SESAR Engage KTN – PhD final report

PhD title:	Trajectory Planning for Conflict-free Trajectories: A Multi Agent Reinforcement Learning Approach
Candidate's name:	Alevizos Bastas
Lead supervisor's name:	George Vouros
Co-supervisor's name (if applicable):	-
Proponent institute:	University of Piraeus
Consortium institutes (if applicable):	CRIDA (Spain)
Thematic challenge:	TC2 Data-driven trajectory prediction
Edition date:	29 June 2022
Edition:	1.1
Dissemination level:	Public

This project has received funding from the SESAR Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No 783287.

1. Abstract
The objective of this Engage KTN PhD study is to explore and present state of the art AI/ML algorithms towards planning conflict-free trajectories in computationally efficient ways, for a large number of trajectories in airspaces comprising multiple FIRs, following a methodology combining data-driven and agent-based approaches.
In the context of this study the conflicts-free trajectory planning task is defined to incorporate *trajectory prediction* and *conflicts detection and resolution*. While *trajectory prediction* concerns predicting the spatiotemporal evolution of the aircraft state along a trajectory (also called, trajectory evolution), *conflicts detection and resolution* concerns the detection of conflicts that breach separation minima (loss of separation) between flights and their resolution by appropriate actions. Therefore, the objective of the conflicts-free trajectory planning task is to predict the evolution of trajectories, and regulating flights to avoid loss of separation.
While trajectory planning may take place at the pre-tactical phase of operations, we expect the methods developed in this study to have a large impact in the tactical phase of operations.
Aiming to model stakeholders' decisions to planning conflict-free trajectories, the major emphasis of this study is to imitate flights' trajectories and air traffic controller's behavior according to demonstrations provided by historical data.

The challenges that this study addressed are as follows:
1. Plan trajectories, considering complex ATM phenomena and operational constraints regarding traffic and conflicts among trajectories.
2. Follow a data-driven approach to learn stakeholders' preferences on the evolution of trajectories and on resolving conflicts: stakeholders include airspace users (for trajectory prediction) and air traffic controllers (for conflicts' detection and resolution actions).
3. Address optimization in trajectory planning w.r.t. multiple objectives, preferences and constraints of stakeholders involved, as these are demonstrated by historical data.
4. Address scalability: demonstrate the efficiency of the methods to be applied in settings with a large number of flights.

Contributions that this study makes are as follows:
1. The problem of modelling air traffic controllers' behavior has been split into two well-defined problems: modelling air traffic controllers' reactions on *whether and when* conflicts' resolution actions should be applied, and modelling air traffic controllers' reactions on *how* conflicts should be resolved, i.e. what resolution actions should be applied.
2. The problem of trajectory planning (either with or without considering conflicts) has been formulated as an imitation learning problem, based on historical flown trajectories.
3. AI/ML methods have been developed and tested on learning models regarding the evolution of 4D trajectories, using data-driven approaches, i.e. based on historical real-world data.
4. AI/ML methods have been developed and tested on learning models regarding air traffic controllers' reactions and policy using data-driven approaches, i.e. based on historical real-world data.
5. This study has proposed an elaborated evaluation method for data-driven imitation learning techniques predicting air traffic controllers' reactions, considering the uncertainties involved in the evolution of trajectories, in the assessment of conflicts, and in the reactions of ATCO.
6. Challenging issues due to inherent data limitations have been addressed and thoroughly discussed.
7. The study provides an integrated trajectory planning approach, where data-driven trajectory predictions are intertwined with data-driven conflicts detection and resolution.

2. Objective of the study

The main objective of this study is the following:

Develop AI/ML methods towards planning conflicts-free trajectories.

This objective is aligned with the Engage thematic challenge 2 "Data-driven trajectory prediction".

The concrete objectives of this study are as follows:

A. Develop a formulation of the trajectory planning problem (and subproblems), considering prediction of trajectories and resolution of conflicts.
B. Develop an AI/ML learning method incorporating reinforcement learning for the prediction of trajectories per Origin-Destination pair, without explicitly considering conflicts.
C. Develop AI/ML learning methods incorporating reinforcement learning for the resolution of conflicts.
D. Combine AI/ML models for the intertwined prediction of trajectories and the detection and resolution of conflicts, towards a method for the planning of conflicts-free trajectories.

Methodological objectives towards these main objectives are the following:

1. Gather and process high-quality data sets for data-driven planning of trajectories, revealing behavior of airspace users and air traffic controllers in different circumstances. Specifically,
   - 1.a Specify and gather high-quality data sets for training AI/ML methods towards planning conflicts-free trajectories.
   - 1.b Process and associate data from the different data sets, so as to
     - 1.b.1 understand the phenomena regarding conflicts and their resolution, as these are revealed by the datasets,
     - 1.b.2 detect data imperfections, resolve challenging issues and address inherent data limitations,
     - 1.b.3 provide AI/ML methods with concrete cases to which they will be trained and tested.

2. Review and compare thoroughly state of the art techniques on
   - 2.a Data-driven conflict-free trajectory planning (i.e. trajectory prediction and conflicts detection and resolution), providing evidence on the novelty and significance of the developments in this study.
   - 2.b AI/ML methods for imitating experts' policy to perform tasks (in our case, to imitate the evolution of flight trajectories and the resolution of conflicts).

3. Formulate and test alternative formulations using state-of-the-art AI/ML methods, regarding
   - 3.a The behavior of airspace users on executing trajectories per Origin-Destination pair.
   - 3.b The behavior of air traffic controllers on resolving conflicts.

Contributions that this study aims to make to the ATM Master Plan:

1. Imitate air traffic controllers to resolve conflicts, supporting better planning of operations for Airspace Users.
2. Improved operations productivity via contributions to improved collaborative planning tools accounting for complex phenomena due to traffic.
3. Increasing predictability via efficient operation plans, reducing buffers and uncertainty.

3. Motivation

Europe has a complex airspace, where 30.000 daily flights usually overfly its sky. Therefore, it is one of the airspaces with most activity in the world. While this number is expected to increase significantly in the coming years, Air Traffic Management (ATM) needs to handle greater complexity and larger volumes of traffic. While ATM in Europe is shifting towards the Trajectory-Based Operations (TBOs) paradigm, it aims to rely on 4-D trajectories to increase the efficacy of planning at pre-tactical phase (i.e. a few hours to a few days before operation, before the trajectory becomes a reference business trajectory (RBT)), and also during the tactical phase (i.e. during operations) towards addressing complexity issues.

During the planning phase of operations the airspace users agree with ANSPs, airport operators, the airspace user's preferred trajectory, where the various constraints of airspace and airport capacity are fully taken into account. Once agreed, the business trajectory (BT) becomes the Shared Business Trajectory (SBT), the trajectory that the airspace user agrees to fly and all the service providers agree to facilitate. Exploiting this plan in conjunction to the evolution of the actual trajectory, air traffic controllers detect and resolve conflicts during the pre-tactical (as done by the planning controller) and the tactical phase (as done by the executive controller) of operations.

The main objective of conflict-free planning of trajectories is to protect the Air Traffic Control (ATC) service from overload[1], enabling controllers (ATCO) to deal with complex traffic situations. Given the uncertainties during the planning phase, as well as while executing a plan, reliable planning of conflict-free trajectories is not that straightforward. While the transition to TBOs will progressively improve availability of information, our aim is to increase accuracy for trajectory planning and execution, providing great benefits to the ATM system, mainly in terms of predictability, which is the main driver for improvement in other KPAs, such as capacity.

In addition to the above, while planning of flight trajectories involves multiple stakeholders (Airspace Users (AUs), Air Navigation Service Providers (ANSPs), Network Manager (NM), Airport Operators (AOs)), planning of conflict-free trajectories also brings the preferences/best practices of Air Traffic Controllers in perform of their duties.

Based on the above, this study is motivated to present methods for the planning of conflict-free trajectories, either at the pre-tactical phase towards an SBT, or at the tactical phase of operations, incorporating into the process preferences/practices and constraints of stakeholders (mainly, air space users and air traffic controllers), building models that are close to their objectives and their behavior, as these are revealed by historical data on executing flight trajectories and resolving conflicts.

Following an agent-based approach, this study targets towards addressing complex phenomena occurring due to traffic, thus, resolving co-occurring conflicts simultaneously, dealing with the effects of conflicts' resolutions applied, w.r.t. stakeholders' preferences and interests, as well as w.r.t. operational constraints.

By doing so, this study contributes towards collaborative decision making by imitating conflict-free trajectory planning (i.e. trajectory prediction and conflicts detection and resolution), accounting for complex phenomena due to traffic, increasing predictability via efficient operation plans, reducing buffers and uncertainty as much as possible, and reducing flight inefficiencies due to tactical ATC actions, supporting better planning of operations for Airspace Users.

[1] University of Westminster; Eurocontrol, 'European airline delay cost reference values', p. 86, 2011.

4. Advances this work has provided with regard to the state of the art

Given that the main objective of this study is to develop AI/ML methods towards planning conflicts-free trajectories, *the main contribution is the development of data-driven AI/ML models for (a) the prediction of trajectories, (b) the resolution of conflicts among flights, as well as (c) the combination of such models towards devising a method for planning conflict free trajectories*.

To achieve this main objective, this study advances the state of the art in three major and challenging topics:

Develop an AI/ML learning method incorporating reinforcement learning for the prediction of trajectories per Origin-Destination pair, without explicitly considering conflicts.

Specifically, this study has formulated the trajectory prediction problem as a data-driven imitation learning problem and developed imitation learning algorithms for learning trajectory prediction models for different origin destination pairs. The study reports on extensive experimental results regarding the efficacy of these models.

Specific, major contributions made are as follows:

- The trajectory prediction problem has been formulated as an imitation learning process, where models of trajectories are learnt from historical trajectories provided as “expert" demonstrations, considering that these trajectories have been “shaped" by aggregating stakeholders' policies, preferences and objectives: This, in synergy with the Engage KTN Catalyst Project on Data-driven Trajectory Imitation with Reinforcement, is the first work that has done so (as far as we know).
- State of the art imitation learning methods have been studied, towards learning trajectory models without making any assumption on the form of a cost function, in continuous state-action spaces, with no specific requirements on specifying trajectory constraints (e.g. without requiring information on flight plans), and with minimal data pre-processing requirements.
- Extensive experimental results are provided that concern trajectories between Origin-Destination (OD) airports' pairs with different characteristics, demonstrating the prediction abilities of the method, either at the pre-tactical or at the tactical stage of operations.

Develop an AI/ML method incorporating reinforcement learning for the detection and resolution of conflicts.

This study contributes to conflict detection and resolution (CD&R) tasks executed as part of the Air Traffic Control (ATC) service, promoting safe, orderly and expeditious flow of air traffic, by modelling Air Traffic Controllers' (ATCO) behavior in resolving conflicts using data-driven AI/ML techniques. In general, according to the problem specifications made in this study, this implies learning “*when*” the ATCO will react to resolve a detected conflict, and “*how*” he/she will react: The first is *the ATCO reaction problem* specifying “whether” and “when” the ATCO will react , while the second is the *problem of learning the ATCO policy*, specifying “how” he/she will react in the presence of conflicts.

The specific contributions made towards *the ATCO reaction problem* are as follows:

- The problem of CD&R has been formulated as an imitation learning problem, aiming to learn ATCO behavior in a hierarchical manner. In so doing, ATCO reaction prediction problem is formulated.

- A supervised deep learning method employing a Variational Auto-Encoder (VAE) for predicting ATCO reactions has been devised, in the context of a methodology to model ATCOs behavior;
- A data-driven method for simulating the uncertainty in the evolution of trajectories and for detecting the potential conflicts that *may have triggered ATCOs reactions* (this is a challenging issue due to inherent data sources limitations), has been proposed;
- A methodology for evaluating data-driven methods to resolve the ATCO reaction problem has been devised, taking into account uncertainties involved in the process;
- The proposed method has been evaluated comparatively with baseline methods towards modelling ATCOs reactions, using real world data.

Regarding the *problem of learning the ATCO policy,* the contributions made towards this objective are as follows:

- The problem of learning the ATCO policy has been formulated as an episodic, single-stage imitation learning problem and as a classification task (alternatively), aiming to learn the actual ATCO behavior, as it is revealed in historical data.
- Supervised machine learning methods for predicting the ATCO actions have been devised in the context of the overall methodology to model ATCO: These methods include (a) a single stage episodic imitation learning method based on the Generative Adversarial Imitation Learning (GAIL), and (b) classification methods using a neural network, random forests, gradient boosting and support vector machines.
- The proposed methods have been evaluated comparatively towards modelling ATCO behavior, using real world data.

As far as we know, this is the first study towards modelling ATCO behavior and ATCO reactions in a data-driven way. Indeed, most of the CD&R approaches are trained, validated and tested in simulated settings, where agents (flights) learn autonomously while acting in their environment, without considering actual ATCO behavior. We conjecture that in safety critical domains such as ATC, the actions proposed by automated systems should be “similar” to those taken by humans: By “similar” we mean with a short distance in any of the temporal, spatial dimensions at which actions are decided and applied along the trajectory, as well as with strong similarity to conflict resolution actions performed by ATCO. This implies safety in the automation process, taking into account human expertise, (human-like) flexibility and tolerance in reacting to situations. We believe that such an approach, is capable of developing trust to an automated system: Actions that are close to the human rationale are more understandable or self-explanatory to human operators, and the system objectives can be made intuitively transparent, given that the system models the ATCO objectives and preferences. Although this does not imply lack of need for explainability/transparency of AI/ML methods, the provision of explanations is out of this study scope.

Indeed, data-driven techniques for conflict resolution have the potential to reveal and incorporate in the decision-making process the preferred behavior of the various stakeholders, as this information lies implicit in the demonstrated historical data, and is being represented in a machine-crafted model, learnt by exploiting the appropriate data sources.

A challenging issue of such a data-driven imitation process, as experienced by this study, is that historical expert samples (i.e. flown trajectories annotated with ATCO resolution actions) do not indicate, together with the resolution actions, the observations perceived by ATCOs before the resolution action, driving the specific action. Such observations include features concerning the

evolution of the trajectories perceived/assessed by the ATCO before their “intervention”, the features of conflicts assessed, as well as the evolution of conflicts after the instruction of a resolution action. However, historical data sets indicate in the best case the effect of ATCO resolution actions, but neither the potential evolution of the trajectories before the resolution action, nor how trajectories would evolve if the ATCO resolution action had not been applied. This is a challenging issue in the learning process, since imitating the “*when*” and “*how*” of the ATCO behavior necessitates *recovering* the specific state, and the important observations that the ATCO perceived or predicted, driving decisions. Neither of this information is provided in historical data (data provide only the action type instructed). To reveal such details from historical data is not a trivial task as the evolution of the trajectories is uncertain.
This study deals with this challenging problem to a large extent.

Develop an AI/ML method for planning conflict-free trajectories.

Addressing this challenge involves combining models for ATCOs behavior for resolving conflicts with models of predicting trajectories, into a single method for planning conflict-free trajectories.
In this context, we advance the trajectory prediction problem formulated as an imitation learning problem to take into account models of ATCO behavior in resolving conflicts.
Results from experiments on trajectories among different OD pairs in Europe, aim to show the effectiveness and efficiency of the overall approach and show its effectiveness in terms of accuracy of predictions.

The specific contributions (this is still in progress when publicizing this report) towards this objective are as follows:

- Enhance the imitation learning method for predicting trajectories without considering conflicts, to a conflict-free trajectory planning framework, where models of ATCO behavior are incorporated to predict the evolution of trajectories while resolving conflicts.
- The data-driven framework for conflicts-free trajectory planning is generic, being able to incorporate different trajectory prediction and ATC behavior models.
- Different instantiations of the proposed framework are evaluated comparatively, using real world data.

5. Methodology

The overall methodology to address the objectives of this study, whose structure in terms of work-packages is shown in Figure 1, comprises the following steps:

Step 1: Develop an AI/ML learning method incorporating reinforcement learning for the prediction of trajectories per Origin-Destination pair, without explicitly considering conflicts.

Specifically, this step formulates the trajectory prediction problem as a data-driven imitation learning problem. Aiming to imitate the experts "shaping/evolving" trajectories, this study devises AI/ML methods that learn policy models incorporating preferences, strategies, practices etc. in an aggregated way, as revealed by historical data.

In this context, the trajectory prediction problem has been formulated as an imitation learning problem and the Generative Adversarial Imitation Learning (GAIL) state of the art imitation learning method has been selected to learn the models.

To evaluate the effectiveness and efficiency of the approach, experiments on trajectories among different OD pairs report on the following measures regarding the accuracy of the predictions: (a) Root Mean Square Error (RMSE) in meters in each of the 3 dimensions, as well as in 3D, (b) Along-Track Error (ATE), (c) Cross-Track Error (CTE), and (d) Vertical deviation (V), between predicted and historical trajectories.

Results show the effectiveness and efficiency of this approach, and show that GAIL can be effective (in terms of accuracy of predictions) even with a small number of historical trajectories, able to provide accurate long-term predictions, compared to state of the art trajectory prediction approaches.

Step 2: Develop an AI/ML method incorporating reinforcement learning for the detection and resolution of conflicts.

This step models the Air Traffic Controllers' (ATCO) behavior in resolving conflicts using data-driven AI/ML techniques. In general, according to the problem specifications made in this study, this implies learning "*when*" the ATCO will react to resolve a detected conflict, and "*how*" he/she will react. Timely reactions, focus on "*whether*" and "*when*" do reactions happen as the trajectory evolves, aiming to predict the trajectory points that the ATCO issues a conflict resolution action.

More specifically, towards this goal, this study proposes a two-stages data-driven methodology towards meeting the following two objectives:

1. Formulate the ATCO reaction prediction problem, towards building a model of ATCO reactions for resolving conflicts. The aim is to answer "*whether*" and "*when*" the ATCO decides to apply an action to resolve a conflict. Towards predicting the ATCO timely reactions to resolve conflicts, this study trains a Variational Autoencoder (VAE) imitating the demonstrated ATCO policy in a supervised way. The proposed method has been evaluated in two different operational settings (sector-related and sector-ignorant), reporting on the precision, recall and f1-score of predictions. A weighted version of these measures is introduced, to deal with the inherent uncertainties regarding (a) the evolution of trajectories, (b) the detection of conflicts (which are not specified in the dataset), and (c) the ATCO reaction.

2. Formulate the ATCO policy modelling problem, towards building a model of ATCO behavior for resolving conflicts. The aim is to answer "*how*" the ATCO reacts (i.e. what resolution actions he/she applies) in the presence of conflicts. Towards predicting the ATCO policy, thus, predicting the resolution action the ATCO prescribes in case that he/she reacts in a potential detected conflict, this study evaluates comparatively (a) an imitation learning method based on the Generative Adversarial Imitation Learning (GAIL) framework, and (b) classification methods using neural networks (NN), random forest (RF), gradient boosting (GB), and support vector machines (SVM). To evaluate the different methods we report the precision, recall, f1-score and the Matthews Correlation Coefficient between the predictions and the resolution actions of the dataset.

Step 3: Develop an AI/ML method for planning conflict-free trajectories.

This step combines the models for ATCO behavior for resolving conflicts with models of predicting trajectories, into a single method for planning conflict-free trajectories. To do so, it proposes an integrated framework that incorporates models trained for predicting ATCO reactions, models of ATCO policy, and trajectory models learnt from GAIL without considering conflicts.
Results from experiments on trajectories among different OD pairs in Europe, aim to show the efficiency of the overall approach and its effectiveness in terms of accuracy of predictions. We do so by comparing results from the overall conflicts-free trajectory prediction method to the results from training GAIL without considering conflicts and the application of resolution actions.

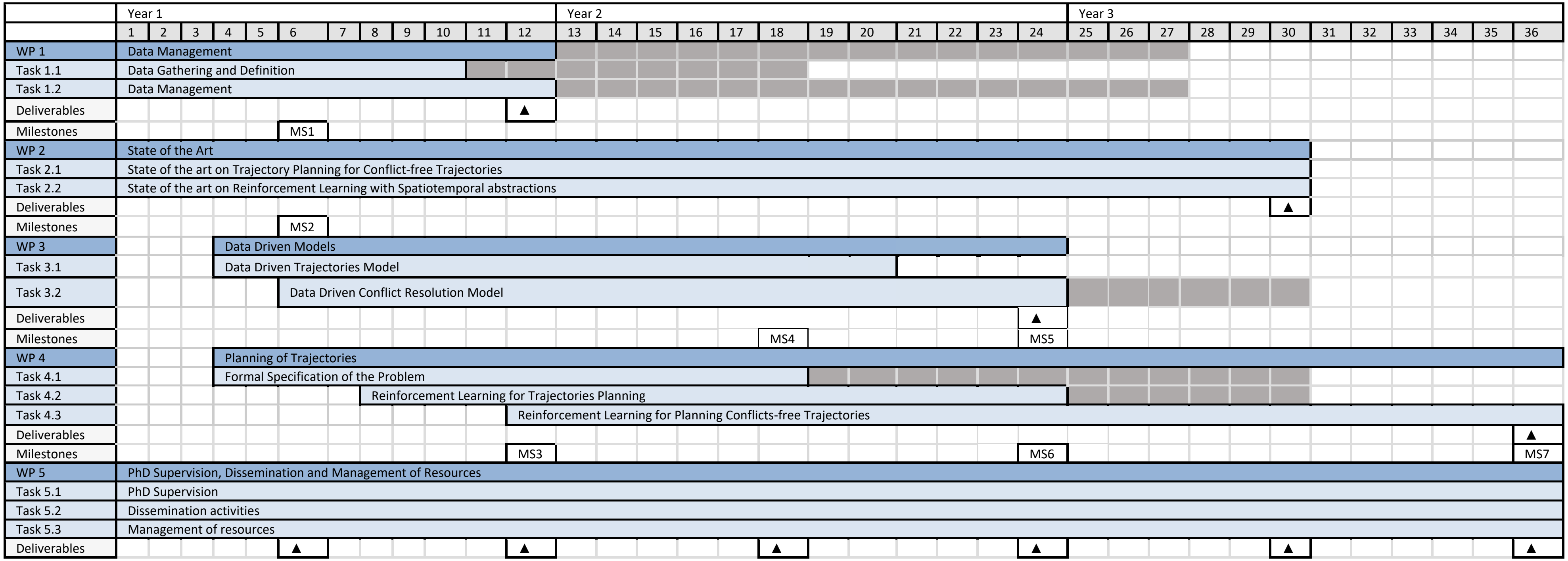


Figure 1 Gantt chart: Original plan (in blue) and revisions made (in gray).

6. Description of the data the study relies on

The methodology proposed in this study exploits data from the Spanish airspace, considering flights over Spain, without sacrificing the generality of the methods introduced. The data sources comprise:

- **Surveillance data:** operational quality data with actual flights (raw) trajectories (Spanish ATC Platform SACTA)
- **Flight plan data:** all flight plan updates for any given flight, since flight plan creation, allowing continuous snapshots (Spanish ATC Platform SACTA)
- **Sector configuration data:** the schedule of deployed sector configurations, as well as the catalog of possible sector configurations (Spanish ATC Platform SACTA)
- **Weather data:** weather forecast information regarding the area corresponding to the trajectories considered (provided by the National Oceanic and Atmospheric Administration (NOAA) platform)
- **Aircraft identification data:** provides specific information on the aircraft flying a particular trajectory. (World Aircraft Database and ICAO Doc8643[2])
- **ATCO events:** provides actions taken by the Air Traffic Controllers in order to ensure safety of flights (provided by Automated NORVASE Takes (ATON)).

In the following sections we describe the datasets in more detail and also their spatial and temporal coverage.

Surveillance data

This data set provides radar tracks of the Spanish airspace controlled by the Spanish ATC provider ENAIRE. A radar track is reported in tabular form, with a timestamp key and geospatial information. Tracks are updated with an interval of 5 seconds. The spatial area coverage of the data is the whole Spanish airspace. The temporal coverage of the data includes the years 2016, 2017, 2018. For this study we have used radar tracks over the Iberian Peninsula for the year 2017. The AI/ML methods have been trained using trajectories between 5 different origin destination pairs Malaga (LEMG) - Gatwick (EGKK), Malaga (LEMG) - Amsterdam (EHAM), Lisbon (LPPT) - Paris (LFPO), Zurich (LSZH) - Lisbon (LPPT) and Geneva (LSGG) - Lisbon (LPPT).
In addition, we consider only trajectories that have at least one ATCO resolution action corresponding to a detected conflict. This results to 668 trajectories from 2017.

In addition to these datasets, for the purposes of evaluating trajectory imitation methods we exploited (a) radar tracks between 3 OD pairs: Barcelona to Madrid (BCN-MAD) during July 2019 (308 trajectories), London Heathrow to Rome Fiumicino (LHR-FCO) during July 2019 (219 trajectories), and Helsinki to Lisbon (HEL-LIS) during July 2019 (44 trajectories).

Flight Plan data

The Flight Plan data set is essential for the aviation domain, as it contains information that triggers a lot of operational decisions, both in the planning and execution phases. The data source that provides the data is a subsystem of the Spanish ATC platform (GIPV, Flight Plan Information Management System). The GIPV is a Flight Plan Report Manager Subsystem that

[2] See also https://www.icao.int/publications/DOC8643/Pages/default.aspx

contains information about flight plans that are being flown or going to be flown soon (to 15 hours), in the part of the airspace that is being operated under the responsibility of the Flight Plan Central Treatment. The spatial area coverage of the data is the whole Spanish airspace. The temporal coverage of the data includes the years 2016, 2017, 2018.
For this study we have used radar tracks over the Iberian Peninsula for the year 2017.

Sector Configuration Data

The Airspace data set describes the existing airspace organization, with no gaps or overlaps, and all the possible ways of combining volumes to generate different operational sector configurations. This data set describes the schedule of sector configurations that have been effectively put in place in Spanish airspace. The temporal coverage of the data includes the years 2016, 2017, 2018.
For this study we have used radar tracks over the Iberian Peninsula for 2017.

Weather Data

This data set provides the forecast of the weather conditions, at the position of an aircraft at any given time during its flight. Specifically, for each 4D position (latitude, longitude, altitude and time) it reports the values of the weather variables describing the weather conditions at that position. The most frequently used variables in the aviation domain, are the Temperature, the Pressure, and the two horizontal components of the Wind Speed, u and v. The available data cover the Iberian Peninsula and Canary Islands for the whole 2016 and July 2019.

For the purposes of evaluating trajectory imitation methods we have used weather data obtained from National Oceanic and Atmospheric Administration (NOAA) for 2019.

Aircraft Identification and Models

For the identification of aircraft reported in surveillance data set, the World Aircraft Database is exploited[3]. This data set provides specific information on the aircraft flying a particular trajectory (thus enriching the information available in the surveillance and flight plans data sets).

ATCO Events Dataset

As ATCO events we consider regulations assigned by the air traffic controllers to flights, in order to ensure that the minimum separation minima are not violated, and thus, aircraft fly safely. An ATCO event contains information about the callsign of the regulated flight, the origin airport, the destination airport, the timestamp of the event, the type of the event and the sector in which the event took place. This dataset is in .csv format and contains regulations assigned by the Air Traffic Controllers to flights that pass over the Spanish FIR. It contains several types of events made by the controller from which we consider as relevant to the conflict resolution problem the following:

- Flight level clearance due to traffic
- Speed adjustment due to traffic
- Direct to waypoint clearance due to traffic

[3] Sun, J. (2017). World Aircraft Database [Data file]. Retrieved from http://junzis.com/adb

The spatial area coverage of the data is the whole Spanish airspace. The temporal coverage of the data includes the years 2017, 2018.

For this study we exploit ATCO events over the Iberian Peninsula for the year 2017.

7. Computational experiments

This section is divided in three subsections, each for the different major objectives addressed.

7.1. Prediction of trajectories per Origin-Destination pair, without explicitly considering conflicts.

To predict aircraft trajectories via imitation learning, this study uses the Generative Adversarial Imitation Learning (GAIL) imitation learning framework. GAIL employs a generative trajectory model G that models the trajectory-evolution policy of the aircraft and a discriminative classifier D that distinguishes between the distribution of state-action pairs generated by the policy and the demonstrated data. Both the policy and D are represented by function approximators with weights **θ** and **w**, respectively. Following the implementation described in [1], GAIL alternates between an Adam [2] gradient step on **w** to increase the GAIL objective function with respect to D, and a step on **θ** using the Trust Region Policy Optimization (TRPO) algorithm [3] to decrease the objective function with respect to the policy.
The input for G corresponds to the position and temporal variables per aircraft state, and other variables enriching a trajectory state (specified below). D takes as additional input the three action variables.
G has a dense output layer with size equal to the number of action variables, while the output layer of D has one node.

Datasets exploited in experiments for imitating trajectories include (a) radar tracks for flights between 3 OD pairs: Barcelona to Madrid (BCN-MAD) during July 2019 (308 trajectories), London Heathrow to Rome Fiumicino (LHR-FCO) during July 2019 (219 trajectories), and Helsinki to Lisbon (HEL-LIS) during July 2019 (44 trajectories); (b) weather data obtained from National Oceanic and Atmospheric Administration (NOAA); and (c) aircraft models' ids.

Trajectories in these datasets have been pre-processed, cleaned and enriched with five (5) numerical variables corresponding to 4 meteorological features at any trajectory state position and time, provided by NOAA, and the aircraft model of each trajectory. The NOAA features are temperature, geopotential height, u-component of wind, v-component of wind.

The prediction accuracy is measured at the pre-tactical phase (starting from a position in the origin airport) and at the tactical phase (starting from any point en-route), introducing a parameter M in {0,0.2, 0.5, 0.7}. M determines the initial state of the prediction, i.e. the state in the actual trajectory after (M x FlightDuration) minutes, starting from t_0.

Results report on the trajectory prediction accuracy using the following measures: (a) Root Mean Square Error (RMSE) in meters in each of the 3 dimensions, as well as in 3D, (b) Along-Track Error (ATE), (c) Cross-Track Error (CTE), and (d) Vertical deviation (V). ATE and CTE are

computed according to the methodology proposed in [4]. The along track error is measured parallel to the predicted trajectory, while the cross track error is measured perpendicular to the predicted course. V measures the difference in altitude between the predicted and the corresponding test (actual) trajectory.
Finally, results report on the estimated time of arrival (ETA) error, given the predicted ETA and the arrival time of test trajectories. All errors ATE, CTE, V and ETA are signed errors, but results reported use their absolute values in order to report on average scores from multiple experiments, providing a clear indication of the errors.

The RMSE error is computed for each predicted trajectory point after computing its corresponding point in the test trajectory using the Dynamic Time Warping (DTW) method.

Specifically, results report on the average of RMSE for the longitude, latitude and all three dimensions (3D), as well as the average of ATE, CTE, V and ETA reported by 20 independent experiments per experimental case.
The division of the historical trajectories for training and testing purposes is done randomly for each of the individual experiments using 90% of them as expert trajectories and 10% as test trajectories.

7.2. Develop an AI/ML learning method incorporating reinforcement learning for the resolution of conflicts.

7.2.1 Model ATCO reactions

Towards predicting the ATCO's timely reactions to resolve conflicts, this study trains a Variational Autoencoder (VAE) imitating the demonstrated ATCO policy in a supervised way. Modes of behavior (corresponding to answering the question "whether" the ATCO reacts in every trajectory state) are decided by the encoder and exploited by the policy (modelling the actual resolution action the ATCO prescribes), which is represented by the decoder network and prescribes sequences of conflict resolution actions. However, the emphasis here in on building a model of ATCO reactions, rather than a model of ATCO policy (addressed subsequently).

The encoder and decoder networks are trained jointly by exploiting enriched trajectory points, the associated ATCO reaction modes and ATCO resolution actions. The errors regarding the predicted actions propagate backwards from the decoder. The encoder aims to minimize the categorical cross entropy loss between the distribution of modes in the dataset and the distribution predicted by the encoder.

To train the VAE for the continuous low-level actions, the Mean Squared Error (MSE) is minimized, and for the categorical actions the categorical cross entropy between the distribution of actions in the data set and the distribution of the decoder predictions, is minimized.

The proposed method is evaluated in two different types of settings w.r.t. the area of responsibility (AoR) chosen: a) The sector-related and b) the sector-ignorant settings.

In the sector-related case the area of responsibility (AoR) corresponds to a sector crossed by the trajectory of the ownship. Neighboring flights are all flights in AoR with a (predicted) conflict to the ownship. In the sector-ignorant case a flight-centric setting is simulated, ignoring sectors, considering a wide area covering the Iberian Peninsula, and assessing conflicts independently from any specific sector.

The trajectory points for all historical trajectories have been annotated using the modes C_0 ("No conflicts detected, and no resolution action has been applied"), C_1 ("At least one conflict is detected, and a resolution action has been applied"), and C_2 ("At least one conflict is detected but no resolution action has been applied").

The dataset contains trajectories between 5 different origin-destination pairs, all from 2017: Malaga (LEMG) - Gatwick (EGKK), Malaga (LEMG) - Amsterdam (EHAM), Lisbon (LPPT) - Paris (LFPO), Zurich (LSZH) - Lisbon (LPPT) and Geneva (LSGG) - Lisbon (LPPT). We study only ATCO resolution actions issued at the en-route phase of operations and filter out the climb and descent parts of the trajectories. In addition, we consider only trajectories that have at least one ATCO resolution action corresponding to a detected conflict. This results to 255 enriched trajectories corresponding to 344 resolution actions for the sector relevant case and 668 trajectories corresponding to 791 resolution actions for the sector-ignorant case. It must be noted here that the available ATCO events dataset covers the Spanish airspace and thus we consider the points of the trajectories that are in this airspace. However, the proposed method is generic, goes beyond sectors, as we show in the sector-ignorant case, and can be applied in any airspace.

To deal with the inherent uncertainties regarding (a) the evolution of trajectories, (b) the detection of conflicts (which are not specified in the dataset), and (c) the actual ATCO reaction, we have devised an evaluation method for data-driven reaction prediction methods, introducing a weighted variation of precision, recall and f1-score.

7.2.2 Model ATCO decisions on resolution actions

Towards predicting the ATCO policy, thus, predicting the resolution action the ATCO prescribes in case that he/she reacts in a potential detected conflict, this study evaluates comparatively (a) an imitation learning method based on the Generative Adversarial Imitation Learning (GAIL) framework, and (b) classification methods using neural networks (NN), random forest (RF), gradient boosting (GB), and support vector machines (SVM).

While the GAIL objective here is to learn a model of the ATCO policy (i.e. the mapping of the particular conflicts detected to the resolution action types), the imitation learning process is a single-stage prediction: I.e. GAIL learns to imitate the ATCO only for the particular states where conflicts are detected. No particular trajectory (i.e. sequences of states-actions) is generated by the GAIL generator G, although the discriminator takes into account the evolution of the conflicts depending on the resolution action prescribed. In doing so, the method learns to predict the resolution actions so as the trajectory to evolve according to the demonstrated conflict-free trajectories.

Alternatively, the classification methods, learn the ATCO policy in a supervised way, classifying states with conflicts to resolution actions, according to the historical data.

The dataset contains the trajectories between the 5 different origin-destination pairs, all from 2017, considered previously on modelling ATCO reactions: Malaga (LEMG) - Gatwick (EGKK), Malaga (LEMG) - Amsterdam (EHAM), Lisbon (LPPT) - Paris (LFPO), Zurich (LSZH) - Lisbon (LPPT) and Geneva (LSGG) - Lisbon (LPPT). We study only ATCO resolution actions issued at the en-route phase of operations and filter out the climb and descent parts of the trajectories. In addition, we consider only trajectories that have at least one ATCO resolution action corresponding to a detected conflict. We consider the sector ignorant case as this is the most general case. The dataset comprises 635 trajectories corresponding to 722 resolution actions. It must be noted here that the available ATCO events dataset covers the Spanish airspace and thus we consider the points of the trajectories that are in this airspace. However, the proposed method is generic, does not consider specific sectors, and can be applied in any airspace.

7.3. Develop an AI/ML method for planning conflict-free trajectories.

The proposed framework for planning conflicts-free trajectories in a data-driven way is depicted in Figure 2.

This framework incorporates models of ATCO behavior and models for the prediction of trajectories. Currently, this framework is being instantiated by models trained for predicting ATCO reactions, models of ATCO policy, and trajectory models learnt from GAIL without considering conflicts.
In particular, the “Reaction prediction model” predicts the mode of ATCO reactions at every trajectory point, and in any case (i.e. when a conflict is detected, and when no conflict is detected). This is a model trained according to Section 7.2.1, and whose output is used by the “resolution actions prediction model”.
Specifically, this later model takes as input the state including features of the detected conflicts and prescribes the type of resolution action (if any) to be applied. This is a model trained as described in Section 7.2.2. Its output is being used by the trajectory prediction module to predict the aircraft state evolution in 4D, taking into account the potential resolution action instructed.

In so doing, the overall method models trajectories according to historical data, to predict conflicts-free trajectories, while it incorporates data-driven models that predict ATCO behavior, according to historical data.

GAIL, as it has been originally designed, employs a generative trajectory model G that models the trajectory-evolution policy of the aircraft and a discriminative classifier D that distinguishes between the distribution of state-action pairs generated by the policy and the demonstrated data.
Both, the generator G and the discriminator D have been trained to model trajectories according to demonstrated trajectories, as discussed in Section 7.1. However, here G rolls-out trajectories taking into account the resolution action type prescribed. Historical trajectories (expert data) have been annotated with the ATC resolution actions and demonstrate the course of the flight after the application of the prescribed resolution action. The objective of GAIL here is to learn to imitate the flown trajectories w.r.t. the resolution action prescribed.

Any of these three models, can be replaced by others, given that the functional requirements (input/output specifications) are satisfied. For instance, the resolution actions prediction model may be the one constructed via imitation learning in 7.2.2, or via the classification task using a neural network. Both of these options will be tested in the experimental study of the method.

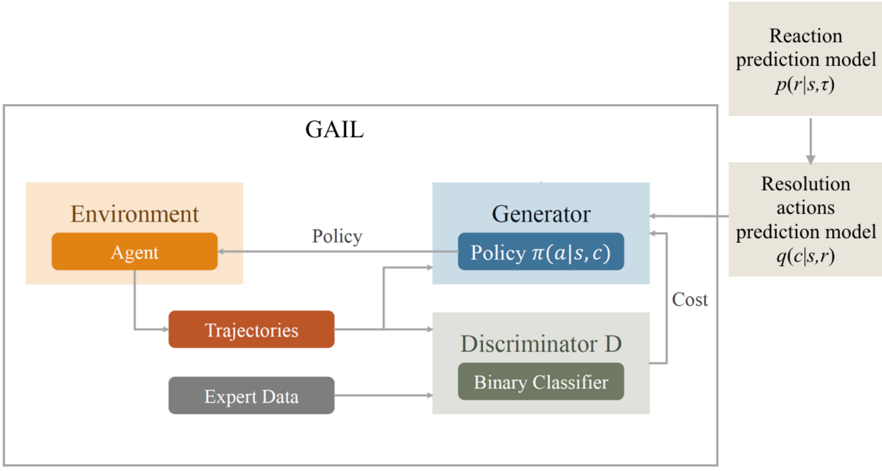


Figure 2 AI/ML framework for planning conflict-free trajectories.

The evaluation dataset contains the trajectories between the 5 different origin-destination pairs, all from 2017, considered previously on modelling ATCO reactions: Malaga (LEMG) - Gatwick (EGKK), Malaga (LEMG) - Amsterdam (EHAM), Lisbon (LPPT) - Paris (LFPO), Zurich (LSZH) - Lisbon (LPPT) and Geneva (LSGG) - Lisbon (LPPT). It includes only trajectories with ATCO resolution actions issued at the en-route phase of operations, and the climb and descent parts of the trajectories have been filtered out. In addition, we consider only trajectories that have at least one ATCO resolution action corresponding to a detected conflict. It must be noted here that the available ATCO events dataset covers the Spanish airspace and thus we consider the points of the trajectories that are in this airspace.
Results report on the trajectory prediction accuracy using the following measures: (a) Root Mean Square Error (RMSE) in meters in each of the 3 dimensions, as well as in 3D, (b) Along-Track Error (ATE), (c) Cross-Track Error (CTE), and (d) Vertical deviation (V). V measures the difference in altitude between the predicted and the corresponding test (actual) trajectory. All measures are computed using the methodology used for evaluating trajectory prediction without considering conflicts (Section 7.1).

Results from the overall conflicts-free trajectory prediction method are compared to the results from training GAIL without considering conflicts and resolution actions prescribed.

8. Results

8.1. Prediction of trajectories per Origin-Destination pair, without explicitly considering conflicts.

Table 1 shows the average RMSE error of the predicted vs the actual (test) trajectory in meters, for each of the three dimensions and in 3D, together with the average absolute ATE, CTE, and VE, in meters. It also reports the average error of the expected arrival time (ETA) in seconds for each case. The table is split to parts corresponding to the different origin destination pairs examined, starting from the short trajectories and going into the longer ones with fewer samples, and for each pair the table reports the results for different values of M (showing the percentage of the trajectory towards the destination airport).

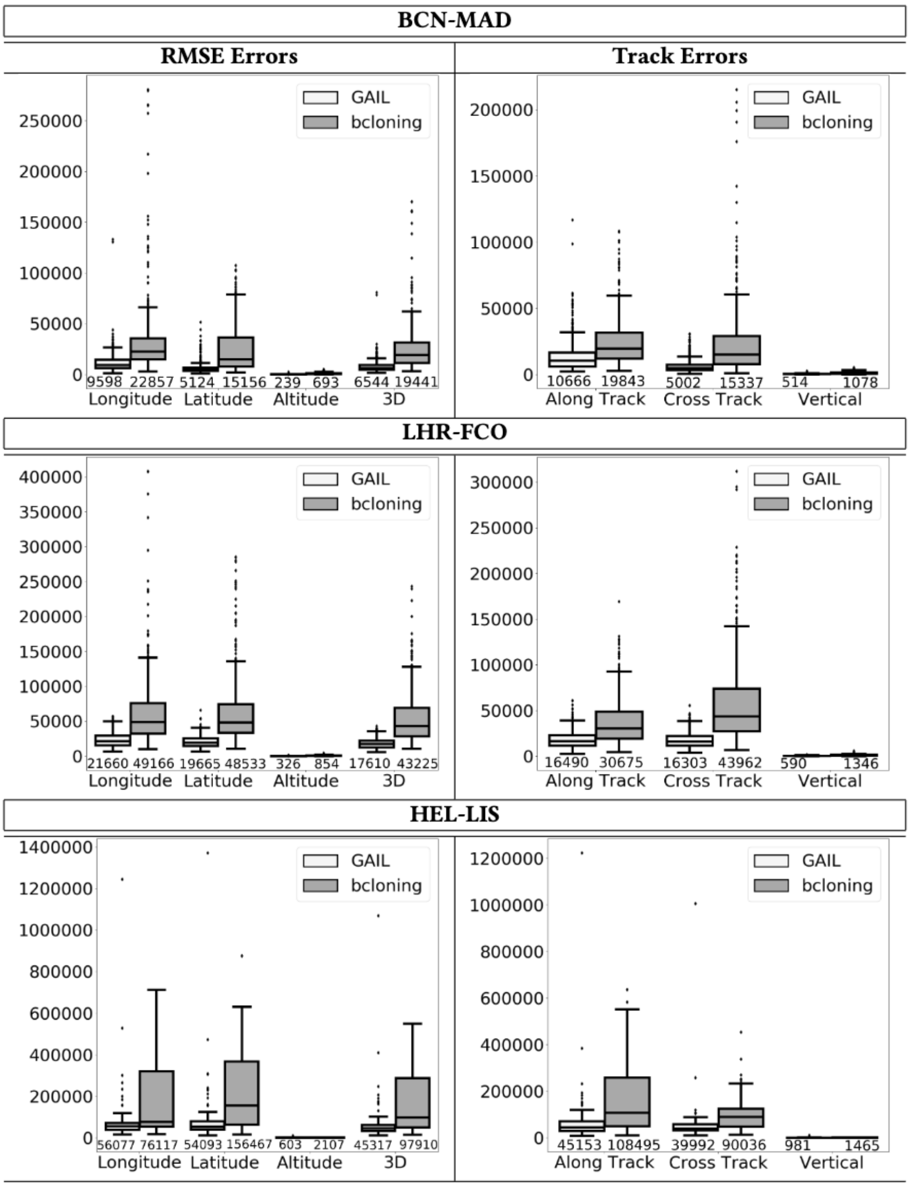
Table 1 Prediction Errors (in meters) and ETA (in seconds).

BCN-MAD									
	M	**Long**	**Lat**	**Alt**	**3D**	**ATE**	**CTE**	**VE**	**ETA**
GAIL	0	11994.99	6214.72	282.63	8005.49	13923.3	6392.61	574.13	263.06
	0.2	10021.46	5549.65	317.87	6774.39	11541.14	6159.43	516.15	259.81
	0.5	8680.78	5103.39	391.5	5977.16	11360.9	6330.23	510.78	239.35
	0.7	6327.37	4078.96	273.41	4519.16	9751.45	5284.18	375.73	158.9
LHR-FCO									
	M	**Long**	**Lat**	**Alt**	**3D**	**ATE**	**CTE**	**VE**	**ETA**
GAIL	0	23371.12	20888.65	372.33	18351.99	18689.42	17874.3	636.2	457.1
	0.2	24427.65	20568.25	359.35	18733.01	19273.79	19362.4	621.78	615.12
	0.5	20274.53	18497.25	370.98	16209.15	15758.31	16399.46	629.06	791.83
	0.7	14313.57	14444.13	539.25	12126.93	13205.18	11294.95	659.35	910.28
HEL-LIS									
	M	**Long**	**Lat**	**Alt**	**3D**	**ATE**	**CTE**	**VE**	**ETA**
GAIL	0	88448.14	95173.02	1096.41	75950.75	77341.27	59731.61	1074.71	801.44
	0.2	91184.7	100957.56	1062.17	79921.51	81309.09	52941.16	1052.04	978.19
	0.5	90334.3	92006.24	1090.32	76575.08	81468.87	49669.47	1252.01	1080.75
	0.7	77587.38	76998.23	1966.88	64771.15	81990.64	46206.48	1691.44	1113.12

Figures in Table 2 show box plots for all the measures. The x axis specifies the error measured. Horizontal lines of each box plot represent the 25th, the 50th, the 75th and the 100th percentile. Dots indicate outliers and the numbers indicate the medians. The left column provides RMSE and the right the along-track and cross-track errors. These box plots correspond to the cases where M=0.

Not surprisingly, the GAIL method provides consistently better results compared to the Behavioral Cloning (BC) baseline: BC minimizes the Mean Square Error between demonstrated actions and the policy actions, over the training set. Indeed, Table 2 shows that GAIL reports smaller errors with narrower deviations, and very small number of outliers compared to BC. In addition to that, low deviations of predicted from the actual trajectories, compared to state of the art methods provide evidence of the imitation learning approach efficacy, even in very long trajectories spanning the European continent and with few training examples.

Table 2 Prediction Errors box plots: Numbers below the boxes indicate the medians.



8.2. Develop an AI/ML learning method incorporating reinforcement learning for the resolution of conflicts.

8.2.1 Model ATCO reactions

Subsequent paragraphs succinctly report on the results achieved by the VAE model for modelling ATCO reactions. These are compared to the results achieved by training only the encoder network of VAE (baseline). This shows the difference in performance between the two methods, caused by the effects of decoder's error backwards propagation in VAE.

To evaluate the VAE and the baseline method we have performed 10 experiments with two times repeated 5-fold cross validation, training the models for 1000 epochs at each experiment.

Results report the 95% confidence interval (CI) of the non-weighted precision, recall and f1-score, in conjunction to the weighted versions of these measures, as proposed by this study.

Specifically, Table 3 reports the 95% confidence interval for the precision, recall and f1-score, achieved by the VAE and the Encoder (Enc) for the ATCO modes of behavior and the resolution actions, for the sector-ignorant case. Columns “modes non-weighted”/ “actions non-weighted" and “modes weighted” / “actions weighted” report respectively on the non-weighted and the weighted versions of the measures for modes and resolution actions. As the encoder does not predict resolution actions, corresponding columns in the second row are empty.

Table 3 Experimental Results of the sector-ignorant case achieved by the VAE and the Encoder (Enc). Columns report the 95% confidence interval of precision, recall and f1-score w.r.t. the modes and the resolution actions of ATCO, for the non-weighted and weighted measures.

model	modes non-weighted	modes weighted	actions non-weighted	actions weighted
	precision	precision	precision	precision
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$	$A_0 : 0.975 \pm 0.004$	$A_0 : 0.998 \pm 0.003$
	$C_1 : 0.976 \pm 0.006$	$C_1 : 0.982 \pm 0.005$	$A_1 : 0.635 \pm 0.028$	$A_1 : 0.640 \pm 0.026$
	$C_2 : 0.934 \pm 0.012$	$C_2 : 0.990 \pm 0.000$	$A_2 : 0.646 \pm 0.035$	$A_2 : 0.653 \pm 0.035$
	recall	recall	recall	recall
VAE	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$	$A_0 : 0.990 \pm 0.000$	$A_0 : 0.993 \pm 0.003$
	$C_1 : 0.936 \pm 0.014$	$C_1 : 0.989 \pm 0.002$	$A_1 : 0.549 \pm 0.026$	$A_1 : 0.589 \pm 0.028$
	$C_2 : 0.976 \pm 0.006$	$C_2 : 0.983 \pm 0.005$	$A_2 : 0.670 \pm 0.022$	$A_2 : 0.709 \pm 0.021$
	f1-score	f1-score	f1-score	f1-score
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$	$A_0 : 0.985 \pm 0.004$	$A_0 : 0.993 \pm 0.003$
	$C_1 : 0.956 \pm 0.008$	$C_1 : 0.985 \pm 0.004$	$A_1 : 0.588 \pm 0.017$	$A_1 : 0.610 \pm 0.017$
	$C_2 : 0.954 \pm 0.008$	$C_2 : 0.986 \pm 0.004$	$A_2 : 0.656 \pm 0.021$	$A_2 : 0.679 \pm 0.023$
	precision	precision		
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$		
	$C_1 : 0.950 \pm 0.010$	$C_1 : 0.959 \pm 0.009$		
	$C_2 : 0.870 \pm 0.038$	$C_2 : 0.975 \pm 0.009$		
	recall	recall		
Enc	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$	-	-
	$C_1 : 0.863 \pm 0.053$	$C_1 : 0.969 \pm 0.017$		
	$C_2 : 0.951 \pm 0.011$	$C_2 : 0.961 \pm 0.011$		
	f1-score	f1-score		
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$		
	$C_1 : 0.904 \pm 0.032$	$C_1 : 0.964 \pm 0.009$		
	$C_2 : 0.909 \pm 0.020$	$C_2 : 0.967 \pm 0.006$		

Results for the prediction of ATCO resolution actions are not so good as those achieved on the prediction of modes (i.e. the prediction of ATCO reactions that this objective aims). The prediction of ATCO resolution actions are further explored in addressing the subsequent PhD study objective (shown in the next subsection).

Similarly to the above, Table 4 reports the 95% confidence interval of the non-weighted and weighted versions of precision, recall and f1-score, achieved by the VAE and the Encoder (Enc) for the ATCO modes and the resolution actions, for the sector-related case. The structure of the table is similar to that of Table 3.

Table 4 Experimental Results of the sector-related case achieved by the VAE and the Encoder (Enc). Columns report the 95% confidence interval of precision, recall and f1-score w.r.t. the modes and the resolution actions of ATCO, for the non-weighted and weighted measures.

model	modes non-weighted	modes weighted	actions non-weighted	actions weighted
VAE	precision	precision	precision	precision
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$	$A_0 : 0.952 \pm 0.012$	$A_0 : 0.993 \pm 0.003$
	$C_1 : 0.919 \pm 0.029$	$C_1 : 0.929 \pm 0.028$	$A_1 : 0.604 \pm 0.087$	$A_1 : 0.611 \pm 0.088$
	$C_2 : 0.791 \pm 0.048$	$C_2 : 0.960 \pm 0.011$	$A_2 : 0.439 \pm 0.099$	$A_2 : 0.446 \pm 0.100$
	recall	recall	recall	recall
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$	$A_0 : 0.981 \pm 0.007$	$A_0 : 0.983 \pm 0.006$
	$C_1 : 0.835 \pm 0.045$	$C_1 : 0.965 \pm 0.012$	$A_1 : 0.566 \pm 0.070$	$A_1 : 0.661 \pm 0.068$
	$C_2 : 0.893 \pm 0.036$	$C_2 : 0.919 \pm 0.033$	$A_2 : 0.362 \pm 0.076$	$A_2 : 0.428 \pm 0.093$
	f1-score	f1-score	f1-score	f1-score
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$	$A_0 : 0.966 \pm 0.005$	$A_0 : 0.986 \pm 0.004$
	$C_1 : 0.873 \pm 0.025$	$C_1 : 0.945 \pm 0.015$	$A_1 : 0.569 \pm 0.039$	$A_1 : 0.620 \pm 0.036$
	$C_2 : 0.835 \pm 0.025$	$C_2 : 0.940 \pm 0.014$	$A_2 : 0.384 \pm 0.075$	$A_2 : 0.419 \pm 0.076$
Enc	precision	precision	-	-
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$		
	$C_1 : 0.863 \pm 0.020$	$C_1 : 0.874 \pm 0.021$		
	$C_2 : 0.740 \pm 0.037$	$C_2 : 0.950 \pm 0.009$		
	recall	recall		
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$		
	$C_1 : 0.805 \pm 0.031$	$C_1 : 0.955 \pm 0.011$		
	$C_2 : 0.809 \pm 0.024$	$C_2 : 0.857 \pm 0.022$		
	f1-score	f1-score		
	$C_0 : 1.000 \pm 0.000$	$C_0 : 1.000 \pm 0.000$		
	$C_1 : 0.833 \pm 0.022$	$C_1 : 0.913 \pm 0.013$		
	$C_2 : 0.774 \pm 0.024$	$C_2 : 0.902 \pm 0.011$		

8.2.2 Model ATCO decisions on resolution actions

To evaluate the effectiveness of the AI/ML methods in predicting the type of the ATCO's resolution actions we report the precision, recall and f1-score for resolution actions types, A_1 (speed change) and A_2 (direct to waypoint) and also the Matthews Correlation Coefficient (mcc). Considering the true and predicted classes as two random variables, the mcc is the correlation coefficient between these random variables with values in [-1, 1], with 1 indicating a perfect prediction, -1 complete disagreement between the true and predicted classes and 0 random prediction.

Table 5 reports the experimental results achieved by the imitation learning algorithm GAIL, exploiting an attention mechanism (GAIL+att) and without attention (GAIL), the neural network classifier with an attention mechanism (NN+att) and without attention (NN), the Random Forests (RF), the Gradient Boost (GB) and the SVM algorithms. Columns report the 95% confidence interval of precision, recall, f1-score and the mcc w.r.t. resolution action types of ATCO. GAIL and NN variants are trained for 1500 mini-batches and epochs respectively.

Table 5 Experimental Results achieved by the one stage GAIL imitation with (GAIL att) and without (GAIL) attention, the neural network classifier with (NN att) and without (NN) attention, the Random Forests (RF), the Gradient Boost (GB) and the SVM algorithms. Columns report the 95\% confidence interval of precision, recall, f1-score and the Matthews correlation coefficient w.r.t. resolution action types of ATCO. GAIL and NN variants are trained for 1500 mini-batches and epochs respectively.

method	dataset	precision		recall		f1-score		mcc
		A_1	A_2	A_1	A_2	A_1	A_2	
GAIL	train	0.73 ± 0.02	0.86 ± 0.04	0.86 ± 0.05	0.73 ± 0.04	0.79 ± 0.02	0.79 ± 0.02	0.60 ± 0.04
	test	0.66 ± 0.10	0.78 ± 0.10	0.78 ± 0.10	0.67 ± 0.07	0.71 ± 0.09	0.72 ± 0.07	0.44 ± 0.16
GAIL att	train	0.81 ± 0.02	0.84 ± 0.02	0.81 ± 0.04	0.84 ± 0.03	0.81 ± 0.02	0.84 ± 0.01	0.65 ± 0.02
	test	0.72 ± 0.10	0.76 ± 0.07	0.72 ± 0.07	0.76 ± 0.08	0.72 ± 0.05	0.75 ± 0.05	0.48 ± 0.09
NN	train	0.82 ± 0.02	0.87 ± 0.03	0.85 ± 0.04	0.85 ± 0.01	0.84 ± 0.02	0.86 ± 0.02	0.70 ± 0.04
	test	0.73 ± 0.12	0.75 ± 0.07	0.70 ± 0.09	0.78 ± 0.09	0.71 ± 0.08	0.76 ± 0.05	0.48 ± 0.13
NN att	train	0.88 ± 0.06	0.92 ± 0.04	0.90 ± 0.06	0.90 ± 0.06	0.89 ± 0.02	0.90 ± 0.02	0.80 ± 0.03
	test	0.74 ± 0.04	0.81 ± 0.04	0.81 ± 0.06	0.74 ± 0.07	0.77 ± 0.02	0.77 ± 0.03	0.55 ± 0.04
RF	train	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
	test	0.74 ± 0.10	0.82 ± 0.06	0.80 ± 0.08	0.77 ± 0.07	$\mathbf{0.77 \pm 0.08}$	$\mathbf{0.79 \pm 0.05}$	$\mathbf{0.56 \pm 0.13}$
GB	train	0.97 ± 0.01	0.98 ± 0.00	0.98 ± 0.00	0.97 ± 0.01	0.97 ± 0.00	0.98 ± 0.01	0.95 ± 0.01
	test	0.75 ± 0.07	0.80 ± 0.05	0.78 ± 0.05	0.78 ± 0.05	0.76 ± 0.03	0.79 ± 0.01	0.55 ± 0.03
SVM	train	0.61 ± 0.03	0.65 ± 0.02	0.55 ± 0.07	0.71 ± 0.05	0.58 ± 0.04	0.68 ± 0.03	0.26 ± 0.04
	test	0.62 ± 0.10	0.65 ± 0.07	0.56 ± 0.11	0.69 ± 0.13	0.58 ± 0.02	0.66 ± 0.05	0.26 ± 0.07

9. Analysis of the results

9.1. Prediction of trajectories per Origin-Destination pair, without explicitly considering conflicts.

Table 1 shows that the proposed method is quite effective to predict the whole trajectory at the pre-tactical stage (M=0). Also, error measures are reduced in almost all cases, while increasing M, i.e. while we select a starting point towards the destination airport, simulating the tactical stage: This happens for instance in the prediction of very long trajectories regarding HEL-LIS. As an exception to that, the average along and cross track errors may increase while increasing M in cases, due to the complexities of the trajectories while approaching the destination airport (i.e. due to holding patterns, manoeuvres, etc.). Thus, it seems that a more refined approach must be used to address the landing part of the trajectory more accurately. This is also the case for the ETA error: If we eliminate the holding patterns while measuring errors for the LHR-FCO pair, we get unsigned ETA errors of 67.82, 61.47, 45.94, 35.22 (signed -14.77, -24.22, -17.51, -8.74) seconds, for M in {0, 0.2, 0.5, 0.7}, respectively. Similar patterns are recorded for the other error measures, providing evidence to our conjecture about the difficulty of predictions in destination airports with complex holding patterns and multiple modes of approach.

9.2. Develop an AI/ML learning method incorporating reinforcement learning for the resolution of conflicts.

9.2.1 Model ATCO reactions

Regarding the modes of ATCO reaction, results reported in Table 3 show that both the VAE and the Encoder networks achieve an f1-score greater to 0.9 on all modes, for the non-weighted and the weighted measures, with VAE achieving the best results with a weighted f1-score greater or equal to 0.985±0.004 on all modes. Also the VAE outperforms the encoder on all measures, weighted or not, although the Encoder is really competitive.

Similarly, for the sector-related case, as shown in Table 4, weighted measures are higher than the non-weighted. This shows that in many cases the model makes false predictions that are penalized lightly by the weighted measures, given that, as it also happens in the sector-ignorant case, they are not critical.

Regarding the predictions of resolution actions, in the sector-related case results are not good: For instance, the f1-score of the A_2 resolution action is 0.384$\pm$0.075 for the non-weighted and 0.419$\pm$0.076 for the weighted measure. As already pointed out, this is further explored in the continuation of the work, as reported in the next subsection.

9.2.2 Model ATCO decisions on resolution actions

As shown in Table 5 the RF method achieves best results on the testing set with mean mcc value of 0.56, mean f1-score 0.77 for resolution action type A1 and mean f1-score 0.79 for resolution type A2. Slightly reduced mean mcc and f1-score is reported by the GB and the NN+att algorithms. The GB algorithm achieves a mcc mean value of 0.55 and mean f1-score 0.76 for resolution action type A1 and 0.75 for resolution type A2. The NN+att achieves a mcc mean value of 0.55 and performs slightly different regarding the f1-score with a value of 0.77 for both A1 and A2 resolution action types.

This said, all three methods are greatly competitive with each other as the difference w.r.t. the mean mcc and f1-score is small (0.01). GB and NN+att report a narrower confidence interval compared to RF, implying smaller standard deviation between independent experiments. When considering the precision and recall measures reported, the differences between these algorithms are very small except for the significantly reduced confidence interval achieved by GB and the NN+att compared to the RF algorithm. The next most competitive method is the GAIL+att algorithm, reporting an mcc value of 0.48 for the testing dataset, mean f1-score 0.72 for resolution action type A1 and mean f1-score 0.75 for resolution action type A2.
The SVM algorithm performs significantly worse compared to all other methods, achieving a mcc value 0.26 and f1-scores 0.58 and 0.66 for resolution action types A1 and A2, respectively.

Considering the capacity of the models to learn, we observe that all methods, except the SVM, achieve a strong positive correlation between true and predicted A1 and A2 resolution action types on the training set with mcc values ranging from 0.6 (GAIL) to 1 (RF). F1 scores are also high for all methods, except for the SVM, with values in the interval [0.79, 0.98]. RF and GB achieve mean mcc values 1 and 0.95, respectively, significantly outperforming the GAIL and NN variants on the training set. GAIL+att and NN+att methods achieve mean mcc 0.65 and 0.8 on the training set, respectively.
The SVM performs poorly on the training set achieving 0.26 mean MCC value and mean f1-score 0.58 and 0.66 for the A1 and A2 resolution action types.

Finally, although classification methods outperform the GAIL method implementing the single-stage imitation learning task formulated, it must be noted that during learning, GAIL learns the trajectory evolution w.r.t. to the resolution action that it applies. This has the following benefits: a) GAIL can incorporate further trajectory optimization objectives in a straightforward way by augmenting its reward function with other terms i.e. added nautical miles, fuel consumed, CO_2 emissions, etc., b) models learned by GAIL can be exploited by reinforcement learning methods that aim to solve conflicts and evolve the trajectories according to demonstrated maneuvers. These are issues to be investigated in the future, also in comparison to the models learned by other methods, Specifically, the latter benefit will be studied while experimenting with the conflict-free trajectory planning framework proposed in this study.

10. Conclusions and look ahead

The concrete objectives of this study are as follows:

A. Develop a formulation of the trajectory planning problem and of the constituent subproblems, considering prediction of trajectories and resolution of conflicts.
B. Develop an AI/ML learning method incorporating reinforcement learning for the prediction of trajectories per Origin-Destination pair, without explicitly considering conflicts.
C. Develop AI/ML learning methods incorporating reinforcement learning for the resolution of conflicts.
D. Combine AI/ML models for the intertwined prediction of trajectories and the detection and resolution of conflicts, towards a method for the planning of conflicts-free trajectories.

As already reported above, these objectives have been achieved to a large extent with innovative problem formulations and AI/ML methods, which have been evaluated to real-world cases.
This study still progresses on performing an experimental study towards achieving objective D. The emphasis in the upcoming months will be on this last objective.

Therefore, no deviation is expected to what has been committed to within the PhD contract.

All objectives are expected to have been achieved until the end of 2022, where the PhD study is expected to be awarded.
Towards the completion of this PhD we also aim to investigate the role of flight plans on detecting conflicts, similarly to how ATCO use them: Initial results towards this were not encouraging, given the deviations of flown trajectories from flight plans and the uncertainty on the evolution of trajectories.

As a follow-up of this work, interesting paths to follow are as follows:

(a) Investigate the gathering of data sets that provide the maximum possible information towards alleviating the limitations of the data sets exploited in this study. These new data sets may come from simulated settings, however with valid features regarding ATCO observations and actions. Then the proposed methods can be further evaluated using this high-quality data recording important ATCO observations, without making specific assumptions on the detection of conflicts and the evolution of trajectories, as done in this study. These results will further show the potential of the models to be used in operational settings.
(b) Investigate the combination of data-driven models learnt from the methods devised in this study, with active reinforcement learning techniques for the resolution of conflicts: This combination may result into methods that balance between ATCO reactions, as demonstrated in datasets, and active methods assessments on the reactions necessary to resolve conflicts, as learnt from the interaction with the operational setting.

(c) Investigate the effect of the models learnt to the explainability / transparency of the AI/ML methods exploiting them, as well as their effect to the trustworthiness on the automation using them.
(d) Study an end-to-end method aiming to learn the evolution of trajectories, together with maneuvers for resolving conflicts: As said, this method must work in two state-action spaces: One regarding the evolution of the trajectory states themselves, and the other on the evolution of the conflicts. This is a very challenging problem.

11. References

11.1 Link to PhD thesis / repository

The PhD dissertation will be published in Dione (https://dione.lib.unipi.gr/xmlui/discover?filtertype=type&filter_relational_operator=equals&filter=Doctoral+Thesis). Dione contains the intellectual production of the University of Piraeus, including undergraduate and graduate dissertations and the doctoral theses that have been elaborated in the University of Piraeus.

It will also be published in the repository of the National Documentation Centre (EKT) https://www.ekt.gr/en/library/didaktorika.

11.2 Associated outputs and publications

1. *Presentation in TC2 "Data-driven Trajectory Prediction" Workshop. https://drive.google.com/drive/u/0/folders/1cotqSwfNXbTn-zqEqNxaAhKyL0YHT67K*
2. Alevizos Bastas and Theocharis Kravaris and George A. Vouros, "Data Driven Aircraft Trajectory Prediction with Deep Imitation Learning", arXiv , cs.LG , 2005.07960, 2020, https://arxiv.org/abs/2005.07960 (in synergy with Alevizos Bastas' Engage KTN PhD project)
3. *Presentation in Engage Summer School 2020 virtual event. https://drive.google.com/drive/u/0/folders/1cotqSwfNXbTn-zqEqNxaAhKyL0YHT67K*
4. *Presentation at the Engage Thematic Challenge 2 Workshop on Automation, AI and ML, held as a virtual event at Sep 3, 2021. https://drive.google.com/drive/u/0/folders/1cotqSwfNXbTn-zqEqNxaAhKyL0YHT67K*
5. *The article "Data-driven prediction of Air Traffic Controllers reactions to resolving trajectory conflicts" has been submitted to a top scientific journal (a preprint also in Arxiv http://arxiv.org/abs/2205.09539).*
6. *The article "Data Driven Modeling of Air Traffic Controllers Policy to Resolve Conflicts" will be submitted to a top scientific journal.*

11.3 References cited in this report

[1] J. Ho and S. Ermon, "Generative adversarial imitation learning," in *Advances in neural information processing systems*, 2016.
[2] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *CoRR*, 2014.

[3] J. Schulman, S. Levine, P. Abbeel, M. Jordan and P. Moritz, "Trust region policy optimization," in *International conference on machine learning*, 2015.

[4] C. Gong and D. McNally, "A methodology for automated trajectory prediction analysis," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2004.

Annex I: Acronyms

Term	**Definition**
ATM	Air Traffic Management
AI	Artificial Intelligence
ATC	Air Traffic Control
ATCO	Air Traffic Controller
ANSPs	Air Navigation Service Providers
AoR	Area of Responsibility
AOs	Airport Operators
ATE	Along-Track Error
AUs	Airspace Users
BT	business trajectory
BC	Behavioral Cloning
CD&R	Conflict Detection and Resolution
CTE	Cross-Track Error
CI	Confidence Interval
DTW	Dynamic Time Warping
ETA	Estimated Time of Arrival
FIR	Flight Information Region
GAIL	Generative Adversarial Imitation Learning
GB	Gradient Boost
ML	Machine Learning
MCC	Matthews Correlation Coefficient
MSE	Mean Squared Error
NM	Network Manager
NOAA	National Oceanic and Atmospheric Administration
NN	Neural Network
OD	Origin-Destination
RMSE	Root Mean Square Error
RF	Random Forest

Term	Definition
RBT	Reference Business Trajectory
SBT	Shared Business Trajectory
SVM	Support Vector Machines
TBOs	Trajectory-Based Operations
VAE	Variational Auto-Encoder
V	Vertical deviation
WP	Work Package