

Final Project Results Report

Deliverable ID:	D4.5
Dissemination Level:	[PU]
Project Acronym:	[DART]
Grant:	[699299]
Call:	[H2020-SESAR-2015-1]
Topic:	[ER-2-2015]
Consortium Coordinator:	UPRC
Edition date:	[19 July 2018]
Edition:	[04.00.00]
Template Edition:	02.00.02

Founding Members



EUROPEAN UNION



EUROCONTROL





Authoring & Approval

Authors of the document

Name/Beneficiary	Position/Title	Date
George Vouros / UPRC	Project Coordinator / Professor	18.05.2018
Georg Fuchs/Fraunhofer	Fraunhofer Group Leader	18.05.2018
Enrique Casado/BR&T-E	BR&T-E representative, Dissemination and Exploitation Manager	18.05.2018
Jose Manuel Cordero/CRIDA	CRIDA Representative, Project Member	18.05.2018

Reviewers internal to the project

Name/Beneficiary	Position/Title	Date
Georg Fuchs/Fraunhofer	Fraunhofer Group Leader	18.05.2018
Enrique Casado/BR&T-E	BR&T-E representative, Dissemination and Exploitation Manager	18.05.2018
Jose Manuel Cordero/CRIDA	CRIDA Representative, Project Member	18.05.2018

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

Name/Beneficiary	Position/Title	Date
George Vouros/UPRC	Project Coordinator	18.05.2018
Enrique Casado/BR&T-E	Project Member	18.05.2018
George Fuchs/FRHF	Project Member	18.05.2018
Jose Manuel Cordero/CRIDA	Project Member	18.05.2018

Rejected By - Representatives of beneficiaries involved in the project

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

Document History

Edition	Date	Status	Author	Justification
00.06.00	25/04/2018	First Draft	George Vouros	
00.08.00	03/05/2018	First Draft with contributions from work package leaders	George Vouros	
01.00.00	18/05/2018	Final with contributions from partners	George Vouros, Jose Manuel Cordero, Pablo Costas, Georg Fuchs	



02.00.00	11/06/2018	Revised to address SJU comments	George Vouros, Jose Manuel Cordero, Pablo Costas, Georg Fuchs	
03.05.00	03/07/2018	Revised to address SJU comments	George Vouros, Jose Manuel Cordero, Pablo Costas, Georg Fuchs	
04.00.00	06/07/2018	Final Version after internal review	George Vouros, Jose Manuel Cordero, Pablo Costas, Georg Fuchs	
04.00.00	19/07/2018	Final Version	George Vouros, Jose Manuel Cordero, Pablo Costas, Georg Fuchs	Minor corrections after SJU comments and indication on publicly available deliverables

Copyright Statement

Copyright 2018 DART

This document has been produced within the scope of the DART project. The utilisation and release of this document is subject to the conditions of the Grant Agreement no.699299 within the H2020 Framework Programme, and the Consortium Agreement signed by partners.

© – 2018 – DART Consortium.

All rights reserved. Licensed to the SESAR Joint Undertaking under conditions.

Founding Members



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.



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 final DART results report describes work that has been done in the course of the project, developments and outcomes of project activities, answering major questions posed and specific research objectives targeted. Contributions to SESAR objectives and ATM Master Plan, as well as maturity of results are assessed. The document concludes with major findings, lessons learnt and plans for further activities.



Table of Contents

1	<i>Executive Summary</i>	8
2	<i>Project Overview</i>	11
2.1	Operational/Technical Context	11
2.2	Project Scope and Objectives	12
2.3	Work Performed	13
2.3.1	Overview of structure of work plan	14
2.3.2	DART Operational Context of Research and Scenarios for WP2 and WP3	15
2.3.3	Data management (WP1)	16
2.3.4	Single Trajectory Prediction (WP2)	22
2.3.5	Collaborative Trajectory Predictions (WP3)	29
2.4	Key Project Results	36
2.5	Technical Deliverables	40
3	<i>Links to SESAR Programme</i>	42
3.1	Contribution to the ATM Master Plan	42
3.2	Maturity Assessment	44
4	<i>Conclusion and Lessons Learned</i>	60
4.1	Conclusions	60
4.2	Technical Lessons Learned	61
4.3	Recommendations for future R&D activities (Next steps)	63
4.3.1	Roadmap of research goals	63
4.3.2	Proposed exploratory and industrial research activities per DART work package	65
5	<i>References</i>	68
5.1	Project Deliverables	68
5.2	Project Publications	69
5.3	Other	70
<i>Appendix A</i>		72
A.1	Glossary of terms	72
A.2	Acronyms and Terminology	72
<i>Appendix B</i>		73

List of Tables

Table 1: Project Deliverables	41
Table 2: Project Maturity	43
Table 3: ER Fund / AO Research Maturity Assessment for Data-driven single trajectory prediction	45

Founding Members



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 4: ER Fund / AO Research Maturity Assessment for Agent Based Methods for DCB optimization	52
Table 5: Glossary	72
Table 6: Acronyms and terminology	72

List of Figures

Figure 1- DART work package structure	14
Figure 2 – Synthetic Trajectory Reconstruction Infrastructure	17
Figure 3: DART Data Pool Evolution	17
Figure 4 : The Visual Analytics Loop followed by DART’s Visual Analytics toolset, adapted from [34].	19
Figure 5 : Tools for visual exploration of contextual data such as airspace sectorization schemes (top left) and various types of aircraft trajectory data such as actual tracks and flight plans (bottom row).	19
Figure 6 – 3D visualizations of two different configurations the LECM sector. Since sectorizations are 3D constructs comprised of several airblocks, a 2D map cannot always convey all relevant information. Here, two configurations vary in the inclusion of airblocks defining the lower airspace, which does not however change their 2D boundary.....	20
Figure 7 – Visual exploration of spatial data coverage: spatial gaps in aircraft tracks from a surveillance data set missing a number of stations.	20
Figure 8 – Visual exploration of flight dynamics for flights between Barcelona and Madrid: integrated display of 2D trajectory shapes with flight level changes. Orange color encodes increases of flight levels (climbing), while blue indicates descend. The line thickness encodes the absolute change between to subsequent aircraft positions.....	21
Figure 9 – 34 clusters representing the main approaches to the airports of London represented by coloring of the relevant parts of the trajectories, while a density surface summarizes irrelevant trajectory parts (here: cruise phases, holding patterns).....	21
Figure 10 – Training and test data for routes LEAL-LEBL and LEMD-LEIB.	23
Figure 11 – Qualitative assessment Actual vs Predicted.	23
Figure 12 – HMM vs Eurocontrol ETA prediction – Zoom in.	24
Figure 13 – 2D map comparison view of a pair of trajectories. Blue: ground truth data, orange: HMM prediction. Note the different sampling rates of positions along the track. View has been filtered to show only a selected pair of trajectories for illustration purposes.....	24
Figure 14 – Tabular view of the aggregate statistics of the matching point pair differences in space, time, and positional attributes. This view can be used for detail quantitative assessment of trajectory differences, but its primary use is to select user-defined instances of interesting values or value combinations (e.g., above a user-selected threshold) for further aggregation and visual analysis when applied to larger sets of trajectory comparisons.	25

Founding Members



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.

Figure 15 –Overview of the multi-stage approach, using HMM as predictive model.....	25
Figure 16 –Overview LEMD/LEBL dataset, April 2016, IFS tracks (red) and flight plans (blue).	26
Figure 17 – Half-width confidence intervals for HMM accuracy estimations per spatial dimension and in 3-D (y-axis provides the mean squared error estimations).	26
Figure 18 – Example NN-MLP distribution of prediction errors (signed MAPE)(m) for Lat for one waypoint.	27
Figure 19 – Performance of different prediction models for clustered and non-clustered datasets. ..	27
Figure 20 – MOME R^2 scores (y-axis) vs. normalized flight duration (x-axis).	28
Figure 21 – Trajectory Prediction accuracy by the Reinforcement Learning approach, by flight phase.	29
Figure 22 – Average delays for regulated flights from CFMU (dark blue line), Hierarchical (light blue line), IndLearners (red line), EdgeBased (grey line), AgentBased (yellow line). The x axis shows evaluation cases and the y axis the average delays.	32
Figure 23 – Number of regulated flights from Hierarchical, IndLearners, EdgeBased, AgentBased compared to CFMU regulated flights per case. The x axis shows evaluation cases and the y axis the number of flights. Evaluation cases correspond to those in Figure 22).	32
Figure 24 – x-axis: evaluation case, y-axis: difference of the average delay per flight from each of the methods, with the average delay per flight reported in CFMU.	33
Figure 25 –The maps represent the events of sector capacity excess based on CFMU-regulated (blue) flight data. The circle sizes are proportional to the excess amounts.	34
Figure 26 –The distribution of delays to flights by all methods and CFMU, for one experimental case. The x axis shows the delay imposed while the y axis corresponds to the number of flights.	34
Figure 27 –The space-time cubes show the spatio-temporal distribution of the delays. The time axis is oriented upwards. From top to bottom and from left to right: CFMU, AgentBased, EdgeBased, Hierarchical, and IndLearners.	35
Figure 28 –Roadmap of proposed short / medium / long term goals (from bottom to top) starting from DART contributions.	63

1 Executive Summary

DART (Data-driven Aircraft Trajectory prediction research) addresses the topic “**ER-02-2015 - Data Science in ATM**” exploring the applicability of **data science** and **complexity science** techniques to the ATM domain. DART delivers an understanding on the suitability of applying big data and agent –based modelling techniques for predicting aircraft trajectories based on data-driven models and accounting for ATM network complexity effects, considering multiple correlated trajectories.

DART has been motivated by the fact that the current Air Traffic Management (ATM) system worldwide has reached its limits in terms of predictability, efficiency and cost effectiveness. Nowadays, the ATM is based on an airspace management paradigm that leads to demand imbalances that cannot be dynamically adjusted. This entails higher air traffic controllers’ workload, which, as a final result, determines the maximum system capacity.

With the aim of overcoming the ATM system deficiencies, different initiatives, dominated by SESAR in Europe and Next Gen in the US, have promoted the transformation of the current environment towards a new trajectory-based ATM paradigm. This paradigm-shift changes the old-fashioned airspace management to the advanced concept of Trajectory Based Operations (TBO).

[Intentionally emphasized] The proposed transformation requires high-fidelity aircraft trajectory prediction capabilities, supporting the trajectory life-cycle at all stages efficiently. Making accurate predictions about individual trajectories in an early phase of operations, should allow predicting ATM network status, evolution of demand for resources, and thus hotspots, accounting for complex phenomena of the ATM system, when predictions are combined. Thus, predictions should allow assessing the overall system status and the impact of traffic on individual trajectories, w.r.t. operational constraints. Consequently, advances towards this direction should support effective decision making and optimization of resource exploitation during operations time. Motivated by these objectives, and given the existing and growing wealth of ATM data, as well as advances in machine learning, DART focuses on data-driven approaches to increasing predictability, and agent-based approaches for accounting for complex phenomena in the overall ATM system due to traffic and congestion of resources: These are major areas in which DART developments contribute.

In particular DART focused on the following questions:

- What are the supporting data required for accurate trajectory predictions?
- What is the potential of machine learning algorithms to support high-fidelity aircraft trajectory prediction?
- How the complex nature of the ATM system impacts trajectory predictions?
- How can this insight be used to optimize the ATM system?

Answering these questions, DART has gathered and managed in the DART Data Pool voluminous datasets from disparate sources, which have been identified as necessary to trajectory prediction and to accounting for complex phenomena that impact trajectory predictions. The DART Data Pool has been made accessible to DART beneficiaries through an advanced DART Transaction Pipeline. Feedback to the quality of data sources, erroneous/noisy data, as well as additional requirements identified, advanced further the quality of data gathered, tuning them to be unique – as far as we know – to advancing research in ATM. Additional datasets, computed from raw data, such as enriched surveillance data, Aircraft Intent and reconstructed trajectories were provided to the consortium,

deemed to be appropriate towards exploring various features that could improve predictability of trajectories. Interactive visualizations, developed and customized to the purposes of the project, played an important role towards assessing the quality of data provided.

Having devised these datasets, DART partners had a common background to train and test the data-driven methods for trajectory prediction, as well as the agent-based machine learning methods for assessing the delays of flights due to DCB problems. In doing so, results from the different methods developed are comparable, and reaching the end of this project we can draw useful conclusions to the potential of machine learning methods to improve predictability and prescribe regulations towards resolving DCB problems.

Therefore, towards exploring the potential of machine learning algorithms to support high-fidelity aircraft trajectory prediction, we devised and evaluated several machine learning methods that – after a thorough review of the state of art – assessed to be promising. These methods could be categorized in three major categories according to the data that they exploit: (a) Those exploiting raw data; (b) those exploiting enriched surveillance data (e.g. data enriched with weather variables or with variables from computed Aircraft Intent); and (c) Hybrid ones, combining data-driven with model-based approaches. Orthogonally, we may distinguish methods being “unconstrained” – which in order to predict a trajectory they do not exploit any information on the intended trajectory or flight plan; and constrained ones, exploiting information from the flight plan. Given this wide range of possibilities, this report resumes the useful conclusions drawn to the potential of methods evaluated. Overall, DART managed to reach predictions with unconstrained methods with errors in 3D predictions of approximately 7 km, and predictions with constrained methods exploiting enriched surveillance data with errors of approx. 3 km. These methods, compared to the model-based ones seem very promising, although in some cases they cannot offer prediction that are more accurate than model-based predictions. A major conclusion thus, is that hybrid methods – combining data-driven and model-based methods - can potentially offer more accurate predictions and DART stresses the necessity for further exploration in this direction.

Exploring the ways the complex nature of the ATM system impacts trajectory predictions, we focused on assessing flight delays towards resolving Demand Capacity Valance (DCB) problems at the planning stage of operations. This is one of the critical problems in ATM, concerning the predictability and planning of flights, the efficacy of the ATM network, towards supporting decision making among stakeholders for the efficiency of operations and utilization of airspace resources. To address this problem in DART, we provided a rigorous formulation of the problem as a Markov Decision Process (MDP). To solve DCB problems we explored multi-agent collaborative reinforcement machine learning approaches, where each agent represents a single trajectory (or aircraft executing a flight). In doing so, it is important to note that, we provide a shift-of-paradigm from regulating flights in a first-come-first-regulated basis, as it happen today: Each agent, independently from the others, decides its own delay w.r.t to operational constraints (e.g. capacities of sectors and sectorizations being active through the flight), own constraints (e.g. the maximum delay preferred), and of course, according to the cost of strategic delay imposed. A major issue here is “traffic”, i.e. the ways different trajectories do interact and thus, for the agent-based system, how agents mutually affect their decisions towards reaching a joint solution. It must be noted that by “solution” we mean that agents should reach a state where, due to the delays assessed and given their trajectories, there are no hotspots in the airspace. Evaluation results from the devised methods show that indeed, multi-agent collaborative reinforcement learning methods can manage to reach solutions in DCB problems very effectively, assessing less average delay for regulated flights, compared to CFMU reported regulations, and with computational efficiency, learning with few exploration rounds w.r.t. the size of the state space.



Specifically, methods manage to reduce the average delay for the regulated flights, while distributing delays to flights with fairness, keeping the number of regulated flights with more than 10' of strategic delay at a low number.

Trajectory predictions and resolution of demand-capacity imbalances are supported by interactive visualizations, providing overview of predictions and solutions at varying levels of space and time: Doing so, they provide the necessary means to justify outcomes and compare alternative solutions.

[Intentionally emphasized] The major outcome of this exploratory project comprises data-driven single trajectory prediction methods and agent-based methods accounting for the complexity of the ATM network due to traffic. While data-driven single trajectory prediction methods provide the means to make accurate predictions about individual trajectories and consequently about ATM network status and evolution of demand for resources, agent-based methods provide effective means to assess regulations to resolve hotspots, at the planning phase of operations. These capabilities, either individually, or combined, with the support of advanced visualization tools, provide the potential to advance stakeholders' collaborative decision making at the planning phase of operations.

Delving into these issues, this final project results report is structured as follows:

Section 2 provides a project overview, the specific research objectives targeted during the project, and detailed reference to work that has been completed per work package during the course of the project, elaborating on the problems addressed and methods explored, also providing indicative results and insights obtained through exploration. Key results per targeted objective are detailed, together with references to the deliverables produced.

Section 3 links project outcomes with SESAR objectives and ATM Master Plan objectives, providing a description of contributions per objective and references to specific OIs and ENs. DART answers to maturity assessment questions are provided also in the same section.

Section 4 provides conclusions per WP, answering the questions on which DART focused, also proposing – according to lessons learnt – plans for further exploratory and application-oriented activities.

Section 5 provides references to project published outcomes and deliverables, as well as references to bibliography considered.

In addition to main Sections there is an appendix (Appendix B) presenting the work, structure, presentations and results of the final DART workshop, held in conjunction to ICRAT 2018.



2 Project Overview

2.1 Operational/Technical Context

DART (Data-driven Aircraft Trajectory prediction research) addresses the topic “ER-02-2015 - Data Science in ATM” exploring the applicability of **data science** and **complexity science** techniques to the ATM domain. DART delivers an understanding on the suitability of applying big data and agent-based modelling techniques for predicting aircraft trajectories based on data-driven models and accounting for ATM network complexity effects, considering multiple correlated trajectories.

DART has been motivated by the fact that the current Air Traffic Management (ATM) system worldwide has reached its limits in terms of predictability, efficiency and cost effectiveness. Nowadays, the ATM is based on an airspace management paradigm that leads to demand imbalances that cannot be dynamically adjusted. This entails higher air traffic controllers’ workload, which, as a final result, determines the maximum system capacity.

With the aim of overcoming the ATM system deficiencies, different initiatives, dominated by SESAR in Europe and Next Gen in the US, have promoted the transformation of the current environment towards a new trajectory-based ATM paradigm. This paradigm-shift changes the old-fashioned airspace management to the advanced concept of Trajectory Based Operations (TBO). In the future ATM system, the trajectory becomes the cornerstone upon which all the ATM capabilities will rely on. The trajectory life-cycle describes the different stages from the trajectory planning, negotiation and agreement, to the trajectory execution, amendment and modification. The envisioned advanced Decision Support Tools (DSTs) required for enabling future ATM capabilities will exploit trajectory information to provide optimized services to all ATM stakeholders (airlines, Air Navigation Service Providers (ANSPs), Air Traffic Control (ATC), etc.).

[Intentionally emphasized] The proposed transformation requires high-fidelity aircraft trajectory prediction capabilities, supporting the trajectory life-cycle at all stages efficiently. Making accurate predictions about individual trajectories in an early phase of operations, should allow predicting ATM network status, evolution of demand for resources, and thus hotspots, accounting for complex phenomena of the ATM system, when predictions are combined. Thus, predictions should allow assessing the overall system status and the impact of traffic on individual trajectories, w.r.t. operational constraints. Consequently, advances towards this direction should support effective decision making and optimization of resource exploitation during operations time: These are the main points motivating work in DART. Increasing predictability and accounting for complex phenomena in the overall ATM system due to traffic and congestion of resources, are major areas in which DART developments contribute.

Current Trajectory Predictors (TPs) are based on deterministic formulations of the aircraft motion problem. Although there are sophisticated solutions that reach high levels of accuracy, all approaches are intrinsically simplifications to the actual aircraft behaviour, which delivers appropriate results for a reasonable computational cost. TP outputs are generated based on a-priori knowledge of the flight plan, the expected command and control strategies released by the pilot or the Flight Management System (FMS) - to ensure compliance with ATC restrictions - (known as Aircraft Intent [1]), a forecast of weather conditions to be faced throughout the trajectory, and the aircraft performance. This model- or physics-based approach is deterministic: It returns always the same trajectory prediction for a set of identical inputs.



Although the use of the concept of Aircraft Intent together with very precise aircraft performance models such as BADA [32] (Base of Aircraft Data) has helped to improve the prediction accuracy, the model-based approach requires a set of input data that typically are not precisely known (i.e. initial aircraft weight, pilot/FMS flight modes, ...). In addition, accuracy varies depending on the intended prediction horizon (look-ahead time).

[Intentionally emphasized] In summary, one can identify current TP as an area of improvement with consequent benefits supporting TBO.

Recent efforts in the field of aircraft trajectory prediction have explored the application of statistical analysis and machine learning techniques to capture non-deterministic influences that arise when an aircraft trajectory prediction is requested by a DST. For instance, linear regression models [33] [34] or neural networks [35] [36], have returned successful outcomes for improving the trajectory prediction accuracy on the vertical plane and for traffic flow forecasting. Generalized Linear Models [37] have been applied for the trajectory prediction in arrival management scenarios and multiple-linear regression [38] [39] for predicting estimated times of arrival (ETA).

The DART approach has been motivated by the outcome of the recent efforts, also reviewed during the project to assess their potential: These provide promising results in terms of accuracy prediction [41] [42] [43].

[Intentionally emphasized] DART has been challenged by the lack of a global vision on how to apply data-driven approaches to real ATM scenarios, and what the expected improvement will be.

Although most of existing efforts include as input dataset the available surveillance data, there is no consensus on the additional supporting data required for accurate trajectory predictions. Such additional supporting data may include filed or amended flight plans, airspace structure, ATC procedures, airline strategy, weather forecasts, etc. The disparity of the datasets used for validating different methods makes difficult the comparison among those studies, and therefore, prevents from extending the applicability of such techniques to more realistic and complex scenarios.

A strong limitation found in the current state-of-the-art research is that the proposed data-driven approaches are limited to individual trajectory predictions. The trajectories are predicted one-by-one based on the information related to them, ignoring the expected traffic at the prediction time lapse; hence disregarding contextual aspects on the individual predictions. Consequently, the network effect resulting from the interactions of multiple trajectories is not considered at all, which may lead to huge prediction inaccuracies. The complex nature of the ATM system impacts the trajectory predictions in many different manners. DART is motivated to capture this complexity and aims to devise agent-based methods towards assessing the impact of traffic to individual trajectories.

Given these motivating points, the major questions that DART aims to answer are as follows:

- [Q1] What are the supporting data required for accurate trajectory predictions?
- [Q2] What is the potential of machine learning algorithms to support high-fidelity aircraft trajectory prediction?
- [Q3] How the complex nature of the ATM system impacts trajectory predictions?
- [Q4] How can this insight be used to optimize the ATM system?

2.2 Project Scope and Objectives

The main research objective of DART (Data-driven AiRcraft Trajectory prediction research) is to explore the application of data-driven techniques to the aircraft trajectory prediction problem, and multi-agent learning approaches accounting for ATM network complexity effects. As part of this objective DART emphasizes the role modern visualization techniques can have in facilitating trajectory predictions.

To achieve this high-level main research objective, and to answer the major questions Q1-Q4 stated in the previous section, the following specific research objectives have been defined:

- [SRO1] Definition of requirements for the input datasets needed. The requirements will consider the trajectory prediction accuracy expected.
- [SRO2] Study of the application of big-data techniques to trajectory related data gathering, filtering, storing, prioritization, indexing or segmentation to support the generation of reliable and homogenous input datasets.
- [SRO3] Study of different data-driven learning techniques to describe how a reliable trajectory prediction model will leverage them.
- [SRO4] Formal description of the complexity network to support correlated multiple trajectory predictions¹.
- [SRO5] Study of the application of agent-based models to the prediction of multiple correlated trajectory predictions considering complexity network.
- [SRO6] Description of visualization techniques to enhance trajectory data management capabilities.
- [SRO7] Exploration of advanced visualization processes for data-driven model algorithms formulation, tuning and validation, in the context of 4D trajectories.

Specifically, answers to Q1 are provided by achieving SRO1, SRO2, SRO3, SRO6, SRO7, while Q2 is answered by achieving SRO3. Q3 is answered by the results produced towards SRO4 and SRO5. Finally, Q4 is answered by the direct and indirect contributions of DART developments to ATM Master Plan objectives and to contributions on specific Operational Improvements (OIs) and Enablers (ENs), as well as by the potential of further research activities inspired by DART.

Overall, DART aims to deliver understanding on the suitability of applying data-driven techniques for predicting multiple correlated aircraft trajectories. The outcomes of these techniques will be comparable thanks to the common dataset infrastructure developed in the project. DART specifies test case scenarios, giving insight about (a) the influence of the available data to predicting trajectories and (b) the efficacy of novel methods to assessing the correlated influence among trajectories: Different sets of inputs produce different improvements to the final prediction algorithms and correlations among trajectories impact the trajectory prediction output.

2.3 Work Performed

¹ Concerning “multiple/collaborative trajectory predictions”, the goal in DART is not to imitate the NM, but to build models that assess how interacting trajectories are affected due to traffic by considering predictions of individual trajectories.



2.3.1 Overview of structure of work plan.

DART project has been structured in four Work Packages following a layer-based approach. The core WP1 provided datasets to the remaining WPs. Upon this WP, the WP2 basically evaluates the suitability of machine learning techniques to elucidate which is the best alternative to enable robust and accurate data-driven trajectory prediction capabilities, and under which conditions. WP3 devises and evaluates mechanisms for detecting the influence of surrounding traffic on a trajectory prediction, enhancing the capabilities of prediction algorithms. This mechanism, based on collaborative reinforcement learning techniques accounts for network complexity effects and aims to assess the impact of traffic to individual trajectories by considering the complexity of the actual ATM environment. Finally, WP4 provides the project management activities required for the overall coordination of the defined WPs, including the dissemination and project impact activities.

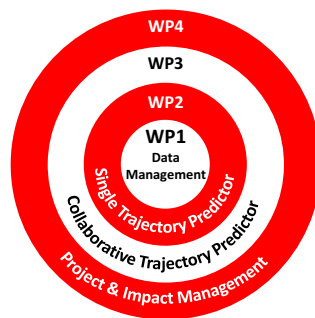


Figure 1- DART work package structure

Figure 1 shows a high-level organization of the WP's planned for the DART project execution. The work packages structure has inspired the project logo. The following list describes how this project structure has been designed for meeting the project specific research objectives:

- **WP1 Data Management** is devoted to create useful, reusable and high quality datasets of recorded aircraft trajectory information. These datasets, in conjunction to others (weather, airspace structure, flight plans etc.) are themselves a valuable deliverable of the project, since the difficulties existing in any data-driven efforts to start with a high quality and well documented dataset from operational source, are well known. WP1 generated datasets iteratively along the project and provided them to the other two research WPs (WP2 and WP3), starting with small and simple datasets. The size and complexity of the datasets increased accordingly to the needs and feedback from WP2 and WP3. Regarding trajectories, the datasets comprised raw data obtained from the ATM surveillance infrastructure and synthetic data generated offline thanks to the exploitation of the initial set of raw data. Visualization techniques eased inspection, evaluation and exploitation of each dataset by data scientists, facilitating the identification of data errors and/or omissions and exploration of data correlations.
- **WP2 Single Trajectory Prediction** is devoted to the analysis and evaluation of a wide range of data-driven techniques to the aircraft trajectory prediction problem. The most promising prediction algorithms obtained in WP2 have been evaluated in order to get knowledge about the pros- and cons- of their extensive usage.

- **WP3 Collaborative Trajectory Prediction** is devoted to unveiling the complexity to be considered due to the influence of the surrounding traffic. WP3 focused on the application of reinforcement learning algorithms in an agent-based framework in order to obtain assessment of the impact of traffic and operational constraints regarding the use of the airspace, on individual aircraft trajectories, accounting for ATM network effects.

WPs 2 and 3 explore interactive visual interfaces for supporting the sensitivity analysis of the proposed algorithms in each WP, providing also overview of outcomes in space and time.

- **WP4 Project and Impact Management** devoted to project management tasks plus two specific tasks related to the impact management (Dissemination & Exploitation and Communication activities), promoting WP2 and WP3 results across the ATM research community.

2.3.2 DART Operational Context of Research and Scenarios for WP2 and WP3.

Of the first tasks for DART (delivered in deliverable D3.1 [10]) was to clarify the operational context of its research activities, as well as the scope of the project.

Therefore, considering the objectives and requirements of both WP2 and WP3, two types of scenarios have been designed.

The first considers the airspace users' point of view and aims at predicting a single trajectory, isolated from other trajectories, according to WP2 purposes. This WP2 aims at analyzing and evaluating machine learning algorithms for trajectory prediction from an individual trajectory perspective (i.e. without considering traffic) and from the airspace users' point of view. Specifically, the objective of the scenario for WP2 is to demonstrate how DART predictive analytics capability can improve trajectory prediction at the planning phase of operations (i.e., during three days before operation). The objective is to compute the predicted trajectory that an aircraft will fly during an operation day. While increasing the accuracy of trajectory prediction has a value on its own, DART assumes that predictions happen in the context of a DCB process at planning phase that is developed in Spain. The ANSP role is represented by CRIDA (local level). Airspace users' role is represented by Boeing Research & Technology – Europe (BR&T-E). The separation between aircraft is guaranteed, thus there won't be conflicts in the proposed scenario. Resolutions adopted by ATCO won't be part of the scope of this operational scenario. The scenario also assumes that there won't be any regulation applied by the NM.

The second scenario assumes the Air Navigation Service Provider's (ANSP) point of view, in which multiple trajectory predictions are considered, taking into account traffic, according to WP3 purposes. Therefore, the WP3 operational scenario takes into account multiple trajectories, their interactions with respect to the DCB problems and aims to assess the regulations to be applied to these trajectories to resolve imbalances. The operational scenario in WP3 concerns the planning phase during the DCB process (three days before operation). The scenario develops in Spain, where the ANSP role is represented by CRIDA (local level). On the other hand, airspace users' role is represented by Boeing Research & Technology – Europe (BR&T-E). The separation between aircrafts is guaranteed; therefore, the scenario does not consider conflicts: Resolutions adopted by ATCO won't be part of the scope in the operational scenario WP3.



In this case, regulations of type C (i.e. delays) will be applied to the trajectories due to the imbalance between demand and capacity, so DART will have to obtain the final trajectories taking into account surrounding traffic.

2.3.3 Data management (WP1)

Data management focuses on the creation and management of datasets, that are the essential pieces or “bricks” for the success of the project. In particular, with respect to data sets definition, gathering, management and exploitation, DART followed an iterative approach: it provided WP2 and WP3 with suitable, reusable, and high-quality datasets that support accomplishing the objectives of these two WPs. Additionally, WP1 received feedback from WP2 and WP3 in order to improve the adequateness of the datasets to the objectives, subsequently and iteratively refining the datasets.

The necessity of work devoted on data management relies on the data required (or assumed to be required) by data-driven methods, and on the remarkable amount of data available for research, also in variety and complementarity. The hypothesis is that data from the different sources, when taken together can provide a comprehensive view of ATM trajectories and related operational data (such as weather, flight plan, regulations,...). This should be supported by an extensive amount of data, covering several years of operation for Spanish airspace.

The following table defines, for every data category considered, the primary data source identified in DART.

Data category	Primary Data source
Weather	<i>NOAA</i>
	<i>METAR</i>
	<i>SIGMET</i>
	<i>TAF</i>
Radar	IFS
Airspace	Spanish Operational data – Sector Configurations
Synthetic Trajectories	BR&T-E Trajectory Predictor using Spanish operational Flight Plans as input
Flight Plans	GIPV
Network Management	CFMU

In addition to these sources, synthetic trajectories are provided by the BR&T-E Trajectory Predictor, which processes operational information (flight plans). The intention of the synthetic reconstruction process is twofold:

- Populate the trajectory dataset with additional information with a higher sampling. This increases the available data providing an enriched aircraft state vector.

- Enhance the aircraft state vector by the addition of supplementary kinematic (e.g., airspeeds) and kinetic (e.g., fuel rate) variables.

The construction of synthetic trajectories relies on the definition of two models:

- **Aircraft performance model (APM)**, representing the performance of the aircraft (i.e., combination airframe and power plant).
- **Weather model (WM)**, representing a model of the weather conditions of the day of operation along the executed trajectory.

Figure 2 provides an overview of the synthetic trajectory reconstruction infrastructure, elaborated in deliverable D1.4 [4].

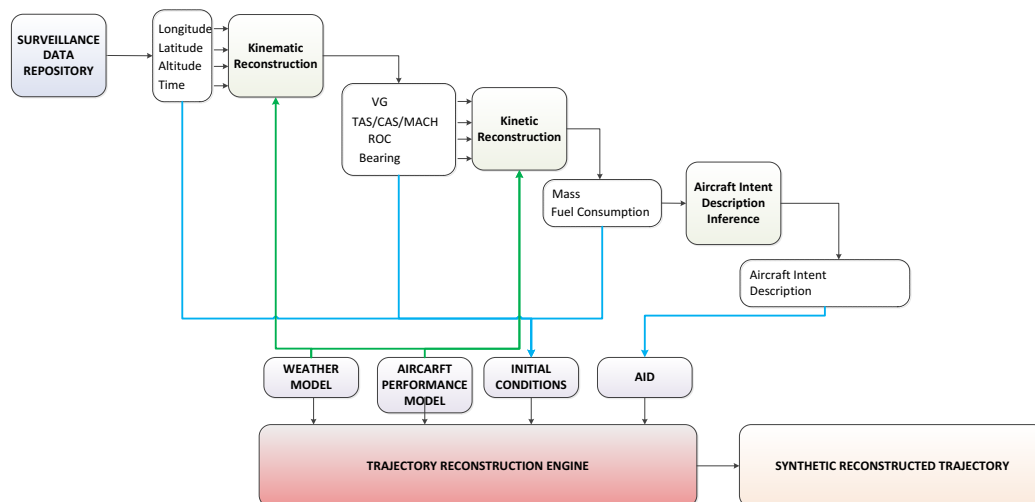


Figure 2 – Synthetic Trajectory Reconstruction Infrastructure

At the end of the project, the completed DART data pool has the following picture (showing also the timeline evolution of datasets during the project), detailed also in deliverable D1.3 [3].

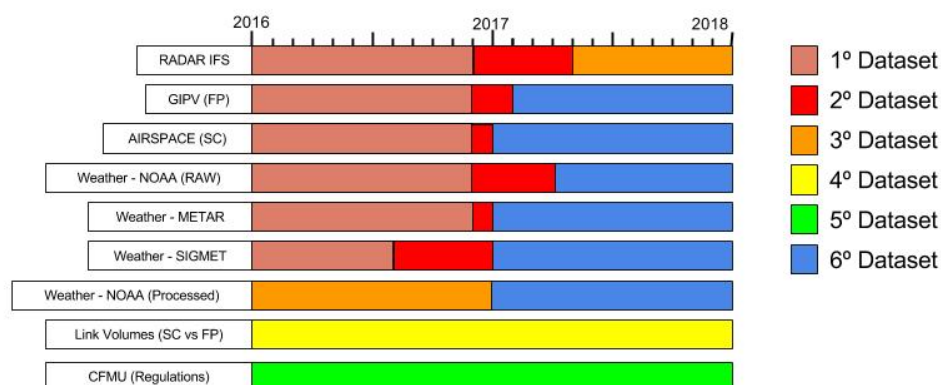


Figure 3: DART Data Pool Evolution

SIZE SUMMARY:

AirspaceStructures-SC

354.6 MB

Founding Members



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)

Additionally, WP1 defines and provides access to the data. All the necessary datasets are accessible to the rest of WPs, through a secure connection channel (due to the operational nature of the data, no physical copies of them can be done out of CRIDA premises, nor access to the raw operational files will be provided): Indeed, this WP provides the Data Transaction Pipeline, detailed in deliverable D1.2 [2]), as a suitable way for accessing this data.

In addition to these, synthetic reconstructed trajectories classified in different folders by months, and with related information regarding the Aircraft Intent (provided in separate files per trajectory due to Aircraft Intent nature and complexity) are provided to the DART consortium, covering 1-year traffic within the Spanish airspace.

Visual analysis methods and techniques developed in the context of DART for *Visual Exploration for Data Validation and Hypothesis Formulation*, as specified and detailed in deliverable D1.5 [5], included two key aspects, namely, the creation of a set of interactive visual interfaces that enable (1) identification of most common types of errors and omissions in data, and (2) exploration of cleaned data from multiple perspectives, namely focusing on locations in air space, time moments and intervals, and trajectories of single and multiple inter-related aircraft.

The results are part of a more comprehensive suite of visualization techniques, interactive filtering, and coupled analysis tools developed and implemented over the course of the DART project: The underlying conceptual model is the Visual Analytics Loop adapted from [34] (Figure 4). Due to the nature of the subject matter underpinned by a well-defined structure of interrelated principal data types, most constituent visualizations are employed both for the exploratory and data curation phases of analysis, as well as during confirmatory analysis during algorithm design and evaluation, which is the focus of DART work packages WP2 and WP3.

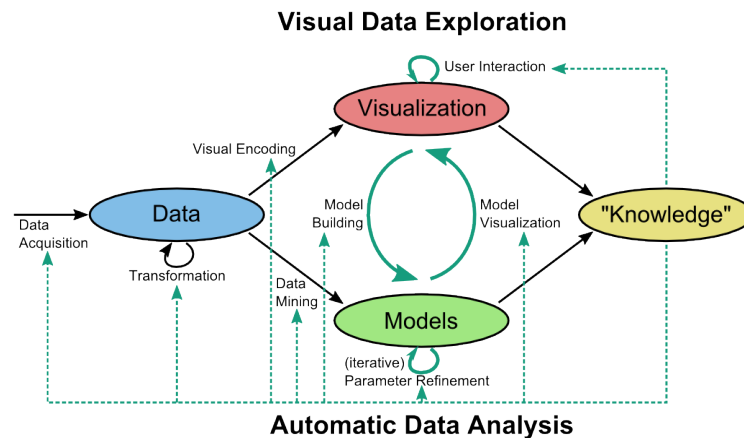


Figure 4 : The Visual Analytics Loop followed by DART's Visual Analytics toolset, adapted from [34].

DART expanded an existing set of standard visualization techniques, such as line plots and 2D map displays, by task-specific visualizations. These additions focus on the visual exploration of 3D aircraft trajectories (i.e., including the altitude and airspeed components): Examples from visualizations of sectorization schemes and aircraft trajectory data (e.g. actual tracks and flight plans) are shown in Figure 5.

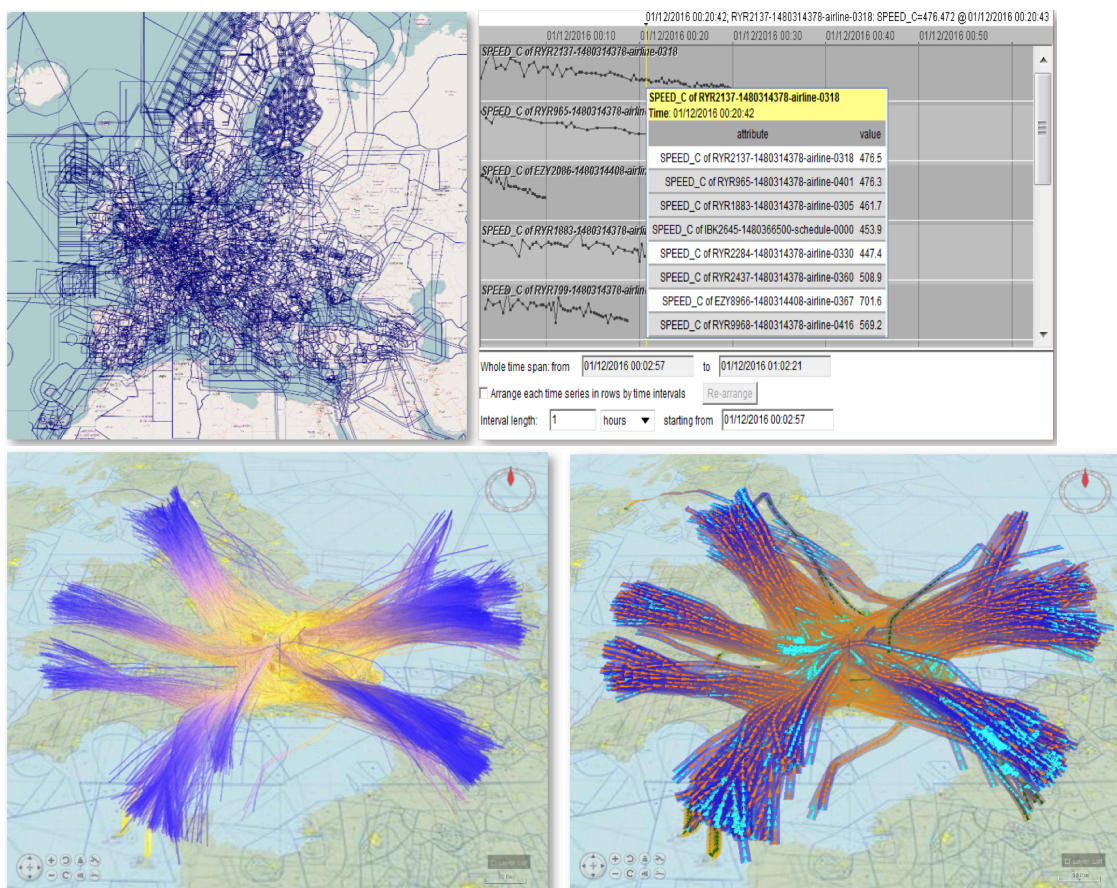


Figure 5 : Tools for visual exploration of contextual data such as airspace sectorization schemes (top left) and various types of aircraft trajectory data such as actual tracks and flight plans (bottom row).

Visual-interactive exploration of aviation data sources, include

Founding Members

- Visualizations of sectorizations in 2D maps and 3D volumetric representations (e.g. those provided in Figure 6), as well as updated temporal displays for the visualization and analysis of temporal dynamics/cyclicity of airspace configuration schemes. Interactive view manipulation and filtering functionality is provided in all cases, as well as an interface to the analytical functions, such as temporal clustering of airspace configurations.

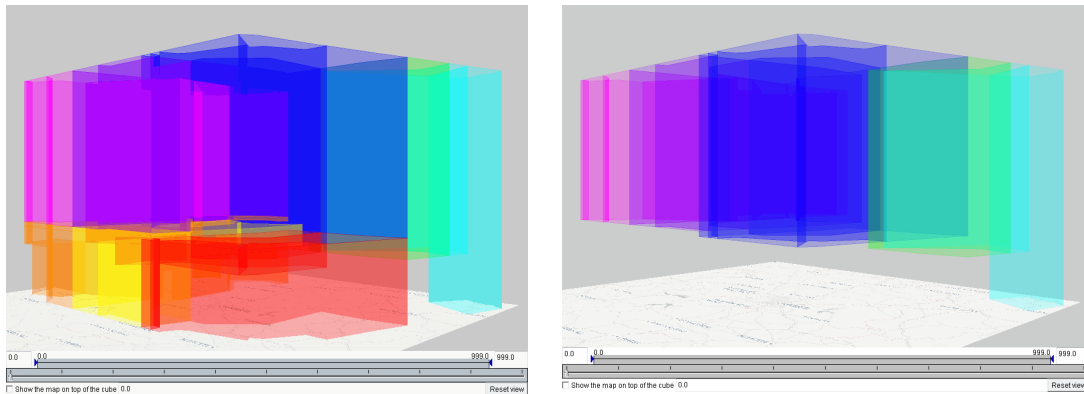


Figure 6 – 3D visualizations of two different configurations the LECM sector. Since sectorizations are 3D constructs comprised of several airblocks, a 2D map cannot always convey all relevant information. Here, two configurations vary in the inclusion of airblocks defining the lower airspace, which does not however change their 2D boundary.

- Visualizations that provide support for checking surveillance data coverage gaps in space and time. This is facilitated through 2D and 3D map displays that either display trajectory shapes directly (e.g. in Figure 7), or that visualize task-specific transformations derived from trajectories (Figure 8).

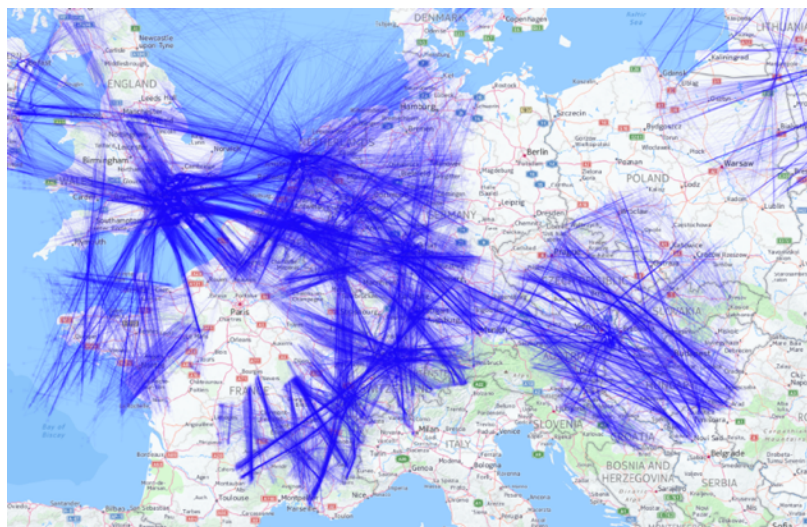


Figure 7 – Visual exploration of spatial data coverage: spatial gaps in aircraft tracks from a surveillance data set missing a number of stations.

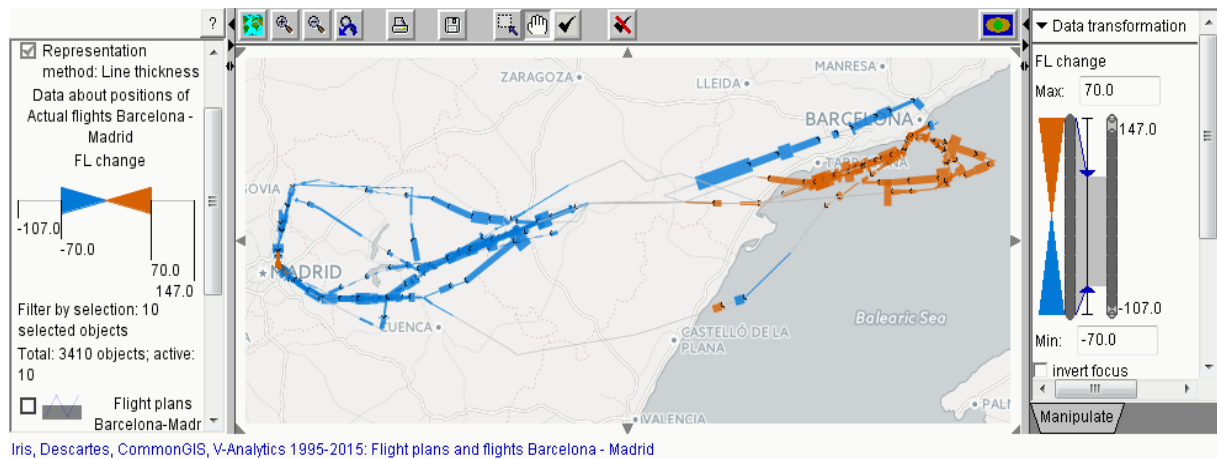


Figure 8 – Visual exploration of flight dynamics for flights between Barcelona and Madrid: integrated display of 2D trajectory shapes with flight level changes. Orange color encodes increases of flight levels (climbing), while blue indicates descend. The line thickness encodes the absolute change between to subsequent aircraft positions.

- Duplicated IDs, inspection of outliers or suspicious values in data (e.g. speed or altitude profiles) are further facilitated via appropriate visualizations combined with filtering.
- Techniques and visualization guidelines for supporting the use of relevance-aware clustering in visual exploration and analysis of movement data: This includes summarization of trajectory clusters and visual representation of the clusters in the context of the original data with visual distinction between relevant and non-relevant parts. At a high level of abstraction, the proposed approach supports an analytical workflow that consists of (1) selecting task-relevant parts of trajectories, (2) filter-aware clustering of the trajectories by the similarity of their relevant parts, and (3) exploiting the clustering results in subsequent analysis with the help of interactive visual displays.

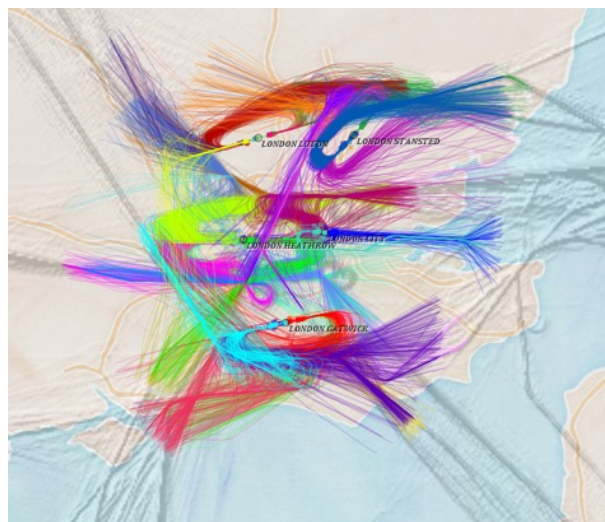


Figure 9 – 34 clusters representing the main approaches to the airports of London represented by coloring of the relevant parts of the trajectories, while a density surface summarizes irrelevant trajectory parts (here: cruise phases, holding patterns).

Implementations of the presented approaches have been integrated into a broader visual analytics framework comprising visualization techniques, interactive filtering, and coupled analysis tools. The framework's design follows a well-defined structure of interrelated principal data types and transformations between these types. These are presented in detail in deliverable D1.5 [5].

2.3.4 Single Trajectory Prediction (WP2)

As already pointed out above, the concept of data-driven trajectory prediction is a completely different approach than the model-based approaches (kinematic and kinetic approaches). It does not consider any representation of any realistic aircraft behaviour, only exploits trajectory information recorded from the ground-based surveillance infrastructure or by onboard systems (e.g., Flight Recorded Data or Quick Access Recorder Data) and other contextual data that may impact the final trajectory. This decoupled solution from the mathematical formulation of the aircraft motion should capture variations of the trajectory that cannot be derived directly from the filed flight plans (both during the strategic and tactical phases). These discrepancies usually come from Air Traffic Control (ATC) interventions to ensure optimum traffic management and safe operations. If these interventions respond to a pattern, big data analytics and machine learning algorithms might potentially identify them once the proper system features are considered.

Thus, the preparation of available trajectory data is crucial to train the algorithms in accordance to the expected data-driven Trajectory Prediction (TP) performance. Several solutions aim at predicting some aircraft state variables, for a representative scenario. The DART objective is to assess generic prediction methods to be applied in different possible scenarios envisioned in the future Trajectory Based Operations (TBO), in which the ATM paradigm will evolve from current tactical-airspace based to a strategic-trajectory based traffic management.

After a thorough review of state of the art methods for trajectory predictions – presented thoroughly in D2.1 [6] - and given the raw data available in DART, the most promising approaches to be explored in DART have been identified and specified in detail in D2.3 [8].

Aiming to explore various features that may influence the accuracy of the trajectory prediction process, different inputs to TP have been considered. For instance, information about aircraft performance is not necessary because the aircraft motion will be predicted by learning from historical recorded tracks, not by solving a mathematical formulation of the aircraft motion problem. In addition, data related to the day of operation, airline, airspace sectorization or average delay at departure airport could be of interest to obtain accurate data-driven predictions. To overcome this, enhanced datasets generated from the original raw data have been obtained: As a result, as already described in Section 2.3.3 “WP1 - data management”, synthetic reconstructed trajectories and information regarding the Aircraft Intent are provided per trajectory, adding more points and features to the representation of a trajectory, all being compatible to the realities of the flights.

The potential three candidates chosen to be assessed throughout the execution of DART for data-driven aircraft trajectory prediction- based on the current state-of-the-art, as well as the specifications of the problem - are as follows (all methods are presented in detail in D2.3[8] and evaluation results are discussed in D2.4 [9]).

The first is a Hidden Markov Models (HMM) approach, being one of the most popular and well-known approach for studying the state transitions of a system. This approach is already being tested for trajectory prediction from raw data and some very recent case studies show that its results on real data are very promising. This method exploits flight and meteorological information in order to define

the spatio-temporal cubes that are the basis of 3D prediction, as well as airport and flow features, that can be calculated using raw radar data, for timestamp prediction. Airspace information is commonly used in both prediction cases. Figure 10 provides a visual overview of training and testing data for two routes used in evaluation, detailed in deliverable D2.4 [9]. Our evaluation on the trajectory dataset verified that this prediction system achieved horizontal and vertical accuracy of 7.692nmi and 1589.452ft, respectively. This shows that in many cases data-driven trajectory prediction can perform better than model based trajectory prediction, but not in all cases. Figure 11 provides a qualitative assessment in order to understand how predictions look like vs the actual trajectory flown.



Figure 10 – Training and test data for routes LEAL-LEBL and LEMD-LEIB.

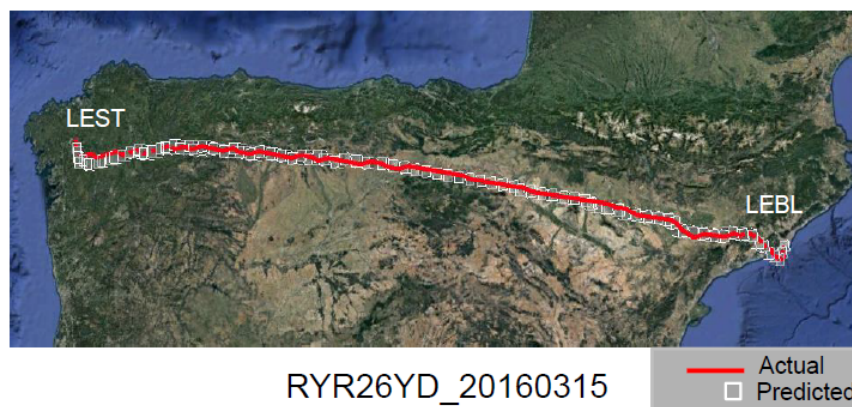


Figure 11 – Qualitative assessment Actual vs Predicted.

Comparing our final results with the Estimated time of Arrival (ETA) values, Eurocontrol, next figure illustrates RMSE values in minutes for each route between our predictions versus Eurocontrol's prediction. The following figure is a close look to the results, focusing at the box plots rather than to outliers, where the median values are visible. From the results, we make the following observations:

1. Our prediction yields better median scores on eight routes, while the Eurocontrol's ETA shows better median scores on two routes (LEBL-LEVX and LEBLLEZL).
2. The standard deviation values in Eurocontrol's ETAs are much larger, resulting in larger windows of predictability at arrival times.
3. Boxplots representing Eurocontrols's ETAs show extreme outliers.

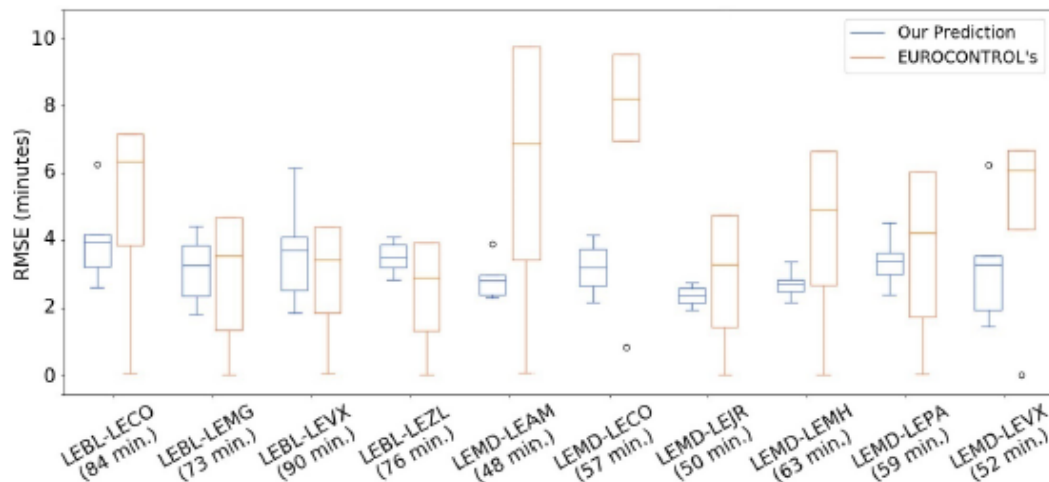


Figure 12 – HMM vs Eurocontrol ETA prediction – Zoom in.

Visual Analytics capabilities, as presented in detail in deliverable D2.2 [7], have been specifically designed to plot HMM trajectories and specially to perform visual comparison of the trajectories generated with this model with any other trajectory (typically actuals). The core idea of the solution chosen in DART to compare trajectories with different numbers and distributions of positions (trajectory points) is an algorithm that finds best matching points along pairs of trajectories, with the relaxation that not all points of one trajectory must have a matching point in the other trajectory. In addition, the algorithm automatically calculates core statistics for each pair of matching points – such as spatial distance, difference in time, and differences in values of positional attributes, such as altitude or flight level – to allow for fine-grained analysis and visualization of trajectory differences. In the next figures a single trajectory pair is selected for illustration purposes. The visualization techniques, especially the statistics view, are designed to handle larger sets of trajectories.

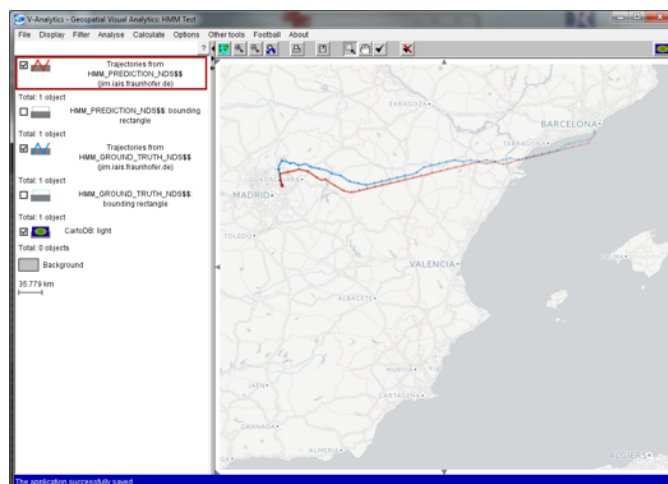


Figure 13 – 2D map comparison view of a pair of trajectories. Blue: ground truth data, orange: HMM prediction. Note the different sampling rates of positions along the track. View has been filtered to show only a selected pair of trajectories for illustration purposes.

Table view: Compared trajectories: actual and predicted

	N matched points	N unmatched own points	N unmatched points in buddy	% matched points	% matched own points	% matched points in buddy	Mean distance between matched points	Max distance between matched points	Mean distance for unmatched points	Max distance for unmatched points
LEBL-LEMD_20160412_actual	48	6	6	80.00	88.89	88.89	6233.16	17336.51	10862.28	17731.75
LEBL-LEMD_20160412_predicted	48	6	6	80.00	88.89	88.89	6233.16	17336.51	11183.21	17736.77

Sort by: No selection Ascending TableLens Attribute...

Figure 14 – Tabular view of the aggregate statistics of the matching point pair differences in space, time, and positional attributes. This view can be used for detail quantitative assessment of trajectory differences, but its primary use is to select user-defined instances of interesting values or value combinations (e.g., above a user-selected threshold) for further aggregation and visual analysis when applied to larger sets of trajectory comparisons.

Kernel-based distance metrics, as an alternative to the HMM approach, provide the basis of a “constrained” method that makes use of the flight plan itself as the feature vector, testing its similarity with other tracks. The input vector for that method can include several other properties associated with any trajectory segment but not necessarily derived from the spatio-temporal data of the trajectory (e.g. weather variables, aircraft type, etc). So, a multi-stage approach comprising clustering, constructing a predictive model (e.g. HMM) representing each cluster, flight plan assignment to a cluster, refinement of prediction within the assigned cluster, has been designed and implemented.

It must be noted that since this method requires the flight plan as input, it assumes that flight plans are available as earlier as possible during the planning phase of operations.

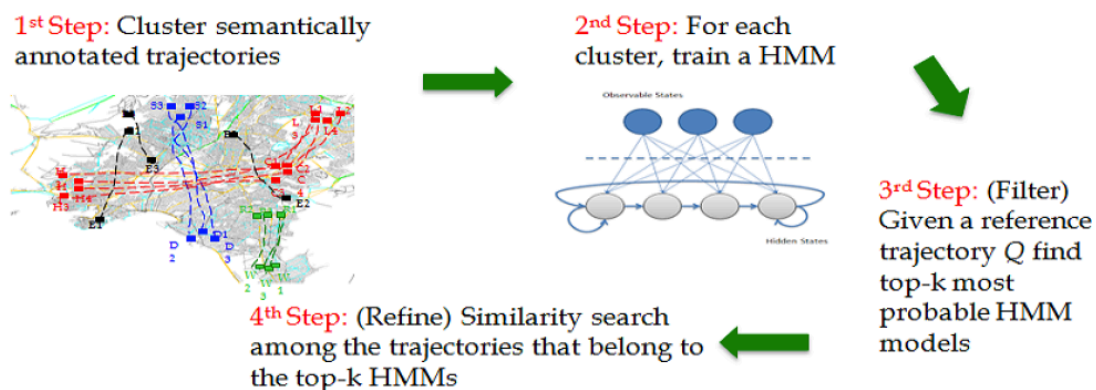


Figure 15 –Overview of the multi-stage approach, using HMM as predictive model.

As a result of clustering, Figure 16 provides an overview of clusters and cluster medoids detected for a dataset.

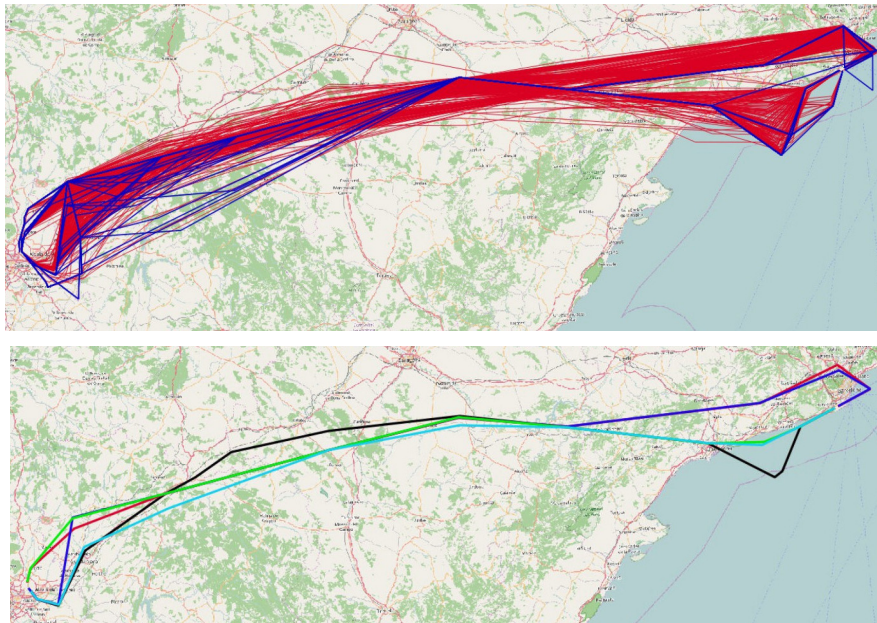


Figure 16 –Overview LEMD/LEBL dataset, April 2016, IFS tracks (red) and flight plans (blue).

Figure 17 presents the Half-Width Confidence Interval (HWCI) estimations for the best clustering result (stage-1) using HMM In the 2nd stage (constructing a predictive model for representing each cluster), consisting of 4 main clusters of 696 flights and one of 7 outliers (excluded).

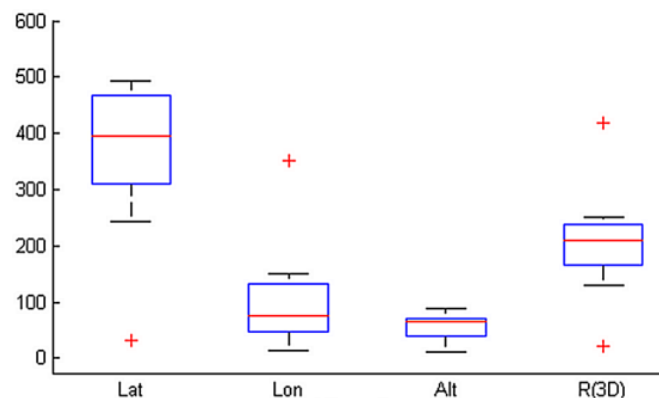


Figure 17 – Half-width confidence intervals for HMM accuracy estimations per spatial dimension and in 3-D (y-axis provides the mean squared error estimations).

For the purposes of constructing a predictive model per cluster, instead of HMM, the use of Neural Networks – Multi Layer Perceptron (NN-MLP) regressors, which are the best-performing model for stage-2 among those explored (Hidden Markov Models (HMM), Linear Regressors (LR), Classification and Regression Trees (CART), Neural Networks – Multi Layer Perceptron (NN-MLP)), provides the distribution of prediction errors (signed MAPE) for one waypoint shown in Figure 18.

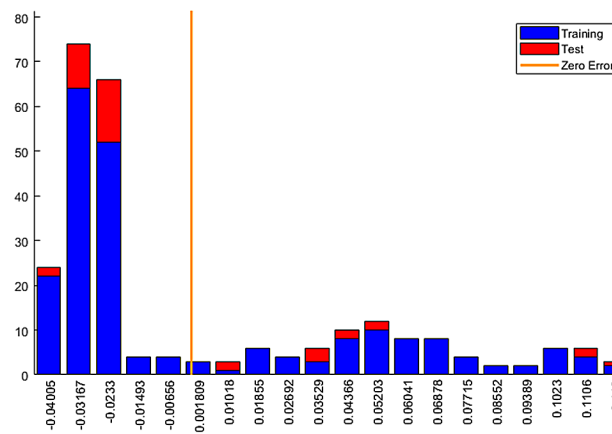


Figure 18 – Example NN-MLP distribution of prediction errors (signed MAPE)(m) for Lat for one waypoint.

Finally, Figure 19 presents the summary of the performance of all stage-2 predictor models for non-clustered and clustered dataset.

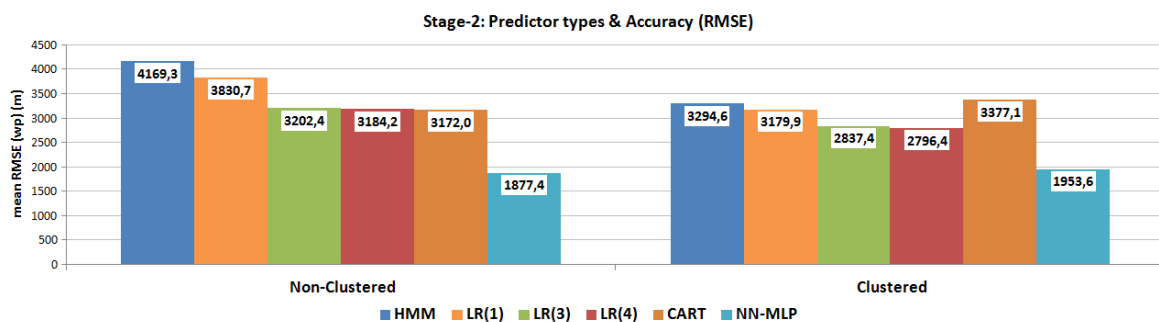


Figure 19 – Performance of different prediction models for clustered and non-clustered datasets.

As a result, and compared to the current “unconstrained” data-driven methods, e.g. the HMM approach, which produce cross-section 3-D prediction errors in the order of 7-13 km, using flight plans for constrained-based trajectory prediction produces per-waypoint 3-D prediction errors consistently in the order of 2-3 km, especially when NN-MLP regressors are used. However, it must be noted that these approaches are not directly comparable.

Statistical regression and machine learning models exploiting higher-than-linear parametric polynomials and other non-linear kernel functions in the error estimation are further explored towards single trajectory predictions. In contrast to the above-mentioned approach, the trained machine learning model itself acts as a functional mapping between input parameters (initial conditions, reference points) and corresponding output (flight path) according to some optimality criterion, which is normally the minimum prediction error based on a set of training data.

In addition to the above prediction methods and towards more informed and hybrid (machine learning in combination to model-based) methods, approaches exploiting Aircraft Intent inferred from raw surveillance data and reconstructed trajectories (using a subset of independent Aircraft Intent variables) have been explored: These include Reinforcement Learning methods, Random Forests, Hierarchical Agglomerative Clustering, Multi-Output Meta Estimators (MOME). These are presented in D2.3 [8] and evaluated in D2.4 [9].

As an indication of results obtained, Figure 20 shows the results obtained by one of these methods as function of fraction of the trajectory with respect to the duration (e.g., 0.4 = 40% of trajectory duration). It can be clearly concluded that this method returns accurate predictions in the central part of the trajectories (i.e., cruise phase), while diverges in both the initial and final parts (i.e., take-off and climb phases and descent and approach phases respectively). These results are similar to those obtained from Random Forest approach.

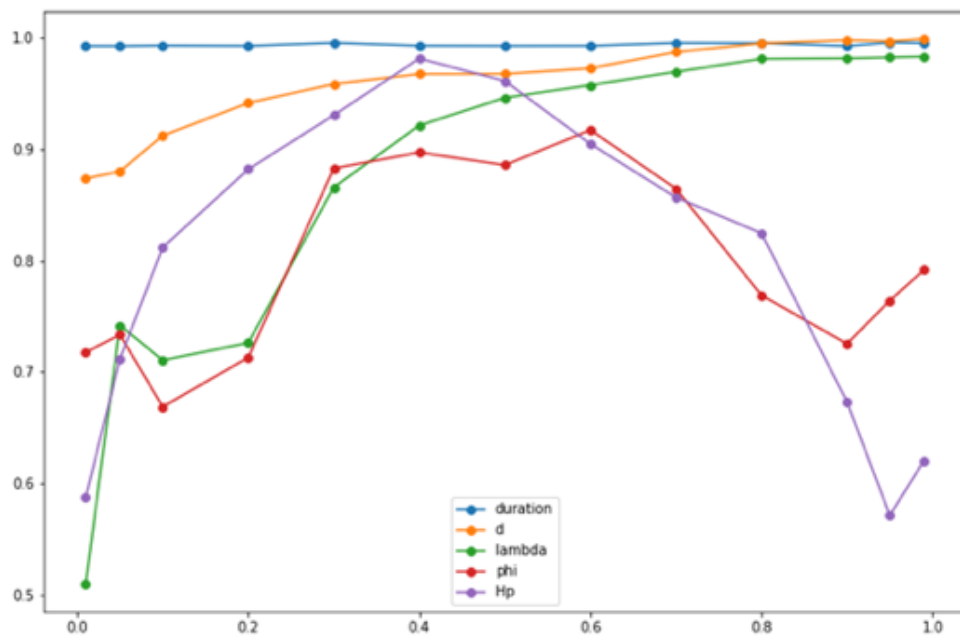


Figure 20 – MOME R^2 scores (y-axis) vs. normalized flight duration (x-axis): The diagram shows results on predicting Flight Duration (duration), Flown Distance (d), Longitude (lambda), Latitude (phi), Pressure Altitude (Hp).

Figure 21 presents results obtained from the Reinforcement Learning approach interacting with a model based trajectory predictor generating a trajectory iteratively. Prediction accuracy is not constant, it really depends on the flight phase being more accurate far from the airports:

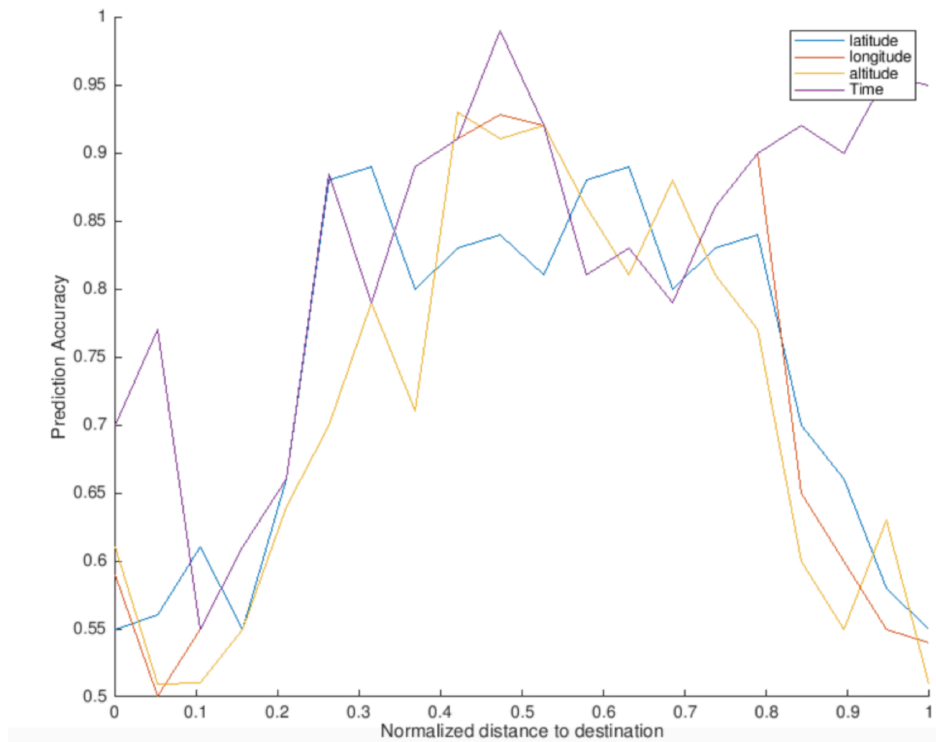


Figure 21 – Trajectory Prediction accuracy by the Reinforcement Learning approach, by flight phase.

As a conclusion, comparing the results and the complexity of the single trajectory prediction techniques, clustering and Hidden Markov algorithms based in raw data and enriched trajectory are the best performers and the most promising.

In addition, other processes such as clustering methods are more suitable to be used in combination with other machine learning algorithms.

As a general conclusion for all algorithms, these perform well in the cruise phase of flight where accuracy of prediction is very close to 100%, however algorithms have lower performance when predicting the trajectory for the phases of departure and arrival. It should be noted, that while we may think that it can be easy to predict the position of an aircraft at the moment of take-off, this is not straightforward, unless we use as feature in our algorithms the real-time flight position and we predict when the aircraft is about to take off.

Finally, validation results of data-driven approaches in comparison to model-based ones help to understand that the combination of model-based and data-driven approaches is the correct way to evolve current operational systems towards the implementation of Trajectory Based Operations (TBO).

2.3.5 Collaborative Trajectory Predictions (WP3)

One of the major goals of DART is applying agent-modelling approaches to account for complex phenomena due to the nature of the complex ATM sociotechnical system. In doing so, as already pointed out above, WP3 aims to devise agent-based methods that would allow ATM actors to assess regulations on trajectories, taking into account *interactions of multiple trajectories*.

Against this background, and according to the WP3 scenario specified, the main objective is to demonstrate how agent-oriented machine learning methods can help in refining single trajectory

Founding Members

predictions and/or flight plans at the planning phase, considering cases where demand of airspace use exceeds capacity, resulting to *hotspots*, accounting for the *Demand and Capacity Balance (DCB) problem*.

Therefore, the goal is to deliver an understanding on the suitability of applying agent-based models to regulate multiple correlated aircraft trajectories in the pre-tactical stage of operations with respect to operational constraints concerning the use of the airspace, focusing on the DCB problem in Air Traffic Management. In doing so, we aim to assess delays imposed to flights, as a result of influence by other flight trajectories (i.e. traffic), already in the pre-tactical phase of operations.

To achieve this goal we explored the suitability of agent-based collaborative reinforcement learning methods for supporting the Network Manager (NM) to regulate flights: It must be noted that such an agent-based approach provides a shift of paradigm regarding the methods applied today by the NM, contributing to a collaborative decision-making process where aircraft operators can assess own flights' delays collaboratively with others, considering how ATM network effects affect their scheduled flights. Therefore, as already noted in Section 2, the aim is not to imitate the decisions of the NM.

One of the key objectives towards applying agent-based collaborative reinforcement learning methods is to specify the problem and formulate it as a Markov Decision Process (MDP): This is specified in deliverable D3.2 [11].

As it is well known, the DCB problem considers two important types of objects in the ATM system: *trajectories* and *airspace sectors*. The *capacity of sectors* is of utmost importance. This quantity determines the maximum number of flights flying within a sector during a specific time interval.

Aircraft trajectories are series of spatio-temporal points of the generic form $(long_i, lat_i, alt_i, t_i)$, denoting the longitude, latitude and altitude, respectively, of the aircraft at a specific time point t_i . Such trajectories may be enriched and/or represented with other information such as specific events on interest. For instance, a *flight plan*, is an intended trajectory consisting of events of flights crossing air blocks and sectors and flying over specific waypoints. Thus, for the purposes of WP3 objectives we consider a trajectory to be a series of events for entering/existing sectors.

$$T = [(sector_1, entry_{t1}, exit_{t1}) \dots (sector_m, entry_{tm}, exit_{tm})],$$

where $sector_i, i=1, \dots, m$, is a sector in the airspace of interest, and *entry/exit* are time points.

Interacting trajectories are those that co-occur in space (i.e. in a sector) and time period (i.e. they co-exist during a time interval in space). Given a trajectory, all trajectories interacting with it, constitute the traffic experienced along that trajectory. Based on traffic we define

1. The coordination graph among agents; where each agent represents an aircraft executing a specific trajectory. Edges in this graph denote "interactions" among trajectories.
2. The demand $D_{s,p}$ per sector s and time period p , measured by the Hourly Entry Count measure.

Given the demand $D_{s,p}$ and the capacity C_s of a sector s , an imbalance of sectors' demand and capacity (and thus, hotspot) for any period p occur, when $D_{s,p} > C_s$.

Each agent i has to choose among a number of options $D_i = \{0, 1, 2, \dots, MaxDelay_i\}$, for its delay (in minutes), towards resolving DCB problems.

Based on those specifications, we formalized the problem as an MDP (specified in D3.2 [11]), where agents' states comprise the delays and the number of hotspots in which agents are involved, and actions (or strategies) comprise adding one further time point of ground delay or not, up to a maximum



delay (*MaxDelay*), specified either as a global problem parameter (i.e. for all flights), or as a parameter for each flight.

The reward for each agent is a linear function combining an assessment of the cost of participating in hotspots, and strategic delay costs.

As usually, a policy is a mapping from states to actions, while the objective is for any agent to find an optimal policy that maximises the expected discounted future reward while it executes its trajectory *jointly with others*.

Based on this problem formulation, we have devised four collaborative reinforcement learning algorithms, detailed in deliverable D3.2 [11]:

- An independent Q-learners algorithm, where each agent learns its own policy independently from others and treats other agents as part of the environment.
 - An edge based sparse cooperative Q-learning method, and
 - An agent based sparse cooperative Q-learning method,
- where – in both methods- agents learn jointly with their peers in the coordination graph; and finally,
- A hierarchical reinforcement learning method: In order to make reinforcement learning methods computationally efficient and to further explore advancing their effectiveness to produce solutions in large-scale complex problems, we apply abstraction or generalization operators. The idea behind state abstraction is that, instead of working in the (*original*) *state space*, the decision maker usually finds solutions in the *abstract state space* much faster by treating groups of states as a unit, ignoring irrelevant state information.

Based on this idea, we designed and implemented a hierarchical collaborative reinforcement learning framework – detailed in D3.2 [11]- which comprises two levels: The ground level and an abstracted level.

To evaluate the proposed methods, as shown in deliverable D3.3 [12], we have constructed evaluation cases of varying complexity/difficulty: Each case corresponds to a specific day of operations in 2016 above Spain, considering flight plans from all flights crossing Spain at any time during a 24 hours period. Cases have been chosen by inspecting problem parameters (number of flights, average number of interacting flights, maximum delay according to CFMU data, number of regulated flights according to CFMU data, average delay for regulated flights according to CFMU data, number of hotspots).

The results reported for the four methods show the following qualities:

- All methods manage to find solutions – i.e. they do manage to regulate flights crossing an operational space in a day so as to resolve all hotspots.
- All methods manage to find solutions effectively: They do converge to solutions quite fast, few rounds after exploration, in all of the cases.
- Methods manage to reduce the average delay for the regulated flights considerably, compared to the average delay for the regulated flights reported by CFMU. This is shown clearly in the figure below in 10 cases comparing average delays achieved by all methods, also compared to the average delay reported by CFMU for these cases.

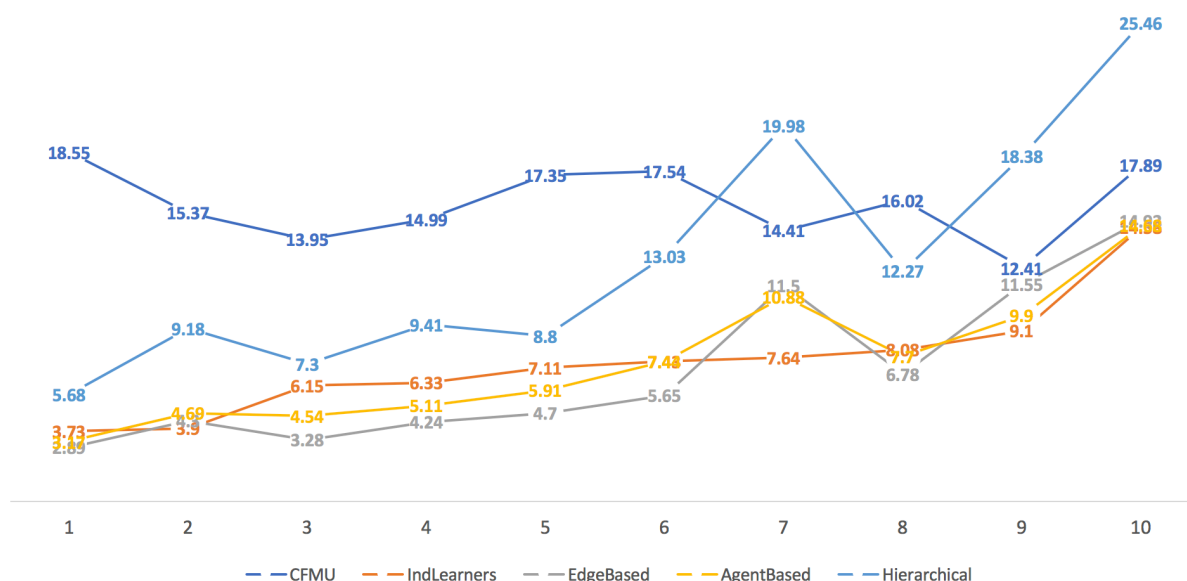


Figure 22 – Average delays for regulated flights from CFMU (dark blue line), Hierarchical (light blue line), IndLearners (red line), EdgeBased (grey line), AgentBased (yellow line). The x axis shows evaluation cases and the y axis the average delays.

In addition to that, the number of regulated flights per evaluation scenario is shown in the following diagram:

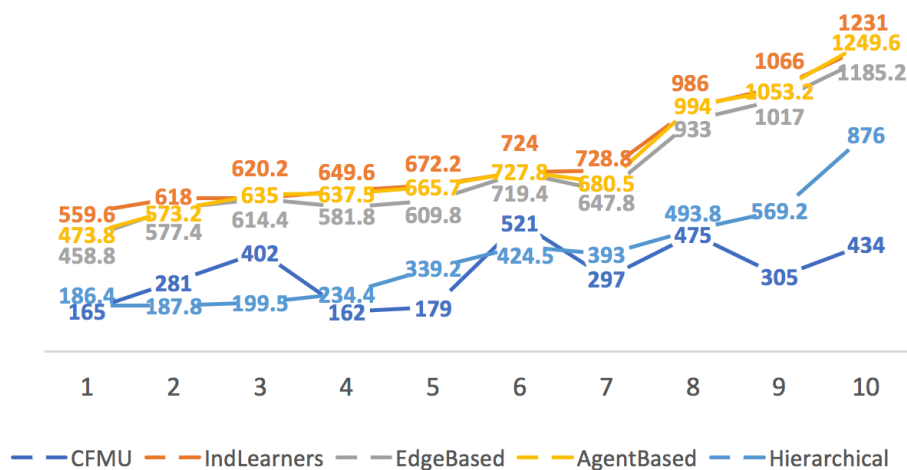


Figure 23 – Number of regulated flights from Hierarchical, IndLearners, EdgeBased, AgentBased compared to CFMU regulated flights per case. The x axis shows evaluation cases and the y axis the number of flights. Evaluation cases correspond to those in Figure 22).

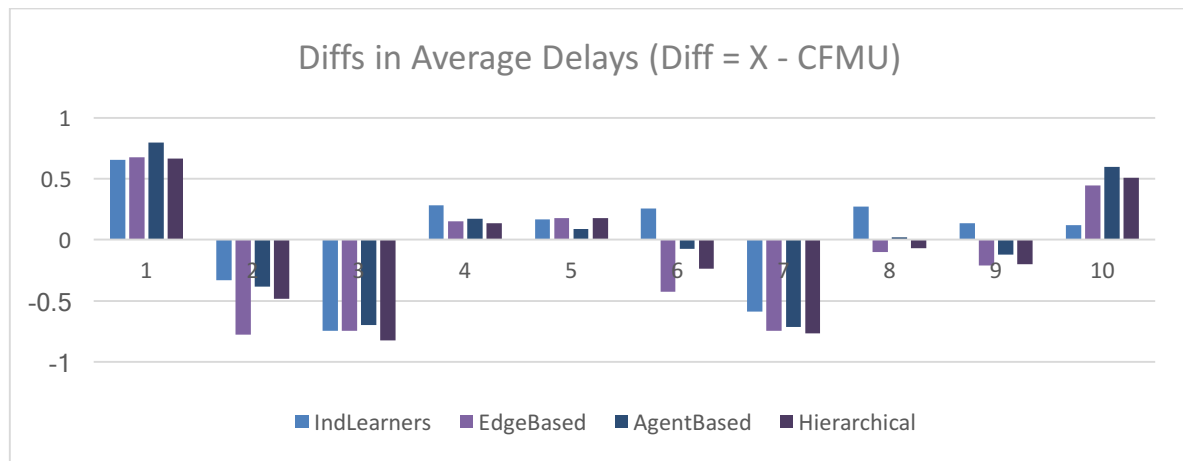


Figure 24 – x-axis: evaluation case, y-axis: difference of the average delay per flight from each of the methods, with the average delay per flight reported in CFMU.

It must be noted that delays imposed by CFMU do not resolve the hotspots for any of the evaluation cases.

- The Hierarchical method, has the potential to reduce significantly the regulated flights, and thus the average delay for all flights, compared to the other methods and of course CFMU. However, in its current implementation is not that effective, given that the average delay to the regulated flights is much higher than that reported by the other methods.
- Considering the average delay per flight, as Figure 24 shows, in most of the evaluation cases, the DART methods managed to reduce the average delay compared to CFMU average delay per flight, while only in one of the methods the DART methods have average delay per flight greater than 0.5 min than the CFMU average delay per flight: Among the methods, the Edge-based and the Hierarchical show to be more effective. It should be noted again that DART methods manage to resolve all hotspots in any evaluation case, in contrast to what happens to CFMU regulations.
- Finally, all methods manage to incorporate, as one of their inherent features, airlines preferences to the delays of some of the flights: This is a significant issue explored in D3.3 [12], showing that our methods can contribute to prescribing solutions to DCB problems, taking into account stakeholders' preferences and constraints on flights' delays.

Visualizations, as presented in D3.3, provide a comprehensive way to summarize results in space and time, while – and more importantly- they provide to a certain degree justifications of the “reasoning” behind decisions on regulations: For instance, while aggregated results on average delays and regulated flights show the potential of the methods, delving into details of their efficacy requires inspecting the spatio-temporal distribution of capacity excess events and their intensity (an example is shown in Figure 25), as well as the spatio-temporal distribution of delays (an example is shown in Figure 26). In addition, showing the spatiotemporal distribution of delays in space-time cubes (as shown in Figure 27), provides a comprehensive way to compare solutions and understand how delays are distributed in space and time.



Figure 25 –The maps represent the events of sector capacity excess based on CFMU-regulated (blue) flight data. The circle sizes are proportional to the excess amounts.

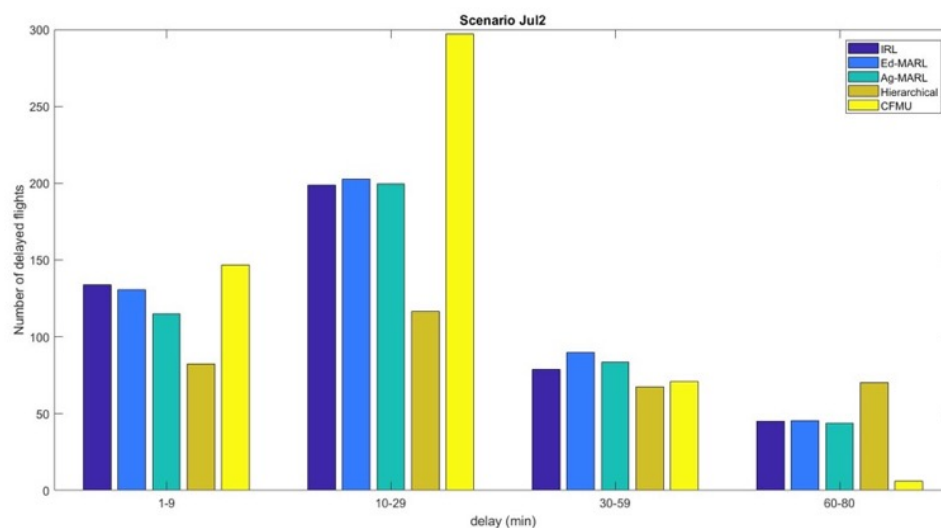


Figure 26 –The distribution of delays to flights by all methods and CFMU in one of the most difficult cases. The x axis shows the delay imposed while the y axis corresponds to the number of flights.

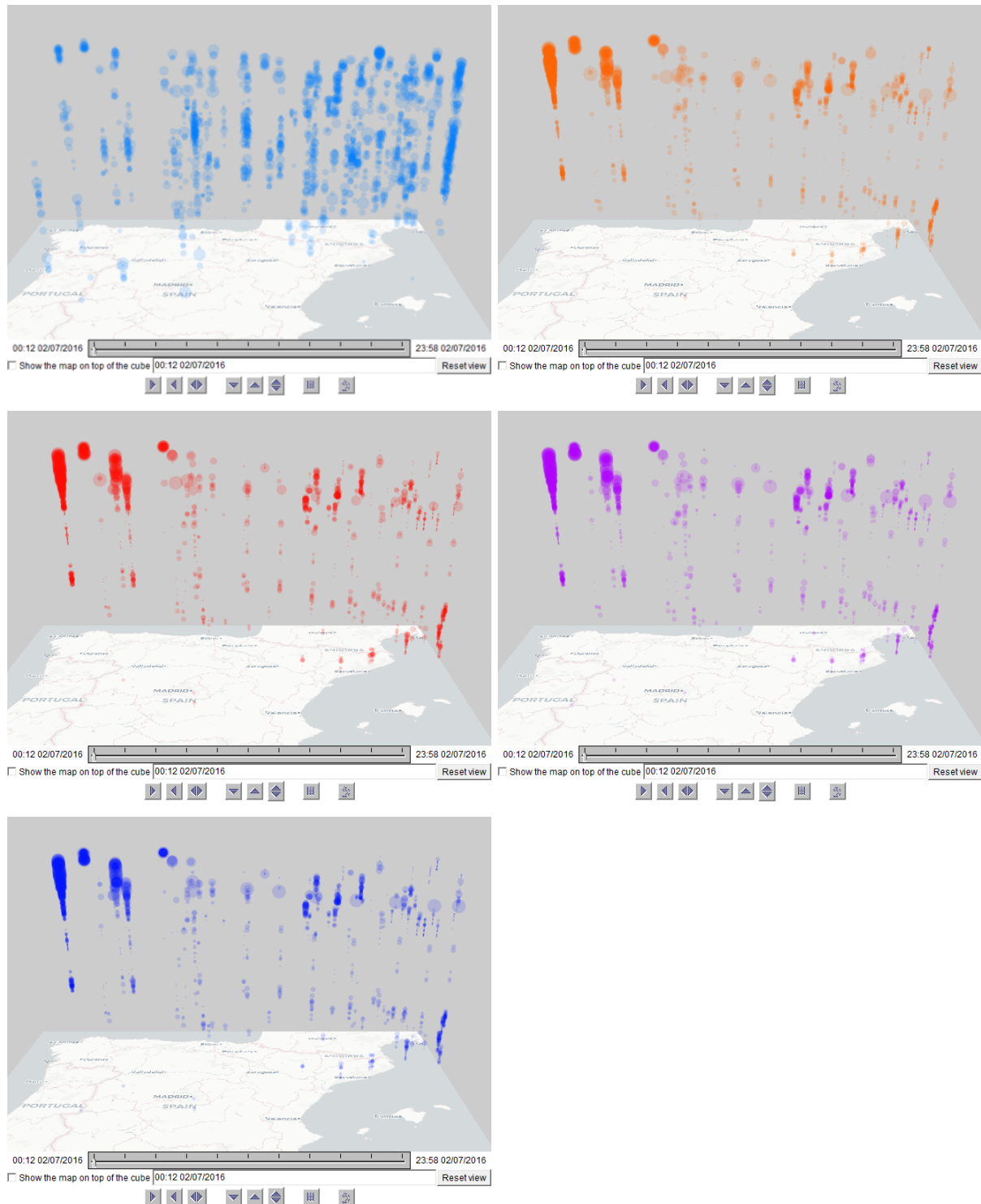


Figure 27 –The space-time cubes show the spatio-temporal distribution of the delays. The time axis is oriented upwards. From top to bottom and from left to right: CFMU, AgentBased, EdgeBased, Hierarchical, and IndLearners.



Nevertheless, such visualizations provide the means to compare and justify solutions, and guide human decision and/or preferences on solutions generated.

2.4 Key Project Results

Based on the description of the work performed and results achieved, as described in the previous section, below we provide description of key project results per project objective: Each bullet item describes the objective and developments are described in the paragraphs following it, citing the deliverables in which further information could be found.

- [SRO1] Definition of requirements for the input datasets needed. The requirements will consider the trajectory prediction accuracy expected.

DART has allowed to explore different ATM datasets (D1.3 [3], D1.4 [4]), addressing the linkage among them and also the reliability of information provided (D1.5 [5]). The iterative workflow between WPs has allowed that the different WPs set their needs with respect to data, that have been tackled by exploitation of the available data in different ways according to needs. This included the accuracy needs for trajectory prediction, that have been successfully addressed by the usage of high-quality operational datasets, including operational surveillance information.

- [SRO2] Study of the application of big-data techniques to trajectory related data gathering, filtering, storing, prioritization, indexing or segmentation to support the generation of reliable and homogenous input datasets.

Data gathering, filtering, storing, curation and management have been addressed through the development of Data Transaction Pipeline (D1.2, [2]), a system able to manage in an efficient way a large dataset such as the one in DART, with over 14 Tb of information accessible in real time, for which a Big-Data architecture has been deployed.

- [SRO3] Study of different data-driven learning techniques to describe how a reliable trajectory prediction model will leverage them.

DART reviewed and implemented the most promising data-driven algorithms to perform single data driven trajectory prediction (D2.1, [6]).

Multiple algorithms implemented exploring different directions for predicting trajectories: Either by ingesting raw data from datasets provided by WP1, or enriched surveillance datasets with additional variables, or derived datasets such as enriched trajectories and AIDL datasets (D2.4 [9]). In doing so, DART provides a comprehensive evaluation of state of the art machine learning algorithms for single trajectory prediction, also combined with clustering and model-based approaches, exploiting varying features concerning 4D trajectories.

Algorithms benchmarking (D2.4, [9]) activities included comparison between data-driven predictions versus flown trajectories, and data driven predictions versus Eurocontrol Network



Manager pre-flight prediction. Results show that data-driven methods can achieve high accuracy in predicting trajectories, also when they exploit information about flight plans.

Validation results help to understand that the combination of models and data-driven approaches is the correct way to evolve current operational systems towards the implementation of Trajectory Based Operations (TBO).

- [SRO4] Formal description of the complexity network to support correlated multiple trajectory predictions.

DART provided two formulations of the demand-capacity balance (DCB) problem using multi-agent Markov Decision Processes (MDP), modelling flights as agents whose decisions range in the space of their preferred/allowed delays:

- (a) A “flat” model, where states comprise agents’ own delays and hotspots in which they participate.
- (b) A hierarchical model, where in addition to the initial “ground” level, states are abstracted to an abstract level.

In both cases, agents’ reward depends on their participation in hotspots and the strategic costs of their delays.

These formulations have been reported in D3.2 [11].

- [SRO5] Study of the application of agent-based models to the prediction of multiple correlated trajectories considering complexity network.

DART designed and devised four multi-agent reinforcement learning methods towards assessing the impact of traffic to individual trajectories for resolving the DCB problems at the planning phase:

- (a) An independent learners approach, where each agent considers the others as part of its environment affecting its own decisions;
- (b) Two collaborative approaches that factorize the problem according to correlated/interacting trajectories: Interacting trajectories form neighbourhoods of connected agents in a coordination graph. Connected agents have to find a joint policy to resolve joint problems (i.e. hotspots in which they jointly participate):
 - a. An Edge-Based Sparse Q-learning method, which shares agents’ local rewards via edges to peers,
 - b. An Agent-Based Sparse Q-learning method, which considers the effect of each neighbourhood in the decision of connected agents.
- (c) A hierarchical method that solves the hierarchical multi-agent MDP problem: At each level of state abstraction one of the previous methods may be used, while results are mapped to levels of finer-states. These levels refine solutions, also by using one of the previous methods.



All these methods have been implemented (their technical description has been delivered in D3.2 [11]) and thoroughly evaluated (methodology and results are provided in D3.3 [12]). These methods, as already pointed out, provide a shift-of-paradigm from regulating flights in a first-come-first-regulated basis, as it happens today: Each agent, corresponding to a trajectory, and independently from the others, decides its own delay w.r.t to operational constraints (e.g. capacities of sectors and sectorizations being active through the flight), own constraints (e.g. the maximum delay preferred), and of course, according to the cost of strategic delay imposed.

Evaluation results for all methods show that they are capable to resolve hotspots, while keeping the average delay for regulated flights at very low levels, also compared to CFMU regulations, while regulating few flights with fairness. It is important to stress the potential of the methods to incorporate stakeholders' constraints on flight delays.

- [SRO6] Description of visualization techniques to enhance trajectory data management capabilities.

DART designed, implemented and employed a suite of interactive visualization techniques integrated into a common framework that facilitate the visual exploration and evaluation of trajectory data and associated context data, such as airspace sectorizations, in space and time (D1.5 [5]):

- (a) 2D and 3D map-based visualizations of individual aircraft trajectories (both actual flights and predicted) as first-class citizen, complex data type for the spatial aspect of trajectory data; complemented by
- (b) different forms of time displays (time series, time histograms, 2D histograms) for temporal aspects including linear and cyclic patterns of trajectory data; as well as
- (c) linked juxtaposition of the above as well as integrated space-time cube visualization to support exploration and analysis of spatio-temporal patterns and outliers in trajectory data.
- (d) Visualization techniques have been complemented by interactive filtering and clustering tools over spatial, temporal, and thematic attributes (e.g., speed, altitude, a/c type, ...) in the same framework to enable visual exploration and assessment of patterns, outliers, and spatio-temporal dynamics of user-specified subsets of trajectory data.

A major aspect of data management is data quality assessment and data curation (D1.5 [5]). The developed visual analysis toolset allows domain experts to efficiently leverage their expertise in addressing these tasks.

- [SRO7] Exploration of advanced visualization processes for data-driven model algorithms formulation, tuning and validation, in the context of 4D trajectories.

DART has designed and tested visual analytics workflows supporting model development and



evaluation on the basis of the aforementioned visualizations as their building blocks (D1.5 [5], D2.2 [7], D3.3 [12]):

- (a) Linked comparative visualizations on 2D maps, 3D maps, and Space-Time cubes, and
- (b) linked comparative visualizations in the temporal domain, including linear and cyclic time frames; complemented by
- (c) user-defined, interactive aggregation and comparison functions to enable analysis task-specific drilldowns in combination with the basic filtering and clustering methods provided.

The description of the proposed workflows, the results obtained, as well as a technical description of the comprising building blocks have been reported in D2.2 [7] & D3.3 [12].

Given the above results per research objective, the major questions of DART are answered as follows:

- [Q1] What are the supporting data required for accurate trajectory predictions?

A direct answer to this question is provided by the inventory of datasets gathered and curated in DART [SRO1], providing various features towards exploring the potential of data-driven methods to achieve high accuracy in trajectory prediction [SRO3].

These would not be achieved that effectively,

- (a) without the necessary means to providing data to the DART consortium with consistency and reliability [SRO2]; and
- (b) without the appropriate visual means to assess data quality and manage trajectories [SRO6], further supporting data-driven model algorithms formulation, tuning and validation, in the context of 4D trajectories [SRO7].

- [Q2] What is the potential of machine learning algorithms to support high-fidelity aircraft trajectory prediction?

This question is answered by the results reported by data-driven methods to trajectory prediction, developed in DART towards achieving [SRO3], as reported above.

Again, appropriate visualization methods provide the necessary means to support data-driven model algorithms formulation, tuning and validation, in the context of 4D trajectories, as reported above towards achieving objective [SRO7].

- [Q3] How the complex nature of the ATM system impacts trajectory predictions?

Answers to this question, considering the DART operational context, are provided by the results reported towards providing formal descriptions of the complexity ATM network [SRO4] and the potential of agent-based models (in particular, multi-agent reinforcement learning methods) to assessing delays considering complex network phenomena due to multiple interacting trajectories [SRO5], towards resolving DCB problems at the planning phase of operations.

- [Q4] How can this insight be used to optimize the ATM system?

This particular question is related to all research objectives, and concerns the impact of DART developments and how these are linked to the SESAR programme: Therefore, answers are provided in Section 3.



2.5 Technical Deliverables

Reference	Title	Delivery Date ²	Dissemination Level ³
D4.1	Project Management Plan	19.06.2016	Confidential
This deliverable documents the main project procedures. These procedures will be presented and agreed at the kick-off meeting chaired by the coordinator.			
D1.1	Data Management Plan	27.07.2016	Confidential
The DART Data Management Plan (DART DMP) details all data the project will collect and generate, how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved.			
D4.2	Dissemination Plan	26.09.2016	Public
The dissemination plan states the dissemination and exploitation goals of the project as well the means for achieving these goals. DART underpins all of its dissemination actions with the desired needs of the identified target audiences, project results and the impact of past dissemination efforts. The dissemination plan and any dissemination collateral (website, posters etc.) should be considered as 'living deliverables' and should be shaped to complement the exploitation of project results during the lifetime of DART. The most relevant dissemination means that will be adopted are technical reports, articles in journals and conference proceedings, and presentation of the project outcomes specially focused on the benefits of the use of the data-driven methods for the ATM community. The dissemination and exploitation plan will also include contact with stakeholders to support the prototype implementation and get a feedback from end-users.			
D4.3	Project website, wiki, social media channels	03.10.2016	Public
Social media platforms and project website comprise this deliverable. These will gather and disseminate all the information related to DART. They facilitate direct feedback and discussions among partners, as well with experts or other parties.			
D1.2	Data Transaction Pipeline Description	29.11.2016	Confidential
This deliverable provides the design and deployment of the data access mechanism (data transaction pipeline) for secure and reliable access to data, ensuring both integrity and accessibility for all the WPs.			
D2.1	Initial Set of Data-Driven Trajectory Prediction Algorithms	19.12.2016	Confidential
This deliverable reports on the initial set of data-driven single trajectory prediction algorithms to be developed and evaluated: This set comprises algorithms with the most promising data mining, machine learning and neural network techniques for predicting aircraft trajectories useful in the current and future ATM environments.			
D3.1	Collaborative Trajectory Prediction Scenarios and Requirements Specification	19.12.2016	Confidential
This deliverable reports on the scenario(s) for considering collaborative trajectory predictions, specifying geographical areas to be considered, actual states/stages to be considered when performing trajectory prediction, and data to be considered. In conjunction to scenarios specification; the specification of requirements for the algorithms to be developed are reported and algorithms' evaluation criteria will be specified.			

² Delivery data of latest edition

³ While this column specifies the dissemination level according to the Grant Agreement, the DART Consortium took the decision of publishing confidential deliverables via its website at <http://dart-research.eu/2018/07/10/dart-final-deliverables/>, so that all its results will be available to the research community.

Publicly available deliverables do not include deliverables explaining details on the data-driven (i.e. D2.1 and D2.3) and agent-based methods (D3.2) developed, although these methods are succinctly presented together with evaluation results in documents shared (i.e. D2.4 and D3.3).

Founding Members



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.



D2.2	Visual Interface for Algorithms Analysis	19.06.2017	Confidential
This deliverable provides interactive visual interfaces for algorithm tuning, supporting sensitivity analysis in respect to algorithm parameter, and providing an overview of modeling results in space and time.			
D1.3	DART Data Pool	19.02.2018	Confidential
The Dart Data Pool comprises all the suitable, reusable, and high-quality datasets that will support accomplishing the objectives of WPs 2 and 3. Internal updates of DART Data Pool from feedback of WP2/WP3 have been released periodically from T0+3 to T0+16; this final deliverable covers the whole process, description and results.			
D2.3	Enhanced Set of Data-Driven Trajectory Prediction Algorithms	19.02.2018	Confidential
This deliverable reports on the enhanced and final set of data-driven trajectory prediction learning algorithms. These Have been selected from the initial set of algorithms and are thoroughly evaluated. The selection is based on criteria such computational efficiency or algorithm performance according to requirements.			
D3.2	Collaborative Trajectory Prediction Algorithms	19.02.2018	Confidential
This deliverable reports on collaborative RL algorithms accounting for multiple trajectory predictions, allowing agents to learn offline and in batch-mode, effectively and in a totally distributed way, while being able to support the on-line joint assessment of trajectory delays towards resolving DCB problems, taking into account other agents' decisions and overall contextual information.			
D1.4	Synthetic Data Package	03.03.2018	Confidential
Synthetic datasets delivered are new datasets generated and linked with the original ones containing, among others, aircraft intent, flight intent and reconstructed and predicted aircraft state information.			
D3.3	Evaluation and Validation of the Collaborative Trajectory Prediction Algorithm	19.04.2018	Confidential
This deliverable reports on evaluation and validation of algorithms using actual and synthetic data gathered and/or generated in WP1. Different evaluation/validation criteria are considered concerning assessment of delays for resolving DCB problems.			
D1.5	Visualization exploration report	19.05.2018	Confidential
This deliverable provides a set of interactive visual interfaces that enable (1) identification of most common types of errors and omissions in data, and (2) exploration of cleaned data from multiple perspectives, namely focusing on locations in air space, time moments and intervals, and trajectories of single and multiple inter-related aircrafts.			
D2.4	Evaluation and Validation of Algorithms for Single Trajectory Prediction	19.07.2018	Confidential
This deliverable reports on evaluation and validation of algorithms using actual and synthetic data gathered and/or generated in WP1. Different evaluation/validation criteria will be considered: (a) Precision of the predictions made, (b) goodness of the predictions in comparison to the actual trajectories.			
D4.4	Dissemination Report	19.05.2018	Public
This report specifies the DART dissemination achievements, supporting communication with targeted audiences, communities and businesses by developing information material, articles and reports on project achievements, using different kinds of communication strategies and mediums.			
D4.5	Project Results Final Report	19.07.2018	Public
This is the final report for DART.			

Table 1: Project Deliverables



3 Links to SESAR Programme

3.1 Contribution to the ATM Master Plan

DART links to the SESAR vision

“The [SESAR] vision builds on the notion of trajectory- based operations and relies on the provision of air navigation services (ANS) in support of the execution of the business or mission trajectory - meaning that aircraft can fly their preferred trajectories without being constrained by airspace configurations. This vision is enabled by a progressive increase in the level of automation support, the implementation of virtualisation technologies and the use of standardised and interoperable systems.”

Towards this vision DART explored the potential and benefits of data-driven and agent-based methods contributing mainly to the following features, as described in ATM MasterPlan, Chapter 2:

- Offering Improvements across ATM:

- **Improved operations productivity** via increased trajectory predictability and contributions to improved (collaborative) planning tools accounting for complex phenomena due to traffic

Impacted KPAs/Focus Areas (aligned with SESAR Performance Framework⁴):

- **Predictability:** Enabling better forecast at an earlier pre-tactical stage will allow having a more efficient operation plan, reducing buffers (indirectly improving Capacity). In particular, it refers to KPI PRD1 within SESAR Performance Framework.
- **Punctuality: Minimizing the impact of delays due to ATM factors** via assessment of delays at the planning phase and contributions to collaborative decision making among stakeholders. This benefit refers to KPI PUN1 within SESAR Performance Framework.
- **Environment:** Better trajectory prediction is expected to reduce the number of flight inefficiencies due to tactical ATC actions, as well as to allow better planning of operations for AUs. This refers to KPIs FEEF1 (Actual Average fuel burn per flight) and FEEF2 (CO2 emissions), with also a potential positive impact in FEEF3 (reduction in average flight duration).

- Supporting change in ATM:

- **Addressing known critical network performance deficiencies** via the provision of tools that can support decision making and collaboration among NM, ANSPs and airlines for solving DCB problems at the planning stage.

Impacted KPAs/Focus Areas (aligned with SESAR Performance Framework):

- **Punctuality:** Reducing departure delay, as reflected in KPI PUN1 within SESAR Performance Framework.

⁴ PJ19.04 D4.1 Performance Framework, 2017.



- **Capacity:** By increasing network traffic throughput mainly en-route, as reflected in KPI CAP2 in SESAR Performance Framework (with also potential positive impact in TMA throughput, CAP1).

Code	Name	Project contribution	Maturity at project start	Maturity at project end
New EN	Data-driven single trajectory prediction	DART identified, implemented and tested state of the art machine learning algorithms, identifying the more suitable to implement data driven trajectory prediction methods, able to outperform, at least in some scenarios, model based trajectory prediction engines. Algorithms evaluation results show that data-driven methods can achieve high accuracy, while they help us understand that combination of model-based and data-driven approaches is the correct way to evolve current operational systems towards the implementation of Trajectory Based Operations (TBO).	TRL-0	TRL-1 (problem formulated, methods developed, and their potential has been assessed)
New EN	Agent Based Methods for DCB optimization	DART contributes to providing individual traffic predictions, which in combination can be used for assessing the impact of traffic on the network and on individual flights' delays: These developments contribute directly to enhanced Traffic/Demand Forecast services from long term planning to execution phases in 4D Trajectory Management context. Agent-based methods developed in DART provide the basis for tools for assisting NM optimized decision-making, by identifying, arbitrating and resolving multiple imbalances and hotspots, towards optimizing delays and utilization of airspace resources. These developments are directly connected to OI DCB-0211⁵.	TRL-0	TRL-1 (problem formulated, methods developed, and their potential has been assessed)

Table 2: Project Maturity

⁵

The aim of this OI is to benefit from the shared iterative SB/MT development and provides enhanced and continuous Traffic/Demand Forecast services from long term planning to execution phases in 4D Trajectory Management context. It includes the development of 4D trajectory based forecast methodology (build on 2D, 3D and 4D trajectory data provided by the AUs), operational workflows, and the appropriate infrastructure which provide European airspace planners and airspace users with a common and consistent picture of European air traffic demand that will meet their planning and monitoring needs. It includes as well improved traffic predictability thanks to the elaboration of the probabilistic demand, the consideration of planned DCB measures and the consideration and processing of airspace users shared flight information (including a set of confidence indexes and their interpretation) in support of the traffic demand enrichment.

Founding Members



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.



3.2 Maturity Assessment

[Intentionally left blank]



Table 3: ER Fund / AO Research Maturity Assessment for Data-driven single trajectory prediction

EN: Data-driven single trajectory prediction

ID	Criteria	Satisfaction	Rationale - Link to deliverables - Comments
TRL-1.1	Has the ATM problem/challenge/need(s) that innovation would contribute to solve been identified? Where does the problem lie?	Achieved	Yes. For WP2 (Single Trajectory Prediction) the main problem identified is the improvement on the accuracy of trajectory prediction at the pre-tactical phase. The main issues related to the improvement in the accuracy of trajectory prediction is the uncertainty related to the factors affecting a flight trajectory, and generalizing over different pairs of departure/destination airports. Advances towards addressing these problems should result in improvements in planning, flight management and traffic control. These issues are identified and addressed In D2.3.
TRL-1.2	Has the ATM problem/challenge/need(s) been quantified?	Achieved	Yes. For WP2 (Single Trajectory Prediction), individual trajectory predictions need to be as accurate as possible, given that current trajectory predictions



			are not accurate enough. This problem has been quantified, given that DART quantified the prediction accuracy of model-based methods in evaluation cases. It must be noted that while DART addresses the problem at the planning phase (i.e. days before operation) it is connected to the availability of accurate forecasted information and to the availability of historical data of high quality, volume and spatio-temporal coverage. The quantification of the problems and challenges addressed are described in D2.4.
TRL-1.3	Are potential weaknesses and constraints identified related to the exploratory topic/solution under research? - The problem/challenge/need under research may be bound by certain constraints, such as time, geographical location, environment, cost of solutions or others.	Partial – Non Blocking	For WP2 (Single Trajectory Prediction), towards data-driven trajectory prediction methods, the most crucial factors presenting constraints are (a) the availability of large volumes of high-quality data to train learning algorithms, (b) the heterogeneity of data to be linked so as to enrich trajectories with information, and consequently, train the algorithms incorporating a rich set of features needed to advance predictions. In addition to those, (c) trajectory prediction methods need to be trained for different pairs of airports w.r.t. all different prediction factors/conditions.



			For these purposes, gathering of data, assessing its quality and linking it requires effort and time (including processing time).
TRL-1.4	Has the concept/technology under research defined, described, analysed and reported?	Achieved	Yes. For WP2 (Single Trajectory Prediction), D2.1 provides a review of state of the art methods for individual trajectory prediction methods, while D2.3 reports on the specific methods that DART explores regarding this problem.
TRL-1.5	Do fundamental research results show contribution to the Programme strategic objectives e.g. performance ambitions identified at the ATM MP Level?	Partial – Non Blocking	DART, as far as developments in WP2 (Single Trajectory Prediction) are concerned, directly contribute towards (a) offering improvements across ATM via increased trajectory predictability, (b) R&D results for network prediction and performance, assessing traffic complexity and potential overload situations. DART indirectly contributes towards: (c) Performance efficiency via increased predictability of flight arrivals at the planning phase;



			(d) Provision of Common Support Services for stakeholders for data-driven trajectory prediction capabilities and flight planning The potential and limitations of data-driven methods for trajectory predictions have been documented in D2.4.
TRL-1.6	Do the obtained results from the fundamental research activities suggest innovative solutions/concepts/capabilities? -What are these new capabilities? - Can they be technically implemented?	Achieved	Innovative capabilities developed in DART include: (a) from WP2 (Single Trajectory Prediction) innovative data-driven and hybrid (data driven in combination to model-based) methods for predicting trajectories, (b) visual analytics methods and visualizations for assessing data quality, facilitating modelling and algorithms' parameters tuning, and providing comparative results and overviews in space and time of predictions. Technically, all these solutions can be implemented. These capabilities have been described, analysed and evaluated in D2.2, D2.3, D2.4, D1.5.
TRL-1.7	Are physical laws and assumptions used in the innovative concept/technology defined?	Achieved	The only assumption behind machine learning methods explored either in WP2 is that all the necessary features, patterns and ways to assess and compare strategies are there (i.e. In the



			historical data exploited during training of algorithms) and can be identified, or can be computed accurately: This implies requirements on data to be used for training algorithms, as well as expertise to tune algorithms' parameters. This has been reported in WP1 and WP2 deliverables.
TRL-1.8	Have the potential strengths and benefits identified? Have the potential limitations and disbenefits identified? - Qualitative assessment on potential benefits/limitations. This will help orientate future validation activities. It may be that quantitative information already exists, in which case it should be used if possible.	Achieved	Concerning WP2 on single trajectory predictions, the strengths and limitations of the state of the art methods developed are thoroughly examined and reported in D2.4 with quantitative and qualitative assessment of benefits and limitations. This happened also in comparison to model-based methods, showing the benefits and limitations of data-driven methods developed in DART. In addition to that, DART has specified a collaborative decision-making framework involving ATM stakeholders, showing the exact points where developed innovations can contribute in that collaborative process at the planning phase of operations.
TRL-1.9	Have Initial scientific observations been reported in technical reports (or journals/conference papers)?	Achieved	Concerning WP2 (Single Trajectory Prediction) , DART deliverable D2.4 and technical papers submitted and presented/accepted in workshops and conferences provide results and observations on the potential of proposed methods to solve the problems addressed.



TRL-1.10	Have the research hypothesis been formulated and documented?	Partial – Non Blocking	For WP2 (Single Trajectory Prediction) , the hypothesis for data-driven trajectory prediction has not been formulated in any DART document, but as also mentioned in D2.1, it is based on the well-known hypothesis for machine learning algorithms: Re-occurring patterns in historical data can be learnt and be used for making accurate predictions compared to conventional model-based predictions.
TRL-1.11	Is there further scientific research possible and necessary in the future?	Partial – Non Blocking	Yes. Further validation and evaluation of machine learning algorithms for achieving the objectives of WP2 (Single Trajectory Prediction) , is necessary to fully understand their potential and limitations, also in synergy with model-based approaches: Towards this, scalability (for training and exploitation) is one of the issues, while the modularity of implementations allows replacement of modules with alternatives to evaluate different approaches thoroughly. Section 4.3 provides specific detailed proposals for further research activities.
TRL-1.12	Are stakeholder's interested about the technology (customer, funding source, etc.)?	Achieved	DART contacted several meetings with the DART working group comprising people with network management expertise, expertise with airlines operations and ATM research and data



			exploitation. These meetings showed stakeholders' high interest in DART developments as they directly contribute to increasing predictability and addressing critical network management problems. As pointed out DART developments in any of its WPs can be used as a source of inspiration in the solution proposed to the SJU for the SESAR2020 Industrial Research framework programme, namely "SESAR Solution 45: Enhanced Network Traffic Prediction and shared complexity representation". "This SESAR Solution aims at improving the accuracy of the network manager traffic prediction from medium-term planning phase (D-2) to execution. The solution continues and completes the work done in Wave 1 by solutions #09-01 and #09-03. It shall also consider the outcome of the relevant ER projects such as COPTRA". Also, DART outcomes may be considered to feed SESAR 2020 Wave 2 activities (in particular those SESAR Solutions dealing with Advanced DCB as a continuation of current PJ09).
--	--	--	--



Table 4: ER Fund / AO Research Maturity Assessment for Agent Based Methods for DCB optimization

EN: **Agent Based Methods for DCB optimization**

ID	Criteria	Satisfaction	Rationale - Link to deliverables - Comments
TRL-1.1	Has the ATM problem/challenge/need(s) that innovation would contribute to solve been identified? Where does the problem lie?	Achieved	Yes. For WP3 (Collaborative Trajectory Predictions) the problem identified is the assessment of flight delays to trajectories (predicted/planned) due to traffic and DCB problems at the planning stage of operations. The main issues to the assessment of delays due to DCB problems is the complexity involved due to interacting trajectories and the incorporation of stakeholders' interests. However, problem factorization in space and time proved to advance methods' efficiency and scalability. Advances towards addressing these problems should result in improvements in planning, flight management and traffic control. These issues are identified and addressed In D3.1 and D3.2.
TRL-1.2	Has the ATM problem/challenge/need(s) been quantified?	Achieved	Yes. Considering the assessment of delays towards resolving DCB problems in WP3 (Collaborative

Founding Members



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.



			Trajectory Predictions), the problem is to reduce the number of regulated flights and the average delay imposed to flights, also with respect to strategic delay costs, dynamic airspace configurations, and airlines preferences on flights' delays. These developments aim to support NM to reach optimal decisions on flights delays. The quantification of the problems and challenges addressed are described in D3.3, where DART provided information about the delays imposed by the NM in evaluation cases.
TRL-1.3	Are potential weaknesses and constraints identified related to the exploratory topic/solution under research? - The problem/challenge/need under research may be bound by certain constraints, such as time, geographical location, environment, cost of solutions or others.	Partial – Non Blocking	For WP3 (Collaborative Trajectory Predictions), regarding the assessment of delays towards resolving DCB problems, distribution of algorithms to operate at European level and incorporation of all stakeholders' interests into the process are of importance, while being challenging. For the purposes of both WPs, gathering of data, assessing its quality and linking it requires effort and time (including processing time).
TRL-1.4	Has the concept/technology under research defined, described, analysed and reported?	Achieved	Yes. For WP3 (Collaborative Trajectory Predictions), D3.1 provides the overall DART operational context and specific scenarios for individual and collaborative trajectory prediction. Regarding



			multiple co-occurring (in space and time) trajectories that result to DCB problems, the problem has been formalized as a multi-agent Markov Decision Process in D3.2 and D3.2 describes thoroughly the algorithms implemented.
TRL-1.5	Do fundamental research results show contribution to the Programme strategic objectives e.g. performance ambitions identified at the ATM MP Level?	Partial – Non Blocking	DART, as far as agent-based modelling approaches explored in WP3 are concerned, directly contributes towards (a) supporting change in ATM by addressing known critical network performance deficiencies and the provision of tools that can support stakeholders to jointly decide on solving DCB problems at the planning stage; (b) Operational efficiency via advanced dynamic capacity balancing, incorporating the preferences and constraints of stakeholders to flights' delays; contributing R&D results for network prediction and performance, assessing traffic complexity and potential overload situations, providing mitigation strategies to be applied towards resolving DCB problems. Also, DART, as far as agent-based modelling approaches explored in WP3 are concerned, indirectly contributes towards:



			(a) offering Improvements across ATM via improved (collaborative) planning tools and improved flight trajectories, allowing dynamic adjustment to capacity and demand needs; (b) Minimizing the impact of delays due to ATM factors via assessment of delays at the planning phase; (c) Provision of Common Support Services for stakeholders for flight planning and assessment of delays due to DCB problems; The potential and limitations of agent-based methods for assessing flights' delays to resolve DCB problems have been documented in D3.3.
TRL-1.6	Do the obtained results from the fundamental research activities suggest innovative solutions/concepts/capabilities? -What are these new capabilities? - Can they be technically implemented?	Achieved	Innovative capabilities developed in DART include: (a) collaborative multi-agent reinforcement learning methods for regulating flights towards resolving DCB problems, and (b) visualizations for assessing data quality, facilitating modelling and algorithms' parameters tuning and providing comparative results and overviews in space and time of solutions. Technically, these solutions can be implemented.



			These capabilities have been described, analysed and evaluated in D3.2, D3.3, D1.5.
TRL-1.7	Are physical laws and assumptions used in the innovative concept/technology defined?	Achieved	The only assumption behind machine learning methods explored in WP3 is that all the necessary features, and ways to assess and compare strategies are there (i.e. In the historical data exploited during training of algorithms) and can be identified, or can be computed accurately: This implies requirements on data to be used for training algorithms, as well as expertise to tune algorithms' parameters. This has been reported in WP3 deliverables.
TRL-1.8	Have the potential strengths and benefits identified? Have the potential limitations and disbenefits identified? - Qualitative assessment on potential benefits/limitations. This will help orientate future validation activities. It may be that quantitative information already exists, in which case it should be used if possible.	Achieved	Concerning WP 3 , the strengths and limitations of the state of the art methods developed are thoroughly examined and reported in D3.3 with quantitative and qualitative assessment of benefits and limitations. This happened also in comparison to delays imposed by the NM, showing the benefits and limitations of methods developed in DART. In addition to that, DART has specified a collaborative decision-making framework involving ATM stakeholders, showing the exact points where developed innovations can



			contribute in that collaborative process at the planning phase of operations.
TRL-1.9	Have Initial scientific observations been reported in technical reports (or journals/conference papers)?	Achieved	Concerning WP3 , DART deliverable D3.3 and technical papers submitted and presented/accepted in workshops and conferences provide results and observations on the potential of proposed methods to solve the problems addressed.
TRL-1.10	Have the research hypothesis been formulated and documented?	Partial – Non Blocking	For WP3 (Collaborative Trajectory Predictions) , the problem of DCB problems' resolution has been specified in detail and has been formulated in D3.2.
TRL-1.11	Is there further scientific research possible and necessary in the future?	Partial – Non Blocking	Yes. Collaborative reinforcement algorithms for prescribing delays towards resolving DCB problems, explored in WP3 (Collaborative Trajectory Predictions) , may incorporate further constraints (e.g. capacities of airports and other congested resources that need to be regulated), interests of stakeholders (policies for adapting airspace capacities dynamically, according to demand, constraints, etc.), options for alternative routes to be followed – also incorporating relevant costs, as well as additional types of entities (e.g. airspace configurations).



			Section 4.3 provides specific detailed proposals for further research activities.
TRL-1.12	Are stakeholder's interested about the technology (customer, funding source, etc.)?	Achieved	DART contacted several meetings with the DART working group comprising people with network management expertise, expertise with airlines operations and ATM research and data exploitation. These meetings showed stakeholders' high interest in DART developments as they directly contribute to increasing predictability and addressing critical network management problems. As pointed out DART developments in any of its WPs can be used as a source of inspiration in the solution proposed to the SJU for the SESAR2020 Industrial Research framework programme, namely "SESAR Solution 45: Enhanced Network Traffic Prediction and shared complexity representation". "This SESAR Solution aims at improving the accuracy of the network manager traffic prediction from medium-term planning phase (D-2) to execution. The solution continues and completes the work done in Wave 1 by solutions #09-01 and #09-03. It shall also consider the outcome of the relevant ER projects such as COPTRA". Also, DART outcomes may be considered to feed SESAR 2020 Wave 2 activities (in particular those



			SESAR Solutions dealing with Advanced DCB as a continuation of current PJ09).
--	--	--	---





4 Conclusion and Lessons Learned

4.1 Conclusions

WP1 – Data Management

WP1

- Gathered, prepared and made available the different datasets through the DART Data Transaction Pipeline.
- Refined the different datasets as per the requirements and needs from the other WPs, in an iterative workflow which ended in 6 different datasets released.
- It developed capabilities for Visual Analytics in the context of Data Quality assessment, producing new methods.
- It produced a set of reference dataset definitions, able to become a reference for the use cases addressed.

Overall, WP1 provided the necessary data layer supporting the rest of WPs in an adaptive way to support their different needs. This was done by managing a dataset with over 14 Tb size, comprising two years of operational data (2016-2017) including 4 Millions of Flights (4106320 flights), 2.5 Billions of Radar points (2714042496 points), 1 Billion of Flight Plan Route points (1003734563 points), and 89 Millions of Flight Plan messages (89903772 messages).

WP2 – Single Trajectory Prediction

WP2:

- Reviewed and implemented the most promising data-driven algorithms to perform single data driven trajectory prediction.
- Data-driven single trajectory predictor engines developed in WP2 where trained and benchmarked using WP1 datasets.
- Algorithms implemented were able to ingest raw data and derived datasets such as enriched trajectories and AIDL datasets.
- Algorithms benchmarking activities included comparison between data-driven predictions versus flown trajectories, and data driven predictions versus Eurocontrol Network Manager pre-flight prediction.

Overall, WP2 identified, implemented and tested state of the art and promising machine learning algorithms (detailed in section 2) with a variety of incomes/outcomes, identifying which of those are more suitable to implement data driven TP able to outperform, at least in some scenarios, the state of the art model-based trajectory prediction engines.

Algorithms evaluation results show that data-driven methods can achieve high accuracy in predicting trajectories, in particular when they exploit information about flight plans. As a general conclusion for

Founding Members



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.



all algorithms, these perform well in the cruise phase of flight where accuracy of prediction is very close to 100%, however algorithms have lower performance when predicting the trajectory for the phases of departure and arrival.

Evaluation results help to understand that the combination of models and data-driven approaches is the correct way to evolve current operational systems towards the implementation of Trajectory Based Operations (TBO).

WP3 – Collaborative Trajectory Predictions

WP3:

- Identified the DART operational context and scenarios of interest;
- Formulated the demand capacity problem (DCB) at the planning stage of operations, towards regulating flights and resolving hotspots.
- It analysed and exploited data sources provided by WP1 towards constructing real-world experimental cases with comparable results.
- It devised and evaluated four multi-agent collaborative reinforcement learning methods which have been thoroughly evaluated and validated, identifying their potential and limitations: Indeed, the best of the methods managed to reduce significantly the delays imposed to regulated flights, while resolving all hotspots and incorporating preferences of airlines concerning the maximum preferred delay of flights.

Overall, the solution proposed is novel and can take into account dynamically adjusted airspace configurations, stakeholders preferences on maximum flights' delays, while it can be expanded to include capacities of other facilities beyond airspace sectors. Evaluation results for all four methods show that they are capable to resolve hotspots, while keeping the average delay for regulated flights at very low levels, also compared to CFMU regulations, while regulating few flights – compared to the number of flights per day - with fairness.

Agent-based methods can solve DCB problems in computational efficient ways, due to the factorization of the problem imposed.

4.2 Technical Lessons Learned

WP1 – Data Management

- Storage, combination, curation for different data sources have been applied, including different data quality assessment techniques. The logical data management layer has required significant expertise and effort to tackle the different needs and challenges. Data management does not just means having different data sources available, but a much higher degree of complexity, which needs to be managed properly.
- In terms of technical data management, the DART Data Transaction Pipeline has proved to be robust and efficient, while very challenging as the size, heterogeneity and needed throughput of the dataset was wider than usual. Managing a Big Data system, or Data Warehouse, implies



many other aspects than available disk space, for which careful system architecture design needs to be put in place in earlier stages for successful applications research and development.

- The datasets developed have the potential to become a reference in the defined usecase/context of operation, as demonstrated by the different experiments.

WP2 – Single Trajectory Prediction

- Data-driven trajectory prediction engines must be based in different algorithms depending of the training datasets. In DART we have researched a wide range of features going from RAW data, to enriched datasets to dataset's based on formal languages, like AIDL.
- There is a drawback that is common with other data-driven prediction methods that has to be studied: All algorithms learn and predict for city pairs. Learning phase is quite resource demanding and generalizing it to all airport pairs for all flights crossing or flying within the European Airspace is a big challenge.
- Looking at performance times, data-driven algorithms are quite heavy in the training phase, but this is not a problem as this process can be run offline. On the contrary, response times are really low (as the model is already computed, prediction is basically searching in a table) what makes data-driven TP suitable to be run in real time operations.

WP3 – Collaborative Trajectory Predictions

- Visualizations of results provide insights and justifications about the delays assessed, being the necessary tools for comparing solutions and exploring “what-if” alternatives.
- Comparison of results provided by the proposed methods with real-world type C regulations is not that straightforward: This has been done in a very meticulous way, requiring a number of rounds of validating results with WP1, requiring expertise and effort.
- Validation of algorithms has been done in various cases, representing a spectrum of problems to be solved from easy to hard ones. However, our study did not conclude to which are the exact features that determine a problem to be hard. Large delays imposed by CFMU is an indication, but we need in-depth-analysis to reveal what drives these large delays.

Visualizations & Visual Analytics

- The principal approach of using a flexible visualization framework that provides multiple linked views in combination with composable filters to flexibly define subsets in space (airspace entities), time, and entities (flights, routes) has proven valuable both for data exploration and comparative analysis of algorithmic results.
- The ability to select and adjust different aggregates for visualization of these subsets greatly facilitates the assessment of specific areas in space and time with regard to algorithm performance.
- The composable nature of the implemented methods afford great flexibility in defining task-specific visual analytics workflows; this high degree of freedom makes it more suited for analytical experts at the expense of a steeper learning curve for general users.

4.3 Recommendations for future R&D activities (Next steps)

This section comprises two parts: (a) a proposed roadmap of short / medium / long term research activities based on the DART contributions, and (b) specific proposals for exploratory and industrial research based on the findings of individual DART WPs.

4.3.1 Roadmap of research goals

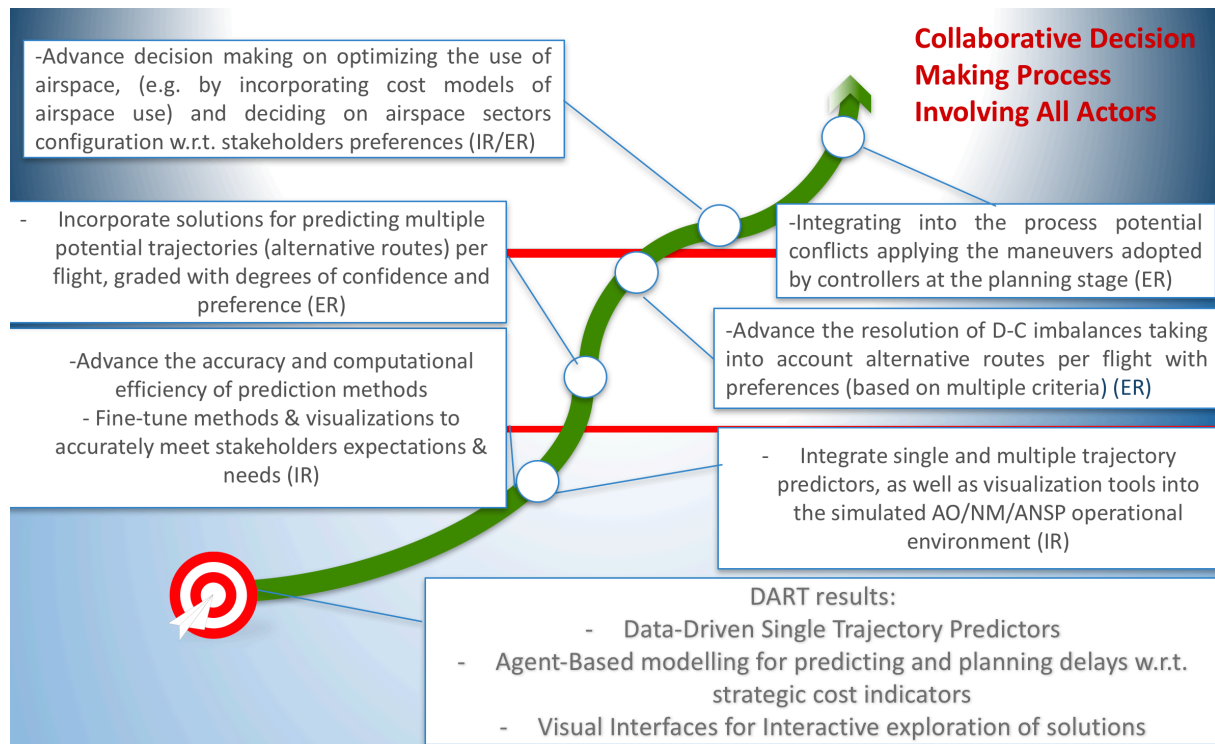


Figure 28 –Roadmap of proposed short / medium / long term goals (from bottom to top) starting from DART contributions.

Figure 28 shows the proposed short, medium and long term goals, starting from DART findings and developments: The figure is “splitted” into three levels by horizontal red lines, indicating the short term goals (at the bottom), medium (in the middle) and long term (at the top) goals.

It must be noticed that research goals at all levels are characterized to be goals to be achieved by either exploratory (ER) or industrial (IR) research activities: For instance, short term goals that are based directly on the DART outcomes and aim to explore the deployment of these developments in operational context, can be achieved by IR activities, while further, more ambitious, developments need further exploration, and thus these should be achieved by ER activities. The overall orientation of proposed goals and activities is to contribute towards “collaborative decision making involving all actors” of the complex ATM system of systems.

It should be noted that combinations of goals mentioned can be achieved by one or a combination of research activities (thus, there is not an one-to-one correspondence between proposed goals and research activities achieving them).



The first set of goals that can be achieved via IR activities (Figure 26: lower level, left) include:

- Advancing the accuracy, computational efficiency and potential for deploying methods for single trajectory prediction, based on the weaknesses and limitations of state of the art methods explored in DART.
- The computational efficiency and distribution of agent-based methods for computing the effects of traffic to individual trajectories, need to be explored towards the deployment of these methods at European level.
- Computational methods and visualizations, need to be tuned to accurately meet stakeholders needs and requirements.

A second set of short term goals that can be achieved by industrial research activities (Figure 26: lower level, right) concern integrating single trajectory predictors and agent-based methods developed in DART, as well as visualization tools into the simulated AO/NM/ANSP operational environment: This tool will provide further advise to the NM (in addition to existing tools and methods) to regulate flights based on predictions for individual flights and on methods for exploiting these predictions and/or flight plans towards assessing the delays of flights for resolving DCB problems, at the planning stage. This first deployment effort of DART developments in an operational context, in conjunction to existing tools, shall provide further experience on assimilating DART-related research outcomes into decision making tools.

Medium term goals correspond to ER activities, and are the following ones:

A first goal is to explore the potential to advance and/or combine DART state of the art methods (Figure 26: medium level, left) on single trajectory prediction with probabilistic methods that provide multiple predictions per trajectory (such those explored in COPTRA SESAR project). A detailed comparison between these methods and an in-depth exploration of the benefits offered to stakeholders towards (collaborative) decision making should be explored.

A second goal (Figure 26: medium level, right), as a further consequence of the first medium term goal, is to explore the incorporation into DART multi-agent methods multiple predictions per trajectory, towards resolving demand capacity imbalances: This is not straightforward as multiple predictions and their possible combinations explode the state space for reinforcement learning methods. However, such methods may advance the operational status of NM towards a collaborative decision-making process, incorporating into the DCB process airlines' preferences on routes to be followed per flight, taking also into account cost indicators.

Finally, as far as long terms goals are concerned, these can be addressed mostly by ER activities as follows:

A first long-term goal (Figure 26: upper level, left) is to advance DART multi-agent reinforcement learning techniques by incorporating, in conjunction to agents representing individual flights, other types of agents: For instance, agents corresponding to air space sectors, so as to optimize the airspace sectorization, in conjunction to optimizing the flights' delays, based on stakeholders' preferences and cost indicators. Results from the COCTA SESAR project, aiming to introduce marketing models for NMs to provide the use of airspace, can be incorporated into such multi-agent models. We expect that DART multi-agent methods can be extended towards incorporating such models, so this goal can be achieved by IR activities, in conjunction to ER activities.

A second more ambitious goal for single trajectory predictions and for multi-agent models towards resolving DCB problems in the planning phase of operations, it to advance these methods towards incorporating predictions for resolving conflicts, already at the planning phase. This will allow making fine predictions and assessment of trajectory evolution, incorporating air traffic controllers' interests and concerns into the overall process.

4.3.2 Proposed exploratory and industrial research activities per DART work package

WP1 for Data Gathering, Curation and Provision

Exploratory Research:

- Data Quality assessment is an essential aspect for successful Big Data technologies deployment. Data Quality assessment is an essential aspect for successful Big Data technologies deployment. The actual maturity level for ATM information needs to be further refined with the development of adequate techniques. In particular:
 - While the existing ATM information is mainly considered for operational purposes, it presents some flaws that make difficult the direct exploitation of data in data-driven processes (e.g. imperfections in data, lack of common references among data sources complicating their integration, etc.). Nowadays, little (if any) post-operational curation is applied to operational data.
 - There is a lack of integrated data which is usable for research purposes. Integrated data sources (i.e., DDR), when available, do not present the granularity that is necessary for advanced processing & analytics tasks, which require large volume of data of high detail/frequency. To mitigate this issue a lot of different data sources need to be considered (with potential lack of access to some of them).
 - The information currently available in planning phase is not reliable for accurate optimization or forecasting. This situation will be improved with RBT/SBT and SWIM, providing more reliable information in the planning phase.

Industrial Research:

- In terms of enabling, the system developed for data provision shows sufficient maturity to be considered in the application of Big Data techniques in a SWIM environment. The findings of DART in terms of Big Data architecture and Data Transaction Pipeline suggest a sufficient maturity level to become an input for corresponding Enabling Projects.

WP2 for Single Trajectory Predictions

Exploratory Research:

- Explore clustering methods to be used in combination with other machine learning algorithms. This would help for instance classification methods to assign trajectory to be predicted to the "correct" clusters.

Industrial Research:

Founding Members



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.



- Study the situations where Eurocontrol NM and other state of the art model-based trajectory predictors can be improved by using data driven TP's. A quick win for the industry won't be switching directly to a data-driven environment, but combining the best of both worlds (data driven – model driven).

WP3 for Multi-Agent Models Accounting for the Complexity of DCB Problems

Exploratory Research:

- Multi agent methods have the potential to include the capacities of other resources/facilities (e.g. airports or specific routes) in addition to the capacities of sectors, and this is something to be explored in order to decide on delays, taking into account all the facilities/resources required by any flight.
- Alternatives flight plans or alternative routes within a single flight plan (i.e enhancing the form of flight plans to include multiple routes) should be considered while deciding on delays due to traffic, taking into account cost indicators for re-routing, as well.
- While the multi-agent methods proposed take into account adaptations of sectorizations at scheduled times, their capability to adapt solutions in more dynamic environments should be explored.
- Multi-agent methods where agents represent flights (as it is done now) but also other entities of interest (e.g. sectors/sectorizations) can be considered, towards addressing decisions regarding adaptations of capacities to demand, in conjunction (jointly) to the regulation of flights due to DCB problems.
- Multi-agent methods should be explored towards (a) analysing equity, and (b) relaxing the DART assumption that the set of flights are known in advance (i.e before solving the DCB problem). Towards this, agent-based methods should be adapted to handle a rolling plan, where flight plans are submitted continuously even when other interacting flights are already airborne.

Industrial Research:

- Distribution of algorithms to solve problems at regional / FIR levels in a coordinated way, resulting to solutions for the entire European FIR is an issue that requires further research.
- Development of a decision-making tool as an aid for Network Manager to resolve DCB problems at the planning phase: This tool can be used in conjunction to existing facilities, to provide advice towards resolving DCB problems using the DART methods' paradigm, also in conjunction to appropriate interactive visualizations that can provide justifications to methods' results and allow NM to ask what-if questions.

Visualizations & Visual Analytics

Exploratory Research:

- Visualization methods focus on the exploration and visual analysis of the results of the various algorithms developed in WP2 and WP3, after these have been built/trained. Future work



should examine possible benefits from visually instrumenting the modelling/learning process itself to better utilize tacit domain expert knowledge.

- The aforementioned exploration of dynamic environments and additional entities for agent-based simulation need suitable means to define, manipulate and monitor these environments as well as agent behaviour through novel interactive visualizations.

Industrial Research:

- The proposed ensemble of visual analytics methods aims at maximizing analytical flexibility. An operational tool used by Network Managers likely requires a more streamlined composition aligned with prescribed workflows, after further corresponding requirements engineering.
- For further streamlined “what-if” analysis, e.g. in arbitration of trajectory management solutions, further software development should achieve tight integration of the selected analysis algorithm and interactive visualization components.



5 References

5.1 Project Deliverables

- [1] **D1.1: DART- Data Management Plan**, ed. 00.01.03, CRIDA, 27 Oct 2016, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [2] **D1.2: DART- Data Transaction Pipeline Description**, ed. 00.01.00, CRIDA, 25 Nov 2016, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [3] **D1.3: DART- DART Data Pool**, ed. 00.01.01, CRIDA, 15 Jan 2018, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [4] **D1.4: DART- Synthetic Data Package**, ed. 01.00.00, BR&T-E, 02 May 2018, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [5] **D1.5: DART- Visualization exploration report**, ed. 01.00.00, FRHF, 18 May 2018, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [6] **D2.1: DART- Initial Set of Data-Driven Trajectory Prediction Algorithms**, ed. 02.00.00, BR&T-E, 02 Feb 2017.
- [7] **D2.2: DART- Visual Interface for Algorithms Analysis**, ed. 02.01.00, FRHF, 29 Aug 2017, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [8] **D2.3: DART- Enhanced Set of Data-Driven Trajectory Prediction Algorithms**, ed. 01.01.00, BR&T-E, 19 Feb 2018.
- [9] **D2.4: DART- Evaluation and Validation of Algorithms for Single Trajectory Prediction**, ed. 04.020.00, BR&T-E, 03 July 2018, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [10] **D3.1: DART- Collaborative Trajectory Prediction Scenarios and Requirements Specification**, ed. 02.00.00, UPRC, 18 Apr 2017, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [11] **D3.2: DART-Collaborative Trajectory Prediction Algorithm**, ed. 00.02.00, UPRC, 15 Mar 2018.
- [12] **D3.3: DART- Evaluation and Validation of the Collaborative Trajectory Prediction Algorithm**, ed. 02.00.00, UPRC, 16 May 2018, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [13] **D4.1: DART- Project Management Plan**, ed. 02.05.00, UPRC, 19 Sept 2016, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>



- [14] **D4.2: DART-Dissemination Plan**, 02.00.00, BR&T-E, 26 Sept 2016, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [15] **D4.3: DART- Project website, wiki, social media channels**, ed. 01.01.00, BR&T-E, 26 Oct 2016, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [16] **D4.4: DART- Dissemination Report**, ed. 01.00.00, BR&T-E, 19 May 2018, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>
- [17] **D4.5: DART- Project Results Final Report**, ed. 04.00.00, UPRC, 9 Jul 2018, available at <http://dart-research.eu/2018/07/10/dart-final-deliverables>

5.2 Project Publications

- [18] **DART Facts Sheet**: http://dart-research.eu/wp-content/uploads/2017/01/DART_fc.pdf, 2016
- [19] **DART White Paper**: <http://dart-research.eu/2017/04/03/dart-white-paper/>, 2017
- [20] T.Kravaris, G.Vouros, C.Spatharis, K.Blekas, G.Chalkiadakis, J-M.Cordero Garcia “**Learning Policies for Resolving Demand-Capacity Imbalances during Pre-tactical Air Traffic Management**”, 15th German Conference on Multiagent System Technologies, Leipzig, Germany, 2017.
- [21] G Andrienko, N Andrienko, W Chen, R Maciejewski, Y Zhao. “**Visual Analytics of Mobility and Transportation: State of the Art and Further Research Directions.**” IEEE Transactions on Intelligent Transportation Systems, 2017.
- [22] J-M Cordero Garcia et al on “**Integration of meteorological information in trajectory prediction (DART Project)**”, 1st “International Workshop on Meteorology and Air Traffic Management, hosted by University of Sevilla and funded by project TBO-MET, 2016.
- [23] Gennady Andrienko, Natalia Andrienko, Georg Fuchs, Jose Manuel Cordero Garcia. **Clustering Trajectories by Relevant Parts for Air Traffic Analysis**. IEEE Transactions on Visualization and Computer Graphics, 2017 (accepted).
- [24] Boeing Research and Technology Europe filed a patent in the European Patent Office: Applicant: The Boeing Company Application title: “**Method and system for autonomously operating an aircraft**”, Filing date: 29th June 2017, Filing number: EP17382412.9
- [25] E.C. Fernández et al., **DART: A Machine-Learning Approach to Trajectory Prediction and Demand-Capacity Balancing**, SID 2017.
- [26] G. Vouros et al., **Multiagent RL for Real-World Interdependent Congestion Problems**, accepted in SETN 2018 (Hellenic Artificial Intelligence Conference).
- [27] C.Spatharis et al: **Multiagent Reinforcement Learning Methods for Resolving Demand-Capacity Imbalances**, accepted in DASC 2018.



- [28] E. Casado, A. Muñoz: **Data-driven Aircraft Trajectory Predictions using Ensemble Meta-Estimators**, accepted in DASC 2018.
- [29] WAC-2018 workshop: **DATA DRIVEN ATM: GOING DIGITAL** (in collaboration with COPTRA, BigData4ATM, UPM, datAcron).
- [30] ICRA 2018 Workshop on **Data Enhanced Trajectory Based Operations**, in collaboration with datAcron.

5.3 Other

- [31] J. Lopez-Leones, M. Vilaplana, E. Gallo, F. Navarro and C. Querejeta, "The Aircraft Intent Description Language: A key enabler for air-ground synchronization in Trajectory-Based Operations," in Digital Avionics Systems Conference, 2007. DASC '07. IEEE/A.
- [32] "<https://www.eurocontrol.int/services/bada>," [Online]
- [33] M. G. Hamed and et al., "Statistical prediction of aircraft trajectory: regression methods vs point-mass model," in ATM 2013, 10th USA/Europe Air Traffic Management Research and Development Seminar 13 Jun. 2013.
- [34] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon. Visual analytics: Definition, process, and challenges. *Information visualization*, Springer, 2008, pp. 154-175.
- [35] W. Kun and P. Wei, "A 4-D trajectory prediction model based on radar data," in Control Conference, 2008. CCC 2008. 27th Chinese 16 Jul. 2008.
- [36] Y. Le Fablec and J.-M. Alliot, "Using Neural Networks to Predict Aircraft Trajectories," in IC-AI, 1999.
- [37] Cheng, Taoya, D. Cui and P. Cheng, "Data mining for air traffic flow forecasting: a hybrid model of neural network and statistical analysis," in Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE 12 Oct. 2003: 211-21.
- [38] A. de Leege, M. van Paassen and M. Mulder, A Machine Learning Approach to Trajectory Prediction., 2013.
- [39] K. Tastambekov and et al, "Aircraft trajectory forecasting using local functional regression in Sobolev space," Transportation research part C: emerging technologies, vol. 39, pp. 1-22.
- [40] S. Hong and K. Lee., "Trajectory Prediction for Vectored Area Navigation Arrivals," Journal of Aerospace Information Systems (2015).
- [41] S. Yue, P. Cheng and C. Mu, "An improved trajectory prediction algorithm based on trajectory data mining for air traffic management," in Information and Automation (ICIA), 2012 International Conference, 6 Jun. 2012.



- [42] Yepes, Javier Lovera, Inseok Hwang, and Mario Rotea. "New algorithms for aircraft intent inference and trajectory prediction." *Journal of guidance, control, and dynamics* 30.2 (2007): 370-382.
- [43] S. Ayhan, H. Samet, "Aircraft Trajectory Prediction Made Easy with Predictive Analytics", *Proceedings of ACM SIGKDD*, 2016, pp. 21-30.



Appendix A

A.1 Glossary of terms

Term	Definition	Source of the definition
------	------------	--------------------------

Table 5: Glossary

A.2 Acronyms and Terminology

Term	Definition
AI	Aircraft Intent
ANSP	Air Navigation Service Provider
ATC	Air Traffic Control
ATFCM	Air Traffic Flow Control Management
ATM	Air Traffic Management
BADA	Base of Aircraft Data
CART	Classification And Regression Tree
CFMU	Control Flow Management Unit
DCB	Demand Capacity Balance
DST	Decision Support Tool
ETA	Estimated Time of Arrival
FMS	Flight Management System
HMM	Hidden Markov Models
LR	Linear Regression
NM	Network Manager
NN-MLP	Neural Networks – Multi-Layer Perceptron
MDP	Markov Decision Process
ML	Machine Learning
MOME	Multi-Output Meta Estimators
MP	Master Plan
SESAR	Single European Sky ATM Research Programme
SJU	SESAR Joint Undertaking (Agency of the European Commission)
TBO	Trajectory Based Operations
TP	Trajectory Predictor

Table 6: Acronyms and terminology



Appendix B

DART Workshop at ICRAT conference

The workshop took place in Barcelona on June 25th during the ICRAT conference. It was organized in collaboration with datAcron (Big Data Analytics for Time Critical Mobility Forecasting), an EU H2020 funded research project that explore the applicability of state-of-the-art applications and most modern data analytics techniques and methodologies into the ATM and Maritime domains.

The main topics discussed during the workshop included:

- ATM Data management
- Data-Driven Trajectory Prediction
- Agent-Based Modelling of the ATM system
- Data-driven predictions for flow management
- Complex Event recognition & forecasting for ATM
- Visual Analytics for ATM

In terms of attendance the workshop exceeded expectations with an average of 50 and peaks of over 60 participants, including Industry, civil aviation authorities, ANSP's, Airlines, Pilots, Eurocontrol and FAA representatives and research institutes all around the globe, with a big presence of researchers and students from Asian universities.

The workshop was organized in three big blocks and a Keynote speaker:

- Keynote speaker – Goran Stojkovic (Boeing Digital Airline and Analytics)

Session 1.- Trajectory Data Management

- Trajectory Data storage (CRIDA)
- Big data integration and management for the ATM domain – (UPRC)

Session 2.- Data-Driven Mobility Predictions & Analytics

- Hybrid data-driven trajectory prediction analytics (BR&T-E)
- Data-driven aircraft trajectory prediction (UPRC)
- Complex Events Recognition (CRIDA)
- Visual Analytics of Trajectory Data (NSCR/FRHF)

Session 3.- Applications of Machine Learning to TBO

- Trajectory based analytics: A Maritime Situation Awareness Perspective – (NARI/CMRE)

Founding Members



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.



- Trajectory-centered agent-based modeling of ATM (UPRC)
- Python Big Data Analytics with Dask (Juan Luis Cano – Python Spain)

All presentations are available in the DART website: <http://dart-research.eu/03-work-plan/work-progress/data-enhanced-trajectory-based-operations-workshop/>

Workshop main discussions and conclusions:

- The importance of the aircraft as a sensor, and the potential benefits of applying analytics and machine learning techniques to data that can be obtained from the aircraft and the one that will be obtained from future models. This especially applies to trajectory prediction.
- The volume of weather information that is by far the biggest dataset you have to deal with when predicting trajectories. The features used from these datasets (i.e. wind magnitude, wind direction, temperature) are necessary to accurately predict trajectories.
- As takeaway, most of the participants agree on the need of visual analytics to check data-driven trajectory predictions. In average (lower cross track, along track and horizontal error) data trajectories may be better but we always have to check visually what data driven algorithms are predicting.
- Great interest in the scope of dataAcron, especially in finding a common architecture and data-driven prediction algorithms and event recognition patterns for Maritime and ATM domain.
- Eurocontrol was interested in DART, to double check that outcomes of DART are used as a baseline in future exploratory or industrial research projects.
- The last presentation, “Visual Analytics of Trajectory Data”, provided a deep dive in specific details of Python implementation of machine learning trajectory prediction algorithms. Attendees greatly appreciated this, so another take-away is to include presentations like this (i.e. hand-on-experience with modern technologies) in any workshop that is related with novel computer implemented techniques.