			Investigate results obtained in scenarios FM01 and FM02.

Performance			Correlation between regulations and imbalances.	Performance of the comparison between results from scenario FM01 and FM02.
Completeness	Assure data completeness.			
Accuracy			Accuracy of the correlation between results from scenario FM01 and scenario FM02.	

Table 4: datAcron WP contribution to FM03

Validation Criteria

The validation criteria will be applied for an experiment with a sufficient wide and representative set of regulations and flight plan data.

- Usability and responsiveness: Ok/no-Ok (detected imbalance).
- Performance: Average time to compare imbalances and regulations is below 10 seconds.
- Completeness: Percentage of lost information is below 1%
- Accuracy: Percentage of imbalances with/without a regulation associated which are correctly predicted is over 99%.

This scenario is based on the two previous one of flow management to get all the information that is needed.

3. datAcron INITIATION PHASE EXPERIMENTS FOR FLIGHT PLANNING SCENARIOS

3.1 Scenario FP01 - Flight Planning - Real Trajectory Reconstruction

Objectives

This scenario objective is to demonstrate how datAcron trajectory reconstruction capability is useful for building the real trajectories of aircraft both off-line and real-time. The trajectories need to be reconstructed from the surveillance data (ADS-B messages and/or radar data). Additional information needed to assign departure and/or destination to each trajectory will be the airport coordinates. When available, additional information about the aircraft can be added to the trajectory, i.e. Model, Owner… The trajectories reconstructed are expected to be spatio-temporal objects.

Experiments description

The experiments for this scenario involve surveillance data, used to create the synopsis of the trajectory, operational context information (airports) and aircraft database. The flight plans associated to each trajectory reconstructed from surveillance data are involved as part of the validation. In the experiments it will be evaluated the following:

- Usability: How easy to use and flexible is the system while reconstructing trajectories from the synopsis and querying under different specific information (airports and aircraft specific information to those trajectories).
- Responsiveness: How responsive is the system while reconstructing trajectories from the synopsis and querying under different specific information (airports and aircraft specific information to those trajectories)
- Performance: Reconstruction time and airport and aircraft linking.
- Realism: The reconstructed trajectories associate the correct information (airport & aircraft).
- Compression: Number of messages discarded when creating the synopsis that will be used for reconstructing the trajectory
- Completeness: Number of original aircraft positions (included in the original surveillance messages) that can be retrieved when reconstructing the trajectory from the synopsis
- Accuracy: RMSE for the positions and time of the reconstructed trajectories when compared with the original messages

datAcron prototype will be used by the user to retrieve reconstructed trajectories from surveillance data (through the synopsis of the trajectories) for different types of queries (e.g., date, departure airport, arrival airport, callsign, airline or aircraft type). A parallel data processing for these queries is expected in order to improve performance and responsiveness of the system.

The reconstructed trajectories are from synopsis of surveillance data. The stream of surveillance data used for this validation will be based on a replay of historic flights in order to be possible to repeat the experiments.

The experiment will be repeated with increasing number of flights to study the scalability of the system

Experiment 1 (stored information)

1. Obtain a unique set of trajectories for:
   - A given time range: local time or UTC time
   - A given set of airports
   - A given identifier: flight id or callsing, hex_id, tail number
   - A given 3D volume (lat,lon,altitude)
2. Retrieve the information associated to each trajectory reconstructed: identifiers (flight id or callsign, hex_id), origin and destination airports, aircraft data (aircraft model,...)
3. Retrieve the original messages and files associated to queried trajectories. This is needed to double check that the creation of synopsis and the reconstruction is executed correctly
4. Retrieve the number of original messages discarded when creating the synopsis
5. Obtain the values (position, time, altitude,…) of the reconstructed trajectory for those points included in the original messages discarded and compare the accuracy of the reconstruction
6. Compare information retrieved from 2) with information obtained from the Flight Plan associated (this will use the datAcron capabilities to support FP02 experiments)

Experiment 2 (real time processing)

1. Obtain a unique set trajectories for:
   - A given time range: local time or UTC time
   - A given set of airports
   - A given identifier: flight_id or callsing, hex_id, tail number
   - A given 3D volume (lat,lon,altitude)
2. Retrieve the information associated to each trajectory reconstructed in real time: identifiers (flight id, callsign, hex_id, aircraft address), origin and destination airports, aircraft data (aircraft model,...)
3. Retrieve the number of original messages discarded when creating the synopsis
4. Obtain the values (position, time, altitude,…) of the reconstructed trajectory for those points included in the original messages discarded and compare the accuracy of the reconstruction
5. Compare information retrieved from 2) with information obtained from the Flight Plan associated (this will use the datAcron capabilities designed to support FP02 experiments)
6. For a subset of streaming trajectories, queries at different instants will be performed to observe if there are any changes in the information retrieved

Data

The data involved in these experiments are:

- Surveillance data: Flight Aware and ADSB-HUB
- Airports Database: AIXM from NM service
- Aircraft Database: CSV file with aircraft information
- Flight Plan (for validation): DDR service. For this scenario, it will be chosen a subset in where there is one flight plan per trajectory reconstructed.

Evaluation Metrics

- Usability: Subjective evaluation is needed here. The user will need to select airport/s, aircraft information (e.g., aircraft type/s, callsign, airline/s) or time and be able to retrieve all the associated reconstructed trajectories.

- Responsiveness: Subjective evaluation is needed here. The user will need to select airport/s, aircraft information (e.g., aircraft type/s, callsign, airline/s) or time and be able to retrieve all the associated reconstructed trajectories and evaluate its user experience
- Performance: Average time to reconstruct the trajectories from surveillance data (i.e. Total time to generate all the trajectories divided by number of trajectories requested) and to answer the different queries.
- Realism: Number of trajectories with the correct data coming from the provided files. Number of reconstructed trajectories with any incorrect information associated
- Compression: Number of original messages (not including outliers)
- Completeness: Number of original messages that cannot be retrieved in the reconstructed trajectory

datAcron WP contribution to the metrics:

	WP1	WP2	WP3	WP4
Usability	Capability to find the selected flight/trajectory according to the filtering parameters			User interface used for finding the trajectories
Responsiveness	Serving the data needed to WP2 may influence experience of the user when using the prototype	Performance of the reconstruction algorithms		
Performance	Serving the data needed to WP2 may influence the performance	Performance of the reconstruction algorithms		
Realism	Accuracy of the linking algorithms to convey the correct data			Can help to result interpretation if visual components are developed that allow to analyze the experiment result. Subjective contribution.
Compression		Creation of the synopsis		
Completeness		Reconstruction algorithms allow the creation of the aircraft positions identified in the original messages Identification of outliers (those surveillance messages containing wrong information about the flight (wrong position, time, aircraft identifier,…)		
Accuracy		Reconstruction		

		algorithms allow the creation of the aircraft positions identified in the original messages at the exact position and time		

Table 5: datAcron WP contribution to FP01

Validation Criteria

The validation criteria will be applied for an experiment with a sufficient wide and representative set of surveillance data. I.e. it must span different months and have enough number of flights in each day. Total number of flights will represent a Big Data challenge, i.e. thousands of flights.

- Usability and responsiveness: Ok/NOT Ok by user experience.
- Responsiveness: Ok/NOT Ok by user experience.
- Performance: Average time to reconstruct the trajectories must allow to reconstruct a set of trajectories for a day in a few hours.
- Realism: Number of trajectories with all the data coming from the provided files: OK if =90-100%. Number of reconstructed trajectories with any incorrect information associated: OK if between 1% and 2%.
- Compression: Number of original messages discarded when creating the synopsis. Outliers will be taken out of this count
- Completeness: Number of original messages which position cannot be retrieved from the reconstruction algorithm. OK if between 1% and 2%.
- Accuracy: For pair of points compared, the RMSE reference values for the difference in time, altitude and position shouldn't be higher than:
    - Altitude: 200-300 feet
    - Horizontal plane distance: 0.01-0.02 nautical miles
    - Time: 1-2 seconds

3.2 Scenario FP02 - Flight Planning - Real Trajectory Enrichment

Objectives

This scenario objective is to demonstrate how datAcron data management capability can help for add (link) new data to real trajectories. The trajectories reconstructed from the surveillance data (ADS-B messages and/or radar tracks) need to be enriched with data from the aircraft (when known), data from the weather, operational context data, and associated Flight Plans. This step will create spatio-temporal objects with more data associated to them that may be key data in later scenarios like event detection, trajectory clustering and/or trajectory prediction.

Experiments description

The experiments for this scenario involve surveillance data, weather, operational context information, and flight plans. In the experiments it will be evaluated:

- Usability and responsiveness: How easy to use, flexible and responsive is the system while querying under different specific information (e.g.,: callsign, runways, etc) and aircraft specific information those trajectories) .
- Performance: Time to add the new information to the reconstructed trajectories.
- Realism: The enriched reconstructed trajectories associate the correct information.
- Completeness: Number of trajectories with no flight plans associated and vice versa. Number of enriched fields empty for each trajectory

datAcron prototype will be used by the user to retrieve and search information from the trajectories and filter them according to different types of queries (e.g., date, departure airport, arrival airport, callsign, airline or aircraft type). A parallel data processing for these queries is expected in order to improve performance and responsiveness of the system.
The enriched trajectories could be from synopsis of surveillance data.. The stream of surveillance data used for this validation will be based on a replay of historic flights in order to be possible to repeat the experiments.
The experiment will be repeated with increasing number of flights, i.e., 5, 25, 100, and 500.

Experiment 1 (stored information)

1. Obtain a unique set trajectories filtered by fields that need to be in the enriched trajectory:
   a. Operational context information: runway, SID/STAR procedure, routes, waypoints, sectors
   b. Weather information: temperature, pressure, wind at any trajectory point
   c. Flight plan information: callsign, aircraft model, airline, origin and destination airports, ETD, ETA, equipage, cruise level, cruise speed
2. Retrieve the new enriched information associated to each trajectory reconstructed: Operational context information, Weather information, Flight plan information
3. Retrieve the files used for the enrichment of any set of trajectories. This is needed to double check that the enrichment is executed correctly
4. Retrieve those trajectories with at least one enriched field empty (those are considered incomplete enriched trajectories)

Experiment 2 (real time processing)

1. Obtain a unique set of trajectories for the new enriched information:
    - Operational context information: runway, SID/STAR procedure, routes, waypoints, sectors
    - Weather information: temperature, pressure, wind
    - Flight plan information: callsign, aircraft model, airline, origin and destination airports, ETD, ETA, equipage, cruise level, cruise speed
2. Retrieve the enriched information associated to each trajectory reconstructed: Operational context information, Weather information, Flight plan information
3. Retrieve those trajectories with any enriched field empty (those are considered incomplete enriched trajectories)
4. For a subset of streaming trajectories, queries at different instants will be performed to observe if there is any changes in the information retrieved

Data

The data involved in these experiments are:

- Surveillance data (stored):Flight Aware and ADSB-HUB
- Surveillance data (streaming): Flight Aware and ADSB-HUB.
- Airports Database: AIXM from NM service
- Weather Data: GRIB files from NOAA
- Aircraft Database: CSV file with aircraft information
- Flight Plan: DDR service

Evaluation Metrics

- Usability: Subjective evaluation is needed here. The user will need to select different enriched trajectory information (e.g., weather, runway, SID/STAR, callsign, airline/s) and be able to retrieve all the associated enriched trajectories.
- Responsiveness: Subjective evaluation is needed here. The user will need to select different enriched trajectory information (e.g., weather, runway, SID/STAR, callsign, airline/s) and be able to retrieve all the associated enriched trajectories in a period of time acceptable for the user
- Performance: Average time to enrich the trajectories from surveillance data (i.e. Total time to generate all the trajectories divided by number of trajectories requested) and to answer the different queries.
- Realism: Number of reconstructed trajectories with any incorrect information associated
- Completeness: Number of trajectories with no flight plans associated and vice versa. Number of enriched fields empty for each trajectory

datAcron WP contribution to the metrics:

	WP1	WP2	WP3	WP4
Usability	Capability to find the selected flight/trajectory according to the filtering parameters			User interface used for finding the trajectories
Responsiveness		Performance of the reconstruction		

		algorithms offline and real time may influence the perception of the user		
Performance		Performance of the reconstruction algorithms offline and real time		
Realism	Accuracy of the linking algorithms to convey the correct data			Can help to result interpretation if visual components are developed that allow to analyze the experiment result. Subjective contribution.
Completeness		Identification of outliers		

Table 6: datAcron WP contribution to FP02

Validation Criteria

The validation criteria will be applied for an experiment with a sufficient wide and representative set of surveillance data. I.e. it must span different months and have enough number of flights in each day. Total number of flights will represent a Big Data challenge, i.e. thousands of flights.

- Usability: Ok/NOT Ok by user experience.
- Responsiveness: Ok/NOT Ok by user experience.
- Performance: Average time to enrich the trajectories must allow to enrich a set of trajectories for a day in a few hours.
- Realism: Number of reconstructed trajectories with any incorrect information associated: OK if between 1% and 2%.
- Completeness: Number of flight plans with no enriched trajectory and vice versa: OK if between 1% and 2%.
- Number of enriched trajectories with any empty fields OK if between 1% and 2%. Number of empty fields per trajectory: Ok if average between 1% and 2%.

3.3 Scenario FP03 - Flight Planning - Event Recognition in trajectories

Objectives

This scenario objective is to demonstrate how datAcron complex event recognition capability can help for detecting relevant events in trajectories. The events are added to the trajectories and, like in previous scenario, this new data can be relevant for next scenarios like trajectory clustering or trajectory prediction.

Experiments description

The experiments for this scenario involve trajectories, either reconstructed from surveillance or generated with a trajectory predictor, i.e., synthetic trajectories. When referring in the description with "trajectories", means it applies to both, reconstructed and synthetic. Notice that NO enriched trajectories should be used for the event recognition, BUT they will be used in the validation. In the experiments it will be evaluated:

- Usability: How easy is to use and how flexible is the system while querying for events and visualizing those events.
- Responsiveness: User experience when querying and visualizing events (encompasses detection time, integration and storage time, retrieval upon querying)
- Performance: Time to identify events for any given set of trajectories and to add them to the trajectories. Time to filter a set of trajectories by events
- Realism: The events are not mixed. No artificial-ghost events are created
- Completeness: All the events per trajectory are detected.
- Accuracy: The events are identified in the correct geolocalization or time instant

datAcron prototype will be used by the user to retrieve and search event information from the trajectories and filter them according to different types of queries (e.g., type of event, localization of the event in 4D, …)
The experiment will be repeated with increasing number of flights, i.e., 5, 25, 100, and 500.
The events that should be identified are classified in the following categories:

- Basic: FIR/Sector crossing points, Turnings (radius and direction) vs straight flight (great circle), Terminal to Enroute / Enroute to Terminal transitions, Surface to Terminal / Terminal to Surface transitions, Hold on ground, Trajectory Change Point (Altitude)
- Medium: Top of Climb , Top of Descent, FlyOver / FlyBy on Waypoints and Navaids, Trajectory Change Point (Speed)
- Complex: SID, STAR, ROUTE, Takeoff Runway, Destination Runway, Crossover Altitude, Transition Altitude

It is expected that datAcron event recognition capabilities will be able to detect basic and medium and some complex events from the provided trajectories

Experiment

1. Retrieve a set of trajectories according with a specific filtering condition, based on FP01 capabilities
2. Identify a set of basic, medium and complex events for that set of trajectories.
3. Filter/query trajectories based on the events detected.
4. Compare events already associated to trajectories retrieved from step (2) with corresponding events in the enriched reconstructed trajectories (resulting from FP02 trajectory enrichment). These events would be a match or not a match. Doing so, one may

detect for instance deviations among events in reconstructed trajectories and events in historic trajectories.

5. Compare the events retrieved from 2) with the events previously detected in the reconstructed trajectories. The 4D position of the events will be compared, as well as the number of unidentified events and the number of false positive events created by 2)
6. Compare the events retrieved from 2) with the events identified by the user in the synthetic trajectories. The 4D position of the events will be compared, as well as the number of unidentified events and the number of false positive events created by 2)

Data

The data involved in this scenario is:

- Reconstructed (enriched and not) trajectories
- Synthetic trajectories

Evaluation Metrics

- Usability: Subjective evaluation is needed here. The user will need to select different trajectories and be able to identify events, visualize them and group those trajectories associated to a common event
- Responsiveness: Subjective evaluation is needed here. The user will need to select different trajectories and be able to identify events, visualize them and group those trajectories associated to a common event in an acceptable period of time
- Performance: Average time to detect events for each trajectory. Average time to filter trajectories by events (i.e. Total time to identify all the events for all the trajectories divided by number of trajectories requested).
- Realism: Number of trajectories with any incorrect event associated
- Completeness: Percentage of events not identified
- Accuracy: RMS for the geographical events detected

datAcron WP contribution to the metrics:

	WP1	WP2	WP3	WP4
Usability	Capability to find the queried event and filter flights/trajectories by events			Geographical representation of the trajectories with visual labels attached to the point where events are detected.
Responsiveness	Serving the data needed to WP3 and Wp4 may influence in the user experience			
Performance	Serving the data needed to WP3 and Wp4 may influence in performance		Performance of the event recognition algorithms	
Realism			No artificial/ghost events are added	Can help the manual inspection of events by the user, displaying 2D, 3D profiles and also speed profile, airspace (waypoints,sectors,route

				s, SID/STAR procedures,...)
Completeness			All the events are detected	Can help the manual inspection of events by the user, displaying 2D, 3D profiles and also speed profile, airspace (waypoints,sectors,route s, SID/STAR procedures,...)
Accuracy			Events are detected in the correct 4D position	Can help the manual inspection of events by the user, displaying 2D, 3D profiles and also speed profile, airspace (waypoints,sectors,route s, SID/STAR procedures,...)

Table 7: datAcron WP contribution to FP03

Validation Criteria

The validation criteria will be applied for an experiment with a sufficient wide and representative set of trajectories spanning different months or having enough number of flights in each day. Total number of flights should represent a Big Data challenge, i.e. thousands of flights, although some validations that might require the user expertise would be done with a small set of flights

- Usability: Ok/NOT Ok by user experience.
- Responsiveness: Ok/NOT Ok by user experience.
- Performance: Average time to identify events must allow processing a set of trajectories for a day in a few hours.
- Realism: Number of trajectories with any incorrect event associated: OK if between 1% and 2%.
- Completeness: % of events no identified: OK if between 1% and 2%.
- Accuracy: For pairs of trajectories compared for the same event, the RMSE reference values for the difference in the location of that event shouldn't be higher than:
    - Altitude: 500-600 feet
    - Horizontal plane distance: 0.5-0.6 nautical miles
    - Time: 30-40 seconds

3.4 Scenario FP04 - Flight Planning - Event Forecasting in trajectories

Objectives

This scenario objective is to demonstrate how datAcron predictive analytics capability in complex event forecasting can help for predicting relevant events in trajectories on flight. After many runs of previous scenarios (event recognition) a baseline for learning will be available.

Experiments description

The experiments for this scenario involve a stream of surveillance data based on a replay of historic flights in order to be possible to repeat the experiments. This replay could be performed from real stored surveillance data or data extracted from synthetic trajectories (useful for certain validation steps). Additionally, operational context information, flight plans and weather will be available
In the experiments it will be evaluated:

- Usability and responsiveness: How easy is to use, flexible and responsive is the system while identifying and visualizing the events.
- Performance: Time to forecast an event before the event occurs
- Realism: An event is not wrongly identified. No false positive events are created
- Completeness: All the events per trajectory are detected.
- Accuracy: The events are detected in the correct geolocalization or time instant
- Robustness: The events are updated accordingly to the evolution of the flight

The experiment will be repeated with increasing number of flights, i.e., 5, 25, 100, and 500.
The classification of the events that need to be identified are described in the experiments for scenario FP03

Experiment

1. Retrieve a set of trajectories according to a specific filtering condition, based on FP01 capabilities
2. Identify in operational any event that has just occurred or might occur in the future in all the trajectories that are being analyzed.
3. Filter/query trajectories based on the events detected.
4. Compare events already associated to trajectories retrieved from step (2) with corresponding events in the enriched reconstructed trajectories (resulting from FP02 trajectory enrichment). These events would be a match or not a match. Doing so, one may detect for instance deviations among events in reconstructed trajectories and events in historic trajectories.
5. Compare the events retrieved from 2) with the events marked previously in the trajectories streamed. The 4D position of the events will be compared, as well as the number of unidentified events and the number of false positive events created by 2)

Data

The data involved in this scenario is:

- Surveillance data real time streaming: Flight Aware, ADS-B Hub and synthetic trajectories
- Context data: AIXM from NM
- Weather: grib2 from NOAA
- Flight plan: DDR

Evaluation Metrics

- Usability: Subjective evaluation is needed here. The user will need to select different trajectories and be able to identify events, visualize them and group those trajectories associated to a common event
- Responsiveness: Subjective evaluation is needed here. The user will need to select different trajectories and be able to access to the forecasted events, visualize them and group those trajectories associated to a common event
- Performance: Average time to forecast a type of event for each trajectory before the event occurs. Average time to filter trajectories by events (i.e. Total time to identify all the events for all the trajectories divided by number of trajectories requested).
- Realism: Number of trajectories with any incorrect or artificial event associated
- Completeness: Percentage of events no detected
- Accuracy: RMSE for the geographical events detected

datAcron WP contribution to the metrics:

	WP1	WP2	WP3	WP4
Usability	Capability to find the selected event and filter flights/trajectories by events			Geographical representation of the trajectories with visual labels attached to the point where events are detected.
Responsiveness	Serving the data needed to WP3 and Wp4 may influence the user experience			
Performance	Serving the data needed to WP3 and Wp4 may influence performance		Performance of the event recognition algorithms	
Realism			No artificial/ghost events are added	Can help the manual inspection of events by the user, displaying 2D, 3D profiles and also speed profile, airspace (waypoints, sectors, routes, SID/STAR procedures,...)
Completeness			All the events are detected	Can help the manual inspection of events by the user, displaying 2D, 3D profiles and also speed profile, airspace (waypoints, sectors, routes, SID/STAR procedures,...)
Accuracy			Events are detected in the correct 4D position	Can help the manual inspection of events by the user, displaying 2D, 3D profiles and also speed profile, airspace

				(waypoints, sectors, routes, SID/STAR procedures,...)

Table 8: datAcron WP contribution to FP04

Validation Criteria

The validation criteria will be applied for an experiment with a sufficient wide and representative set of trajectories having enough number of flights to represent a Big Data challenge, i.e. thousands of flights.. Some validations that might require the user expertise would be done with a small set of flights

- Usability and responsiveness: Ok/NOT Ok by user experience.
- Performance: Average time to forecast an event before it occurs
- Realism: Number of trajectories with any incorrect event associated: OK if between 1% and 2%.
- Completeness: % of events not detected: OK if between 1% and 2%.
- Accuracy: For pairs of trajectories compared for the same event, the RMSE reference values for the difference in the location of that event shouldn't be higher than:
    - Altitude: 500-600 feet
    - Horizontal plane distance: 0.5-0.6 nautical miles
    - Time: 30-40 seconds

3.5 Scenario FP05 - Flight Planning - Data Set preparation

Objectives

This scenario objective is to demonstrate how datAcron data management capability (querying of integrated spatio-temporal data) for all the data needed can help to prepare a dataset for subsequent scenarios. The trajectories and context data need to be grouped according to spatio-temporal boundaries.

Before executing subsequent scenarios a dataset must be prepared. Data belonging to dataset may be marked as belonging to that dataset so it can be queried and visually plotted. In case data is not tagged, preparing the dataset before any other scenario execution is an option as well in case the Data Set Preparation can be executed in a reasonable time for the operator.

Experiment Description

The experiments for this scenario involve all information preloaded in datAcron that has to be filtered in a space/time window.

datAcron prototype will be used by the user to input initial and end date/time and to graphically select a volume whose data will be included in the Data Set. At least, the geographical volume will be defined using a polygonal or circular region (normally a rectangle) plus a minimum and a maximum altitude. As volumes are defined in AIXM files representing FIR/Sectors, there should be an option that allows the user to filter data using these predefined volumes instead of manual selection of the volume.

The generated dataset must be assigned an id somehow so they can be retrieved as input in the following use cases.

Experiments will comprise dataset generation going from small datasets covering an airport area for an hour to big ones covering all European region for a month followed by a visual representation.

Once the dataset is generated, additional filtering can be added to the previous spatio-temporal feature. That includes creating the dataset filtering by:

- All aircraft matching an aircraft type.
- List of callsign.
- Origin and destination airport.
- Regions where the winds are in a selected range.
- Flights that have performed a holding pattern.
- Flight duration is in a selected range.
- All synopses of trajectories for a particular time interval and geographical region
- All flight plans for a particular time interval and geographical region
- All weather data for a particular time interval and geographical region
- All static and dynamic context data for a particular time interval and geographical region

Queries over the dataset and dataset filters must be visualized. A helpful visualization in order to validate this capability, would be a geographical representation of all the trajectories, flight plans, weather and context with the option of filtering using a sliding time window.

In the experiments it will be evaluated:

- Usability and responsiveness: How easy to use and responsive is the system while generating a new dataset.

- Performance: Once the new dataset is requested, how much time takes to the system to generate it.
- Accuracy: All data matching the filters must be included in the dataset. Random checkings of flightAware, network manager and weather data inclusion will be done, using the same filter over RAW data (externally) and comparing it to the dataset generated in datAcron.

Data

The data involved in this scenario is:

- Reconstructed real trajectories: from surveillance data
- Synthetic trajectories: generated from flight plans
- flight plan data: DDR
- weather data: grib2 from NOAA
- context data: AIXM

Evaluation Metrics

- Generation performance predictability. As the amount of information available for a time interval in a geographical region may be highly variable, establishing performance limit makes no sense, but when preparing a dataset the system must be predictable.
  That means that the response time preparing a dataset must grow linearly (or less) with the number of trajectories, flight plans, weather information and context data in the spatio-temporal window selected.
- Usability. The dataset generation can be a batch process with a predictable generation time but once generated, the visualization time and the filtering of loaded information subjective response time should be within acceptable values for the human operator.

datAcron WP contribution to the metrics:

	WP1	WP2	WP3	WP4
Generation performance predictability	Capability to find the selected flight plans, trajectories context and weather data in predictable response time.			User interface used for preparing the dataset.
Usability				User interface used for finding the Flight Plans and trajectories

Table 9: datAcron WP contribution to FP05

Validation Criteria

To check linearity of response time generating a dataset, different tests will be performed on busy air traffic areas, for instance:

- Query same geographical area for time and 2x time. Check generation time difference.
- Query same for the same timeframe a geographical window and 2x geographical window. Check generation time difference.

To validate usability, for the previously generated datasets and a big one including big areas (ie Europe-wide traffic for a day), human operators will be asked about if loading and filtering time is Ok / No Ok.
Visually, all information filtered will be check to belong to the geospatial window selected when generating the dataset. Filter information, specially geographical volumes should be visualized on screen when showing a dataset.

3.6 Scenario FP06 - Flight Planning - Trajectory Clustering

Objectives

This scenario objective is to demonstrate how datAcron trajectory clustering capability can work for both real and synthetic trajectories of aircrafts. The trajectories with common departure and destination may be clustered based on flight path / aircraft type / time / events…

Experiment Description

The experiments for this scenario involve all trajectory information preloaded in datAcron, that has to be clustered based on a selected filter.

datAcron prototype will be used by the user to visually input the following filters in order to cluster trajectories:

- Aircraft Identifier
- Aircraft Type
- Departure / Destination airport
- Aircraft Route (based on distance to a route obtained from context data)
- SID / STAR procedures followed
- Time
- Events

Routes/SID/STAR/ADEP/ADES can be obtained from context information.
Starting from the datasets created on the previous scenario, for each cluster an Experiment for each filter will be launched and evaluated. The experiment will be repeated increasing the dataset size from a small circular area of 50 miles around an airport for a day to a big polygonal area around Europe for a week.
In the experiments it will be evaluated:

- Usability and responsiveness: How easy to use and responsive is the system while clustering by filter
- Performance: Once the new clustering is performed, how much time takes to the system to generate it.
- Accuracy: Clustering evaluation will be based on trajectories pair comparison for randomly selected trajectories belonging to the same cluster.

Data

The data involved in this scenario is:

- Real Trajectories from surveillance data
- Synthetic Trajectories generated from flight plans
- Context data

Visualizations

One visualization which is helpful to validate this capability will be a geographical representation of the clustered trajectories with the option to switch to visualize the original trajectories included in each cluster.
When several clusters are visualized, the relative size of the clusters (in term of number of original trajectories included in the cluster) should be perceived visually.

Visual pattern recognition may be used to validate the identified outliers.

Evaluation Metrics

- Clustering performance: Time spent clustering trajectories should be always lower than time spent retrieving that set of trajectories individually or by group.
- Number of unassigned trajectories: Must tend to zero (outliers) for frequent routes. An access to a log file or an information window with information about clustering assignments will help to evaluate this capability.
- Visual pattern recognition. A clear pattern in data will be visually identified when doing visual analytics of clustered data. At the same time a graphical representation of the filtering option selected will help to visually identify the patterns.

	WP1	WP2	WP3	WP4
Clustering performance		Capability to associate trajectories		User interface used for finding trajectories
Unassigned trajectories		Performance of the clustering algorithms		
Visual pattern recognition.		Accuracy of the clustering algorithms and accuracy for detecting outliers.		Can help to result interpretation if visual components are developed that allow to analyze the experiment result. Subjective contribution.

Table 10: datAcron WP contribution to FP06

Validation Criteria

- Performance: Average time to cluster information: Ok below 1 minute average time for datasets in European region for 1 day traffic. The average will be done for datasets of different seasons and weather conditions.
- Usability and responsiveness: Ok/NO Ok by user experience.
- Accuracy: Will be measured based on trajectory pair analysis, performing comparison metrics using the datAcron trajectory comparison capability of scenario FP10. For pairs of trajectories clustered for the same route RMSE reference values generally won't be higher than:
  - Altitude: 500-600 feet
  - Horizontal plane distance: 300-400m

3.7 Scenario FP07- Flight Planning - Trajectory prediction - preflight

Objectives

This scenario objective is to demonstrate how datAcron predictive analytics capability can help in trajectory forecasting. For a given flight plan, a forecasted trajectory will be obtained and compared with the real one finally flown (Historical).

Experiment Description

The experiments for this scenario involve a number of Flight Plans for which there is already available a model-based predicted trajectory and real life flown trajectory.

datAcron prototype will be used by the user to find the Flight Plans desired for the evaluation, the Flight Plans need to be "searchable" by callsign, aircraft model, airline, origin and destination airports, ETD, ETA, equipage, cruise level, cruise speed.

datAcron prototype will be used by the user to select a number of Flight Plans of the result set (All typically) and request a predicted trajectory for each of them. For each prediction the end time of the generation is needed to compare it with the time of the request and calculate that way the time needed for each trajectory. Ideally the time used for generation is stored by the prototype together with the trajectory since for scalability evaluation many trajectories will be computed at once.

The generated trajectories will be stored in datAcron prototype and must be marked somehow so they can be retrieved later for the evaluation.

Using scenario FP10 datAcron capabilities the forecasted trajectories will be compared with the real flown trajectories to calculate the accuracy. In order to do that datAcron prototype will allow to assign to each trajectory forecasted another trajectory from the datAcron store (in this case a real flown one) for comparison. To make experiment execution easier it will be nice if the prototype allows to find the real trajectory flown searching by Flight Plan. The results of the comparison will be used for the evaluation of the prediction.
In the experiments it will be evaluated:

- Usability and responsiveness: How easy to use and responsive is the system while finding the Flights Plans required to use in the experiment.
- Performance: Once the prediction is requested, how much time takes to the system to generate the predicted trajectories.
- Accuracy: The predicted trajectories will be compared with the real flown trajectories.

The experiment will be repeated with increasing number of flights, i.e., 5, 25, 100, and 500.

Data

The data involved in this experiment is:

- Flight Plans: Flight Plans from DDR.
- Real Trajectories: Reconstructed real trajectories for a given Flight Plan must be available in datAcron infrastructure.
- Model based (synthetic) predicted trajectories: Model-based trajectories for a given flight plan must be available in datAcron infrastructure.

Evaluation Metrics

- Usability and responsiveness: Subjective evaluation is needed here. The user will need to search and find the flight plans selected for the experiment (based on criteria already mentioned: date, departure airport, arrival airport, callsign, airline and aircraft type)
- Performance: Average time to generate the predictions.(i.e. Total time to generate all the predictions divided by number of Flight Plan prediction requested).
- Accuracy: RMSE of the predicted trajectories for values of Altitude, Bearing, Horizontal plane distance, and Speed.

datAcron WP contribution to the metrics

	WP1	WP2	WP3	WP4
Usability and responsiveness	Capability to find the selected flight plans and relevant synoptic trajectories			User interface used for finding the Flight Plans
Performance	Serving the data needed to WP2 may influence the performance	Performance of the prediction algorithms	Events prediction performance (if used by WP2)	
Accuracy		Accuracy of the prediction algorithms	Events prediction accuracy (if used by WP2)	Can help to result interpretation if visual components are developed that allow to analyze the experiment result. Subjective contribution.

Table 11: datAcron WP contribution to FP06

Validation Criteria

The validation criteria will be applied for an experiment with a sufficient wide and representative set of Flight Plans. I.e. it must span different months and have enough number of flights in each day. Total number of flights will represent a Big Data challenge, i.e. thousands of flights.

- Usability and responsiveness: Ok/NO Ok by user experience.
- Performance: Average time to generate the predictions: Ok below 30-40 seconds per trajectory (average of all trajectories in a set for a single day)
- Accuracy: RMSE for all the predicted trajectories: OK if <=Model-Based RMSE.

3.8 Scenario FP08- Flight Planning - Trajectory prediction - preflight schedule based

Objectives

This scenario objective is to demonstrate how datAcron predictive analytics capability can help in trajectory forecasting. For a reduced set of flight plan fields, the airline schedule, a forecasted trajectory will be obtained and compared with the real one finally flown (Historical). The main difference with previous scenario is that in this case there is still not flight plan available, just the schedule, destination and departure.

Experiment Description

The experiments for this scenario involve a number of routes for which there is already available a model-base predicted trajectory and real life flown trajectory.

datAcron prototype will be used by the user to find the Flight Plans desired for the evaluation, the Flight Plans need to be "searchable" by date, departure airport, arrival airport, callsign, airline and aircraft type.

dataAcron prototype will be used by the user to select a number of Flight Plans of the result set (All typically) and request a predicted trajectory for them that will be calculated in parallel. For each prediction the end time of the generation is needed to compare it with the time of the request and calculate that way the time needed for each trajectory. Ideally the time used for generation is stored by the prototype together with the trajectory since for scalability evaluation many trajectories will be computed at once.

The generated trajectories will be stored in dataAcron prototype and must be marked somehow so they can be retrieved later for the evaluation.

Using scenario FP10 datAcron capabilities the forecasted trajectories will be compared with the real flown trajectories to calculate the accuracy. In order to do that datAcron prototype will allow to assign to each trajectory forecasted another trajectory from the datAcron store (in this case a real flown one) for comparison. The results of the comparison will be used for the evaluation of the prediction.

In the experiments it will be evaluated:

- Usability and responsiveness: How easy to use and responsive is the system while finding the Flights Plans required to use in the experiment.
- Performance: Once the prediction is requested, how much time takes to the system to generate the predicted trajectories.
- Accuracy: The predicted trajectories will be compared with the real flown trajectories.

The experiment will be repeated with increasing number of flights, i.e., 5, 25, 100, and 500.

Data

The data involved in this experiment is:

- Airline Schedule
    - Aircraft type
    - Aerodrome of departure
    - Aerodrome of destination
    - Date/time of departure

- Real Trajectories: Real trajectories for a given schedule must be available in datAcron infrastructure.

Evaluation Metrics

- Performance: Average time to generate the predictions.(i.e. Total time to generate all the predictions divided by number of Flight Plan prediction requested).
- Accuracy: RMSE of the predicted trajectories for values of Altitude, Bearing, Distance, and Speed.

datAcron WP contribution to the metrics

	WP1	WP2	WP3	WP4
Performance	Serving the data needed to WP2 may influence the performance	Performance of the prediction algorithms	Events prediction performance (if used by WP2)	
Accuracy		Accuracy of the prediction algorithms	Events prediction accuracy (if used by WP2)	Can help to result interpretation if visual components are developed that allow to analyze the experiment result. Subjective contribution.

Table 12: datAcron WP contribution to FP08

Validation Criteria

The validation criteria will be applied for an experiment with a sufficient wide and representative set of routes. I.e. it must span different months and have enough number of routes for each day. Total number of flights will represent a Big Data challenge, i.e. thousands of flights.

- Performance: Average time to generate the predictions: Ok below 30-40 seconds per trajectory (average of all trajectories in a set for a single day)
- Accuracy: RMSE for all the predicted trajectories: OK if <=reference. Reference values:
  - Altitude: 150-200 mt
  - Bearing: 5º-6º
  - Horizontal Plane Distance: 500-600m
  - Speed (V Ground): 2.5-3 m/s

3.9 Scenario FP09- Flight Planning - Trajectory prediction - real time

Objectives

This scenario objective is to demonstrate how datAcron predictive analytics capability can help in trajectory forecasting in real time. For a given flight plan and the current surveillance data arriving to the platform a forecasted trajectory will be obtained and updated continuously.

Experiment Description

The experiments for this scenario involve a real time stream of surveillance data for a number of flights and the corresponding flight plan.

datAcron prototype will consume a stream of surveillance based on a replay of historic flights in order to be possible to repeat the experiments. This stream will include more or less flights in order to evaluate scalability of the prototype.

datAcron prototype will publish a predicted trajectory for each flight present in the streaming that will be calculated in parallel. The predictions will be updated each X (varying number of) seconds.

The generated trajectories will be stored in datAcron prototype and must be marked somehow so they can be retrieved later for the evaluation.

Using scenario FP10 datAcron capabilities the forecasted trajectories will be compared with the real flown trajectories to calculate the accuracy. In order to do that datAcron prototype will allow to assign to each trajectory forecasted another trajectory from the datAcron store (in this case a real flown one) for comparison. The results of the comparison will be used for the evaluation of the prediction.

In the experiments it will be evaluated:

- Performance: Once the prediction is requested, how much time takes to the system to generate the predicted trajectories.
- Accuracy: The predicted trajectories will be compared with the real flown trajectories.

The experiment will be repeated with increasing number of flights, i.e., 5, 25, 100, and 500.

In the experiments it will be evaluated:

- Performance: Number of simultaneous trajectories the system can handle in parallel.
- Accuracy: The predicted trajectories will be compared with the real flown trajectories at a number of predetermined stages. (i.e. 25%,50%,75% of the flight).

Data

The data involved in this experiment is:

- Surveillance data real time streaming: datAcron lab will have the capability of replay in streaming any flight available in the infrastructure.
- Flight Plans: Flight Plans must be available in datAcron infrastructure. The Flight Plans need to be “searchable” by date, departure airport, arrival airport, callsign, airline and/or aircraft type.

Evaluation Metrics

- Performance: Average time to generate the predictions.(i.e. Total time to generate all the predictions divided by number of Flight Plan prediction requested).
- Accuracy: RMSE of the predicted trajectories for values of Altitude, Bearing, Distance, and Speed.

datAcron WP contribution to the metrics

	WP1	WP2	WP3	WP4
Performance	Serving the data needed to WP2 may influence in the performance	Performance of the prediction algorithms	Events prediction performance (if used by WP2)	
Accuracy		Accuracy of the prediction algorithms	Events prediction accuracy (if used by WP2)	Can help to result interpretation if visual components are developed that allow to analyze the experiment result. Subjective contribution.

Table 13: datAcron WP contribution to FP09

Validation Criteria

The validation criteria will be applied for an experiment with sufficient wide and representative set flights. Total number of flights will represent a Big Data challenge, i.e. hundreds/thousands of flights.

- Performance: Number of simultaneous trajectories to be updated each 5 seconds.
- Accuracy: RMSE for all the predicted trajectories: OK if <=reference. Reference values:
  - Altitude: 150-200 mt
  - Bearing: 5º-6º
  - Horizontal Plane Distance: 500-600m
  - Speed (V Ground): 2.5-3 m/s

3.10 Scenario FP10- Flight Planning - Trajectory comparison

Objectives

This scenario objective is to demonstrate how datAcron data management capability (querying of integrated spatio-temporal data) for all the data needed can help to calculate similarity metrics for trajectories generated in previous scenarios. The validation experiments will need this capability to measure the success of the trajectory prediction algorithms and the trajectory clustering algorithms.

Experiment Description

The experiments for this scenario involve the capability of being able to obtain RMSE comparison metrics any pair of trajectories available in datAcron. That means, calculating RMSE for:

- Lateral deviation (horizontal plane distance)
- Vertical profile deviation (vertical plane distance)
- ETA deviation
- Cross sector ETA deviation
- Cross sector position deviation
- Events deviation
- Bearing
- Speed

For each experiment, datAcron prototype will consume a stream of surveillance based on a replay of historic flights in order to be possible to repeat the experiments. Using external tools the comparison metrics for the injected flights will be calculated.
Based on the developed capabilities for previous scenarios the user must be able to select pair of trajectories to compare. That means:

- Selecting matching trajectories (surveillance and/or synthetic and/or datAcron generated). Usually will be trajectories associated to the same flight plan.
- Similarity of trajectories in a cluster.

Visualizations can be applied to the trajectories compared, the metrics of similarity can be shown attached to both trajectories.
In the experiments it will be evaluated:

- Performance: Once the comparison is requested, how much time takes to the system to generate the metrics.
- Accuracy: The metrics accuracy will be validated using external tools and compared with results obtained in datAcron.

Data

The data involved in this scenario is:

- trajectories

Metrics

- Performance of the similarity calculation. Timed used to calculate all metrics (a. to f.)

- Accuracy of trajectory metrics can be calculated externally (source code totally independent of datAcron implementation) and validated. This process will be done for randomly selected pair trajectories with different RMSE values.

	WP1	WP2	WP3	WP4
Performance	Serving the data needed to WP2 may influence in the performance	Performance of the prediction algorithms		User interface used for finding trajectories and plotting metrics
Accuracy		Accuracy of the prediction algorithms		Can help to result interpretation if visual components are developed that allow to analyze the experiment result. Subjective contribution.

Table 14: datAcron WP contribution to FP10

Validation Criteria

- Performance of the similarity calculation. Once trajectories are retrieved, similarity metrics calculation should be performed in a time lower than 2 seconds for each pair of trajectories.
- Accuracy in metrics evaluation should not deviate more than a 1% from the same metrics calculated with an external tool, independent of datAcron implementation.