



# Effect of Trajectory Prediction Uncertainty on a Probabilistic De- bunching Concept for Inbound Air Traffic

Final Report

Edzer Oosterhof



# Effect of Trajectory Prediction Uncertainty on a Probabilistic Debunching Concept for Inbound Air Traffic Final Report

by

Edzer Oosterhof

to obtain the degree of Master of Science at the Delft University of Technology,  
to be defended publicly on Friday March 18, 2022

Student number: 4387821  
Project duration: November 9, 2020 – March 18, 2022  
Thesis committee: Prof. dr. ir. J.M. Hoekstra, TU Delft, chair, supervisor  
F. Dijkstra, KDC Mainport Schiphol, daily supervisor  
P.C. Roling, Tu Delft, examiner

Front cover Image source <https://airviation.tumblr.com/post/76860264964>. accessed 01-07-2021



# Preface

Before you lies the Final Thesis that concludes my time at the Aerospace Engineering faculty of the TU Delft. It presents the work conducted during a full-time year of research. The research was conducted in cooperation with the Knowledge and Development Centre Mainport Schiphol, with supervision provided by Air Traffic Control the Netherlands (LVNL) and TU Delft. Specifically, I would like to thank Jacco Hoekstra (TU Delft) and Ferdinand Dijkstra (LVNL) for their help and guidance during this project, and Jonah Bekkers for his remarks from the Air Traffic Controller perspective. I would also like to thank my friends with who I have studied during my years at TU Delft, my roommates with who I shared a study spot on the kitchen table while listening to FIP Radio, and my family and girlfriend for their support.

Enjoy!

*Edzer Oosterhof  
Delft, March 2022*



# Contents

List of Figures	vii
List of Tables	ix
1 Introduction	1
1.1 Research Motivation . . . . .	1
1.2 Research Questions . . . . .	2
1.3 Project Structure . . . . .	2
I Literature Review and Preliminary Findings	5
2 Airspace Structure	7
2.1 Airspace Layout . . . . .	7
2.1.1 Upper Airspace . . . . .	8
2.1.2 Control Area . . . . .	8
2.1.3 Terminal Manoeuvring Area . . . . .	8
3 Arrival of Air Traffic	11
3.1 Arrival Manager . . . . .	11
3.2 Freeze Horizon . . . . .	12
3.3 Extended Arrival Manager. . . . .	13
3.4 Linear Delay Absorption Techniques . . . . .	14
4 Air Traffic Arrivals Flow	17
4.1 Air Traffic Flow and Capacity . . . . .	17
4.2 Air Traffic Flow Management . . . . .	18
4.3 Air Traffic Bunching. . . . .	19
4.3.1 Technical Definition for Bunching . . . . .	19
4.4 North Atlantic Traffic . . . . .	21
5 Air Traffic Information Sources	23
5.1 Network Manager . . . . .	23
5.2 Flight Update Messages . . . . .	23
5.3 Enhanced Tactical Flow Management System Flight Data Messages . . . . .	24
5.4 Airport Collaborative Decision Making . . . . .	25
5.5 On-Line Data Interchange . . . . .	26
5.6 Radar . . . . .	27
6 Trajectory Prediction	29
6.1 State Propagation . . . . .	29
6.2 Mathematical Model . . . . .	30
6.3 Trajectory Predictor Uncertainty . . . . .	31
6.3.1 Trajectory Predictor Sensitivity Analysis . . . . .	32
7 Summary of the Preliminary Study	35
II Thesis After Midterm	37
8 Uncertainty Analysis	39
8.1 Uncertainty Analysis . . . . .	39
8.1.1 Data Availability . . . . .	40
8.1.2 Truth Data filtering. . . . .	40
8.1.3 Feature Selection. . . . .	43

8.1.4	Data Analysis . . . . .	44
8.1.5	Curve Fitting . . . . .	49
8.1.6	Curve Fitting Analysis . . . . .	49
8.1.7	Validation . . . . .	51
9	Methodology . . . . .	53
9.1	Introduction . . . . .	53
9.2	Message Class Module . . . . .	53
9.2.1	Input Data . . . . .	54
9.2.2	Algorithmic Steps . . . . .	54
9.2.3	Maximum Delay Absorption in Descent . . . . .	55
9.2.4	Cruise . . . . .	58
9.3	Bunching Detection Module . . . . .	59
9.4	Solution Module . . . . .	61
9.4.1	Genetic Algorithm . . . . .	64
9.4.2	Genetic Algorithm Parameter Selection . . . . .	67
9.5	Verification and Validation of the Bunching Solver . . . . .	68
9.6	AMAN module . . . . .	72
9.6.1	Data input and output . . . . .	72
9.6.2	Algorithmic Steps . . . . .	73
9.6.3	AMAN Verification and Validation . . . . .	74
10	Experimental Setup . . . . .	75
10.1	Experiment 1: Bunching Detection . . . . .	75
10.2	Experiment 2: Debunching . . . . .	76
10.3	Experiment 3: Arrival Manager . . . . .	77
11	Results and Discussion . . . . .	79
11.1	Experiment 1: Bunching Detection . . . . .	79
11.2	Experiment 2: Debunching . . . . .	83
11.3	Experiment 3: Arrival Manager . . . . .	85
12	Conclusion . . . . .	89
13	Operational Potential and Recommendations . . . . .	91
A	Appendix: Uncertainty Analysis . . . . .	93
B	Appendix: Methodology . . . . .	101
C	Appendix: Results . . . . .	105
	Bibliography . . . . .	109

# List of Figures

1.1	Swim lane diagram visually presenting the methodology of this research . . . . .	3
2.1	A schematic breakdown of the airspace structure ( <i>source: Ir Th.J Mulder, 2021</i> ). . . . .	7
2.2	Sector division of the FIRAM without reduced coordination ( <i>source: LVNL, 2021</i> ) . . . . .	9
2.3	Sector division of the FIRAM with fully reduced coordination ( <i>source: LVNL, 2021</i> ) . . . . .	9
2.4	Standard Arrival Routes in the FIRAM ( <i>source: LVNL, 2021</i> ) . . . . .	9
3.1	System architecture of an Arrival Manager . . . . .	13
3.2	Impact of speed on the operating costs ( <i>source: [51]</i> ) . . . . .	15
3.3	Specific Range as a function of the cruise speed ( <i>source: [13]</i> ) . . . . .	15
4.1	Arrival sector division for arrivals in 2019 . . . . .	18
4.2	Airspace capacity and 20 minute moving average of actual arrival times at Schiphol Airport on August 3, 2019 . . . . .	19
4.3	Flying time from COP to IAF for flights arriving at Sector 4 and 5 of Schiphol Airport in September 2019 . . . . .	21
4.4	20 minute moving average of North Atlantic traffic actual arrival times at Schiphol Airport on August 3, 2019 . . . . .	22
4.5	Arrival sector division for arrivals from the North Atlantic in 2019 . . . . .	22
5.1	Distribution of EFD B2B message count per arriving flight in September 2020 . . . . .	24
5.2	A schematic overview of the times at which each DPI is available. <i>Source: [24]</i> . . . . .	26
5.3	EFD B2B message count per arriving flight in September 2020, grouped by airport type . . . . .	26
5.4	SSR Mode S coverage at FL245 ( <i>source: [38]</i> ) . . . . .	28
6.1	Visual representation of nominal (a), worst-case (b), probabilistic (c), and intent based state projection (d) ( <i>sources: [47] [34]</i> ) . . . . .	30
6.2	A schematic representation of five different approaches for modelling of aircraft motion ( <i>source: [3]</i> ) . . . . .	31
8.1	Systematic overview of the data processing steps . . . . .	39
8.2	Closest distance from aircraft to COP for aircraft in September 2019 . . . . .	41
8.3	Flight Level at the COP for aircraft in September 2019 . . . . .	42
8.4	Ground Speed at the COP for aircraft in September 2019 . . . . .	42
8.5	Accuracy of the COP arrival prediction error (Estimated Time Over - Actual Time Over) versus time to fly . . . . .	44
8.6	The total number of messages sent per flight state for a given time to predicted arrival at the COP . . . . .	45
8.7	The share of messages sent with a given flight state for a given time to predicted arrival at the COP . . . . .	46
8.8	Accuracy of the COP arrival predictions (ETO - ATO) versus time to fly for the Air Traffic Control Activated flight state . . . . .	47
8.9	Accuracy of the COP arrival predictions (ETO - ATO) versus time to fly for the Flight Plan Filed flight state, for CDM and non-CDM equipped messages . . . . .	48
8.10	Accuracy of the COP arrival predictions (ETO - ATO) versus time to fly for the Tactical Activated flight state, for CDM and non-CDM equipped messages . . . . .	48
8.11	Probability Density Curves for the AA, TA and SI flight state up to 85 minutes prediction horizon . . . . .	50
9.1	System architecture of the main module . . . . .	53
9.2	System architecture of the message class module . . . . .	56

9.3	Typical descent speed envelope with examples of CAS acceleration and deceleration manoeuvres ( <i>source</i> : [10]) . . . . .	57
9.4	System architecture of the bunching detection module . . . . .	60
9.5	Probability Density Curves for aircraft expected to arrive during a one-runway landing configuration . . . . .	60
9.6	Probability Density Curves for aircraft flying over the ARTIP stack, expected to arrive during a two-runway landing configuration . . . . .	61
9.7	Probability Density Curves for aircraft flying over the RIVER and SUGOL stack, expected to arrive during a two-runway landing configuration . . . . .	61
9.8	The bunching prediction timelines for the runway configurations, combined in one plot . . . . .	62
9.9	An example of Crossover in a GA . . . . .	66
9.10	An example of mutation in a GA . . . . .	66
9.11	System architecture of the debunching module . . . . .	66
9.12	Sample problem 1: day 1, 35% threshold . . . . .	68
9.13	The effect of the bunching probability requirement on the Genetic Algorithm Convergence . . . . .	69
9.14	The Probability Density Curves for all flights on the prediction horizon at 2021-05-01 02:16:27 . . . . .	69
9.15	The Probability Density Curves for all flights, for the solved prediction horizon . . . . .	70
9.16	The bunching detection timeline before and after solving . . . . .	70
9.17	System architecture of the AMAN module . . . . .	74
11.1	Average bunching probability at each prediction horizon for 7 consecutive days for a single runway configuration . . . . .	80
11.2	Average bunching probability at each prediction horizon for 7 consecutive days for a double runway configuration . . . . .	80
11.3	Maximum bunching probability that was reached at each prediction horizon for 7 consecutive days for a single runway configuration . . . . .	81
11.4	Maximum bunching probability that was reached at each prediction horizon for 7 consecutive days for a double runway configuration . . . . .	81
11.5	Maximum bunching probability at each point in time for 7 consecutive days for a single runway configuration . . . . .	82
11.6	Maximum bunching probability at each point in time for 7 consecutive days for a double runway configuration . . . . .	82
11.7	Maximum bunching probability at each point in time for 7 consecutive days for a double runway configuration, for prediction horizons of 18 and 40 minutes . . . . .	83
11.8	The probability density curves for two test cases illustrating the effectiveness of debunching with different uncertainty . . . . .	88
A.1	Accuracy of the COP arrival predictions versus time to fly for the Flight Plan Filed flight state . . . . .	93
A.2	Accuracy of the COP arrival predictions versus time to fly for the Tactical Activated flight state . . . . .	95
A.3	Accuracy of the COP arrival predictions versus time to fly for the Filed Slot Allocated flight state . . . . .	95
A.4	Accuracy of the COP arrival predictions versus time to fly for the Slot Issued flight state . . . . .	96
A.5	Accuracy of the COP arrival predictions versus time to fly for the Suspended flight state . . . . .	96
A.6	Accuracy of the COP arrival predictions versus time to fly for the Filed Slot Allocated flight state, for CDM and non-CDM equipped messages . . . . .	97
A.7	Accuracy of the COP arrival predictions versus time to fly for the Slot Issued flight state, for CDM and non-CDM equipped messages . . . . .	97
A.8	Probability Density Curves for the AA, TA and FI flight state up to 165 minutes prediction horizon . . . . .	98
A.9	Probability Density Curves for the AA, TA and FI flight state up to 245 minutes prediction horizon . . . . .	98
A.10	Probability Density Curves for the FL, sI and SU flight state up to 85 minutes prediction horizon . . . . .	99
A.11	Probability Density Curves for the FL, sI and SU flight state up to 165 minutes prediction horizon . . . . .	99
A.12	Probability Density Curves for the FL, sI and SU flight state up to 245 minutes prediction horizon . . . . .	100
B.1	Sample problem 2: day 1, 60% threshold . . . . .	102
B.2	Sample problem 3: day 4, 30% threshold . . . . .	102
B.3	Sample problem 4: day 4, 55% threshold . . . . .	103
C.1	The AMAN throughput per 20 minutes for each bunching threshold at day 5 . . . . .	108

# List of Tables

3.1	Holding count and percentage per Initial Approach Fix, for arrivals in 2019 . . . . .	16
3.2	Average and median holding time per Initial Approach Fix, for arrivals in 2019 . . . . .	16
4.1	Capacity Declaration for arrivals at AAS . . . . .	17
4.2	Capacity Declaration for Departures at AAS . . . . .	18
4.3	RECAT-EU WTC separation minima on approach [15] . . . . .	20
4.4	Entry COPs for Sector 4 and 5 in 2019 (only showing COPs larger than .2 %) . . . . .	21
5.1	The percentage of all messages and all flights originating from the three airport types, for arrivals in September 2020 . . . . .	26
8.1	Distribution characteristics for the flying time between COP and IAF, the corresponding trajectory predictor error requirement and the corresponding error requirement in time for aircraft in September 2019 . . . . .	42
8.2	The percentage of specific CDM state messages for a given flight state, for arrivals from August 1 <sup>st</sup> (2020) to June 24 <sup>th</sup> (2021) . . . . .	47
9.1	An example of the input table providing nominal airspeeds for Cruise and Final Approach . . . . .	54
9.2	Results of running the bunching solver 10 times for the same scenario . . . . .	71
9.3	Solutions found by the Grid Search for the validation scenario . . . . .	71
9.4	Iterations of the GA for the validation scenario . . . . .	72
9.5	Which IAF feeds which runway in a two runway landing configuration . . . . .	73
9.6	Example of data set . . . . .	73
9.7	Example of the input data set for the AMAN containing the runway configurations . . . . .	73
9.8	A snippet of the frozen planning of the AMAN Module . . . . .	74
11.1	Output table for day 1, with a single runway configuration . . . . .	84
11.2	Output table for day 2, with a single runway configuration . . . . .	84
11.3	Output table for day 3, with a single runway configuration . . . . .	84
11.4	Output table for day 4, with a single runway configuration . . . . .	84
11.5	Output table for day 5, with a single runway configuration . . . . .	84
11.6	Output table for day 1, with a single runway configuration, with the 18 minute solver horizon . . . . .	85
11.7	Output table for day 2, with a single runway configuration, with the 18 minute solver horizon . . . . .	86
11.8	Output table for day 3, with a single runway configuration, with the 18 minute solver horizon . . . . .	86
11.9	Output table for day 4, with a single runway configuration, with the 18 minute solver horizon . . . . .	86
11.10	Output table for day 5, with a single runway configuration, with the 18 minute solver horizon . . . . .	86
A.1	The number of samples and goodness-of-fit statistics for the curve fitting routine . . . . .	94
B.1	An example of the output table for the debunching module . . . . .	104
C.1	Output table for day 1, with a double-runway configuration, fed by the ARTIP stack . . . . .	105
C.2	Output table for day 4, with a double-runway configuration, fed by the ARTIP stack . . . . .	106
C.3	Output table for day 6, with a double-runway configuration, fed by the ARTIP stack . . . . .	106
C.4	Output table for day 7, with a double-runway configuration, fed by the ARTIP stack . . . . .	106
C.5	Output table for day 1, with a double-runway configuration, fed by the RIVER and SUGOL stacks . . . . .	106
C.6	Output table for day 4, with a double-runway configuration, fed by the RIVER and SUGOL stacks . . . . .	106
C.7	Output table for day 6, with a double-runway configuration, fed by the RIVER and SUGOL stacks . . . . .	106
C.8	Output table for day 7, with a double-runway configuration, fed by the RIVER and SUGOL stacks . . . . .	107



# List of Acronyms

**A-CDM** Airport Collaborative Decision Making.

**AA** Air Traffic Control Activated.

**AAS** Amsterdam Schiphol Airport.

**ACID** Aircraft Identification.

**ADES** Airport of Destination.

**AMAN** Arrival Manager.

**ANSP** Navigation Service Provider.

**APLN** Approach Planner.

**ATC-DPI** Air Traffic Control Departure Planning Information.

**ATCO** Air Traffic Controller.

**ATFM** Air Traffic Flow Management.

**B2B** Business-to-business.

**CA** Cancelled.

**CAS** Calibrated Airspeed.

**CDM** Collaborative Decision Making.

**CGS** Corrected Ground Speed.

**CI** Cost Index.

**COP** Co-ordination Point.

**CTA** Control Area.

**CTR** Control Zone.

**DPI** Departure Planning Information.

**E-DPI** Early Departure Planning Information.

**EAT** Expected Approach Time.

**EFD** ETFMS Flight Data Message.

**ELDT** Estimated Landing Time.

**ETA** Estimated time of arrival.

**ETFMS** Enhanced Tactical Flow Management System.

**ETO** Estimated Time Over.

**EXIT** Estimated Taxi-Time.

- FI** Flight Plan Filed.
- FIR** Top of Descent.
- FIR** Flight Information Region.
- FIRAM** Flight Information Region Amsterdam.
- FL** Flight Level.
- FMS** Flight Management System.
- FS** Filed Slot Allocated.
- FUM** Flight Update Message.
- HMI** Human Machine Interface.
- IAF** Initial Approach Fix.
- IAS** Indicated Airspeed.
- IQR** Interquartile Range.
- LIV** Landing Interval.
- LoA** Letter of Agreement.
- LVNL** Air Traffic Control The Netherlands.
- MSL** Mean Sea Level.
- MUAC** Maastricht Upper Area Control Center.
- NAT** North Atlantic Tracks.
- P-DPI** Predicted Departure Planning Information.
- PR** Primary Surveillance Radar.
- SI** Slot Issued.
- SID** Standard Instrument Departure.
- SR** Specific Range.
- SSR** Secondary Surveillance Radar.
- STAR** Standard Arrival Route.
- SU** Suspended.
- T-DPI-s** Target Departure Planning Information - Sequenced.
- T-DPI-t** Target Departure Planning Information - Target.
- TA** Tactical Activated.
- TMA** Terminal Manoeuvring Area.
- TOBT** Target Off-Block Time.
- TOT** Take-Off Time.

**TP** Trajectory Prediction.

**TSAT** Target Start-up Approval Time.

**TTG** Time To Gain.

**TWR** Tower.

**UTA** Upper Airspace.

**UTC** Coordinated Universal Time.

**WTC** Wake Turbulence Category.

**XMAN** Extended Arrival Manager.



# Abstract

The airspace in western Europe is one of the most complex airspaces globally. With the air space already operating near maximum capacity, innovations are needed to increase airspace efficiency. This research aims to design and test a debunching concept for inbound air traffic to increase the efficiency of the arrival traffic stream. A probabilistic debunching concept is created for the Initial Approach Fix using Probability Density Functions fitted on the errors in predicted arrival times provided by the EUROCONTROL Enhanced Tactical Flow Management System Flight Data messages (EFD). A bunching probability is detected when two or more aircraft have a probability of arriving at the Initial Approach fix simultaneously or within the required Wake Category Separation. A debuncher is created using a Constrained Genetic Algorithm that decreases the bunching probability by imposing en-route delay on arriving air traffic. The results show that the bunching probability increases with shorter prediction horizons as the uncertainty in the arrival estimates decreases. It is shown that with the probabilistic debunching method, it is possible to decrease the delay in the Arrival Manager by imposing delay en-route. However, at prediction ranges before 40 minutes before arrival at the IAF, the decrease in the AMAN delay is not consistent, indicating that the trajectory prediction uncertainty at this range is too high. Furthermore, the decrease in the AMAN delay is often lower than the imposed delay by the debuncher, showing that the decrease in AMAN delay comes at the cost of extra delay en-route. It is concluded that when the uncertainty in the trajectory predictions is decreased, the effectiveness of the debuncher increases and the effect on the Arrival Manager is improved, indicating improved arrival efficiency.



# 1

## Introduction

This chapter introduces the Master of Science thesis report titled: "Effect of Trajectory Prediction Uncertainty on a Probabilistic Debunching Concept of Inbound Air Traffic". This thesis is written at the Control Simulation department of the faculty of Aerospace Engineering at the Delft University of Technology. This chapter will present the research motivation and existing research gap in section 1.1, followed by the research questions in section 1.2. This chapter is concluded with the presentation of the project structure, presented in section 1.3.

### 1.1. Research Motivation

The airspace in western Europe is one of the most complex airspaces globally, and with the air space already operating near maximum capacity, innovations are needed to accommodate for potential growth in air traffic demand. Furthermore, with this growth in air traffic, concerns rise about the environmental and noise impact, especially near airports, where aircraft are least efficient, and noise pollution is highest. This research aims to design and test a method for debunching of air traffic that fits within the scope of current operations, such that the efficiency of the arrival traffic stream is increased.

In order to achieve higher efficiency in the arrival phase of a flight, a focus on the Arrival Manager is critical. An extension of the freeze horizon of the Arrival Manager could greatly benefit flight efficiency and support developments in innovative delay absorption techniques. However, the lack of information at long ranges, the reliability of predicted arrival times, the division of the airspace, and the influence of pop-up flights are limiting factors [57][58][62]. It is not expected that all uncertainties can be eliminated; therefore, increasing the AMAN horizon will require ways to perform arrival planning in the presence of uncertainty [57].

Furthermore, the arrival peak modes might result in an inefficient arrival traffic flow in the Control Area as, during such peaks, sector capacity might be exceeded. An inefficient arrival traffic flow might result in air traffic bunching, which is defined as several aircraft arriving within a short time span, such that the capacity of a point or an area in the airspace might be exceeded, and controller workload is significantly increased. This bunching effect occurs during normal traffic operations without unusual weather conditions or disruptions, and current ATFM regulations do not always mitigate this effect [56]. Therefore, there is a need for a locally adaptable and reactive way of debunching. However, little research is currently done on how such a model could be constructed. Furthermore, there is no clear insight into the uncertainties for the trajectory predictions for aircraft flying outside the Radar range. The research on this topic that has been performed is outdated, as the information sources have been updated.

Concluding, to increase the arrival flow's efficiency, more insight should be gained in the current trajectory prediction uncertainties. Then, Decision Support Systems could be developed that incorporate these uncertainties and try to create a more stable stream of arrivals in cooperation with adjacent centres. By doing this, all of the limitations to the extension of the Arrival Manager range could be addressed, and advancements might be made in the development of such a system and the efficiency of arriving traffic.

## 1.2. Research Questions

In this research, the main research question that will be answered is:

**”What is the effect of the inbound air traffic trajectory prediction accuracy on a debunching concept at Schiphol Airport Area Control”**

To guide answering these research questions, five sub-questions have been formulated:

1. What is the accuracy of the available trajectory predictions for inbound air traffic?
2. How can bunching be defined and detected in technical terms?
3. How can a solution be generated for debunching based on the available information?
4. What is the effect of the information accuracy on the debunching model?
5. What is the effect of debunching on the efficiency and the stability of the arrival manager planning?
6. What is the operational potential of the proposed debunching strategy?

The first sub-question serves to gain insight into the current accuracy of the available trajectory information, which is key to the workings of the debunching model. Previous studies on the uncertainty analysis of the available trajectory information are outdated as new, improved sources of information have become available. Analysing the current accuracy of the trajectory predictions is not only useful for use in decision support models, but it also provides insight in the uncertainty within current operations.

The second sub-question explores the topic of bunching and serves to define bunching in technical terms. As bunching can be defined in several ways, the definition for bunching in this research should serve the application that it is used for. Therefore, this research question is not fully answered by just the literature research and data analysis, but the question will play a key role throughout the research.

With the analysis of the trajectory prediction uncertainty from the first sub-question and the analysis of a definition for bunching in the second, a strategy can be defined for debunching of air traffic. The debunching model will be constructed in a modular way, such that it can be iterated and updated throughout the research.

After a final model has been defined and programmed, its performance should be tested. In the fourth sub-question, the uncertainty distributions found in the first sub-question will be used in the debunching model. Here, it will be analysed whether the prediction accuracy at a given horizon is high enough to make an accurate bunching prediction, and what the behaviour of this prediction over the horizon is. Therefore, this sub-question will also dive deeper into the influence of the change of the flight states and the other features selected in the uncertainty analysis of the first sub-question.

The fifth sub-question serves to test the influence of the debunching model on a model of an Arrival Manager. One of the core questions of this research is whether a debunching model will have an effect on the efficiency of air traffic arrivals. An analysis is done by analysing the effect on the traffic stream before entering the FIRAM and analysing the effect on the Arrival Manager. Furthermore, by incorporating the uncertainty of the trajectory prediction information and the Arrival Manager, this study might contribute to the advances in current Extended Arrival Management studies.

The final sub-question serves to explore the operational potential of the proposed debunching strategy. As the answer to this sub-question can be drawn from the conclusions of the previous sub-questions, it will be presented together with potential downsides or recommendations for the proposed debunching strategy.

## 1.3. Project Structure

This section describes the structure of the project. The document starts with the Literature Review and Preliminary study in Part I. The structure of this part is based on a top-down approach, starting at a high level and each chapter becoming more specific. In chapter 2 background information about the structure of the airspace is given. This is followed by chapter 3 where air traffic arrivals and relevant parameters, systems,

and techniques are discussed. chapter 4 dives deeper into the air traffic flow parameters. This is followed by chapter 5, where the traffic information sources and corresponding uncertainty are discussed. chapter 6 focuses on the aforementioned uncertainties by discussing trajectory prediction uncertainties. A summary of the findings in the Literature Review and the Preliminary Study is presented in chapter 7.

The second part of the document, Part II, consists of the body of work performed after the preliminary study. For this, a swim lane diagram was constructed, as shown in Figure 1.1. It can be seen that the project consists of five main steps: data processing, uncertainty analysis, model development, experiments and results. First, the data processing was performed where data was parsed and filtered, followed by the uncertainty analysis. These steps, as well as the corresponding results for the uncertainty analysis are presented in chapter 8. In the next phase of the research, the model development, the data from the first step and the data from the uncertainty analysis were used in the construction of the bunching detection and debuncher modules. Here, an AMAN module was also constructed which will be used in the experiment phase. The methodology of this part of the research is presented in chapter 9. This is followed by chapter 10, where the experimental setup is presented for the three experiments presented in the diagram. The final step is the analysis of the results, which is presented in chapter 11. The conclusions of the research are presented in chapter 12. In chapter 13 the recommendations are presented with final remarks on the operational potential of the debunching concept.

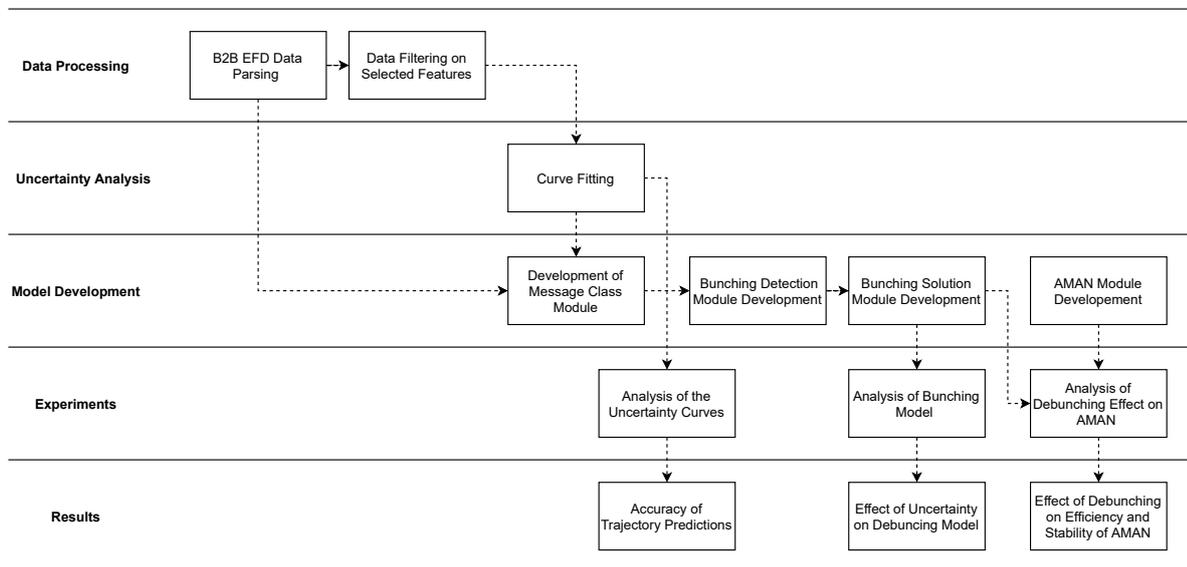


Figure 1.1: Swim lane diagram visually presenting the methodology of this research



# I

## Literature Review and Preliminary Findings



# 2

## Airspace Structure

This chapter will provide background information related to the structure of airspaces encountered by arriving flights at Amsterdam Airport Schiphol. This knowledge serves as a basis for the upcoming chapters, as the division of the airspace structure dictates how aircraft fly towards the airport. The airspace layout is described in section 2.1 followed by sub-sections going more in detail.

### 2.1. Airspace Layout

Airspaces are divided into Flight Information Regions (FIR), which are the largest divisions of airspace. In the Netherlands, this is the Amsterdam FIR (FIRAM), given by the code EHAA. A FIR is divided into several controlled airspaces. A schematic overview of the vertical limits and the ordering of these airspaces can be seen in Figure 2.1. The airspaces are divided and controlled as following:

- Control Zone (CTR): Controlled by Tower Control
- Terminal Manoeuvring Area (TMA): Controlled by Approach Control
- Control Area (CTA): Controlled by Area Control
- Upper Airspace (UTA): Controlled by Upper Area Control

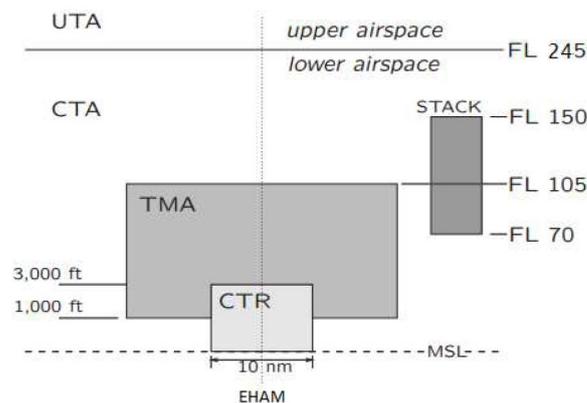


Figure 2.1: A schematic breakdown of the airspace structure (source: Ir Th.J Mulder, 2021).

The Control Zone and the Terminal Manoeuvring Area are only present near the airport. The Upper Airspace and Control Area However, span the full area of the FIR. The airspaces that are the main focus of this research are the Upper Airspace and the Control Area, as this research focuses on the cruise and initial descent phase of a flight. However, the influence of the processes in the Terminal Manoeuvring Area will affect the Control Area, therefore this airspace will be elaborated on as well.

### 2.1.1. Upper Airspace

The UTA is controlled by the Upper Area Control Center. For the Amsterdam FIR this is the Maastricht Upper Area Control Center (MUAC). The role of MUAC is to guide incoming traffic to the CTA, outgoing traffic from the CTA, and en-route traffic above Flight Level (FL) 245. The MUAC airspace is not constrained to the Amsterdam FIR, it extends to parts of Belgium, Luxembourg and Germany.

### 2.1.2. Control Area

The Control Area in the Amsterdam FIR is divided into 5 sectors, each with designated Air Traffic Controllers responsible for the traffic flow in that sector. The sectors are called and labelled as: northeast (1), southeast (2), south (3), southwest (4) and northwest (5). This sector division has been shown in Figure 2.2. From now, the sectors will be named by their corresponding number. Therefore, the sector that sometimes is referred to as sector West will now be referred to as Sector 5. It can be seen that the parts of the Northern, Eastern and Southern regions of the FIRAM are not included in the sector division. These areas are Military Airspaces and are under control of the Military Air Traffic Control. However, this division is flexible, and with reduced coordination, these areas become part of the civil airspace. The corresponding sector division is shown in Figure 2.3.

The role of the Area Control Center is to guide incoming traffic from the Upper Airspace to the Terminal Manoeuvring Area and vice versa for outgoing traffic. The Area Control Center also guides en-route traffic that stays below Flight Level 245. For inbound traffic to Amsterdam Schiphol Airport (AAS), the controller receives advisories from the Arrival Manager (AMAN) on how to guide traffic towards the Terminal Control Area. These advisories consist of a sequence, and an Expected Approach Time (EAT) and will be discussed more in detail later in section 3.1. In order to adhere to these EATs, controllers make use of stacks, which are holding areas designated to cope with a high traffic load or for delaying air traffic before it enters the Terminal Manoeuvring Area. These stacks can be seen in Figure 2.4.

In the same figure, the Standard Arrival Routes (STARs) have been shown. These routes are pre-determined routes that will be flown from one of the FIR entry points to the entry point of the TMA. The entry point of a FIR is called a Co-ordination Point (COP). At this point, traffic will be handed over from one Air Navigation Service Provider (ANSP) to the other. The neighbouring ANSPs are also called 'Adjacent Centres' and will be called so from now on. With each Adjacent Centre, a Letter of Agreement (LoA) has been established. In a LoA the sectorisation, the exact location of COPs, means of communication and exchange of flight data and procedures have been fixed. The procedures mainly consist of a COP location and the corresponding flight level at which an aircraft should be handed over. When looking at Sector 5 Figure 2.4 again, it can be seen that the main COPs with the adjacent centre are TOPPA, MOLIX and LAMSO. The flight levels usually are between FL240 and FL260. As was stated before, a STAR connects the COP with the entry point for the TMA. This entry point is called the Initial Approach Fix and is also the entry and exit point for flying a holding stack. These points are more-or-less located at 30NM from Amsterdam Airport Schiphol (AAS). At the IAF traffic should be flying between FL100 and FL75. The airspeed limit below FL100 is 250 kts Indicated airspeed. On average, the airspeed at the COP will be around 280 kts IAS.

It should be noted that LoAs have been established with every Adjacent Centre. For Sector 4 and 5, which correspond to the Amsterdam Area Control Centre, this means that there are LoAs between this ACC and the London Area Control Swanwick, London Terminal Control Swanwick, Royal Air Force Unit Swanwick, Scottish Area Control Prestwick and Maastricht Upper Area Control.

### 2.1.3. Terminal Manoeuvring Area

The TMA is the airspace connecting the CTA and the CTR and is controlled by Approach and Departure Control. The control tasks mainly consist of separating incoming and outgoing traffic, separating low-flying traffic and sequencing incoming traffic.

Incoming traffic enters the TMA at the IAF, between FL105 and FL70 at 250 kts IAS. This traffic can enter the TMA through one of the three entry points: SUGOL, ARTIP and RIVER. These points are shown in Figure 2.4, and laterally coincide with the entry and exit points of a holding stack. Outgoing traffic from



Figure 2.2: Sector division of the FIRAM without reduced coordination (source: LVNL, 2021)

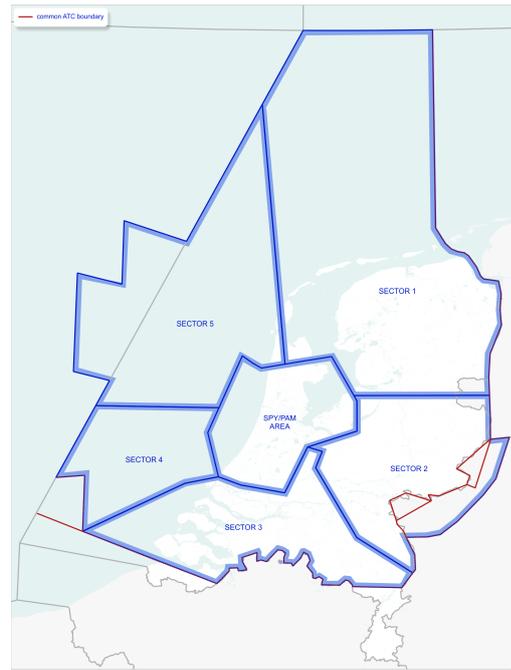


Figure 2.3: Sector division of the FIRAM with fully reduced coordination (source: LVNL, 2021)

the airport follows the Standard Instrument Departures (SIDs) through the TMA and is guided to the CTA until FL130.

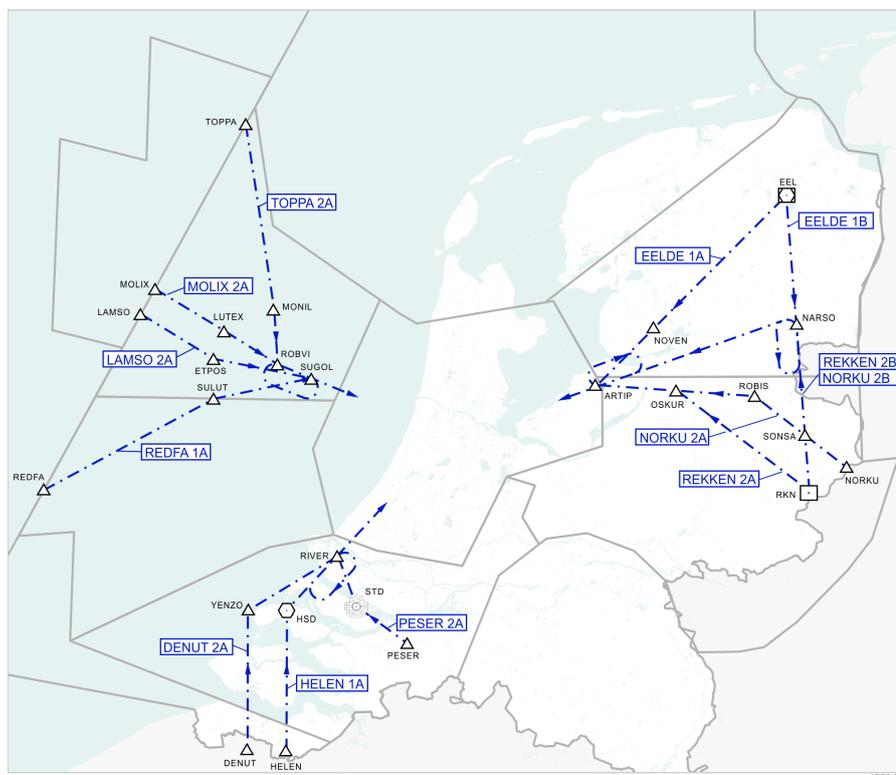


Figure 2.4: Standard Arrival Routes in the FIRAM (source: LVNL, 2021)



# 3

## Arrival of Air Traffic

The previous chapter focused on the airspace structure encountered by arriving flights. This chapter however, focuses on the arrival phase of a flight from the system's perspective that has the most influence on the arrival process: the Arrival Manager. The chapter starts with a general analysis of the Arrival Manager in section 3.1, followed by a more in depth analysis of the planning principles in section 3.2. Then, a discussion on the arrival management range and the extension of this range is presented in section 3.3. The chapter continues with a literature review on Linear Delay Absorption Techniques in section 3.4, which are techniques focusing on delaying flights efficiently. The concept is related to arrival management because it could improve the efficiency of how Arrival Manager advisories are executed.

### 3.1. Arrival Manager

The main focus of this research is on how air traffic can be influenced before it enters the Area Control airspace so that the efficiency of arriving traffic is increased. For this, one needs to analyse the air traffic flow and which parameters have an influence on the flow. The system designated to flow management is the Arrival Manager (AMAN) near the airport. The AMAN is a system that assists an Air Traffic Controller (ATCO) in sequencing, metering and merging air traffic to optimise the runway throughput. When multiple aircraft approach an arrival airport, the AMAN establishes the arrival sequence and the corresponding arrival times. It is the role of the ATCO to safely guide the aircraft to the airport while merging and metering the aircraft as proposed by the AMAN. There are several benefits to using an AMAN, one being the efficiency of the runway throughput (minimising the time for the total landing sequence). Another benefit is the potential decrease in ATCO workload, as controllers only need to follow the AMAN advisory, and therefore less cognitive work is required. The final benefit is the flight efficiency in the arrival phase of a flight due to an optimised runway throughput.

Not all airports use the same AMAN system, but all AMAN systems consist of several general functions. The most basic functions consist of establishing an optimal aircraft landing sequence and the metering of traffic to that sequence. The optimised sequence is established based on several criteria such as the estimated time of arrival (ETA), the aircraft type and the runway capacity [25]. The arrival sequence is dependent on whether there are one or two main landing runways. The AMAN then allocates flights to the runways based on rules set by the Approach Planner (APLN) who is responsible for managing the AMAN system. The flights are allocated to the runways based on their ETA. The landing intervals on the runway between aircraft can be set as a fixed landing interval (LIV) in time or a dynamic LIV. In the case of a dynamic LIV, the following parameters are taken into account: The Wake Turbulence Category (WTC) separation criteria for an aircraft pair, the approach speed (in True Airspeed) of the trailing aircraft over the required WTC-separation, the wind speed on final approach, a buffer distance per WTC-pair to tune operations for the given airport, a buffer distance set by the Approach Planner to tune the planning, and a minimum LIV-distance used in extraordinary weather conditions [37].

The dynamic LIV is calculated as shown in Equation 3.1, where the Calculated Ground Speed (CGS) is given by Equation 3.2, where  $W$  is the average wind on final, typically given at 1000ft,  $Q$  is the angle formed by

the wind and the runway direction and  $pTAS$  is the predicted True Airspeed for the trailing aircraft (as given by the Aircraft Performance Module). The  $LIV_{distance}$  is given by Equation 3.3, with parameters as described in the previous paragraph.

$$LIV = \frac{LIV_{distance}}{CGS} \cdot 3600 \quad (3.1)$$

$$CGS = -W \cdot \cos(Q) + \sqrt{\sqrt{pTAS} - \sqrt{W \cdot \sin(Q)}} \quad (3.2)$$

$$LIV_{distance} = \max(LIV_{minimumdistance}, WTC_{separationdistance} + WTC_{separationbuffer} + LIV_{buffer}) \quad (3.3)$$

Based on the optimal sequence that the AMAN has produced, a landing slot is determined for each aircraft. Since there might be a difference between the landing slot and the ETA, an aircraft might need to absorb delay such that it meets the landing slot given by the AMAN. Therefore, this delay is the difference between the ETA (or 'ETA nominal') and the landing slot. Based on the landing slot, the AMAN also provides an Expected Approach Time (EAT) over a metering point. Usually, this metering point is the Initial Approach Fix (IAF), located at the end of a STAR (see subsection 2.1.2). The EAT is the time that the aircraft needs to fly over the metering point in order to be at the runway at the right time, and is constructed as the Estimated Time Over (ETO) the IAF plus the delay that needs to be absorbed. Note that the ETO of an aircraft can also be later than the EAT, in which case there is Time To Gain (TTG) and an aircraft should increase speed or fly shortcuts when possible. When the ETO of an aircraft is earlier than the EAT, there is Time To Lose (TTL), which is done by reducing speed or by adding extra track miles by vectoring or holding an aircraft. The TTG and TTL that need to be absorbed to meet the landing slot, have to be absorbed before the aircraft flies over the IAF. Therefore, this delay (positive or negative) needs to be absorbed either en-route, or in the Control Area. It is the role of the Air Traffic Controller in the Control Area to guide aircraft to adhere to the AMAN sequence and advisories, and provide instructions such as speed changes and vectoring commands to meet the Expected Approach Time [25]. In current operations, a deviation of plus or minus 120 seconds from the EAT is allowed, even though this might result in a less optimal traffic stream. The reason this deviation is allowed is to give controllers tactical manoeuvring freedom in the metering and sequencing process. The total window size in which traffic can be handed over to the Approach Controllers is therefore 4 minutes. The resulting EAT inaccuracies need to be compensated for in the TMA by vectoring to final approach.

The AMAN consists of several modules. The first module is the Human Machine Interface (HMI), which allows users to interact with the AMAN and allows manual actions to be performed within the AMAN planning process. The second module is the planning core which performs the arrival management functions such as allocating landing runway, landing slot and EAT. The third module is the Trajectory Prediction (TP) module, which provides the planning core with timing information of flights including arrival time estimates over several points in the route. Both the planning core and the TP modules use meteorological information as input, which is the fourth module of an AMAN. The last module is the Aircraft Performance Module, which contains parameters used for the TP module, based on historical data and categorised by aircraft type and operator. Furthermore, inputs consist of runway configuration settings set in the Human Machine Interface Module, flight plan data and surveillance data [37]. A schematic overview of the AMAN system has been given in Figure 3.1.

### 3.2. Freeze Horizon

Before a flight enters the FIRAM, the information that is known in the AMAN system is based on the flight plan. This planning state is called the 'unplanned' flight state, as no runway or planning information is scheduled yet. A flight becomes Activated after an activation (or ACT, see section 5.5) from an adjacent control centre, which contains flight data and transfer conditions. Preceding the receipt of the ACT message, the radar has been correlated with the flight and is fed to the Trajectory Prediction module of the AMAN. With the activation of a flight it will be assigned a landing runway and a landing slot and corresponding EAT. The flight state is now either Planned or Preview-Planned. In the Preview-Planned state, the landing slots and EATs are continuously re-evaluated depending on the traffic situation. A flight is Preview-Planned until it passes a point in time, defined in minutes before the IAF. When the aircraft passes this point, it is fixed (or 'frozen') in the sequence, and the status becomes Planned [37]. This point is also called the freeze horizon, and in current operations at Amsterdam Schiphol Airport it is set at 14 minutes before IAF. This point is determined for each aircraft by the Trajectory Predictor module in the Arrival Manager. Since it is a point in time, and not a

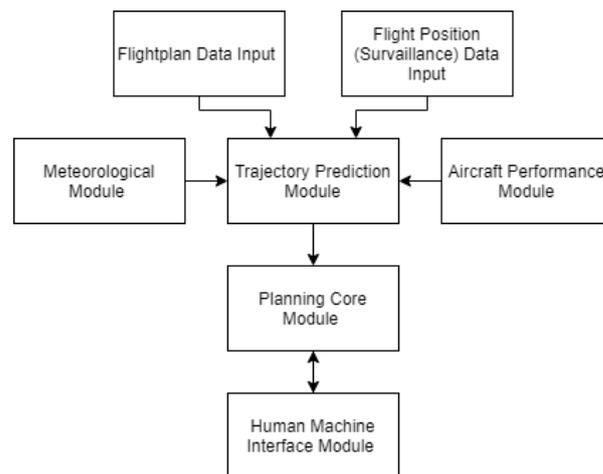


Figure 3.1: System architecture of an Arrival Manager

point in space, it mitigates faster aircraft being sequenced behind slower aircraft when overtaking might still be possible. It should be noted that the freeze horizon depends on the speed of the aircraft, its heading, the wind speed and the wind direction. Furthermore, a lack of an ACT Message or radar range limits the range at which the pre-planning can start and has an impact on the freeze horizon if sent late by the adjacent centre [58]. A flight might skip the Preview-Planned state if it pops-up within the freeze horizon. Such a pop-up flight requires the AMAN planning to be (manually) rescheduled, and thus increases uncertainty in the sequencing and metering process [62].

When the freeze horizon is passed, the sequence is fixed and the AMAN provides controllers with the landing slot and the EAT as discussed in section 3.1. It is important to note that generally, Air Traffic Controllers do not influence an aircraft before it has passed the freeze horizon. Doing this, would mean that the planning itself would be influenced as the Expected Time of Arrival is changed by changing the course of an aircraft. This could result in the aircraft being sequenced behind another aircraft, and thus needing to absorb even more delay (as dictated by the EAT). On the other hand, a delayed aircraft might be required to speed up again to meet the EAT given by the AMAN. Such conflicting speed instructions complicate descent planning for flight crews and compromise flight efficiency. Therefore, a frozen planning is a prerequisite for ATCOs to influence an aircraft. The location of the freeze horizon thus determines how much airspace is left for ATCOs to influence aircraft and how effectively an aircraft can be guided to its EAT over the IAF. With an early freeze, the airspace increases so that delay (the TTL) can be absorbed in an efficient manner. With a smaller airspace, flight and workload efficiency might decrease as a more ad-hoc solution is required by ATCOs. To cope with this issue, extending the range of the AMAN is actively researched, as discussed in the next section.

### 3.3. Extended Arrival Manager

One of the concepts for extending the range of where air traffic can be influenced is the concept of Extended Arrival Management (XMAN). Extended Arrival Management systems use an extended range in order to extend the arrival management up to a 500 NM horizon from the airport [40]. By freezing the AMAN sequence earlier, there is more airspace left for ATCOs to influence air traffic, increasing the EAT adherence. Furthermore, the sooner the arrival sequence is available, the more gradual the merging, metering and delay absorption can be performed by the ATCO. It is important to move from the current tactical manoeuvring techniques such as vectoring and airborne holding, which are fuel inefficient, towards more fuel efficient automation supported procedures.

Because of the size of the FIRAM, extending the range up to 500NM will require cooperation from the Adjacent Centres. The least complex way of cooperation would be for LVNL to send speed requests to these neighbouring ATC centres based on the AMAN planning at LVNL. As traffic upstream is influenced according to the ASAP planning, the workload of the ATCOs in the Control Area decreases. It should be noted that with the added required cooperation, some workload might be increased. However, spreading the workload also

results in a more efficient handling of traffic, and thus an overall decrease in workload.

Another inherent benefit of implementing an Extended Arrival Manager is that bunching (bunches of traffic arriving in a short period of time) can be mitigated. In an XMAN concept, controllers at the Area Control Center and at adjacent FIRs are working with the EAT prescribed by the Arrival Manager. This EAT is the Expected Approach Time over the metering point. By influencing aircraft such that they meet their EAT, bunching is diminished, as the AMAN planning at the metering point is inherently debunched.

However, at AAS no such Extended Arrival Manager is currently in use yet since there are several limitations to how far the AMAN range can be extended. These limitations are the following [57]:

1. Traffic information is limited at long ranges. An example is the limitation of the radar range. Accurate traffic information is required for the AMAN to optimise the landing sequence as accurately as possible [58].
2. Reliability of predicted arrival times used by the AMAN due to the stochastic nature of air traffic. Examples of factors influencing the predictability are flight speed, weather, airspace capacity, and tactical controller and pilot actions.
3. The airspace division. Controllers of adjacent centers might influence the course of an aircraft without regard to the AMAN planning.
4. Pop-up flights from airports within the AMAN horizon [62]. For example, in 2019, 9 of the 10 busiest connections to Schiphol were within a range of 500 nm [1]. This means that errors associated with ground procedures might have influence on the planning and add to the potential disturbances [57].

It is not expected that all uncertainties can be eliminated, Therefore, increasing the AMAN horizon will require ways to perform arrival planning in the presence of uncertainty [57].

### 3.4. Linear Delay Absorption Techniques

When considering an Extended Arrival Manager as discussed in section 3.3, the required time over the Initial Approach Fix is known at a large distance before the airport. Therefore, aircraft can be influenced in a more gradual and efficient way by combining delay absorption and arrival metering. A technique that could be used for this is linear delay absorption.

In Europe, ground delays are imposed by Air Traffic Flow Management (ATFM) in order avoid exceeding airspace capacity and to reduce weather related impact. Imposing ground delays decreases the need for tactical delay absorption techniques such as airborne holdings. The main benefit of these imposed ground delays is that the fuel consumption for ground delay is estimated to be about 6 times lower than that with airborne delays [5].

However, ground delays do not always mitigate the need for airborne holding as aircraft operators often try to make up for the delay by flying faster, and thus increasing airspace congestion [20]. From the operator perspective, airborne delay absorption might be preferred, since ATFM regulations can disappear during flight due to the non-deterministic situations the ATFM system faces. If an aircraft has absorbed delay on the ground, the only way to recover this delay is to fly higher speeds en-route, which increases fuel costs. Also, if an aircraft consumes delay in the air instead of on the ground, an aircraft has more control over how the delay is absorbed, and aircraft could even save fuel by flying at the minimum fuel consumption speed.

Delay absorption can be done in an efficient way by looking at the relation between the flight time costs and the fuel costs. This relation between the cost of flight time and the cost of fuel is described by the Cost Index (CI) with unit [kg/min]. When Flying at maximum CI the cost of fuel is ignored, and the maximum flight envelope speeds are flown. Flying at a CI of zero corresponds with the minimum fuel consumption airspeed, or the maximum range airspeed [49]. The CI setting can affect the whole flight trajectory of an aircraft, as optimal speed settings, flight levels, and the climb and descent profiles might change with the CI setting [51]. Aircraft operators typically fly at CI values ranging between 20 to 70 kg/min [13], meaning that the nominal cruise speed is typically faster than the minimum fuel consumption speed because of the time-based costs for an airline.

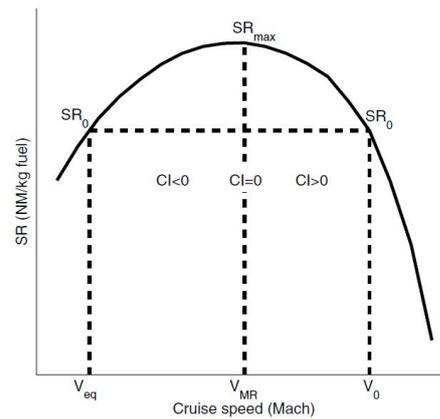
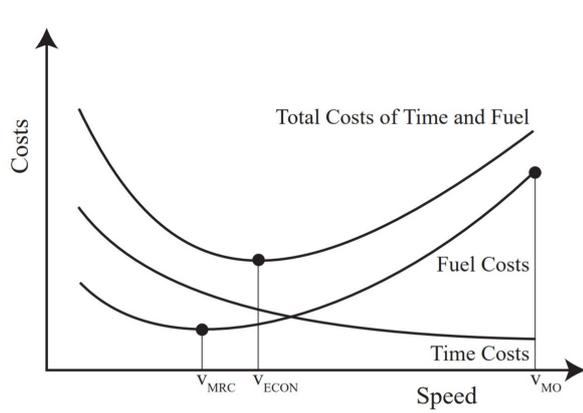


Figure 3.2: Impact of speed on the operating costs (source: [51]) Figure 3.3: Specific Range as a function of the cruise speed (source: [13])

In Figure 3.2 the relation between the fuel cost and the time cost curves are shown. Here,  $V_{ECON}$  corresponds with the Cost Index that results in the lowest operating cost. The maximum range cruise speed is given by  $V_{MRC}$ . It should be noted that the total cost curve varies with the flight conditions. The Fuel cost curve is dependent on the gross weight of the aircraft, the Flight Level, the air temperature and the wind conditions. The time cost curve is dependent on the ground speed, and therefore influenced by the wind speed [51].

The maximum range cruise speed in Figure 3.2 corresponds with the maximum Specific Range (SR). The SR is defined as the distance flown per unit of fuel consumed. Therefore, when the CI is to zero, the SR is maximum, as the aircraft is flying at its minimum fuel consumption airspeed (equal to its maximum range speed). In [13] the equivalent airspeed  $v_{eq}$  was introduced as the minimum speed yielding the same SR for a given nominal airspeed  $v_0 = v_{ECON}$ . This relationship is shown in Figure 3.3. The fuel consumption is equal to or lower than nominal flight for all cruise speeds between  $v_{eq}$  and  $v_0$ . Flying at these lower cruise speeds means that delay can be absorbed without extra fuel costs. Note here that logically, the time related costs for the operator do increase.

If an aircraft slows down to the minimum fuel consumption airspeed, the aircraft can absorb delay while saving fuel. It has been shown that aircraft can slow down further, increasing fuel consumption again, to the point where fuel consumption is the same as the nominal cruise speed. Here, the en-route delay absorption is maximum, without extra fuel costs for the operator. This maximum amount of delay can be increased with a higher fuel allowance, and if cruise altitude is allowed to change with the changed cruise speeds, enabling the aircraft to fly at its optimal cruise altitude for the given airspeed [13]. It should be noted that the lower bound of the airspeed might be bound by the minimum operational airspeed of an aircraft for a given flight level and mass.

The literature above is focused on airborne delay absorption in the cruise phase. However, in [43] it was shown that the most fuel efficient strategy to absorb delays en-route is to first reduce the Calibrated Airspeed for the descent phase as much as possible, before reducing the cruise Mach speed. This study was performed under the assumption of an idle-thrust descent. If still more delay needs to be absorbed, the last option would be to do this with a path stretch at cruise altitude. Furthermore, it should be noted that absorbing delay in the cruise and descent phase does not require a deviation from the initial flight plan, while path stretching requires lateral navigation [63].

The amount of delay that can be absorbed without extra fuel cost is dependant on the margin between  $v_0$  and  $v_{eq}$ . This margin is a function of the CI that is flown at, and the shape of the SR curve, which is aircraft, flight level and mass dependent [13]. In [14] the effect of wind on the maximum delay absorption was analysed. It was found that a head-wind significantly increased the amount of delay that could be absorbed without extra fuel allowance. A simulation was performed for 21 routes to Chicago O'Hare Airport ranging

	Number of Holdings	Percentage of Arrivals
ARTIP	2790	2.6%
RIVER	2415	3.6%
SUGOL	3410	4.3%

Table 3.1: Holding count and percentage per Initial Approach Fix, for arrivals in 2019

	Average Holding Time [min]	Median of Holding Time [min]
ARTIP	8:17	6:43
RIVER	7:50	6:15
SUGOL	7:53	6:33

Table 3.2: Average and median holding time per Initial Approach Fix, for arrivals in 2019

from 200 to 1100 nautical miles. It was found that for the majority of routes the amount of airborne delay for flights with headwind was over 10 minutes, and over 5 minutes for flights with tailwind or perpendicular wind. Interestingly, most of the flights where the amount of airborne delay absorption was over 10 minutes were short haul flights, showing the large influence wind has on the delay absorption [14]. To show the potential of this amount of airborne delay, the average and median time an aircraft spent in an airborne holding pattern is show in Table 3.2. In Table 3.1 the amount of arrivals subject to these holding patterns are shown. These numbers illustrate that the ability to absorb delay en-route of even 5 minutes could have a significant effect on the time an aircraft spends in the inefficient holding patterns.

One major drawback in using the CI and the SR in delay absorption for aircraft arrivals. This drawback is due to the fact that the CI setting of an aircraft is sensitive data from the aircraft operator, as it is based on fuel consumption and weights models of the aircraft. This sensitive data is not shared with the ANSP. As distributed cooperation approaches are currently not in use and are not expected to be in use soon, implementing the CI in aircraft delay absorption might not be a viable option. However, knowing that aircraft can absorb delay without extra fuel costs is key in more efficient delay absorption. It supports the case for long-range delay absorption using minor speed adjustments. Research analysing the maximum amount of delay absorption could assist in a debunching delay absorption technique that is viable and cost-effective, sustainable, and fair from an aircraft operator's perspective.

# 4

## Air Traffic Arrivals Flow

Now that air traffic arrivals have been discussed in the previous chapter, this chapter will focus on the relevant flow parameters related to the arrival of air traffic. The chapter will consist of a mix of data analysis and literature research. The chapter starts with an analysis of the arrival capacity at AAS section 4.1, followed by a brief literature review on the effects of Air Traffic Flow Management. The chapter continues with one of the core topics of this research: An analysis of the bunching effect and a technical definition of bunching that can be used in a technical model. The chapter is concluded with an analysis of the effect of traffic originating from the North Atlantic on Bunching at AAS.

### 4.1. Air Traffic Flow and Capacity

Traffic flying to AAS will follow one of the STARs shown in Figure 2.4. It can be seen that the first waypoint of these STARs commonly originates at the border of one of the sectors. To get insight in the sector load, an analysis has been performed with regard to the arrivals in each sector. In Figure 4.1 the sector division for arrivals in 2019 is shown. It can be seen that most of the traffic arrives through sector 2 and 3. Note that sector 3 is the only sector feeding the RIVER IAF, while sector 1 and 2 combined feed the ARTIP IAF. The arrivals in sector 4 and 5 will be directed toward the SUGOL IAF. This brings the arrivals fed through ARTIP at roughly 42% of all flights, about 32% through SUGOL and 26% of flights through RIVER.

The amount of flights that arrive each day, and the time that these flights can arrive at AAS is determined by a capacity declaration. The capacity of AAS is declared twice a year, once in summer and once in winter. The nominal runway capacity per hour changes with the peak modes, and whether it is day or night. During the departure peak mode, one runway is used for landings and two runways for take-offs. During the arrival peak mode, two runways are used for landings and one for take-offs. At night operations, one runway is used for landings and one runway is used for departures. The declared nominal capacity per hour and per 20 minutes for summer 2019 have been shown in Table 4.1 for the arrivals and in Table 4.2 for the departures [45], clearly showing the alternating effect between the arrival and departure peaks.

The capacity is declared in blocks based on arrival and departure peaks. In Figure 4.2 the arrival capacity blocks can clearly be seen in blue. Furthermore, the 20 minute moving average for the arrival times of flights on August 3 at AAS have been plotted in this figure. It can be seen that, as expected, the arrival flow over the

ATM mode	Period from -to (UTC)	Nominal capacity per hour	Nominal capacity per 20 minutes
		Arrivals	Arrivals
Departure peak (day)	05 : 00 – 19 : 39	36	12
Arrival peak (night)	05 : 00 – 19 : 39	68	23
Off peak (day)	04 : 00 – 04 : 39	24	8
	04 : 40 – 04 : 59	24	8
	05 : 00 – 20 : 39	36	12
Night mode (N)	21 : 00 – 03 : 59	24	8

Table 4.1: Capacity Declaration for arrivals at AAS

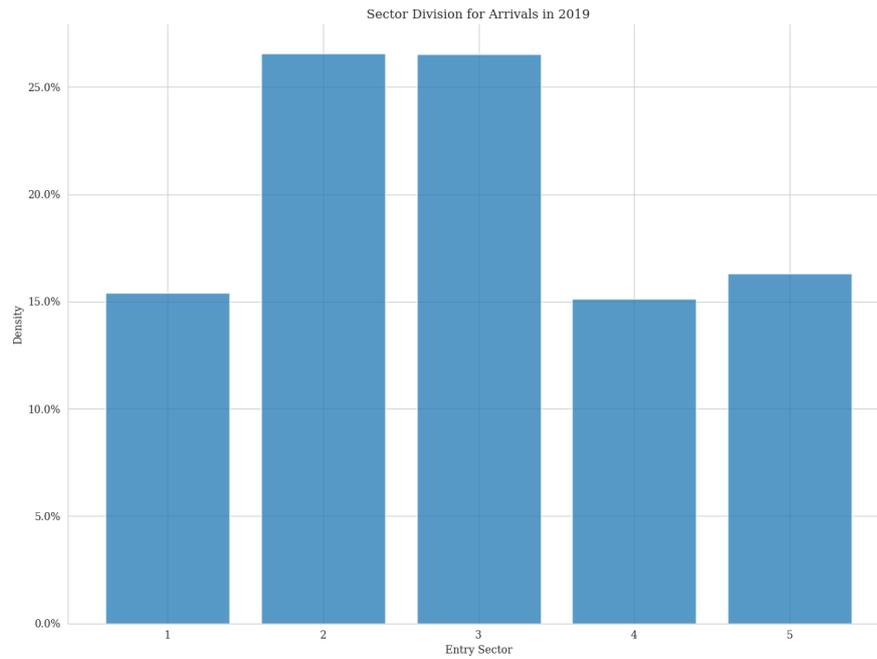


Figure 4.1: Arrival sector division for arrivals in 2019

ATM mode	Period from -to (UTC)	Nominal capacity per hour	Nominal capacity per 20 minutes
		Departures	Departures
Departure peak (day)	05 : 00 – 19 : 39	74	25
Arrival peak (night)	05 : 00 – 19 : 39	38	13
Off peak (day)	04 : 00 – 04 : 39	30	10
	04 : 40 – 04 : 59	40	14
	05 : 00 – 20 : 39	40	14
Night mode (N)	21 : 00 – 03 : 59	25	9

Table 4.2: Capacity Declaration for Departures at AAS

day follows the declared capacity. In some of the peaks the capacity is larger than the declared capacity, such as around 6:00. At 18:00 it can be seen that the arrival peak is shifted with respect to the declared capacity block. Arrival patterns as this might increase complexity for controllers as the arrival peak might coincide with a departure peak. Another implication of the dense traffic peaks is that it might result in traffic bunching, which will be explained in section 4.3.

## 4.2. Air Traffic Flow Management

In the previous section the arrival capacity was given for 20 minute periods during a twenty-four hour period. This capacity serves as a constraint upon which will be acted if the capacity is expected to be exceeded. Exceeding the capacity means that delay needs to be absorbed by traffic through the use of excessive vectoring techniques or even holding stacks. These types of delay absorption techniques are workload intensive, and as such should be minimised. The goal of ATFM is to minimise the extend and impact of these air traffic delays due to congestion, by controlling the flow of air traffic in order to match the demand with the available capacity [11].

The general idea behind ATFM measures is that it is safer and cheaper to let flights absorb delay on the ground instead of in the air. Note that these measures will only be taken when delays are expected to be unavoidable. Examples of ATFM measures are departure delays or re-routing of flights such that overloaded airspaces are avoided [11]. In [17] it was found that the departure delays did not always cause the wanted effect on the arrival time. In fact, it was found that in 2018 about 25% of the ATFM regulation minutes had

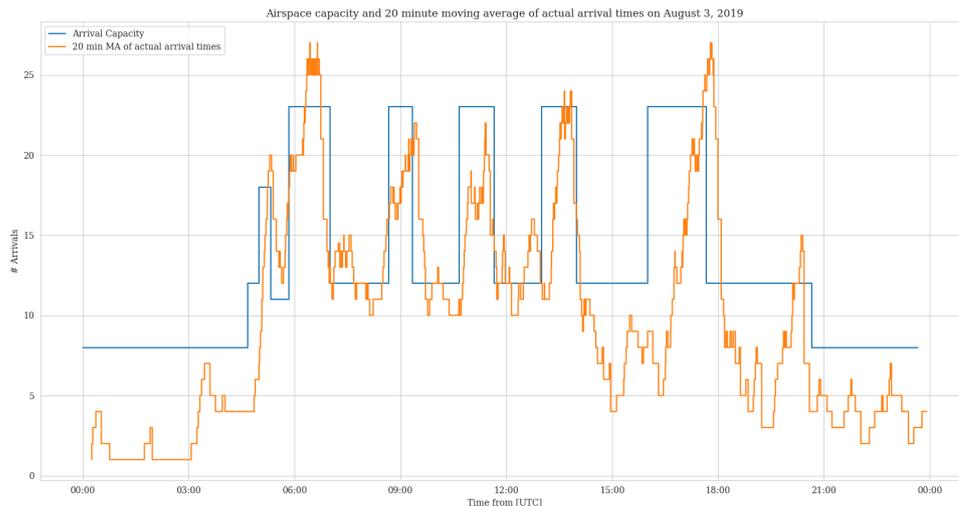


Figure 4.2: Airspace capacity and 20 minute moving average of actual arrival times at Schiphol Airport on August 3, 2019

no effect on the planned arrival time for a flight. Possible explanations for this is that the flights were either ahead of schedule, or due to aircraft operators flying faster en-route in order to make up for the delay due to connecting passengers or the next planned flight for that aircraft, called the 'compensation effect'. Furthermore, in [56] it was shown that ATFM measures might even amplify capacity constraint violations, due to for instance bunching. The reason behind this and more information on the bunching effect will be given in the next section.

### 4.3. Air Traffic Bunching

Bunching occurs when multiple aircraft approach an airspace or a given point in the airspace within a short period of time. In general, the term bunching is used to describe a situation in which the capacity at an airspace or at a given point, is higher than the imposed limits. The goal for ATFM regulations is to achieve a stream of traffic in which there are no such capacity violations. However, margins around imposed ATFM take-off-times and compensation effects for ATFM measures (flying faster en-route) were shown to significantly contribute to the risk of bunching [56]. In [56] it was also concluded that bunching does not necessarily correspond with exceptional traffic situations, but can appear during normal traffic situations. Therefore, an effort should be made to predict and solve bunching, as bunching situations can result in excessive workload as tactically solving it is done through inefficient manoeuvring techniques such as vectoring and airborne holdings. As ATFM regulations do not mitigate the effect of bunching, solving of bunching situations should be done alongside with the current ATFM measures. As a result, debunching should be done in a locally adaptable and reactive way [56].

#### 4.3.1. Technical Definition for Bunching

There is no general numeric definition for when a situation is considered to be bunching. Therefore, several parameters will be discussed in this subsection leading to a definition that will be used in this research. It has been chosen to focus on debunching a point in the airspace, rather than an airspace. The main reason behind this is that from an ATCO perspective, traffic can be considered bunching when a flight needs to be delayed within the airspace. Since flights are merging from the COPs towards the IAF, it makes sense to focus on these two points. On the contrary, from an ATFM perspective, the main focus is on the capacity of a sector. However, even when the hourly capacity of a sector is not exceeded, bunching can still occur as the flights could arrive within a very short time-span.

The definition that is explored in this research is a time-based definition of bunching, so that bunching can be defined as two or more aircraft estimated to arrive at a point in the airspace within a given amount

Leader / Follower	Super Heavy	Upper Heavy	Lower Heavy	Upper Medium	Lower Medium	Light
Super Heavy	3 NM	4 NM	5 NM	5 NM	6 NM	8 NM
Upper Heavy		3 NM	4 NM	4 NM	5 NM	7 NM
Lower Heavy		(*)	3 NM	3 NM	4 NM	6 NM
Upper Medium						5 NM
Lower Medium						4 NM
Light						3 NM

Table 4.3: RECAT-EU WTC separation minima on approach [15]

of time. The time between one aircraft and the next can be based on the WTC of the aircraft involved. The RECAT-EU is a wake turbulence category separation minima scheme for pairs of aircraft [15]. These minima have been established so that arriving aircraft are not influenced by the wake-vortex of its predecessor. As the intensity and the risk of a wake vortex depends on the sizes of both aircraft, several categories have been established based on the Maximum Take-Off Weight of the aircraft [2]. These minima are shown in Table 4.3, where the (\*) denotes the minimum radar separation, set at 2.5 NM [15]. For each aircraft pair, these separation minima can tactically be translated to a separation buffer in time, when the expected ground speed of the trailing aircraft is known. Since these speeds are not known in advance, this could be based on historical data for each aircraft category. Note that at a large prediction horizon it is difficult to predict what the exact sequence of flights at the IAF will be, and therefore also which WTC separation minima needs to be used for the detection of bunching. A solution for this is to take the average separation for all aircraft that could potentially cause bunching at a given time. More on this will be discussed in section 9.3.

The second parameter considered in the definition of bunching is the point in the airspace at which bunching is considered. This research is focused on the initial phases of arrival, therefore, the areas of interest are defined by the airspace layout (as discussed in section 2.1) and the arrival process. Based on these two factors, the area of interest is the Control Area, as this airspace connects the Upper Airspace with the Terminal Control Area. In Figure 2.4 the Standard Arrival Routes were shown. From this, two main points in the airspace can be identified: the Coordination Point (COP) and the Initial Approach Fix (IAF). An aircraft enters the FIRAM at one of the COPs, and descends through the airspace until it reaches the IAF, where its control is handed over to Terminal Control.

Bunching can be either defined at the COP, or at the IAF. However, when measuring no bunching at the COP, it should be noted that due to the converging nature of the traffic, bunching might still occur at the IAF. The decision to evaluate bunching at the COP should therefore still take into account the effects at the IAF, making the definition of bunching more complex. Looking at the STARs of sector 4 and 5 as an example, shown in section 2.1, it can be seen that there are 4 main COPs. When not considering the IAF downstream, bunching could be defined as a number of N aircraft, arriving within X minutes at any one of the four COPs. However, with this option, the difference in length for the flight path between the different COPs and the IAF has not been taken into account, and thus bunching at the FIRAM border (where the COPs are located) does not necessarily have to result in bunching at the IAF. To illustrate this effect, an analysis has been performed of the air traffic entering the FIRAM through sectors 4 and 5. In Table 4.4 the most common entry COPs have been shown for traffic in 2019. For these entry COPs, the average flying time from the COP to the IAF (without potential holding times) has been shown in Figure 4.3. In this plot it can be seen that the median flying time from COP to IAF differs between the COPs by tens of seconds. The median flying time from DIBAL to the IAF was 542 seconds, whereas the median flying time from REDFA to the IAF was 591 seconds.

It is concluded that defining bunching at the COP requires a complex definition of bunching including knowledge of the flown airspeeds. Therefore, the second option of analysing bunching at the IAF is the logical choice. Also, when guiding aircraft from the COPs to the IAF, the ATCO is focused mainly on the IAF, as this is the point where aircraft all merge towards and the point for which EATs are given by the Arrival Manager (see section 3.1). Note that there are three IAFs: ARTIP, RIIVER and SUGOL. Therefore, when detecting bunching tactically, each of these IAFs should be checked for bunching. Each IAF can also be solved separately, as they are fed by separate inbound traffic streams.

Entry COP	% of flights
REDFA	48.1%
DIBAL	41.2%
LARDI	6.7%
RAMID	2.0%
ROKAN	1.7%

Table 4.4: Entry COPs for Sector 4 and 5 in 2019 (only showing COPs larger than .2 %)

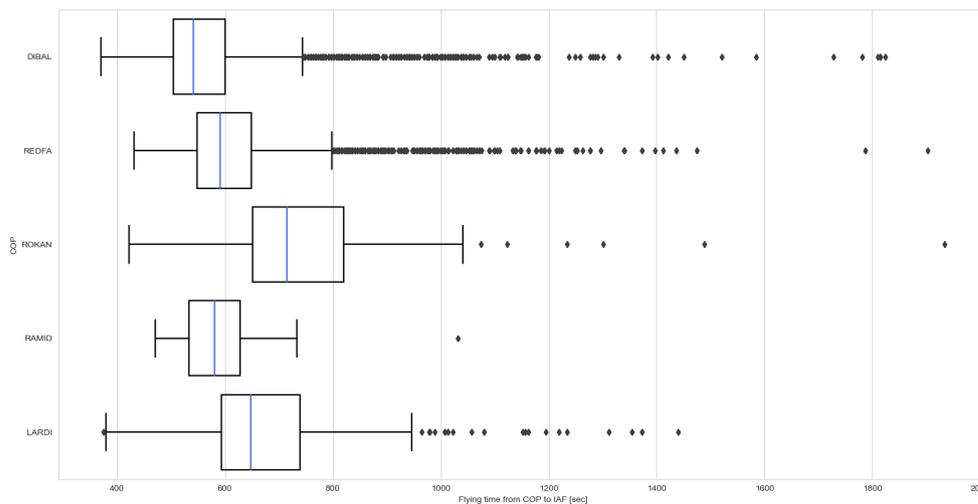


Figure 4.3: Flying time from COP to IAF for flights arriving at Sector 4 and 5 of Schiphol Airport in September 2019

## 4.4. North Atlantic Traffic

Air traffic flying from North America to western Europe and vice versa is commonly flying over the North Atlantic Tracks (NAT) [30]. The North Atlantic Tracks are routes that have been established due to the lack of radar coverage over the Atlantic, and thus assure separation between aircraft. The routes are updated twice a day based on the weather patterns, such that aircraft fly their minimum time tracks and altitude profiles. The presence of jet streams results in different tracks for Eastbound and Westbound flows. The Eastbound routes are constructed to benefit of the present tailwinds, and the Westbound routes try to minimise the exposure to strong headwinds [30]. Because of time zone differences and resulting passenger demands, the North Atlantic air traffic consists of two alternating flows: the Westbound flow departing from Europe in the morning, and the Eastbound flow departing from North America in the evening. The result is an arrival peak at AAS in the morning, which is clearly visible in Figure 4.4.

Besides the fact that arrivals from the North Atlantic cause a peak traffic load, it also is an important flow of traffic since it is not subject to potential ATFM regulations. Therefore, peak loads from the North Atlantic cannot be flattened by EUROCONTROL, and have to be coped with. Furthermore, the presence of jet streams above the North Atlantic might cause aircraft to arrive in the Amsterdam FIR in a bunched manner. Therefore, it was analysed whether flights from the North Atlantic could be linked to an increase in delays within the sector. First, the arrival sectors for traffic arriving from the North Atlantic have been shown in Figure 4.5. As it can be seen that this traffic mainly enters through sector 4 and 5, only these sectors will be analysed. Furthermore, it was found that 24,6% of all traffic entering sector 4 and 5 originated from the North Atlantic Tracks in 2019.

To check whether traffic from the North Atlantic Tracks can be linked to high amounts of delay absorption in the sector an analysis was done on holding delay. For each IAF, it was analysed what percentage of arriving flights was subject to flying in a holding pattern. This was already shown in Table 3.1, where it was shown

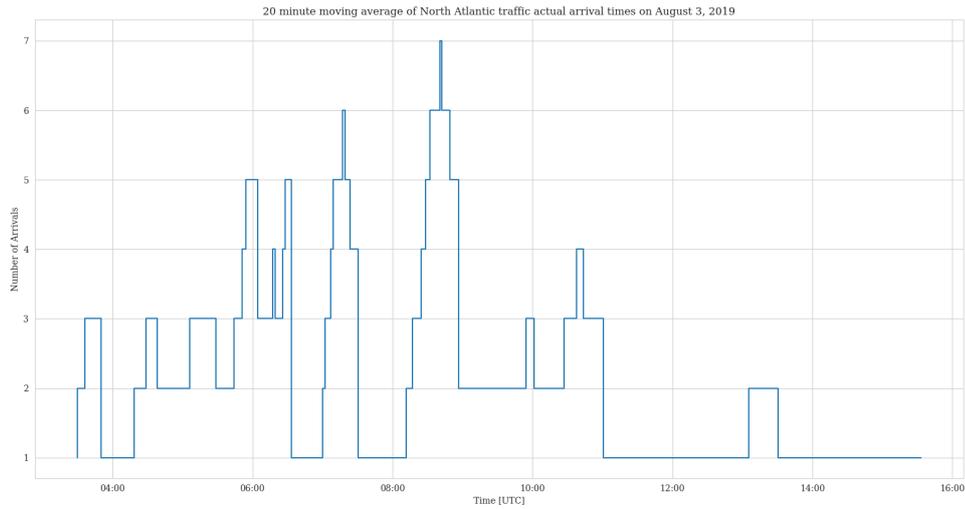


Figure 4.4: 20 minute moving average of North Atlantic traffic actual arrival times at Schiphol Airport on August 3, 2019

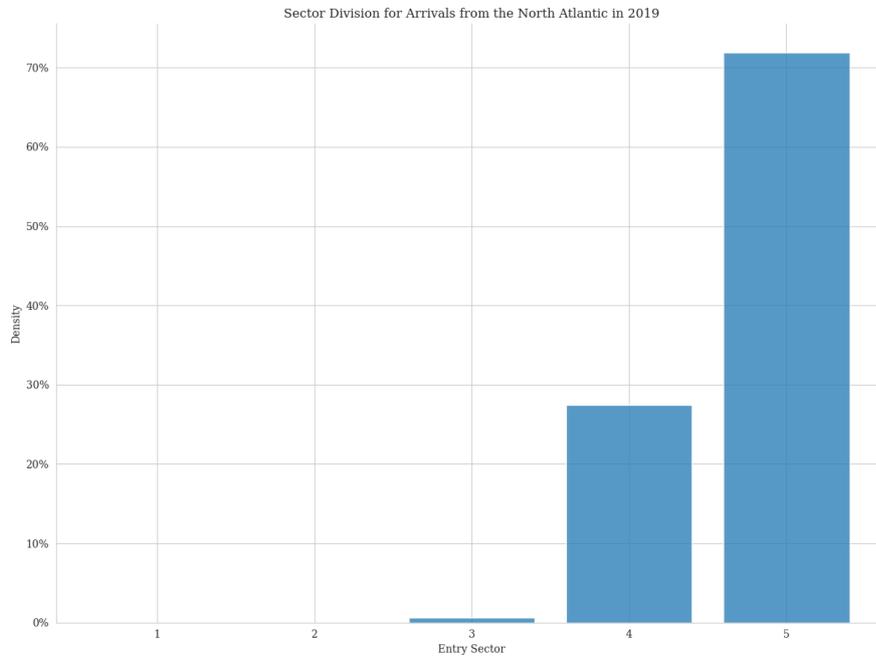


Figure 4.5: Arrival sector division for arrivals from the North Atlantic in 2019

that 4.3% of flights flying over SUGOL were subject to holdings (with a total of 3410 holdings). Furthermore, it was found that 31.2% of the flights in these holdings (1064 in total) were from the North Atlantic. It was also analysed that of all traffic flying over the SUGOL IAF, 25.1% originated from the North Atlantic Tracks. With these numbers it was found that for a flight from the North Atlantic Track in September 2019 there was a 5.3% of being subject to holding. For traffic not from the North Atlantic Tracks this was 4.0%. Therefore, a slight increase can be spotted. It should be noted however, that to get a significant result more months of data should be analysed.

# 5

## Air Traffic Information Sources

The focus of this chapter is on the uncertainty in the arrival flow. This is done by discussing the available traffic information sources and the corresponding uncertainty in the estimates they provide. This chapter starts with an introduction of the Network manager in , followed by the types of information that is (or used to be) provided by the Network Manager in section 5.2, section 5.3 and section 5.4. This is followed by the other relevant data streams during the arrival phase of a flight: the On-Line Data Interchange section 5.5 and the Radar in section 5.6.

### 5.1. Network Manager

In order to make bunching predictions, one should critically look at the information source that is used. In this section the different types of information sources, the range and their accuracy will be discussed. The first source of information about an incoming flight is the Network Manager. The Network Manager (EUROCONTROL in Europe) allows Air Navigation Service Providers, aircraft operators and airports to share and retrieve information. The information is subdivided under Flight Services, Airspace Services, Flow services and General Information Services. In this research, Flight Services is important, as this contains flight plan information and actual flight data updates. The flight plan contains, among other info, the Aircraft Identification (ACID), wake turbulence category, departure aerodrome, intended cruising speed(s), intended cruising level(s) and the intended route to be followed [29].

### 5.2. Flight Update Messages

The flight data updates are received through Flight Update Messages (FUMs). These messages provide an ANSP with updates on the flight progress. FUM messages provide the Estimated Landing Time (ELDT) for the Airport of Destination (ADES) and whether a flight has been delayed due to ATFM measures. The FUM also provides updates on the Estimated Time Over (ETO) the Coordination Point (COP) and the Initial Approach Fix (IAF). A FUM message is sent at each of the following three conditions: when the status of a flight changes (e.g. when it is airborne), when a flight is 3 hours before the ELDT, and when the ELDT changes by more than 5 minutes [32].

Note that ELDT uncertainty can run up to 5 minutes before an update is sent, which might be sufficient to get an indication at long ranges, but it might not be accurate enough to act upon. Also, the flights that are accurately predicted are not known, as no update is sent. In [57] an analysis was performed on the uncertainty of the ELDTs of the FUM messages. It was shown that for airborne flights, the spread of the error was always smaller than 20 minutes, while for the other flight statuses the spread was in the order of 1 hour for prediction horizons up to 180 minutes [57]. However, the descent phase of a flight is more complex to predict due to tactical ATCO influence and airport specific procedures. It was shown that the estimates of ETO predictions are more accurate than the ELDT predictions [32]. It is assumed that the same logic applies to the difference in accuracy for an estimate at the COP, versus an estimate at the IAF. Therefore, an estimate at the COP is assumed to be more accurate than at estimate at the IAF, as more of the complex to predict descent

phase is included in an IAF prediction.

### 5.3. Enhanced Tactical Flow Management System Flight Data Messages

Currently, the FUM messages are slowly replaced by the ETFMS Flight Data Message (EFD). The Enhanced Tactical Flow Management System (ETFMS) provides tactical data to all operational stakeholders. The purpose of the EFD messages is to provide the latest updates and the latest states of a flight within the system. The main focus of the messages is on the flight profile. The first message is sent at the moment of the Flight Plan creation, and the last when the flight is terminated or cancelled. When an aircraft is flying, updates will be sent every 15 minutes, or when the EFD data changes [33]. The EFD messages are also provided as a Business-to-business (B2B) format, which provides a constant stream of data instead of the update frequency given above. This data is currently available through LVNL. Note that this is a great improvement for the ETO estimate accuracy, as the 5 minute error margin is removed, and estimates are now updated continuously. In Tielrooij [57] it was shown that the average number of FUM message per flight varied around a median of 3. For the EFD B2B an analysis was done for arrivals in September 2020, where it was found that there are about 40 messages per flight, yielding a large change. The distribution of the message count is shown in Figure 5.1.

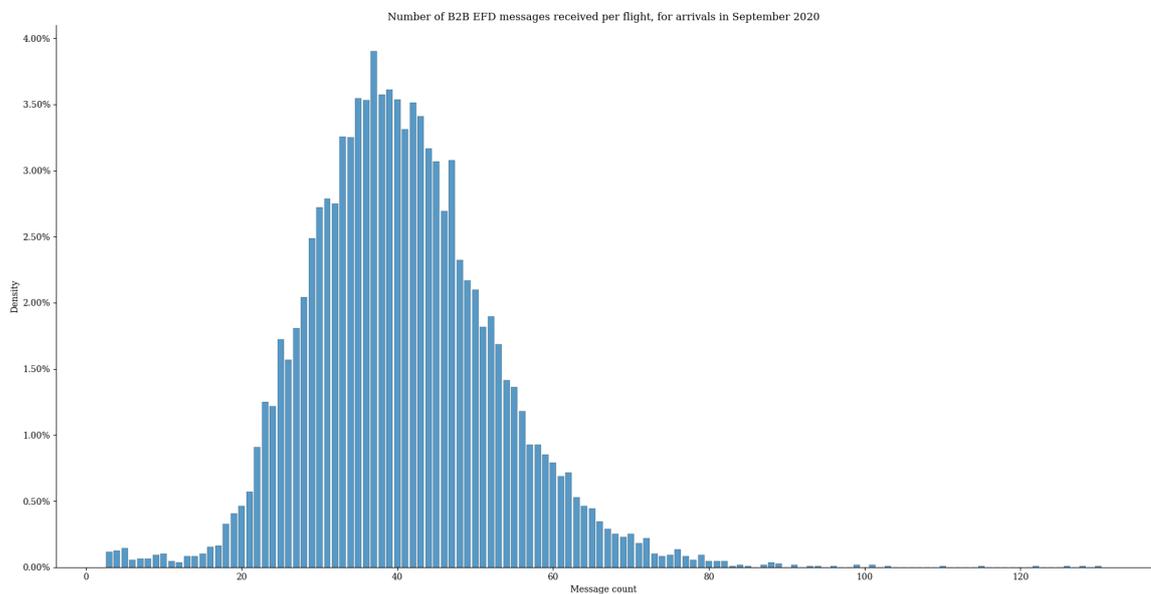


Figure 5.1: Distribution of EFD B2B message count per arriving flight in September 2020

Flights with less than 10 messages include flights from nearby airports such as Eindhoven Airport and Groningen Airport Eelde. It is known that the B2B EFD records of a flight sometimes contain a flight plan ID change around midnight, but these changes have been filtered such that the messages are grouped by the correct flight. It is assumed that the remaining flights with less than 10 messages, and not from nearby airports, are due to recording and processing errors, as no clear pattern was found in either the departure airports, time of day or other information contained in these B2B EFD messages.

Another difference with the FUM messages is that more data is available. The fields that are relevant to the current research are: ADEP, ADES, Aircraft ID, Aircraft type, Airline operator, Airspaces, ATFM delay, Route Points with corresponding ETOs, ETA, WTC, Flight Status, Departure Airport Type, CDM Status. The CDM State will be elaborated on in section 5.4. The accuracy of the messages is assumed to be related with the categorical Flight State field. In this data field there are 8 types of states [33]:

1. *Flight Plan Filed (FI)*: Indicating the basic state for a flight before further information is known
2. *Filed Slot Allocated (FS)*: The flight has been regulated due to ATFM measures, but a slot has not yet been published

3. *Slot Issued (SI)*: The flight is regulated and a slot is issued
4. *Tactical Activated (TA)*: It is assumed that the flight is airborne by ETfMs, but a confirmation has not been received from ATC
5. *Air Traffic Control Activated (AA)*: The flight is confirmed airborne
6. *Cancelled (CA)*: The flight has been cancelled
7. *Terminated (TA)*: The flight is considered to be terminated
8. *Suspended (SU)*: The flight has been temporarily suspended, possibly due to a lack of updates

## 5.4. Airport Collaborative Decision Making

The Airport Collaborative Decision Making (A-CDM) state field in the EFD messages contains Departure Planning Information (DPI) used in the flight profile [33]. The goal of a DPI status is to provide accurate flight data for aircraft that are not airborne yet, and it contains information originating from a wide range of airport systems. The main data it contains is related to estimations of take-off time and taxi-time (EXIT), the Standard Instrument Departure (SID), the Target Off-Block Time (TOBT, the moment an aircraft is ready for start up and push back at the gate and waiting for clearance) and the Target Start-up Approval Time (TSAT, the time an aircraft can expect to receive the start up clearance and push back). There are 5 types of states in the A-CDM status field within the EFD messages, where each message corresponds to an increase in accuracy. The time at which each message will be sent is schematically shown in Figure 5.2. The definition of these messages are as following [24].

1. *Predicted Departure Planning Information (P-DPI)*: The purpose of this message is to provide Airport Data before the A-CDM process starts. It provides relevant data as soon as it becomes available. This state is shown in the CDM field as 'DPIEXPECTED'.
2. *Early Departure Planning Information (E-DPI)*: This type of message is to confirm that an airport slot and flight plan for a flight have been correlated with an airport, in order to prevent duplicated flight plans and eliminating ghost flights. It is shown in the CDM field as 'ESTIMATED'.
3. *Target Departure Planning Information - Target (T-DPI-t)*. The purpose of this message is to provide a Take-Off Time (TOT) based on Estimated Landing Time (ELDT) and the EXIT of the inbound flight, and estimations of the turn around, off-block and taxi time of the outbound flight. This state is shown in the CDM field as 'TARGETED'.
4. *Target Departure Planning Information - Sequenced (T-DPI-s)*: This message contains information on the TOT based on calculations of the Pre-Departure Sequencer, using the TOBT and the taxi time. In the CDM field this state is shown as 'PRESEQUENCED'.
5. *Air Traffic Control Departure Planning Information (ATC-DPI)*: This message serves to inform that the flight has 'off-blocked', which means that the flight is under ATC control and is taxiing to take-off. This message therefore also provides an estimate of the Take-Off Time of higher accuracy than the previous messages. This state is shown as 'ACTUALOFFBLOCK'.

Not all airports have A-CDM capabilities. There are three types of airports, classified by their A-CDM capabilities: Standard Airport, Advanced ATC TWR Airport, and CDM Airport. Standard Airports don't have any A-CDM capabilities, and therefore do not provide any of the mentioned DPI messages. Advanced ATC TWR are airports that have not fully implemented the A-CDM processes, but do provide the ATC-DPI message at the actual off-block time, and therefore provide accurate estimate information on the taxi-time and the take-off time. The only other message that is provided is the P-DPI, therefore, no accurate information of the time between this message and the ATC-DPI is known. The third type of airports are CDM Airports, which have fully integrated the Collaborative Decision Making process and provide all types of DPI messages discussed above [23].

An analysis has been done with the NM EFD B2B data of September 2019 with regards to the airport type the messages originated from. In Table 5.1 it can be seen that the majority of messages is sent from Standard

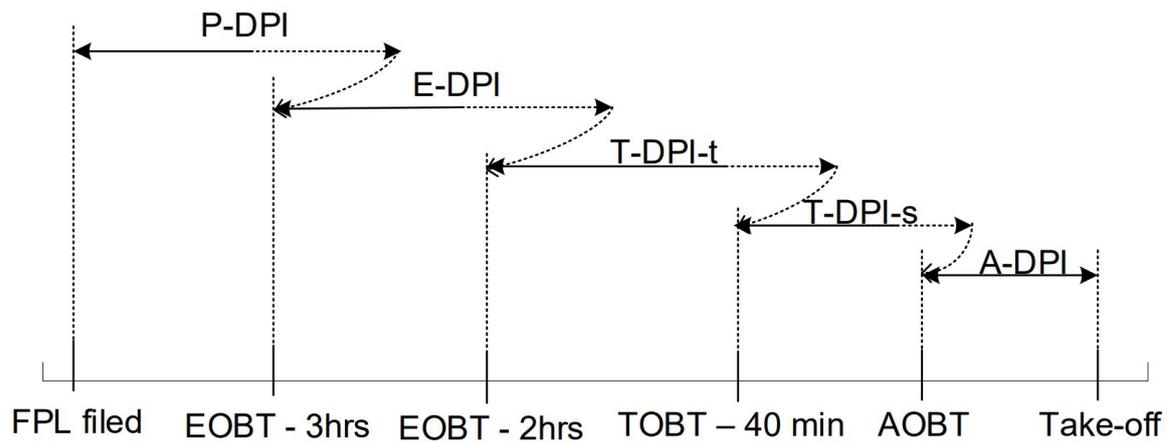


Figure 5.2: A schematic overview of the times at which each DPI is available. Source: [24].

	Percentage of all messages	Percentage of all flights
Standard Airport	46.4 %	47.0 %
Advanced ATC TWR Airport	11.9 %	12.0 %
CDM Airport	41.7 %	41.0 %

Table 5.1: The percentage of all messages and all flights originating from the three airport types, for arrivals in September 2020

Airports, followed by fully integrated CDM Airports and lastly the Advanced ATC TWR Airports. To check if there was a difference in the amount of messages sent per flight, the airport type for all incoming flight was also checked (instead of the origin of each message). It can be seen that the percentages are very similar. For all airport types, the message count per follows a similar distribution as shown in Figure 5.1. The differences between the airport type have been shown in Figure 5.3. It can be seen that for all airport types the median of the amount of messages per flight is 39 or 40, therefore, the airport type does not have a large influence on how many messages are sent, only on the types and thus the accuracy of the messages.

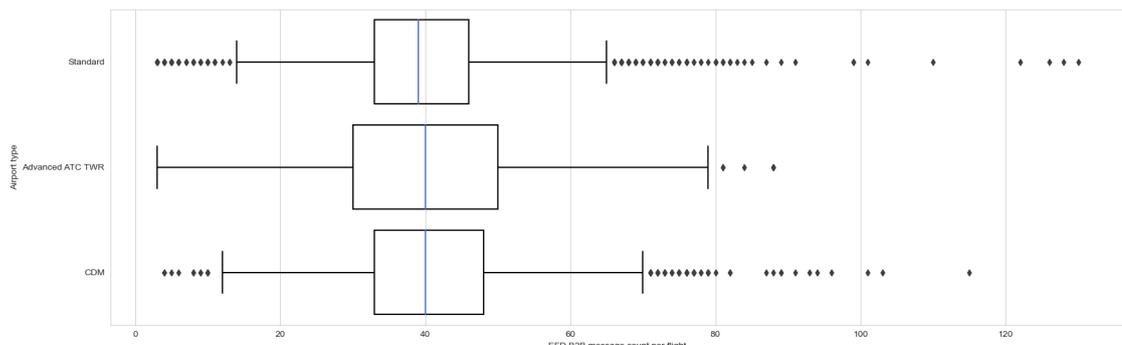


Figure 5.3: EFD B2B message count per arriving flight in September 2020, grouped by airport type

### 5.5. On-Line Data Interchange

Another source of flight information is through the On-Line Data Interchange (OLDI) protocol. The OLDI messages are used for data information exchange between two adjacent centres. These messages are automatically sent in order to reduce the controller workload by reducing verbal coordination and routine actions. There are several messages, each sent at different conditions. The main ones are the Advance Boundary Information Message (ABI), and the Activate Message (ACT).

The ABI is an initial notification message that provides flight plan data, facilitates correlation of radar

tracks (by providing the SSR Code), and provides the entry COP and the corresponding ETO. A message contains the ACID, SSR Code, ETO COP, Aircraft Type and the current flight plan data [16]. An updated ABI message may be sent until the ACT message is received, after which no ABI message is sent anymore. At the receipt of the first ABI message, the radar will be correlated with a flight and the pre-planning of the Arrival Manager starts (see section 3.2). The ACT message is sent to provide details of a flight prior to the transfer of control from one ATC unit to the next. It updates the flight plan data and it provides the transfer conditions to the receiving ATC unit. The fields in the message are the same as in the ABI message stated above, however, as it is sent later, these fields are updated [16].

The time at which an ABI or ACT is sent is dependent on the agreements between two adjacent centres. For aircraft flying from the Scottish Area Control towards the Amsterdam Area Control in Sector 5, the ABI needs to be sent 30 minutes prior to ETO COP, and the ACT at 18 minutes prior to ETO COP. For aircraft flying from the London Area Control to the Amsterdam Area control in Sector 4 and 5, the ABI needs to be sent at 30 minutes prior to ETO COP, and the ACT 12 or 15 minutes prior to ETO COP, depending on which COP is flown over. The ETO in the message is based on radar data from the adjacent centre, and therefore the estimations are of high accuracy and assumed to be better than those of the Network Manager.

## 5.6. Radar

Radar is the primary means of surveillance in ATC, therefore, when a flight is in radar range, all trajectory predictions will be based on this radar data. Two types of radar are used: the Primary Surveillance Radar (PR) and the Secondary Surveillance Radar (SSR). The PR transmits pulses of radio-frequency energy. The signals scattered back by the surface of an aircraft are received back, from which the slant range and the azimuth of the aircraft relative to the radar are determined. The surface range can be determined by using the slant range, the Flight Level and weather information. The SSR Mode A and Mode C transmit signals that interrogate a transponder of an aircraft. In Mode A the transponder of an aircraft replies with the Aircraft ID. In Mode C the aircraft replies with the Flight Level. The SSR Mode S is more advanced as each aircraft has a unique 24 bit address such that specific aircraft can be interrogated. In this mode, ACID, Indicated Airspeed (IAS), Pilot Selected Level (PSL) for altitude and heading can be retrieved. Combined, the position, heading, IAS, PSL, Flight Level and Aircraft ID are identified to a high degree of accuracy.

The clear benefit of using radar data is the accuracy of the acquired data. However, the range is limited to the vicinity of the Flight Information Region Amsterdam (FIRAM). In Figure 5.4 the SSR mode S coverage has been shown for Flight Level 245. As no radar data is exchanged between adjacent centres, and therefore other data sources should be used for information at long ranges.

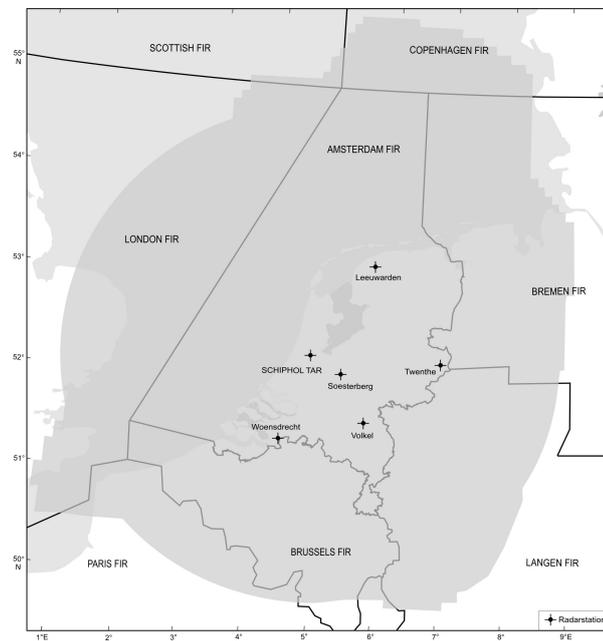


Figure 5.4: SSR Mode S coverage at FL245 (source: [38])

# 6

## Trajectory Prediction

After the information sources and the corresponding accuracy have been discussed in the previous chapter, this chapter will dive deeper into the way the trajectory prediction estimates in these information sources are constructed. The chapter starts with a general overview of state propagation in section 6.1, followed by an analysis of the mathematical models in section 6.2 and concluded with the trajectory predictor uncertainty in section 6.3.

### 6.1. State Propagation

The concept of trajectory prediction is based on predicting the future trajectory of an aircraft based on the current state of the aircraft and meteorological conditions. The ability to properly predict the trajectory of an aircraft is at the core of many systems within the ATM field and is used for different objectives. Even though the applications for the computed trajectories might be different, the underlying prediction processes are based on the same principles. At the core of trajectory prediction is state propagation. Four types of state propagation have been identified, schematically shown in Figure 6.1. The four types of state propagation methods are [34]:

1. Nominal state propagation method. In this method, the current state of the aircraft is projected along a single trajectory without taking into account parameters of uncertainty. An example is the straightforward extrapolation of the velocity vector. As this is a very simple model, the prediction accuracy is low and the look-ahead time consists of a few seconds.
2. Worst-case method. This method considers the maximum range of possible manoeuvres, which can serve as a conservative trigger for conflict alerts or might be used for applications where an aircraft is constrained to remain within specified manoeuvring bounds. This method is also bound by a short look-ahead time as the range of possibilities increases dramatically with the look-ahead time.
3. Probabilistic method. This method models the potential variations of the future trajectory. This can be used to add a positional error to a nominal trajectory, or to determine the probabilities of a set of possible future trajectories. The main benefit of this method is that decisions can be based on the probability that a trajectory will be flown.
4. Intent based method. In this method the state of the aircraft and corresponding uncertainties are modelled. The intent of a flight consists of the goals, constraints and preferences applicable to a flight [3]. In Figure 6.1 the intent based state projection is visualised as an extension of the nominal based state projection method, however, the other methods could be combined with the intent based method as well. The benefit of this method is that it increases the prediction accuracy of the previous methods, and therefore allows for longer look-ahead times.

In the operational context in which the current research operates, the intent information is provided in several forms (see chapter 5): the flight plan contains the route intent, the EFD Messages contains route updates intent in-flight and the OLDI messages contain the Entry COP. The aircraft intent can also be derived from the geometry of the airspace. Examples are the altitude limits at the COPs and at the IAF, as discussed

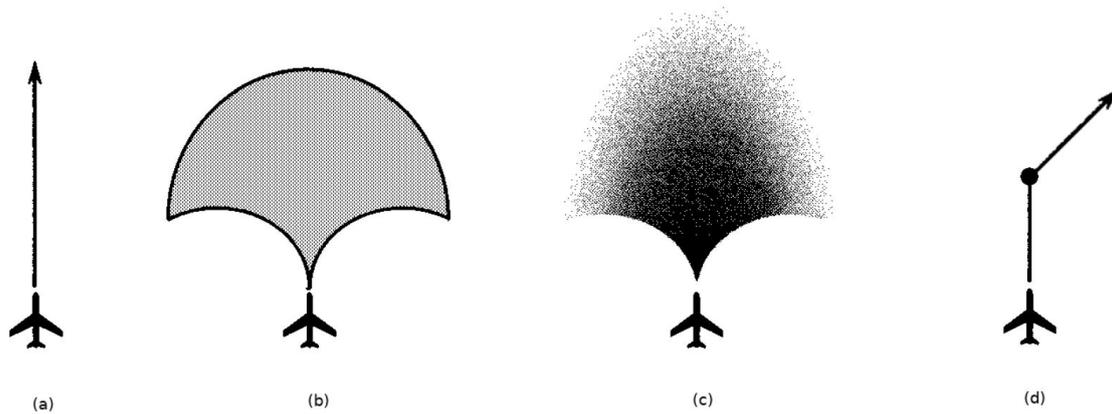


Figure 6.1: Visual representation of nominal (a), worst-case (b), probabilistic (c), and intent based state projection (d) (sources: [47] [34])

in section 2.1. Other forms of intent consist of the arrival procedures, airline preferences and aircraft specific performances. It should be noted that flight intent can be ambiguous, as it does not fully specify how the aircraft is operated during the flight. Commonly, there are several different trajectories that still comply with the intent. Especially for geometry based intent the trajectories might differ significantly, however even for route intent the trajectories might differ due to the way a turn is taken, or how an aircraft flies in between specified waypoints.

## 6.2. Mathematical Model

For all of the aforementioned state prediction models there are six different approaches for modelling of the aircraft motion [3]:

1. Six degree of freedom approach (kinetic). This approach is the approach with the highest fidelity, consisting of all degrees of freedom acting on a rigid-body model of the aircraft.
2. Point-mass approach (kinetic). In this approach the model of the aircraft body is modelled as a point-mass system and only linear motion is considered (3 degrees of freedom), as angular motion is excluded.
3. Energy state approach (kinetic). This method is based on the rate of work and the change of total energy for a 3 degrees of freedom model of the aircraft.
4. Macroscopic approach (kinematic). In this method the motion of an aircraft is based on relations provided by performance tables. An example is the descent rate as a function of altitude. The difference between this approach and the points-mass and the energy state approaches is that this approach is kinematic and does not derive the motion from the forces acting on the aircraft.
5. Hybrid approach (kinematic). This method is similar to the Macroscopic approach, but instead of relying on relations provided in performance tables, it takes performance parameters from historical data based on the aircraft type and carrier. It is hybrid in the sense that it is partly based on kinematics and partly data driven.
6. Data driven approach: For fully data driven methods there is a wide range of options that are based on a model that has been trained on historical data.

A schematic representation of the approaches is shown in Figure 6.2. The figure shows which motions are included in the state vector. It can be seen that only the six degrees of freedom model includes angular motion in the state vector. for the point-mass model the orientation is also required, but it is not included in the state vector and therefore not included in the diagram [3]. The data driven approach is not included in the diagram, as a pure data driven model could also include the energy state. However, it is assumed that

most pure data driven models for trajectory prediction can be represented as the macroscopic and hybrid method representation.

Each of the approaches above differ in mathematical complexity and fidelity. The six degrees of freedom has the highest fidelity, but for this model an extensive aerodynamic model, knowledge of higher order dynamics and the aircraft's inertial tensor are required. This information however, is not (accurately) available. A three degrees of freedom is a simplified approach, where only the aerodynamic forces are required for the aerodynamic model. For a pure data driven approach the fidelity strongly depends on the inputs used in the model, and how the model is trained.

In current ATM systems the six degrees of freedom model is not necessary, as the point-mass approach provides sufficient fidelity for ATM purposes [54] [31]. The benefit of the lower fidelity models is that the trajectory prediction is simplified and the required performance models are easier to obtain than for the high fidelity models. The point-mass is often reduced to the macroscopic approach [3], or to the hybrid approach that is currently in use at LVNL [37]. In the Hybrid approach of LVNL descend speeds, rate of descend and deceleration rates are based on historical data and depend on aircraft type, airline operator and flight segment. Benefits of this method are that it is less sensitive to initial conditions and that it allows to incorporate the subtleties of operator specific performance differences.

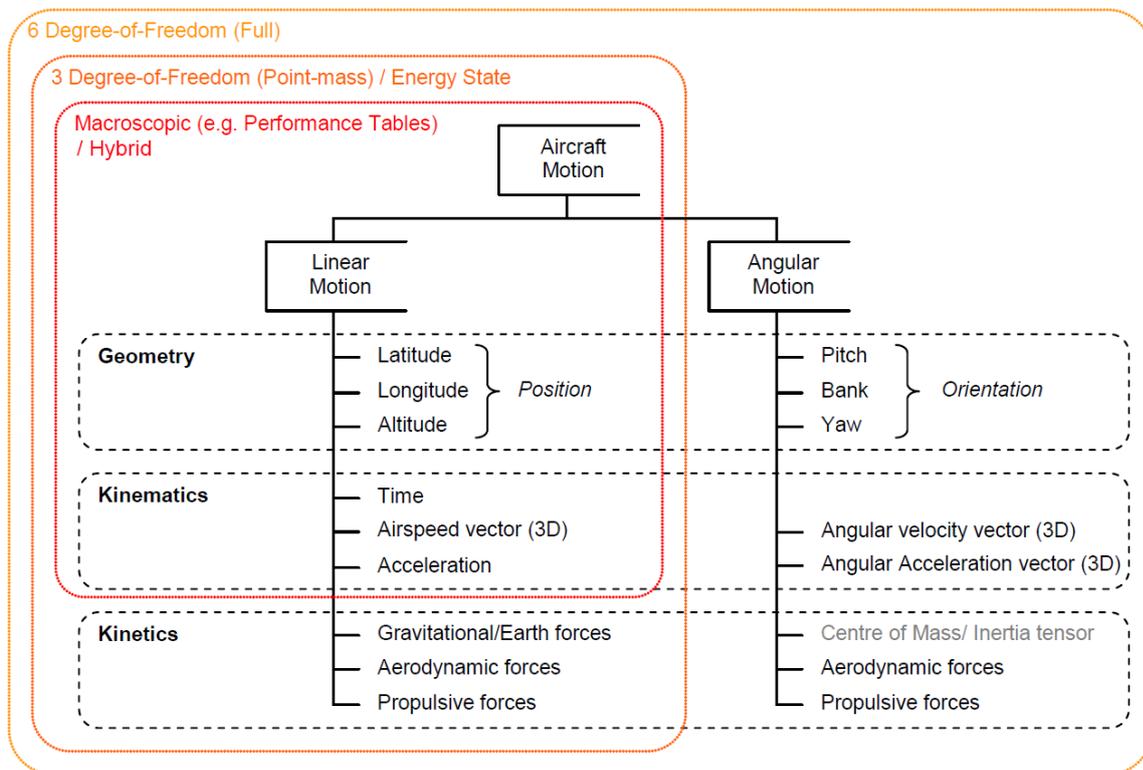


Figure 6.2: A schematic representation of five different approaches for modelling of aircraft motion (source: [3])

### 6.3. Trajectory Predictor Uncertainty

Due to the dynamic nature of the Air Traffic Management system there will always be a degree of uncertainty as to which predictions of air traffic arrival times or trajectories can be made. When considering ground-based predictions for arrival times, it was shown that the degree in prediction uncertainty increases with the distance left to the airport [57]. The same conclusion holds for airborne systems. The standard deviation of the arrival time predictions by the airborne Flight Management System (FMS) increases with the flight time left before arrival [4].

In the current ATM system there is a lack of information sharing between airborne systems and ground based Air Navigation Service Providers (ANSPs). The ground based systems lack information of both the aircraft state and intent, while an airborne aircraft lacks information that ground based systems have, such as accurate weather models, airspace capacity and the ATCO intent. The System Wide Information Management (SWIM) concept addresses the lack of information sharing and aims to increase the predictability of traffic streams by enabling information sharing between airborne aircraft and ground based ANSPs [61]. Through SWIM, Single European Sky (SESAR) aims to move from the current operational infrastructure to Performance-Based Operations. The stages for development are the following [61] [19]:

- *Time-Based Operations*: the time of arrival at specific points is actively monitored by both airborne systems and ground-based systems.
- *Trajectory Based Operations (TBO)*: the (intended) 4D trajectory of an aircraft is known by Air Traffic Management during all phases of flight, and this trajectory is continuously updated and negotiated to achieve optimum system outcome, with minimum deviation from the intended trajectory.
- *Performance-Based Operations (PBO)*: a fully integrated, network-centric, and collaborative air/ground ATM system in which all relevant information is known by both airborne and ground-based systems.

Currently, there is no down-linking trajectory intent or up-linking of data, and therefore real Trajectory Based Operations have not yet been implemented. Since there is limited knowledge available by ground-based systems, trajectories have to be predicted in order for aircraft to be controlled efficiently. The accuracy of arrival time predictions is impeded by several different causes. The main sources of error being the intent, initial conditions, aircraft performance model, trajectory integration, model fidelity, departure uncertainty, and meteorological forecast errors [3]. Intent errors come in several forms, but the main intent error is a deviation from the filed flight plan. An example of this is flying a 'direct' to a waypoint, or a given heading, instead of following a route along waypoints. A direct is either prescribed by the controller, or requested by the pilot. Other examples of intent are the location of the Top of Descent (TOD) and speed settings. The initial conditions, aircraft performance model, trajectory integration and model fidelity errors are model errors, that may change when selecting different types of Trajectory Predictors. Departure uncertainty errors are caused by departure delays which can be caused by a wide range of factors [57]. Meteorological errors are due to prediction errors in the forecasts used by the Trajectory Predictor. It should be noted that logically, the fidelity of the meteorological forecasts has a large impact on these prediction errors.

### 6.3.1. Trajectory Predictor Sensitivity Analysis

To determine the influence of uncertainties on the Trajectory Predictor model output, a sensitivity analysis can be performed. Ideally, a sensitivity analysis should be performed in tandem with an uncertainty analysis [52]. An uncertainty analysis focuses on quantifying the uncertainty in the model output, and helps in identifying the sensitive inputs. In the uncertainty analysis the inputs of the model are sampled, and the uncertainty in these samples is propagated through a model. This allows to estimate the sensitivity of the model, as the sensitivity analysis studies how the uncertainty in the output(s) can be attributed to the uncertainty in the inputs of the model.

Two types of sensitivity analysis have to be explored: Local and global sensitivity analysis. In a local sensitivity analysis one input can be varied at a time, while keeping others at the nominal value. From this, a conclusion can be drawn about the inputs that have an influence on the output, and to what extent. This method is commonly performed as a derivative-based approach, or by taking input steps around a nominal value. Examples of this approach in the context of trajectory prediction are [22] and [39]. The limitations of this approach however, is that it is not suitable for uncertain model inputs and for non-linear models, it also does not provide information about the other input factors. By assuming the other parameters at their nominal value, the interactions between inputs are not explored [52]. In a global sensitivity analysis, all inputs are varied at the same time from a corresponding input distribution. Besides finding the sensitivity of the output to an input, this method allows for examination of the interactions between input parameters, and their effect on the output of the model. A common method for global sensitivity analysis is the Monte Carlo method. This method is based on randomly sampling input values based on distributions acquired from observed data. The output metric that is selected, is used for the uncertainty analysis by inspecting the output distribution and statistics. From this the average output, its standard deviation, the quantiles of the distribution, confidence bounds, scatter plots, and the correlation coefficient can be computed to assist in

inspecting the sensitivity of the output to the input factors [52].

In [42] the effect of uncertainty on an en-route descent advisory model and an aircraft execution model was tested using a Monte Carlo simulation. This Monte Carlo performance analysis was used to analyse the robustness of advisories to input uncertainties. Furthermore, this analysis allowed to identify the parameters whose accuracy was limiting the total advisory model accuracy. Similar to the current research, the key metric to test the model performance was the meter fix crossing time error. The input parameter statistics that were used for the Monte Carlo error analysis were obtained from literature and based on field test measurements, simulation or database analyses. These statistics were assumed to be normally distributed, and from the Monte Carlo runs it was concluded that the parameters that had the highest impact on the meter fix crossing time error were initial aircraft speed estimates and the wind speed estimates. Other examples of a global sensitivity analysis using normally distributed input parameters are [60] and [7], yielding similar results.

However, the assumption of normality in the distributions of the input parameters rarely holds as shown in [50], where the distributions of the input parameters were obtained from data, and not assumed to be normally distributed. The key performance indicators were the along-track error, the cross-track error and the altitude error. The parameters having the most influence on the prediction uncertainty were the ATC intent, the wind conditions, the Calibrated Airspeed (CAS) and Mach number speed settings, the vertical speed settings and the temporary level-offs. However, the ATC intent and the level-off distributions were obtained from general data, although these inputs depend on operational procedures, and thus highly airspace dependent. Therefore, with a specific case-study, these effects could be analysed more accurately.

A final remark is that when performing research on the accuracy of trajectory prediction, the inputs can be based on conclusions of previous research. Selecting inputs which have proven to significantly influence the predictability uncertainty, is in line with the fact that in most uncertainty analyses a few factors create almost all the uncertainty while the majority only making a negligible contribution. Therefore, the addition of extra variables to the uncertainty analysis may add to the completeness of the model, but does not contribute to the variance in the output in a significant way and therefore can be left out [52].

In [42] the analysis was not only focused on the trajectory prediction, but the execution of an advisory was also analysed. For this, a simulation based Monte Carlo analysis was done, where the parameters mainly consisted of Flight Technical Errors, initial condition accuracy, and atmospheric model accuracy of the Flight Management System (FMS). Furthermore, truncation errors of advisories were analysed and the uncertainties of aircraft weight and trust-drag models were taken into account. The wind model used by the FMS, which is less accurate than the wind models used by ground-based tools, was proven to cause the highest uncertainty in the trajectory prediction. Another conclusion was that weight estimates that were too low caused the descent phase of the trajectory to be less accurate when an aircraft was following an advisory. It should be noted that this is a downside of using a trajectory predictor based on a Point-mass model, and that this might be negated by using a Hybrid trajectory predictor as this does not rely on weight estimates. Another downside of the analysis was that intent was not included in the execution phase. When extending the look-ahead time of the trajectory prediction, the aircraft and ATC intent become important parameters as direct headings to way points might be flown, instead of following the original route. Especially in a complex airspace such as in Western Europe, multiple ATC centres might be responsible for an aircraft in the final cruise and initial descent phases.



# 7

## Summary of the Preliminary Study

In the arrival process the Arrival Manager is of key importance as it assists the Air Traffic Controller in sequencing, metering and metering of air traffic. The use of an Arrival Manager results in a high runway throughput efficiency and it lowers the workload for Air Traffic Controllers. It constructs landing slots based on the Expected Time of Arrival for a flight, based on which it provides Expected Approach Times for aircraft over the Initial Approach Fix. As a result of this, flights often need to absorb delay in the Control Area before descending to the Terminal Manoeuvring Area. The Arrival Manager planning at Amsterdam Airport Schiphol is provided at 14 minutes before an aircraft is predicted to fly over the Initial Approach Fix, called the 'freeze horizon'. Air Traffic Controllers do not influence an aircraft before this point in time, as doing this would mean the planning itself would be influenced, as the Expected Time of Arrival is changed in the pre-planning by changing the course of an aircraft. This could result in extra delay for an aircraft and compromise the flight efficiency. Therefore, the location of the freeze horizon determines how much airspace is left for Air Traffic Controllers to influence an aircraft and how efficiently a aircraft can be guided to the Initial Approach Fix at the correct point in time.

There is ongoing research in the extension of the Arrival Manager range, as this might increase the Expected Approach Time adherence and increase flight efficiency. However, there are four main limitations to extending the range: the traffic information is limited at long ranges, the reliability of trajectory predictions decreases with the extension of the range, aircraft might be under control of other Air Traffic Control Centres and flights might pop-up within the Arrival Manager range, resulting in inefficient plannings and requiring manual adjustments. This research aims to address all limitations by proposing a strategy to influence air traffic before it enters the sector, in such a way that the efficiency of the arrival flow is increased.

Traffic peaks in the arrivals are formed due to the peak modes in the capacity declaration, which is established twice a year, and the occurrence of these peaks might result in high workloads and the formation of bunching. In general, excessive peaks in sector demand that exceed the capacity are flattened by imposed Air Traffic Flow Management delay. In literature however, it was shown that margins around the Air Traffic Flow Management take-off-times and compensation effects (flying faster to make up for the delay) might even amplify the effect of bunching. Bunching occurs when multiple aircraft approach an airspace or a given point in the airspace within a short timespan. In general, the term bunching is used to describe a situation in which the capacity at an airspace or at a given point, is higher than the imposed limits. This results in excessive workload as tactically solving it is done using inefficient manoeuvring techniques such as vectoring and airborne holding patterns. In literature it was shown that bunching can also occur during normal traffic patterns, therefore an effort should be made to predict and solve bunching alongside the current Air Traffic Flow Management measures. It was chosen to analyse bunching at the Initial Approach Fixes as this is the point where aircraft are merged towards and the point for which the Expected Approach Times are given by the Arrival Manager. The technical definition for bunching in this research is a time-based definition, so that bunching can be defined as two or more aircraft estimated to arrive at a point in the airspace within a given amount of time, which could be based on the Wake Turbulence Category separations of the involved aircraft.

This research also did a review of the available information sources that provide predictions for the tra-

jectory of arriving aircraft. It was found that the previous literature on analysing the uncertainty of these predictions was outdated since new types of information sources were introduced. The main source identified for aircraft at a range of a few hours was the Enhanced Tactical Flow Management System Flight Data Messages. The data fields in these messages that are expected to have the most influence on the accuracy of the predictions are the flight state, consisting of 8 states, and the Airport Collaborative Decision Making status field consisting of 5 states for aircraft still on the ground. It was found that the median amount of messages per aircraft was about 40 messages.

# II

## Thesis After Midterm



# 8

## Uncertainty Analysis

The overall goal of the model is to achieve higher predictability in the arrival process by creating a stable incoming stream of air traffic. By solving bunching before a traffic stream enters the sector, the flight efficiency can be increased and the controller task load might be decreased. In the the previous chapters, the operational context of such a model and several theoretical frameworks and methods have been researched. This chapter aims to bring the aforementioned operational context, theoretical frameworks and methods together and presents the methodology that has been undertaken in this research. A schematic overview of the steps taken in this chapter are shown in Figure 8.1.

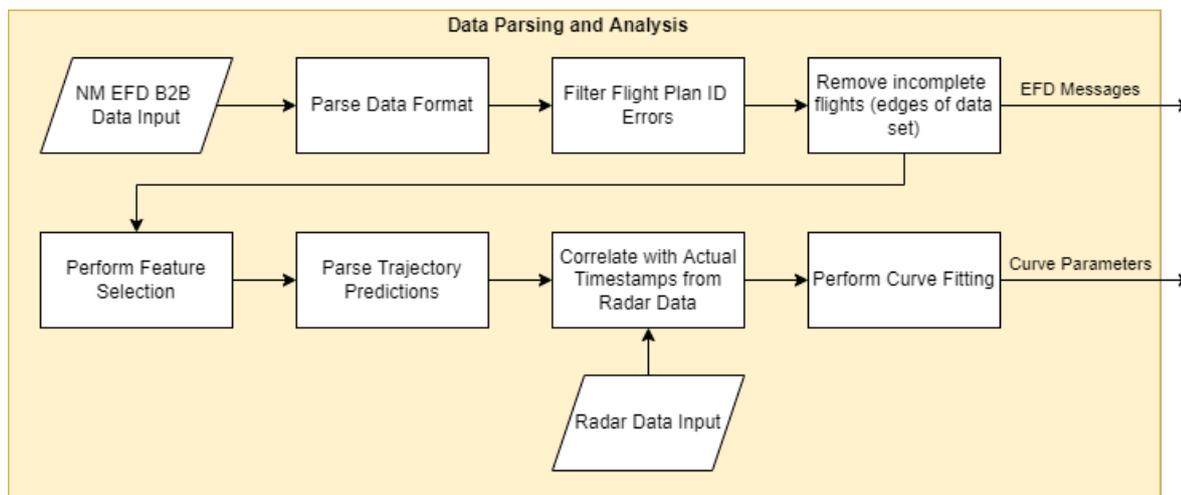


Figure 8.1: Systematic overview of the data processing steps

### 8.1. Uncertainty Analysis

In order to gain insight in the uncertainty in the arrival process, the different streams of information have been analysed in chapter 5. It was concluded that the main information source that is used at long ranges was information from the Network Manager in the form of the EFD Messages. For the predecessor of this information source, the FUM messages, in [57] an analysis was performed for the uncertainty in the arrival times. The goal of that research was to support decision support tools in the presence of prediction uncertainty by predicting that uncertainty itself. It was concluded that input error propagation and empirical Monte-Carlo approaches were not suitable for real-time applications. The author therefore suggested an experience-based approach using empirical information, relating prediction errors to the actual information, based on previous research on sector capacity prediction [59]. The benefit of this method is that the information of the prediction errors can be used in real-time applications and that all relevant uncertainty parameters are captured within the uncertainty predictions [57]. Therefore, the experience-based approach using empirical information will also be used in this research.

### 8.1.1. Data Availability

The data that will be used are the trajectory predictions from the stream of B2B EFD Messages that are sent by the Network Manager to LVNL. This data is shared by LVNL and the first month of data (September 2020) has already been processed and used for analysis throughout this report. There is one downside with using this data however, as LVNL has only started recording the data from August 2020. The data that was available started from August 1<sup>st</sup> (2020) to June 24<sup>th</sup> (2021). This means that the data recording has started in the midst of the COVID crisis, resulting in a decrease of air traffic arrivals of more than 50%<sup>1,2</sup>. To grasp the effect this might have on the data, one could look at the main trajectory prediction error sources as mentioned in section 6.3. These were the aircraft intent, departure uncertainty, initial condition errors, aircraft performance model errors, model fidelity and meteorological forecast fidelity [3]. In section 6.3 it was shown that sensitivity analyses had proven that the intent, (initial) aircraft speed estimates and the wind speed estimates had the most effect on the inaccuracy of the estimates [42] [50]. It is expected that the parameters that are most likely to be influenced by the low traffic density of the data are the departure uncertainty and the intent (for both ATC and the aircraft).

For the aircraft intent one could argue that since more direct routes are flown the error of the predictions would be larger. The reason behind this is that it is assumed that the trajectory predictions are based on the regular track miles when aircraft fly waypoint-to-waypoint. Flying direct routes would result in estimates that are later than the actual arrival times, and therefore have a larger error than during normal traffic patterns. Using this uncertainty analysis during normal traffic (e.g. 2019 traffic) would mean that uncertainty predictions would be too pessimistic, and the trajectory predictions might be more accurate than the uncertainty predictions shows.

For the departure uncertainty however, it is assumed that since there are less aircraft at the airport, the stability of operations is higher and therefore the trajectory prediction estimate accuracy would be higher. Using this uncertainty analysis during normal traffic would mean that the uncertainty predictions would be too optimistic. The same conclusion could be drawn for ATC intent, as in higher traffic densities an ATCO has to intervene with traffic more and therefore the uncertainty would increase in such situations.

### 8.1.2. Truth Data filtering

In [57] a Probability Density Function (PDF) was found for a message based on historical data. The PDF was based on errors between the predictions and the actual arrival times. Using these actual arrival times resulted in an analysis where all tactical uncertainties were included in the analysis. A difference between that mentioned research and the current research is that the former purely focuses on runway arrival time estimates from the perspective of the Network Manager. It should be noted that from the moment an aircraft enters the FIRAM, controllers can affect its landing time based on the planning from the AMAN. Therefore, the actual arrival times used as truth data include deviations from the prediction that are not purely due to uncertainty in the Network Manager estimates, but that are due to tactical sequencing and metering actions at the arrival airport. In the debunching context of the current research, the focus is on the predictions at the IAF instead of the runway. For these predictions at the IAF, prediction errors due to tactical manoeuvring are still included from the freeze horizon to the IAF. This AMAN influence can be considered as a source of uncertainty from the perspective of the Network Manager as the planning is not known in advance. However, in the current context this influence should be excluded, as the uncertainty is analysed from the perspective of the airport of destination and corresponding ATC centre, and for the debunching concept the times at the IAF need to be known, regardless of the potential influence the AMAN will have.

Therefore, the influence from tactical deviations in the sector should be filtered from the actual arrival times at the IAF. To achieve this, two methods were explored. The first method was to adjust the actual arrival times using a filter that identifies the difference between the flown track-miles in the sector, and the actual track-miles between the COPs and the IAFs. Flights with more track-miles than the actual track-miles are assumed to be influenced by ATC. A flight with less track-miles than the actual track-miles is assumed to have flown direct to a waypoint, flying faster than the time that was originally planned. Based on the relation between the actual and the flown track-miles, the actual time over the IAF could be adjusted. However, in order to adjust the flying time accordingly, the location of the deviation in flown track-miles should be known,

<sup>1</sup><https://www.schiphol.nl/en/schiphol-group/page/traffic-review/> (accessed: 30-06-2021)

<sup>2</sup><https://www.eurocontrol.int/covid19> (accessed: 30-06-2021)

as well as the corresponding aircraft speed and altitude. Furthermore, the flown track-miles are influenced by whether an aircraft flies directly over a COP, or at a large distance abeam to it. Flying abeam would result in different values for the actual track-miles, leading to inaccurate results. To get an insight in what effect this might have on the track-miles, a radar data analysis has been performed for flights in September 2019. In Figure 8.2 the closest distances between all aircraft and their corresponding COP has been shown, defined as the distance abeam to the COP.

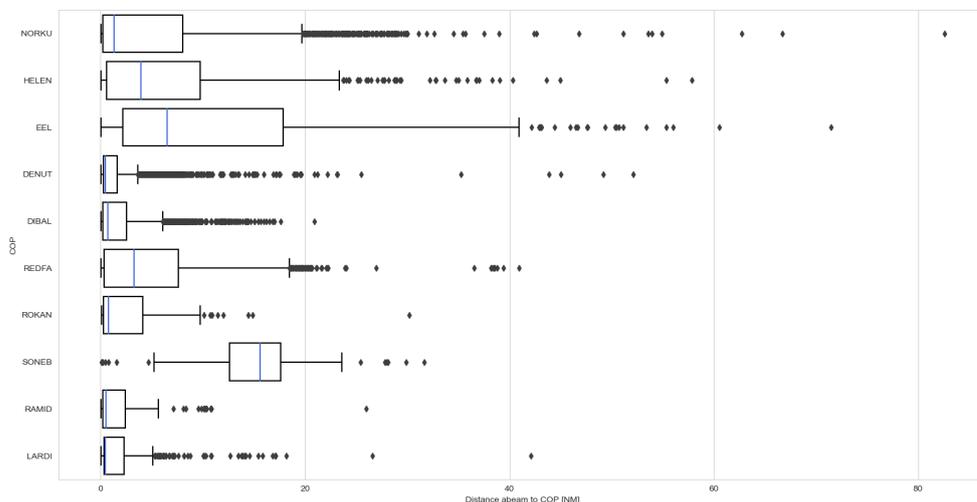


Figure 8.2: Closest distance from aircraft to COP for aircraft in September 2019

It was concluded that applying this filter would require extensive analysis and parsing of radar data, since the location of the deviation, corresponding airspeeds and the altitude needs to be known. To do this for all aircraft in the months of data that are used for the uncertainty analysis is deemed unfeasible. Therefore, another method was used.

The second method relies on using estimates at the COP instead of estimates on the IAF. The reasoning is that at the COP there has not been any influence on the actual arrival time due to AMAN planning and tactical controller actions. The only influence is by controllers of adjacent centres based on instructions of Amsterdam ACC controllers, however is assumed very limited in comparison to the tactical manoeuvring of aircraft guided from the Control Area to the Terminal Manoeuvring Area. Therefore, by comparing actuals and estimates at the COP, a better understanding of the prediction errors is created from the perspective of the arrival airport.

However, since bunching is analysed at the IAF, the prediction errors at the IAF are of interest, instead of those at the COP. Note that the effect of using the estimates at the COP, would result in a gap in the uncertainty analysis of the flight phase from COP to IAF. To analyse what the effect of this gap is an assumption for this uncertainty is taken using an Eurocontrol requirement regarding Trajectory Predictors. This requirement states that for decent, the mean trajectory prediction accuracy requirements for descent (FL300 - FL200) are 0.2 NM/min longitudinally, 100 ft/min vertically and 0.1 NM/min laterally [18].

To estimate the effect of the prediction error requirement in time, only the longitudinal requirement of 0.2 NM/min is used. Furthermore, average flying times from the COPs to the corresponding IAFs are used in order to calculate the maximum allowable longitudinal error. In Table 8.1 the distributions of the flying times are shown for flights in September 2019, acquired through radar data. The goal of the current analysis is to estimate a range of possible prediction time errors, based on the requirements given above. Therefore, the average flying times will be taken (not divided per COP). The median value, the 0.25 and 0.75 percentiles (between which the middle 50% of the dataset is represented) and the maximum and minimum values (using 1.5 times the Inner Quartile Range) are selected. These flying times are shown in the second column of

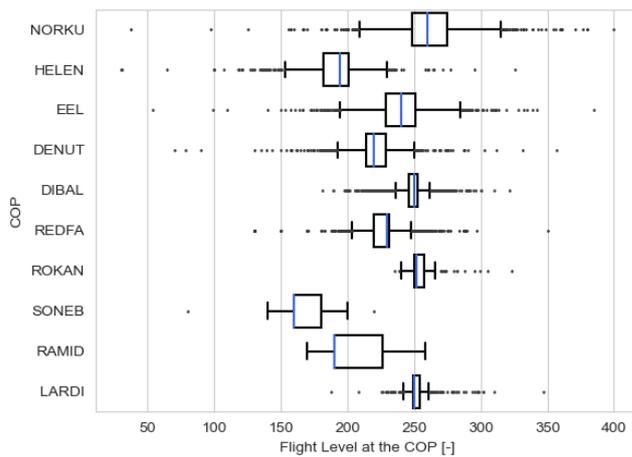


Figure 8.3: Flight Level at the COP for aircraft in September 2019

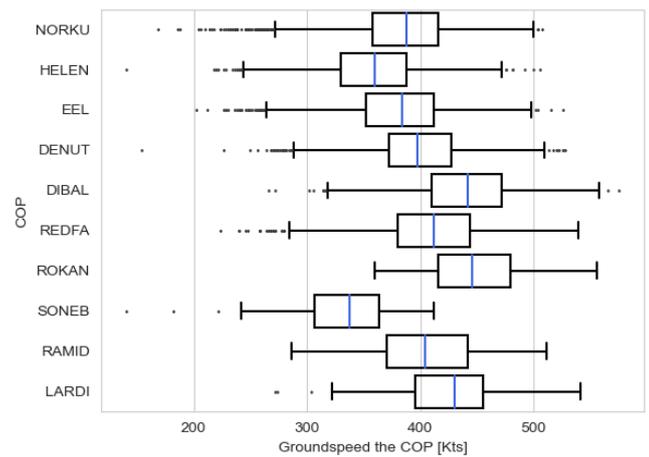


Figure 8.4: Ground Speed at the COP for aircraft in September 2019

Data Characteristics	Flying Time [sec]	Error Requirement [NM]	Error Requirement [sec]
Minimum	279	0.93	5.3
0.25 Percentile	486	1.62	9.3
Median	553	1.84	10.6
0.75 Percentile	624	2.08	12.0
Maximum	831	2.77	15.9

Table 8.1: Distribution characteristics for the flying time between COP and IAF, the corresponding trajectory predictor error requirement and the corresponding error requirement in time for aircraft in September 2019

Table 8.1. Now, using the flying times in minutes, and the longitudinal accuracy requirement of 0.2 NM/min, the lateral error requirement for the flying time between COP and IAF has been calculated in Nautical Miles, and is shown in the third column of Table 8.1

To estimate what the effect of this longitudinal error requirement is on the prediction times at the IAF, average ground speeds for the COP have been plotted in Figure 8.4, acquired through radar data for September 2019. For all COPs combined, it was found that the median ground speed at the COP was 395 kts. The same analysis was performed for ground speeds at the IAF. The median ground speed for all IAFs combined was found to be 292 kts. Assuming a constant deceleration rate, the assumed average ground speed between the COP and the IAF therefore is 345 kts. As the flying times between the COP and the IAF are related to these ground speeds, only the median value is taken for the ground speed instead of again taking the percentiles and maximum and minimum. The influence of wind is already inherently integrated in the representation of the flying time given above, as are the differences in path length from the COPs to the IAFs. By dividing the error requirements given in the third column of Table 8.1 by the average ground speed, the error requirement in seconds is found as shown in the fourth column.

One should keep in mind that the values above have been calculated using a requirement for trajectory prediction accuracy, and therefore using the median value for the flying time, one could assume that the maximum error will be around 10.6 seconds, but as the trajectory predictor should be better than the requirement, the error will likely be lower. Another way the numbers above can be interpreted is by looking at the maximum value of the the flying time between the COP and IAF, which indicates a maximum error requirement of 15.9 seconds. This can be considered as the worst-case scenario.

A check for the numbers given above is the rule-of-thumb used by the Trajectory Predictor operational expert at LVNL, which states that the TP accuracy should be below 1.5% per minute. Given the flying times

from COP to IAF in column 2 of Table 8.1, this would result in error requirements of about 2 seconds lower than those in the fourth column. Given these numbers, it is concluded that performing the uncertainty analysis at the COP and using this for estimates at the IAF will introduce some error. Effectively, it can be assumed that the uncertainty curves used on the bunching timeline should be slightly broader in width. However, this error is accepted as the errors introduced by controller influence for an analysis at the IAF are very likely to be higher.

### 8.1.3. Feature Selection

Feature selection for grouping of data is necessary to improve the accuracy of the PDFs acquired to the experience based approach. In [57] the following message properties were considered for grouping of the error distributions: flight callsign, flight operator, flight status, airport of origin, date of flight, planned date and time of arrival, aircraft type and prediction horizon. It was concluded that the flight callsign and operator were not suitable due to the low amount of samples this would yield for particular flights and operators. The airport of origin and the airport type were not suitable since these properties would yield only a large subset for the most common airports and aircraft. For others there might be a limited sample subset that is too limited. Furthermore, it was found that date or day of the week did not influence the error distribution significantly and was therefore left out. It was concluded that the flight state, arrival time of day and the time to arrival were the best parameters to group the error distribution by. The prediction horizon was evaluated in bins of 10 minutes, and the arrival time of day in bins of 3 hours [57].

Since the EFD data stream contains considerably more messages per flight than the FUM data (as shown in section 5.3, the grouping for the error distributions can be reconsidered. In section 6.3 it was discussed that the main sources of error for trajectory prediction are the flight intent, the initial conditions, the aircraft performance model, the model fidelity and the airport departure uncertainty. It should be reminded here that the type of TP used by the Network Manager is not known, and therefore no insight can be gained in the errors due to initial conditions, the aircraft performance model and the model fidelity. However, to get the effects of flight intent in the accuracy of predictions, the flight operator can be included in the division of the data, since in operations flights belonging to a specific operator typically show similar intent patterns. One of the causes for this could be found in the way an operator instructs their pilots on how to fly particular routes or approaches. Another explanation can be found in the Cost Index, described in Figure 3.2. Aircraft operators might use a specific CI for a given route that is flown each day. This CI has influence on parameters such as (descent) airspeed and TOD, which are forms of intent.

The other mentioned source of error for trajectory prediction was airport departure uncertainty. When an aircraft is still on the ground, the ground processes add to the uncertainty of that flight. To achieve a higher level of accuracy the CDM flight states can be used for data categorisation. This type of flight state was discussed in section 5.4.

Other parameters considered for the feature selection are weather related parameters. Since it is not known what weather model is used by the Network Manager, it is difficult to pinpoint specific shortcomings in the predictions due to weather conditions. However, it is expected that the impact of wind on the prediction accuracy is a function of both wind heading and speed, as higher wind speeds yield to relatively higher impact of weather model errors. The difficulty in using wind magnitude and direction in the feature selection is the bin selection. As the trajectory can span hours, the wind fields encountered en-route can change significantly in both magnitude and direction. Therefore, it is complex to capture this in a way that both contributing to the prediction accuracy and does not cause the data subsets to be too small. A solution would require the current aircraft heading and the wind direction and magnitude at (or close to) the current location of the flight. Aircraft heading can be acquired by an analysis of the way points in the route given in the EFD messages. The geographic location for the wind fields at cruise altitude would also have to be derived from these route points, as the exact aircraft location is not given in the EFD B2B data. However, it should be noted that when wind conditions are not used for the feature selection, the influence of wind conditions is still inherently a factor of uncertainty in the analysis. This means that errors due to wind conditions do contribute to the uncertainty curves, the data is just not split by various wind conditions.

The main limitation to the features above is the availability of the amount of data. Since the method chosen for the current research is an experience based approach, using more selection features results in smaller data buckets, which could yield insignificant results. In other words, the goal of the feature selection for the

uncertainty analysis is to find features for which the data (the EFD messages) can be split up, such that a representative estimate of the uncertainty of the arrival predictions in these messages can be found. Using too many features results in small buckets of data, for which an uncertainty analysis might not be representative for all messages with those features. Using too little features will result in overestimating the uncertainty of the arrival predictions as there is a large generalisation. Going forward therefore, the focus is on finding good features that serve as a baseline for the uncertainty in the predictions of the EFD estimates, and that is representative for the underlying data such that it can be used in a debunching model.

Concluding, the features that will be analysed are the flight state, the CDM state and the time to arrival. Based on this analysis, it will be determined whether the uncertainty analysis can be stretched further with the other parameters such as the intent parameters mentioned above or the time of day used in [57].

#### 8.1.4. Data Analysis

In this subsection the data analysis is performed for the features mentioned above. The goal is to perform a feature selection for the uncertainty analysis that is well balanced such that curve fitting can be performed, and the resulting curves are representative for the underlying data set (with the given features).

##### Prediction horizon

The first feature that is analysed is the prediction horizon. The prediction horizon is determined by the difference between the time at which a message is sent and the estimated time the aircraft will arrive at the COP. As stated before, the prediction horizon is grouped in buckets of 5 minutes. In Figure 8.5 all data (from August 1<sup>st</sup> (2020) to June 24<sup>th</sup> (2021)) is grouped and shown at 20 minute increments until a prediction horizon of 4 hours. In the figure it can be seen that the uncertainty tends to increase with the prediction horizon (the time to predicted arrival at the COP). From the distribution of the box plots it is clear that the time to arrival is well suited as a feature for the division of the data. Even though the median value of the prediction error remains fairly constant, the interquartile range (the box) and the length of the whiskers increase with increasing time to predicted arrival time. For instance, the predicted time at the COP for a flight arriving in 3 hours has a 50% probability of containing an error between -130 and +333 seconds, as 50% of the data is lying in this interquartile range (IQR). For a flight predicted to arrive in 1 hour, this interquartile range is defined by -64 and +138 seconds.

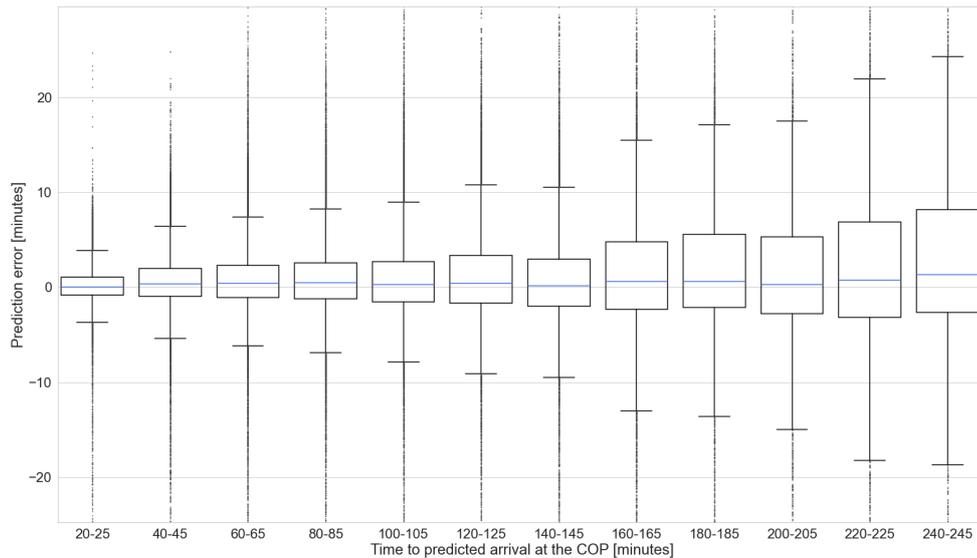


Figure 8.5: Accuracy of the COP arrival prediction error (Estimated Time Over - Actual Time Over) versus time to fly

### Flight State

The second feature that is analysed is the flight state, mentioned in section 5.3. Of the 8 mentioned flight states, only 6 are used in the analysis, as the Cancelled (CA) and the Terminated (TA) flight states do not have to be accounted for as flights with these states are deleted. The remaining flight states are Flight Plan filed (FI), Filed Slot Allocated (FS), Slot Issued (SI), Tactical Activated (TA), Air Traffic Control Activated (AA), and Suspended (SU) (described in section 5.3). As the division of the data in 5 minute buckets was shown to yield significant differences above, these buckets will also be included in the analysis of the flight states.

First, a short analysis is performed based on the number and the share of each flight state at a given prediction horizon in order to gain a better understanding of when and how the flight states are used. In Figure 8.6 the total number of messages sent per flight state was shown for the time to predicted arrival at the COP. Note that as the analysis is performed for the uncertainty up to the COP (and not the IAF, as discussed in subsection 8.1.2), this prediction horizon is not the same as the bunching prediction horizon, where the x-axis represents the time to predicted arrival at the IAF.

From Figure 8.6 it becomes clear that a large amount of updates for the flight state AA are sent just before the arrival at the COP. To gain a better understanding of what the distribution of the flight states is, the share of the messages with a given flight state has also been plotted, shown in Figure 8.7. From this figure it becomes evident that for the given timeline, the greatest share of messages are those with an AA and a FI flight state. This does not directly translate to the bunching prediction timeline, as the figures do not represent the number of aircraft with a given flight state. However, it can be assumed that the share of AA flight states is largest as all aircraft get this state as they move closer to their ETO IAF.

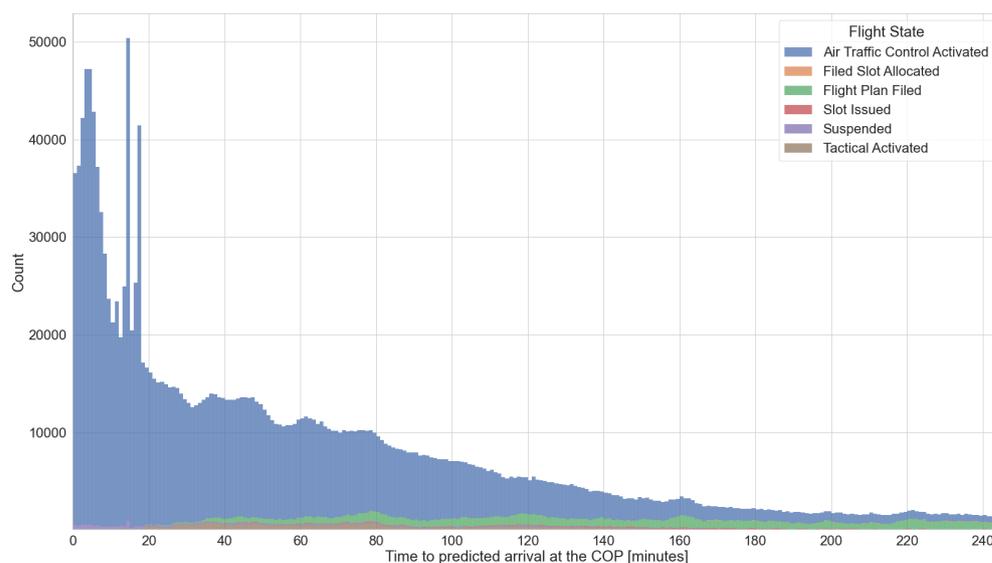


Figure 8.6: The total number of messages sent per flight state for a given time to predicted arrival at the COP

Now the distribution of the flight states is shown, the distribution of their errors is examined. For the most common flight state, Air Traffic Control Activated state, the box plots are shown in Figure 8.8. The AA flight state is the most accurate flight state as aircraft are confirmed airborne, and therefore an accurate arrival prediction can be obtained without influence of ground processes. As expected, the AA flight state has a great decrease in uncertainty compared to the uncertainty for all flight states combined (shown in Figure 8.5).

The plots for the FI, FS, SI, TA and SU flight states are presented in the Appendix, by Figure A.1, Figure A.2, Figure A.3, Figure A.4, Figure A.5 respectively. For the FI state in Figure A.1 it can be seen that the variability of the error (as indicated by the whisker positions) remains relatively constant from 80 minutes prediction

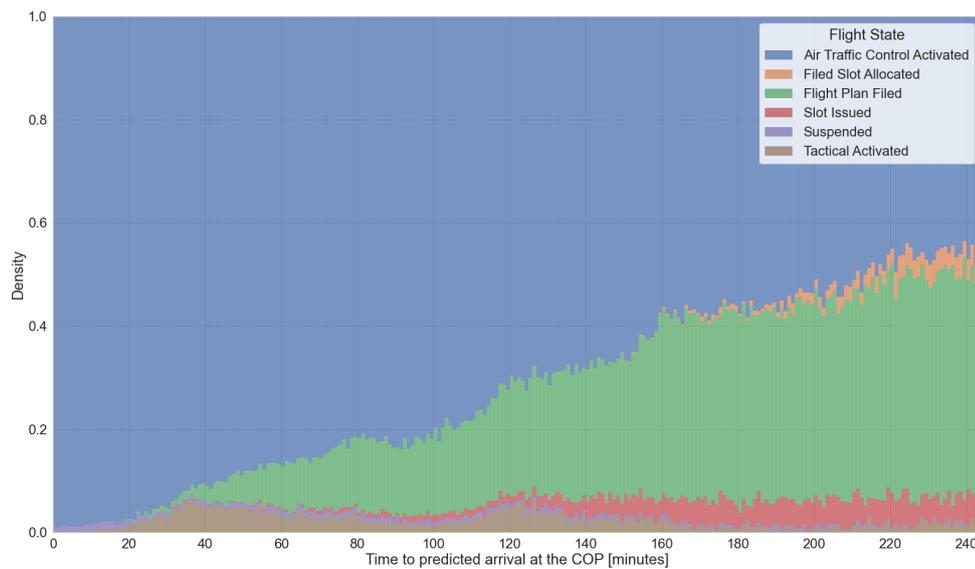


Figure 8.7: The share of messages sent with a given flight state for a given time to predicted arrival at the COP

horizon onward. As expected, the errors in all time buckets are higher as compared to the errors in the AA flight state in Figure 8.8, as the boxes (indicating the IQR) are significantly smaller in the latter.

For the TA flight state in Figure A.2 it can be seen that in general the error is smaller than that at of the FI flight state. However, there is no clear relation between this flight state and the prediction horizon seen by the whiskers and the boxes which is seen in most of the other flight states. For FS flight state in Figure A.3 there is no data in the buckets before 140 minutes prediction horizon. For the higher prediction horizons the maximum number of samples is 156. Due to this small number of samples not many conclusions can be drawn, except from the fact that the flight state is not used often. For the SI flight state in Figure A.1 it can be seen that the error distributions become severely larger for increasing prediction horizon. At increasing prediction horizons the errors are significantly larger than those for all other flight states. The SU flight state in Figure A.5 shows a fairly constant error rate. However, shown in Figure 8.7 only a small fraction of messages is sent with this flight state.

The conclusion from the discussion above is that the flight state can be used as a suitable feature for the division of the data. There are clear differences between the box plot statistics, yielding differences in the uncertainty distributions which can be used in the bunching prediction model.

### CDM State

The third feature that will be considered is the availability of CDM information. As discussed in section 5.4, there are 5 different CDM states. The goal of these states is to provide accurate flight data for aircraft that are not yet airborne. Therefore, the flight states FI, FS, SI and TA are of interest. The flight state TA is included as the CDM state will likely have an effect on the accuracy of the TA. The airborne flight state AA is not included as an CDM state indicating that the flight is still on the ground would be erroneous as the flight is confirmed airborne by ATC. In other words, the flight state is assumed to be leading, in which the CDM state performs a supporting role.

In Table 8.2 it is shown how often a specific CDM state was recorded for each flight state. It should be noted that the CDM states given in the TA flight state column should in theory all be Actual Off-Block, as this is the last CDM state, and the flight is confirmed airborne. However, what is important to note here is that for the FI flight state, there are no messages containing an ATC-DPI CDM state, which leads to inaccuracies

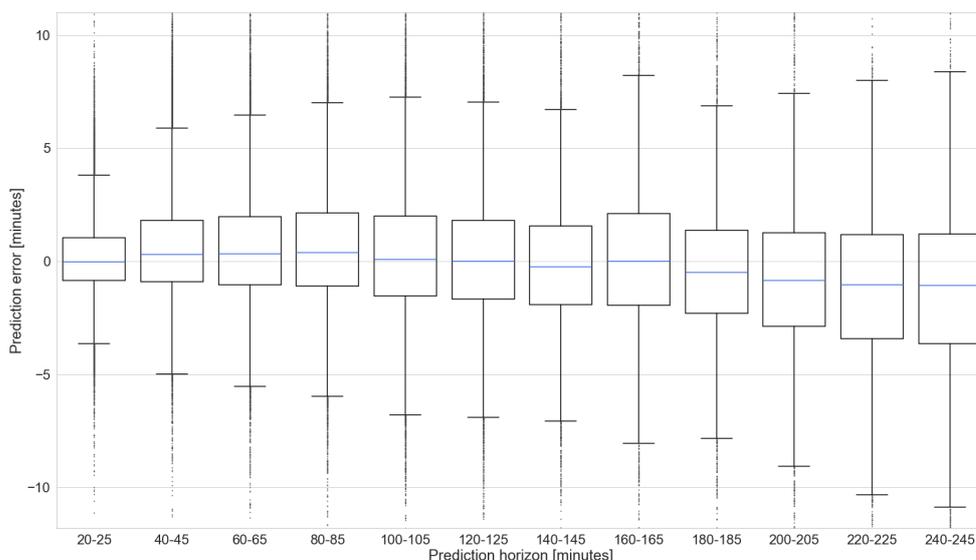


Figure 8.8: Accuracy of the COP arrival predictions (ETO - ATO) versus time to fly for the Air Traffic Control Activated flight state

	FI	FS	SI	TA	AA	SU
DPI Expected (P-DPI)	31.4 %	54.0 %	23.5 %	6.9 %	9.6 %	51.1 %
Estimated (E-DPI)	8.4 %	17.9 %	4.8 %	0.1 %	0.2 %	0.5 %
Targeted (T-DPI-t)	11.7%	0.0 %	26.0 %	0.6 %	0.8 %	2.8 %
Presequenced (T-DPI-s)	11.2 %	0.0 %	23.4 %	1.8 %	0.9 %	0.0 %
Actual Off-Block (ATC-DPI)	0.0 %	0.0%	0.2 %	46.6 %	28.6 %	1.9 %
Standard Airport (No CDM)	37.3 %	28.2 %	22.1 %	44.0 %	59.3 %	43.7 %

Table 8.2: The percentage of specific CDM state messages for a given flight state, for arrivals from August 1<sup>st</sup> (2020) to June 24<sup>th</sup> (2021)

and uncertainties in the estimates. The reason behind this inaccuracy is that the definition of a flight with an ATC-DPI status can still be on the ground, as the definition of the status is that the flight has 'off-blocked', and is taxiing to the runway to take-off [33]. Therefore, part of the messages containing this CDM state are still on the ground, and should not have a TA or AA status, as is the case now. It is concluded that therefore the definition for the TA flight state by EUROCONTROL is erroneous and inaccurate, as flights cannot be assumed airborne yet. Another conclusion is that the CDM states are lacking, as an extra state should be included stating that an aircraft has taken off, as to divide the 'off-blocked' aircraft from aircraft that are in or after take-off.

In Figure 8.9 the prediction error for the flight state FI has been shown using a paired box plot, differentiating between CDM equipped and non-CDM equipped messages. One would expect that the inclusion of CDM information would result in a higher accuracy, as the estimated Take-Off Time can be calculated more accurately using this information. However, no such effect is seen for the FI flight state. In Fact, the CDM equipped messages appear to provide worse estimates than those without CDM information. Both the whisker length and the IQR are similar or larger for the CDM equipped messages at most of the prediction horizons. In Figure 8.10 the same comparison is shown for the TA flight state. There, it can be seen that the CDM equipped messages are more accurate than those without CDM, for the whole timeline (except at 200 minutes, it is assumed that this is due to a lack of observations as there are only 21). Therefore, it can be concluded that the CDM information leads to more accurate predictions in the TA flight state. For the FS, SI flight states shown in Figure A.6 and Figure A.6 respectively, no clear improvements are seen when including the CDM information. Therefore, the conclusion of this third part of the feature selection analysis is to only split the flight state TA by CDM and non-CDM equipped messages.

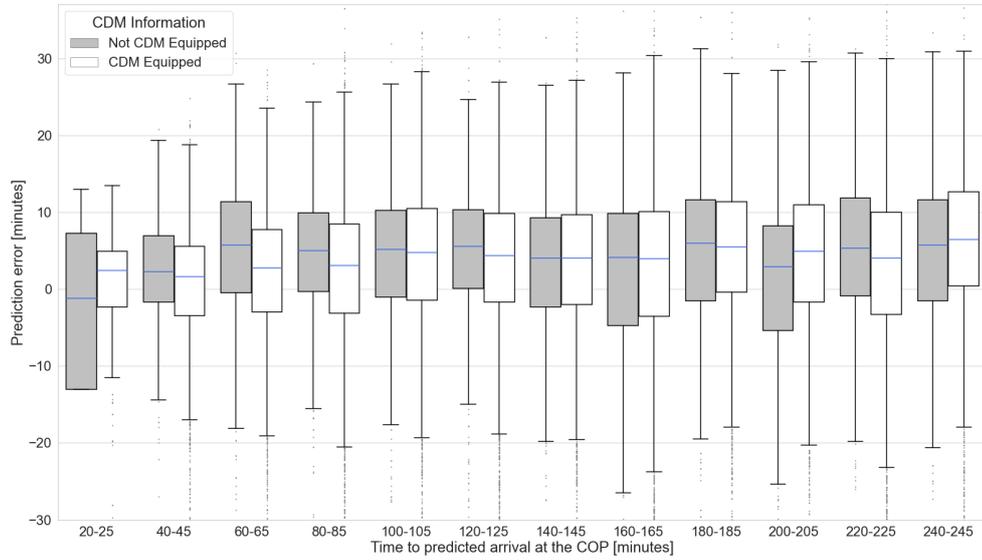


Figure 8.9: Accuracy of the COP arrival predictions (ETO - ATO) versus time to fly for the Flight Plan Filed flight state, for CDM and non-CDM equipped messages

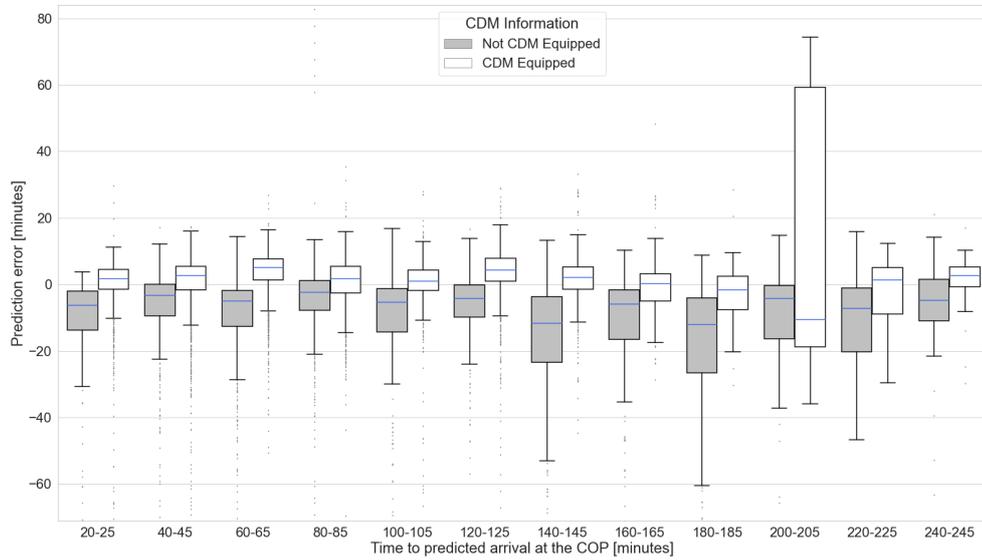


Figure 8.10: Accuracy of the COP arrival predictions (ETO - ATO) versus time to fly for the Tactical Activated flight state, for CDM and non-CDM equipped messages

As stated in the section before, the goal of the feature selection for the uncertainty analysis was to find representative estimates of the uncertainty for groups of data. From the current section it was concluded that relevant features were the time to arrival (in 5 minute buckets), the flight state and the CDM state for the TA flight state. Another proposed feature was the time of day, however, it was found that there was not enough

data for all flight states at all time to arrival horizons to split the data based on this feature. For this reason, the wind and intent factors were left outside of the scope of this analysis.

### 8.1.5. Curve Fitting

In order to present the uncertainty analysis in such a way that it can be used for a probabilistic debunching model, a curve fitting routine is performed on the uncertainty analysis. With this routine, Probability Density Curves (and corresponding Cumulative Distribution Curves) are constructed. In [57] it was found that curve fitting with normal distribution yielded only a reasonable for a small part of the data. Therefore, the use of Johnson curves was introduced, which can be used for non-normal univariate data sets. The Johnson distribution consist of four types of distributions: the normal distribution, the lognormal distribution, the Bounded Johnson (SB) and the Unbounded Johnson (SU) (with upper and lower tails going to infinity). The Bounded and Unbounded Johnson Probability Density Functions are shown in Equation 8.1 and Equation 8.2 respectively, where the function  $\phi$  is the normal PDF, shown in Equation 8.3. The variables  $\gamma$  and  $\delta$  are the Johnson Shape Parameters,  $x$  is a column vector of Johnson variates, the  $\xi$  variable is the location parameter and  $\lambda$  is the scaling parameter.

$$PDF = f(x, \gamma, \delta, \xi, \lambda) = \frac{\delta/\lambda}{\sqrt{((x-\xi)/\lambda)^2 + 1}} \cdot \phi(\gamma + \delta/\lambda \cdot \log((x-\xi) + \sqrt{((x-\xi)/\lambda)^2 + 1})) \quad (8.1)$$

$$PDF = f(x, \gamma, \delta, \xi, \lambda) = \frac{\delta/\lambda}{\sqrt{((x-\xi)/\lambda)^2 + 1}} \cdot \phi(\gamma + \delta/\lambda \cdot \log(\frac{(x-\xi)/\lambda}{1 - ((x-\xi)/\lambda)^2})) \quad (8.2)$$

$$\phi(x) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}} \quad (8.3)$$

The first step in the Johnson curve fitting routine is to determine which of the distributions is used. Hill et. al. [28] developed an algorithm to fit the curves to data based on the mean, standard deviation, skewness and kurtosis of the data, by using standardised moments (quantitative measures uniquely describing the probability distribution and normalised for comparison between curves). However, in [55] it was shown that this method was greatly affected by outliers and not effective for small sample sizes. Therefore, the method of distribution selection by means of quantiles was introduced, by selecting four percentiles of the distribution that satisfy the requirement that their corresponding standard normal values are equally spaced.

For the curve fitting routine the MATLAB Johnson Curve Fitting Toolbox<sup>3</sup> is selected. The algorithm returns which of the curves needs to be fitted based on the quantiles method, and it returns four parameters that define the fitted Johnson distribution: a location parameter, a scale parameter and two shape parameters. Since only these four parameters are necessary to describe the PDF, an uncertainty curve can rapidly be selected for an incoming EFD B2B message based on the selected features in the message by means of a lookup table. This process will be explained in section 9.2. The minimum sample size was set to 20, as curves fitted to smaller sample sizes were deemed unreliable. For the prediction timeline this could mean that for a short period of time a flight would have no curve, as there were less than 20 messages of it's kind. In case this happens, the next available curve is selected with the same features, but with a higher prediction horizon (or if not available, a lower prediction timeline).

### 8.1.6. Curve Fitting Analysis

With the methods mentioned in the previous subsections the curve fitting was performed for prediction horizons up to 300 minutes. For all features the curves were plotted in 20 minute intervals, up to 240 minutes. For the curve fitting the axes were changed such that the curves are oriented in the same manner as presented on the bunchign prediction timelines in the upcoming sections. The error is now computed by: (Actual Time over the IAF) - (Estimated Time Over the IAF). For the AA, TA (without CDM information) and the FI these plots are shown in Figure 8.11, Figure A.8 and Figure A.9. For the flight states FS, SI and SU these curves are shown in Figure A.10, Figure A.11 and Figure A.12. As mentioned in subsection 8.1.3, the TA flight state is split

<sup>3</sup>Dave (2021). Johnson Curve Toolbox (<https://www.mathworks.com/matlabcentral/fileexchange/46123-johnson-curve-toolbox>), MATLAB Central File Exchange. Retrieved September 6, 2021.

up by CDM equippage. The curves for the TA flight with CDM equippage were shown in ??

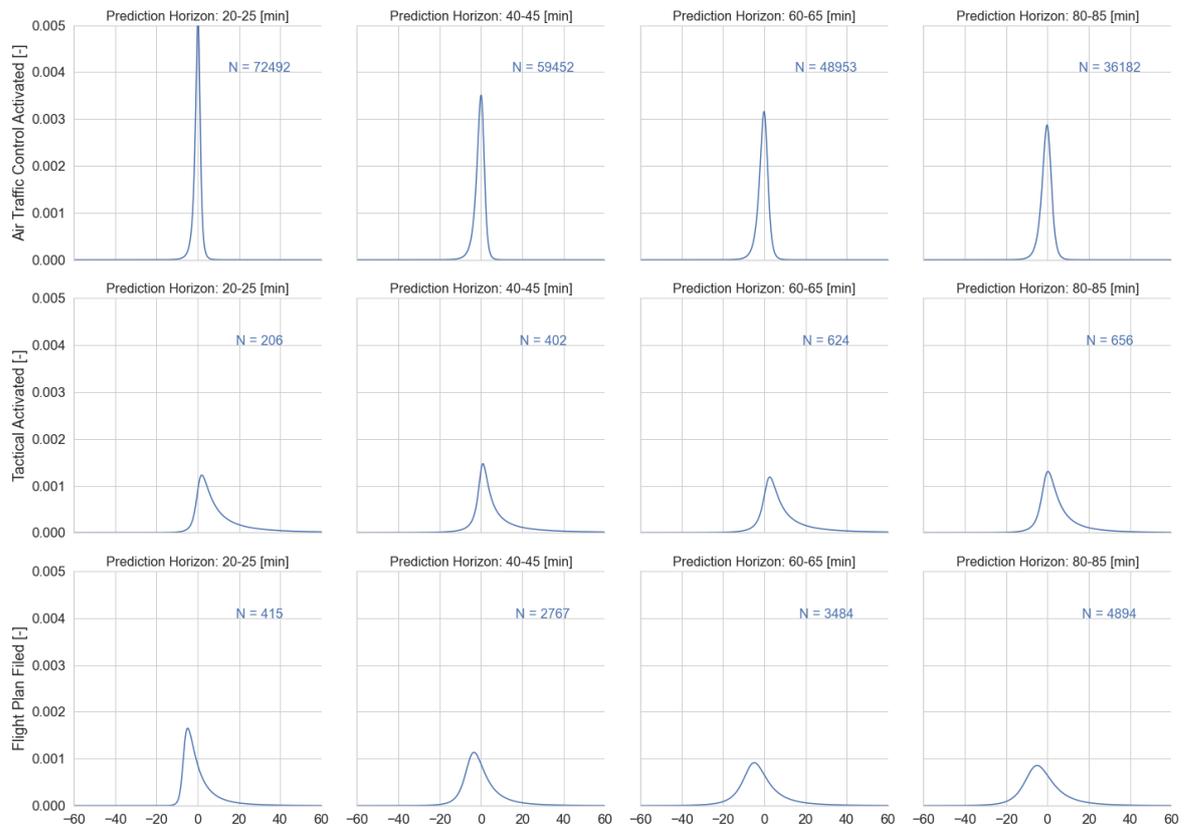


Figure 8.11: Probability Density Curves for the AA, TA and SI flight state up to 85 minutes prediction horizon

For most of the flight states and buckets there is a skew towards negative values, meaning that the ETO COP estimates in the EFD messages are later than the Actual Time Over the COP. An intuitive way of looking at the plots is that with a negative skew, there is a higher probability of the aircraft arriving earlier than the time provided in the EFD message. In Figure 8.2 it is shown that aircraft often do not fly directly over the COP, but instead follow direct routes such that the COP will be bypassed at a distance. This means that the Actual Time Over the COP yields a time earlier than was predicted, explaining the negative skew.

Furthermore, it is shown that the AA flight state is most accurate, with a width of 20 minutes centred around 0, for a prediction horizon between 60 and 65 minutes. For the FI flight state it is shown that the width is about 40 minutes, with a skew towards negative values (arriving earlier than predicted). It is shown that the SI flight state performs worst. For groups of data where there was less than 20 samples, no curve was fitted as it is unlikely that these curves would be accurate.

Note that although a similar analysis is done to [57], the results cannot be directly compared or validated using that research. In the mentioned research the uncertainty analysis was performed for arrival times at the runway instead of at the COP, and thus the delay in the FIR resulted in a skew towards positive values (delay) instead of the negative values in this research. Some of the effects of the negative skew in the current research might be attributed to flying direct over the COPs. Another effect might be due to the effect of low traffic density data, although no definitive conclusions can be drawn on this as this would require in-depth knowledge of the type of TP used by Eurocontrol. A qualitative analysis of the potential effects of low traffic density data was performed in subsection 8.1.1.

### 8.1.7. Validation

To check whether the curves were properly constructed, each curve was plotted over a histogram of the underlying data and checked visually. From this it was concluded that for curves constructed from very little samples (below 20) the curve fitting was not useful. For the others, the fitted curve resembled the histogram visually. However, to test whether the fitted curves represent the underlying data in a proper manner, a validation routine was performed that was proposed in [57]. A goodness-of-fit routine was performed using the Kolmogorov-Smirnov test, which compares two Cumulative Distribution Functions with the hypothesis that both distributions are sampled from the same distribution by measuring the maximum absolute difference between two CDFs [36]. The downside of using this method as a goodness-of-fit measure was that as underlying data was compared to the fitted CDF, the assumption for independence of the samples does not hold and thus the corresponding p-value cannot be used for hypothesis testing. To mitigate this shortcoming, a bootstrapping process was performed on the Kolmogorov-Smirnov statistic such that a confidence interval of the statistic could be found, from which conclusions can be drawn.

In the current study the same bootstrapping method was used with the random sampling size of 1000 (with replacement) and the number of iterations of 1000. In other words, from a given subset of data (with the corresponding fitted curve) 1000 samples were drawn, for which the CDF was compared to the fitted CDF using the Kolmogorov-Smirnov test. This process is repeated 1000 in order to determine the confidence interval of the Kolmogorov-Smirnov statistic. The results for this are shown in the Appendix A, in Table A.1. The bootstrapped confidence interval show the range in which the absolute maximum difference lies for the fitted CDF and the underlying data with 95% confidence.

For the AA flight state it is shown that for all relevant ranges the upper boundary of the confidence interval is in the order of magnitude of 1% (note that the difference between two CDFs represents a probability difference). This indicates that there might be a difference between the fitted curve and the data set. However, this difference is accepted as the goal of the current research is not to perform a perfect uncertainty analysis, but merely to present a baseline for further research and from which the dynamics of the debunching concept can be derived. The second important flight state (as derived from the distribution in Figure 8.7) is the FI flight state. There, the upper boundary of the confidence interval indicates worse performance for the curve fitting, with values ranging between 10 and 20%. What this shows is that the ETO IAF errors for aircraft with this flight state might not be properly represented by these CDFs. Therefore, an improved uncertainty analysis with more discretisation features or different methods might be necessary to model these errors more accurately. The same conclusion can be drawn for the SU and the TA flight state (both with and without CDM information). For FS and SI flight states the goodness-of-fit was shown to worst. However, as these flight states are only a very small portion of the messages sent, this will hardly impact the results.



# 9

## Methodology

### 9.1. Introduction

The goal of the debunching model is to combine the uncertainty analysis with an operational model which detects bunching and suggests a solution. The model should be able to parse the incoming EFD messages sent by the Network Manager, get the estimate uncertainty distribution based on the parameters given in section 8.1, calculate the probability that bunching will occur, and calculate a potential solution. The debunching model for this research has been designed to consist of four separate modules. The model will retrieve the messages and curve parameters as described in section 8.1, and parse the data in the data parsing module, discussed in section 9.2. In a simulation the messages would be sent to the data class module in the order of the timestamps of the messages. The last message that was sent determines the simulation time. From the message class module, the messages are implemented in the bunching detection module, discussed in section 9.3. Finally, a potential solution is calculated in the solution module, discussed in section 9.4. The interdependencies of the modules are shown in Figure 9.1, where each color represents a module.

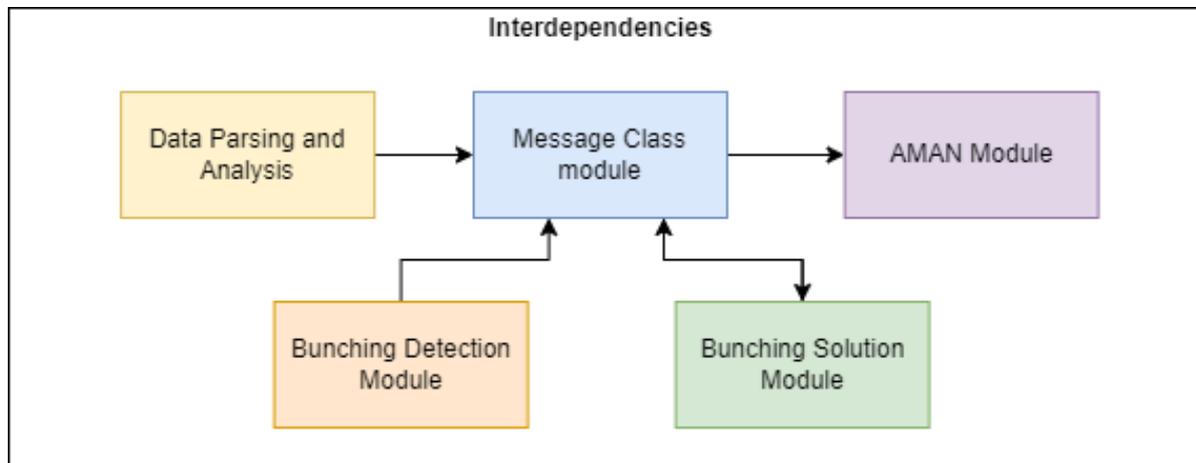


Figure 9.1: System architecture of the main module

### 9.2. Message Class Module

In the Message Class Module all information of a given flight is stored. For this, Object Oriented Programming is used such that the data is structured in an efficient manner. The module is structured such that each aircraft becomes a separate instance based on its flight plan ID. The instance has attributes that hold all information provided by the EFD data, but also information that was computed or parsed within the Message Class Module based on the information provided by the EFD data. Therefore, this module is the core module, and all updates on the data will be sent to and received from this module by the other modules. The module is explained in detail in order to gain a proper understanding of the module structure and processes. A

schematic overview of the module is shown in Figure 9.2

### 9.2.1. Input Data

In this subsection the inputs for the Message Class Module are described in detail:

1. The input data is the unprocessed EFD B2B data that LVNL receives from the Network Manager (EURO-CONTROL). As LVNL receives not only messages for flights with Schiphol Airport as final destination, all other messages were filtered out. Therefore, the input data sets contained only messages for aircraft with destination EHAM. The relevant properties and the structure of the data is shown in ??, the properties that were not used have been left out. As each row (which contains a single message) in the data contains a timestamp, the order at which the messages is sent to the module is determined by the order of the timestamps.
2. The simulation time is determined by the timestamp of the last message (row) that was sent to the module. As all inputs and processes within the module are deterministic, the module is not updated in the experiments at timestamps where there is no new message, as this would only result in a shift on the bunching prediction timeline and no other parameters would be changed.
3. A table containing the conversion from the WTC type to the Recat WTC type is used, as only standard WTC types are given in the EFD Data, but Recat is an improved version which will be used in this research (shown in Table 4.3). This conversion table was received from LVNL.
4. A table of the nominal airspeeds for each aircraft is used as The EFD data does not provide information on the airspeed. However, this information is required in the AMAN computations and the maximum delay absorption computations. Therefore, a table was used that contains nominal airspeeds specified per aircraft type and airline operator, for Cruise (Mach and CAS), Descent (Mach and CAS) and for final approach (the True Airspeeds at various distances (WTC separations) from the runway). This table was provided by LVNL, and constructed by Radar observations of at least 3 years' worth of data. An example of the data is shown in Table 9.1.
5. A table with the runway configuration changes with timestamps attached, for 2020 and 2021. This data was retrieved from a post-processed data set with landing times which contained the runway configurations at the time of landing. This data set was acquired from LVNL.
6. The table of curve parameters which was constructed with the curve fitting analysis explained in subsection 8.1.5. Based on the time to fly, the flight state and the CDM state the 5 curve parameters and the Johnson type are retrieved, which together can be used to plot the PDF and the CDF.

AC Type	Operator	Mach Cruise	Mach Descent	CAS Cruise	CAS Descent	TAS <sub>3</sub>	TAS <sub>4</sub>	TAS <sub>5</sub>	TAS <sub>6</sub>	TAS <sub>9</sub>	TAS <sub>7</sub>	TAS <sub>8</sub>
B738	KLM	0.77	0.76	295	273	150	152	155	158	166	161	164

Table 9.1: An example of the input table providing nominal airspeeds for Cruise and Final Approach

### 9.2.2. Algorithmic Steps

In this section the algorithmic steps shown in Figure 9.2 will be explained more in detail. The numbers in the flow chart correspond to the detailed steps below:

1. The EFD data provides ETOs for each point provided in the route. The relevant ETOs that have to be parsed are the ETOs over the IAF and the COP, which are found by filtering the route data field. With the ETO over the IAF, the flying time to the IAF is calculated by taking the difference between the ETO IAF and the current time (provided as input into the module).

The intended route and the distance to fly can also be acquired through parsing the route points given in the EFD messages. These points are given by either a Latitude and Longitude combination, or the name of the waypoint. The Latitude and Longitude locations of these waypoints have been acquired from a dataset with these values. For EFD message, the distance between all points is summed up in order to get the total distance of the intended route. For this, the distance between two points (given

by Latitude and Longitude) are calculated using the Haversine formula, shown in Equation 9.1. There,  $\varphi_1, \varphi_2$  are the latitude of point 1 and latitude of point 2, and  $\lambda_1, \lambda_2$  are the longitude of point 1 and longitude of point 2. For the radius of the Earth ( $r$ ) is 6371 km is used, as Haversine assumes a spherical earth. The points at which the aircraft is still in the climb phase (determined by the change in altitude between the points) are omitted, so that the only the distance to fly in the cruise and descent phase is left, which are necessary for the maximum delay absorption calculations.

2. In the EFD data, Flight Levels are attached to the points in the route data field. As aircraft typically fly a step climb in cruise flight FL in cruise for fuel efficiency, the maximum FL is considered the FL just before the Top of Descent. In order to differentiate between climb and cruise (which is necessary for the maximum delay absorption in cruise) the minimum cruise FL is also found by finding the difference in FL between all points on the route, and taking the first FL for which the difference in FL between the given point and the next point on the route is zero.
3. Based on the Aircraft type and the Airline Operator the nominal airspeeds discussed in subsection 9.2.1 are parsed.
4. In this step the maximum delay in descent is calculated. The process of this step is described in detail in subsection 9.2.3
5. The maximum delay that can be absorbed in cruise is calculated in this step. The process will be discussed in detail in subsection 9.2.4.
6. Based on the Estimated Time of Arrival at the runway and the IAF provided in the route points data field, the aircraft will be shown on the correct timeline. This was done by correlating the runway configuration at the time of the ETA (from the table explained in subsection 9.2.1).
7. Based on the time to fly, the flight state and the CDM state attributes of a flight, the curve parameters are retrieved from input 5. The Johnson Probability Density Function is calculated, and from this the expectation of the PDF (expected ETA at the IAF based on the curve) is determined.
8. Based on the runway configuration and this expectation, the expected follower of the aircraft is determined, such that the required WTC separation between the leader and the follower aircraft be determined. Now, the  $CDF(t)$  and the  $CDF(t - WTC_{separation})$  are calculated. This is used for the bunching calculation, which will be explained in section 9.3.

$$distance = 2r \arcsin \left( \sqrt{\sin^2 \left( \frac{\varphi_2 - \varphi_1}{2} \right) + \cos(\varphi_1) \cos(\varphi_2) \sin^2 \left( \frac{\lambda_2 - \lambda_1}{2} \right)} \right) \quad (9.1)$$

### 9.2.3. Maximum Delay Absorption in Descent

In order to assign realistic delay values to aircraft, the maximum delay that an aircraft can absorb should be examined. As the detection of bunching (and the debunching) is done based on a prediction timeline, the maximum delay should be known in seconds. The main shortcoming of the EFD data is that no cruise speeds are provided. Therefore, an accurate calculation of the maximum delay for an aircraft at the moment a message is received is not possible. Instead, this subsection serves to provide a method that allows calculating estimated values for the maximum delay in descent that can be absorbed in seconds. This subsection is structured based on the information that is necessary for these calculations: The type of descent, the intended route, the minimum and average operational cruise speed and the corresponding ground speeds.

In section 3.4 it was discussed that the most fuel efficient strategy for delay absorption was to first decrease the target CAS descent as much as possible before decreasing the target Mach cruise. From an operational perspective this is preferred as well, as the delay absorption would be performed at in the Control Area, before contacting adjacent centres for cruise speed decreases. A well known way of approaching descent types is shown in Figure 9.3, which is adapted from [10]. Two cases for descent are shown here, in the first, an aircraft descends with constant Mach number, resulting in an acceleration segment (measured in TAS) from the top of descent. This type of descent is performed when the target CAS in descent is higher than the CAS flown in cruise. As soon as the target CAS is reached, the aircraft descends further along a constant target

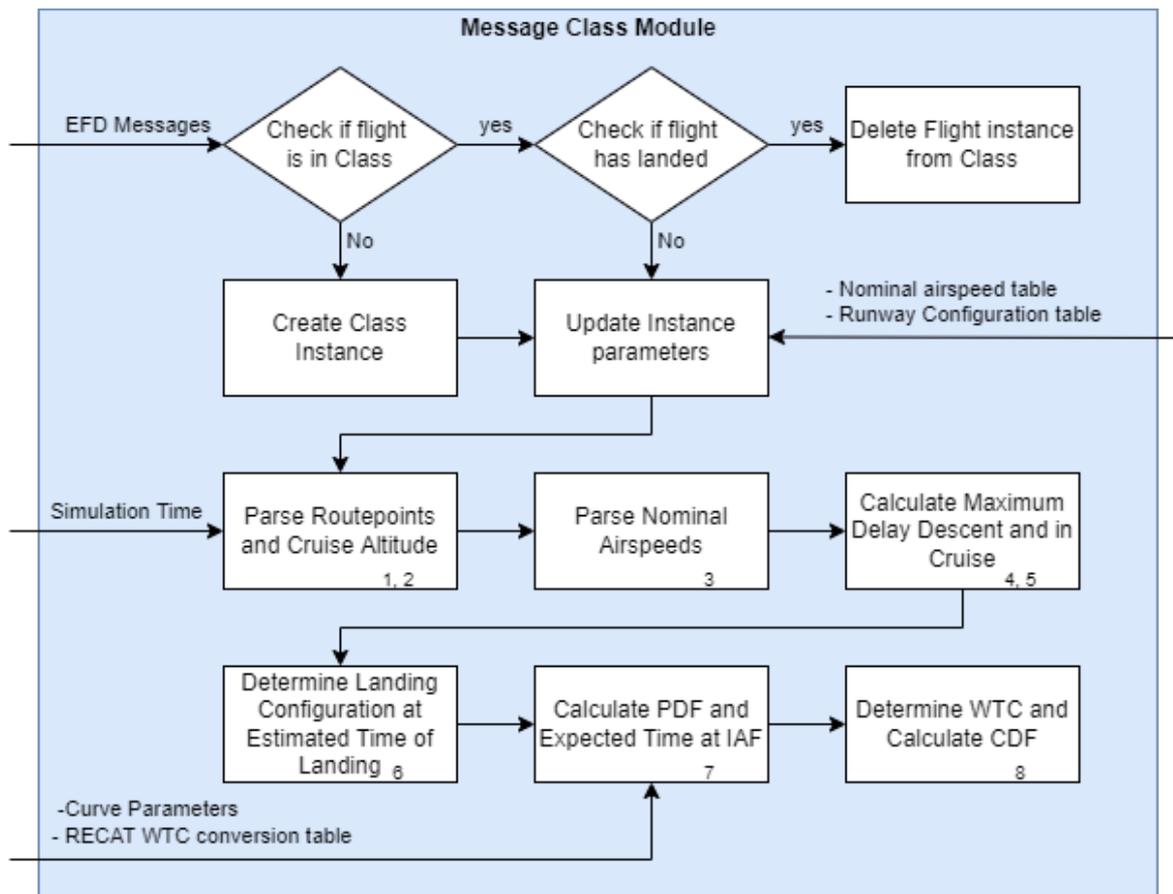


Figure 9.2: System architecture of the message class module

CAS. The second case is where the target descent CAS is lower than the cruise CAS. This is considered to be a slow descent, in which an aircraft first decelerates at idle thrust and constant altitude towards the target CAS, followed by a constant target CAS descent [10]. The minimum value for the target CAS descent is set to 250 knots, which is assumed to be the minimum CAS that is assigned by controllers to aircraft flying above the IAF (at FL100) [43], which was confirmed by an operational expert at LVNL. This value assumes idle thrust settings, but does not take into account the feasibility for each aircraft.

In order to analyse what part of the route distance above is flown in cruise, and what part in descent, the location of the Top of Descent should be known. However, this location is dependent on multiple factors, including the Cost Index that an aircraft is flying at and target descent speeds. In order to estimate the TOD, an accurate trajectory predictor would be required. In this research, the design and implementation of a trajectory predictor is left outside the scope. Instead, a simple estimation is performed using average target CAS descent speeds and constant descent angles for CAS descent and Mach descent. These averages have been acquired through the LVNL dataset shown in Table 9.1, and are known for each aircraft type.

In order to effectively model the descent profile of an aircraft, one needs to know the flight level of the crossover point. The crossover point is the point where the lines of a given Mach number and CAS speed yield the same True Airspeed (shown by intersecting lines in Figure 9.3). In a fast target CAS descent (case 1 in Figure 9.3) this would determine the point from which an aircraft switches from a constant target Mach descent to the constant target CAS descent. When the target CAS descent is lower than the equivalent cruise CAS (case 2 in Figure 9.3), the crossover FL would be above the current cruise FL, and indicate that a level-flight deceleration segment is required.

In the module, the crossover level is determined using the method used in BADA [44], shown in Equation 9.2, defined as the geopotential pressure altitude (in feet) at which CAS and Mach airspeeds present

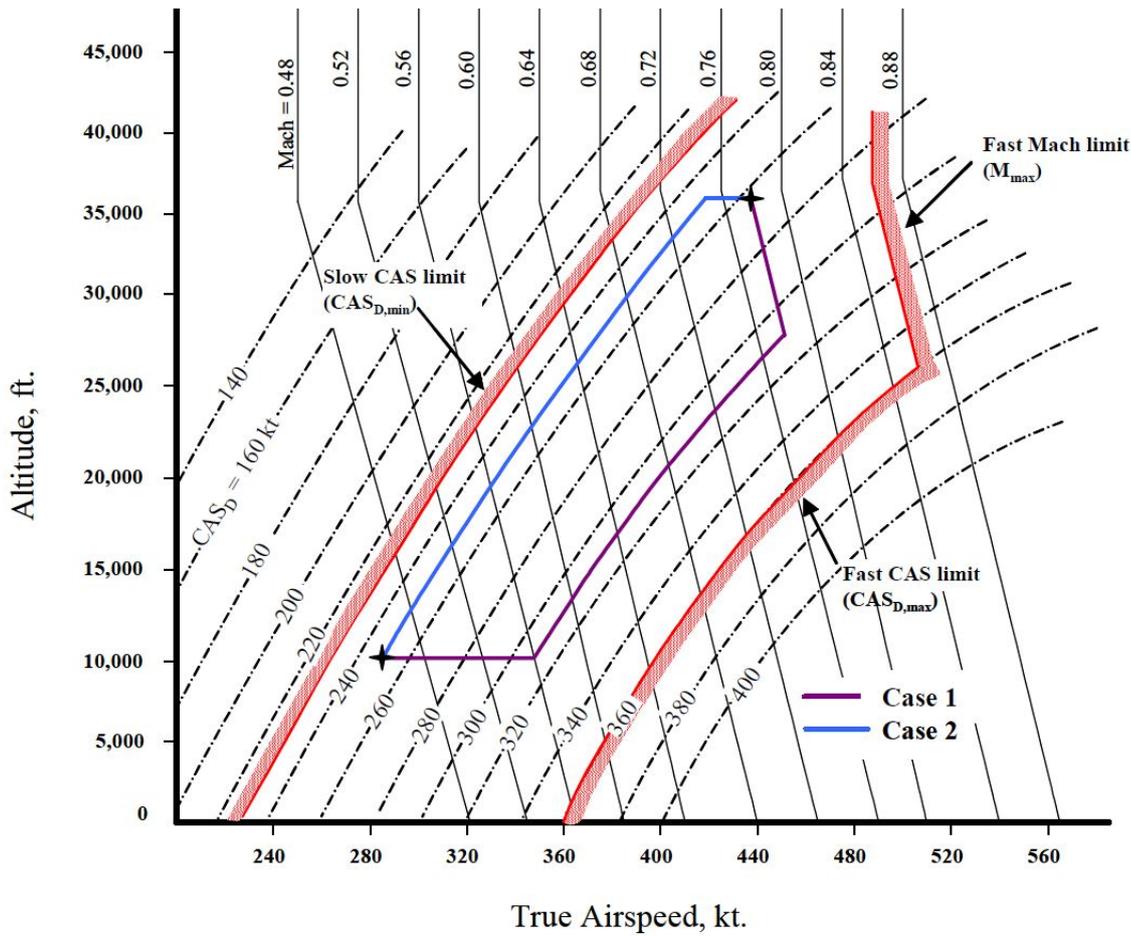


Figure 9.3: Typical descent speed envelope with examples of CAS acceleration and deceleration manoeuvres (source: [10])

the same TAS value. In this equation,  $T_0$  is the standard atmospheric temperature at Mean Sea Level (MSL) (288.15 [°K]) and  $\theta_{trans}$  is the temperature ratio at the transition altitude, given by Equation 9.3. There,  $\tau$  is the ISA temperature gradient with altitude below the tropopause ( $-0.0065$  [°K/m]),  $R$  is the real gas constant for air (287.05287 [m<sup>2</sup>/Ks<sup>2</sup>]),  $g$  is the gravitational constant, and  $\delta_{trans}$  is the pressure ratio at the transition altitude given by Equation 9.2. There,  $a_0$  is the ISA speed of sound at sea level (340.29 [ $\frac{m}{s}$ ]) and  $\kappa$  is the Adiabatic index of air (1.4[-]).

$$H_{p, trans} = \left( \frac{1000}{(.3048) \cdot (6.5)} \right) \cdot [T_0 \cdot (1 - \theta_{trans})] \tag{9.2}$$

$$\theta_{trans} = (\delta_{trans})^{-\frac{\beta_T R}{g}} \tag{9.3}$$

$$\delta_{trans} = \frac{\left[ 1 + \left( \frac{\kappa-1}{2} \right) \left( \frac{V_{CAS}}{a_0} \right)^2 \right]^{\frac{\kappa}{\kappa-1}} - 1}{\left[ 1 + \frac{\kappa-1}{2} M^2 \right]^{\frac{\kappa}{\kappa-1}} - 1} \tag{9.4}$$

Now the crossover level is known, one can calculate the effect on the flying time when decreasing the CAS descent. For this, one needs to convert the CAS in descent to True Airspeed to calculate what the flying time is over the segment (again, assuming zero wind conditions, such that the TAS equals the ground speed). The  $V_{TAS}$  for a given  $V_{CAS}$  is dependent on the air pressure and density at the altitude the aircraft is flying at. To go from  $V_{CAS}$  to  $V_{TAS}$ , Equation 9.5 is used, where  $p_0$  and  $\rho_0$  are the standard atmospheric pressure and density at MSL (101325[Pa] and 1.225 [ $\frac{kg}{m^3}$ ] respectively).

$$v_{CAS} = \left[ \frac{2}{\mu} \frac{p_0}{\rho_0} \left\{ \left( 1 + \frac{p}{p_0} \left[ \left( 1 + \frac{\mu}{2} \frac{\rho}{p} V_{TAS}^2 \right)^{1/\mu} - 1 \right] \right)^\mu - 1 \right\} \right]^{1/2} \quad (9.5)$$

The total delay that can be absorbed in descent is calculated by calculating the time for the descent with the minimum target CAS descent of 250, and subtracting from this the time for the descent with the average CAS descent, given in Table 9.1. The steps for the regular descent, where the CAS cruise (equivalent) is higher than the CAS descent, are given by:

1. Determine the crossover altitude
2. Check if the aircraft is flying above or below the crossover altitude. As no up to date information on the current FL is given in the EFD messages, the FL of the first point in the route is taken as the current FL.

For the part from FL Crossover to FL Cruise:

3. Integrate backwards from the Crossover FL to the Cruise FL (or current FL) in steps of 100ft, determining the TAS at each step.
4. Determine the flying time in each integration step by using a descent angle of 3° (adapted from operational expert at LVNL for Mach descent)

For the part from FL IAF to FL Crossover:

5. Integrate backwards from the IAF (FL100) to the FL Crossover (or the current FL) in steps of 100ft, determining the TAS at each step.
6. Determine the flying time in each integration step by using a descent angle of 2.6° (adapted from operational expert at LVNL for CAS descent)

The flying times for both parts are then added to get the total flying time for descent.

The steps in the module for a type 2, step 1 and step 2 are the same. Then step 5 and 6 are done, from IAF to FL Cruise instead of FL Crossover (as the FL crossover is likely higher than the FL Cruise). Note that the deceleration segment at level-flight was not modelled due to a lack of a kinematic model of the aircraft. Instead that part of the cruise is flown accounted for by a cruise segment at the regular Mach cruise, resulting in a faster time for the total descent that it would be in reality. This makes the delay that is calculated conservative, and thus this inaccuracy is accepted. It should also be noted that for the maximum delay absorption wind conditions were not taken into account. Therefore, the TAS in the steps above equals the Ground Speed, which allows for calculation of the flying times.

#### 9.2.4. Cruise

In this subsection the maximum delay that can be imposed for absorption in cruise is determined. In section 3.4 linear delay absorption techniques were discussed, in which there is an equivalent airspeed ( $v_{eq}$ ) at which an aircraft absorbs delay without extra fuel usage. Between the region of this equivalent airspeed and the regular airspeed, fuel can even be saved. Unfortunately operators do not share fuel consumption, weight and Cost Index data with the Air Navigation Service Provider. Therefore, another approach is necessary to determine the maximum delay in cruise. For this, an operational perspective is chosen in which a fixed Mach cruise speed decrease is chosen. Common Mach speed decreases used in Extended AMAN proposals range between M0.02 and M0.04 [53]. In this research, the conservative value of M0.02 was chosen. Based on the results obtained with this speed reduction, conclusions can be drawn on whether a substantial part of the delay can be absorbed in cruise, or if a higher speed decrease is required.

In the discussion in section 3.4 it was established that the target CAS descent speed should be decreased as much as possible, before decreasing the Mach Cruise when viewing delay absorption from a fuel efficiency standpoint. This strategy is used here as well, so the calculation of the maximum delay is based on the minimum CAS descent case (case 2 in Figure 9.3). The increase in crossover altitude that comes with a decrease in CAS has an effect on the distance of the cruise phase, as different descent angles are used for CAS descent and Mach Descent (2.6° and 3.0° respectively).

The algorithmic steps are fairly straightforward, as given below:

1. Determine the horizontal path length for the CAS descent part
2. Determine the horizontal path length for the Mach descent part (if there still is one)
3. Subtract the descent distance from the cruise distance

4. Calculate the time in cruise based conversion from the Mach number to the TAS given by Equation 9.6, where  $a$  is the mach number,  $\gamma$  is the ratio of specific air (a constant),  $R$  is the gas constant for air and  $T$  is the absolute altitude at the Cruise Altitude. In the current analysis for ground speed of aircraft in cruise, the ISA temperature standard for the tropopause is used:  $216.65^\circ K$
5. Perform the same step as given above, but with the Mach number decreased by 0.02
6. Subtract the time calculated in Step 4 from the time calculated in Step 5. This is the total delay that can be absorbed in the cruise phase

$$V_{TAS} = a\sqrt{\gamma RT} \quad (9.6)$$

### 9.3. Bunching Detection Module

After the information is parsed in the Message Class Module shown in the previous Section, a bunching prediction can be generated, shown in Figure 9.4. In section 4.3 a definition for bunching was given, in which a situation is considered bunching when two or more aircraft arrive at an IAF in a shorter time span than the required RECAT WTC separation. In subsection 8.1.6 the uncertainty curves were shown for estimates at COP were shown, which are the basis of a probabilistic bunching detection technique at the IAF (as discussed in subsection 8.1.2). From the fitted Probability Density Curves, the Cumulative Density Functions can be constructed. These Cumulative Density Functions, imposed on the EFD Estimates at the IAF give the probability of arrival the IAF. In other words, finding the value of the Cumulative Density Function at a given time yields the probability that a flight is exactly at the IAF, or has passed it. Finding the probability earlier in time, a time delta equal to the Wake Turbulence Category separation, yields the probability that the aircraft passed the IAF, and the IAF is now 'free' again from a bunching perspective. Therefore, subtracting this probability from the probability that the aircraft is at the IAF or has passed, yields the probability that the aircraft is currently at the IAF, or it still 'occupies' because of the Wake Turbulence Category separation. This is shown in Equation 9.7.

$$O_i(t) = CDF_i(t) - CDF_i(t - WTC_{separation}) \quad (9.7)$$

The bunching probability was based on the method presented in [41], where a runway capacity model was constructed. The probability that 2 or more aircraft are occupying the IAF at the same time is calculated by calculating the probability that there are exactly zero or exactly one aircraft occupying the IAF. The probability that there are exactly zero aircraft occupying the IAF is shown in Equation 9.8, which is the product of each aircraft not occupying the IAF. The probability that there is exactly one aircraft occupying the IAF is shown in Equation 9.9. This equation consists of the sum of the probabilities of all flights to be the only one occupying the IAF at that moment, times the probability of not being at the IAF for all other flights. Now the final equation, Equation 9.10, shows the probability that there are two or more flights occupying the IAF at the same time or within the WTC separation.

$$P_{0(t)} = \prod_n^{i=0} (1 - O_i(t)) \quad (9.8)$$

$$P_{1(t)} = \sum_n^{i=1} \left( O_i(t) \prod_n^{j=1, i \neq i} (1 - O_j(t)) \right) \quad (9.9)$$

$$P_{>=2(t)} = 1 - P_{1(t)} - P_{0(t)} \quad (9.10)$$

Since there are three Initial Approach Fixes, each with an independent stream of traffic, one could perform the bunching detection as shown above for each of these IAFs. However, when considering the workings of the AMAN, it becomes clear that this is not the most effective way as the independent traffic streams of each IAF might be merged towards the same runway. Therefore, by taking into account runway configurations and information on which IAF feeds which runway, the impact on the AMAN efficiency is increased.

In the current study, the runway configuration is assumed to be known 5 hours in advance. In current operations this is not the case, however, LVNL is working on a Decision Support Tool that would enable establishing accurate estimates of the runway configurations at this timeline. A conversion table which establishes which IAF feeds which runway for each runway configuration is given in Table 9.5, and will be discussed in

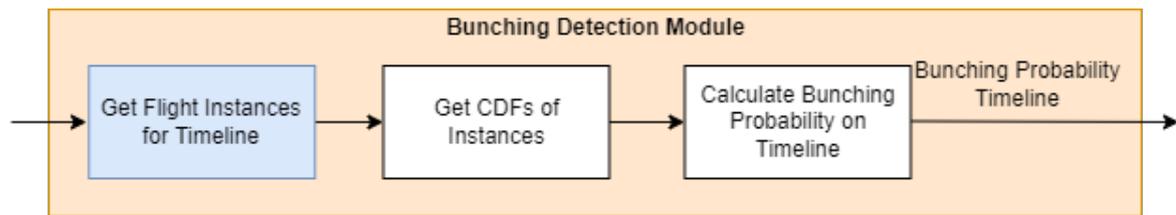


Figure 9.4: System architecture of the bunching detection module

subsection 9.6.1.

Based on this information, three prediction timelines can be constructed. The first prediction timeline shows all flights that are predicted to arrive at a point in time where only one landing runway is active, which means that the aircraft of all three IAFs are shown on this timeline for periods with a single runway configuration. The second and third prediction timelines show aircraft which are predicted to arrive at a point in time where two runways are in use. Based on the conversion table, RIVER and SUGOL will feed one of the runways, and ARTIP will feed the other runway. Therefore, the second and third prediction timeline show the flights based on this division. To illustrate this, the PDFs on the three timelines were plotted in Figure 9.5, Figure 9.6 and Figure 9.7 for the single runway configuration, and the two double runway configurations respectively. The situation shown was on 2021-05-01 04:49:27. The corresponding bunching probability curves were plotted in one plot, shown in Figure 9.8.

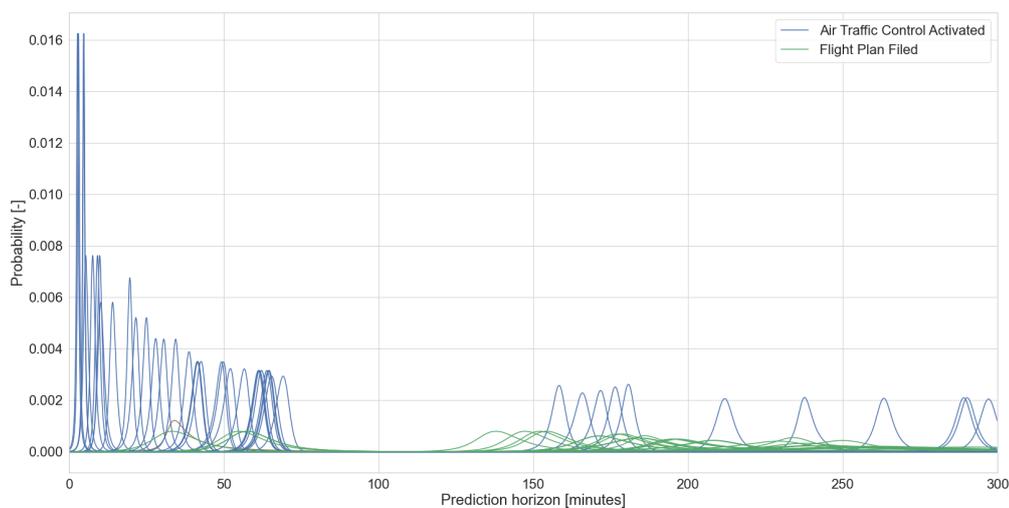


Figure 9.5: Probability Density Curves for aircraft expected to arrive during a one-runway landing configuration

Logically, this division of the aircraft over the timelines has effect on the bunching prediction as is clear in Figure 9.8. As previously shown in the equations above, the RECAT WTC separation is integrated in the technical definition of bunching. Therefore, the sequence of the flights should be known, such that the correct separation is used. As the uncertainty is high at high prediction horizons, the exact order of arrival (the sequence) cannot be constructed. Instead, each time an update is performed in the Message Class Module, a sequence is generated with on the expected value for the arrival time based on the Probability Density Functions. Then, for each given aircraft, the follower is found based on the expected values for all aircraft that are on the same timeline. This means that at each update, the bunching prediction is not only influenced by the change of the Cumulative Density Curves and how these overlap, but also by the change in sequence.

Note that the output of the bunching detection is a timeline which shows the probability that bunching

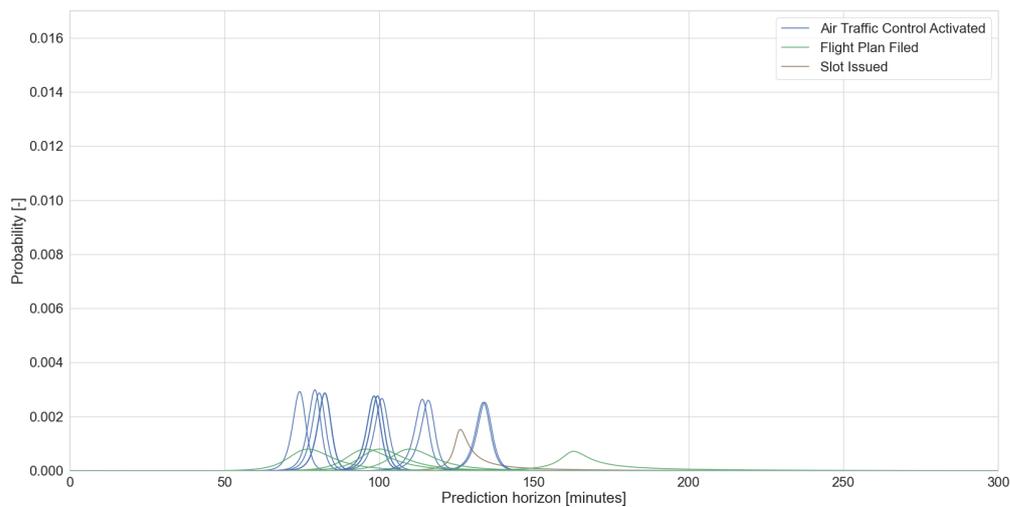


Figure 9.6: Probability Density Curves for aircraft flying over the ARTIP stack, expected to arrive during a two-runway landing configuration

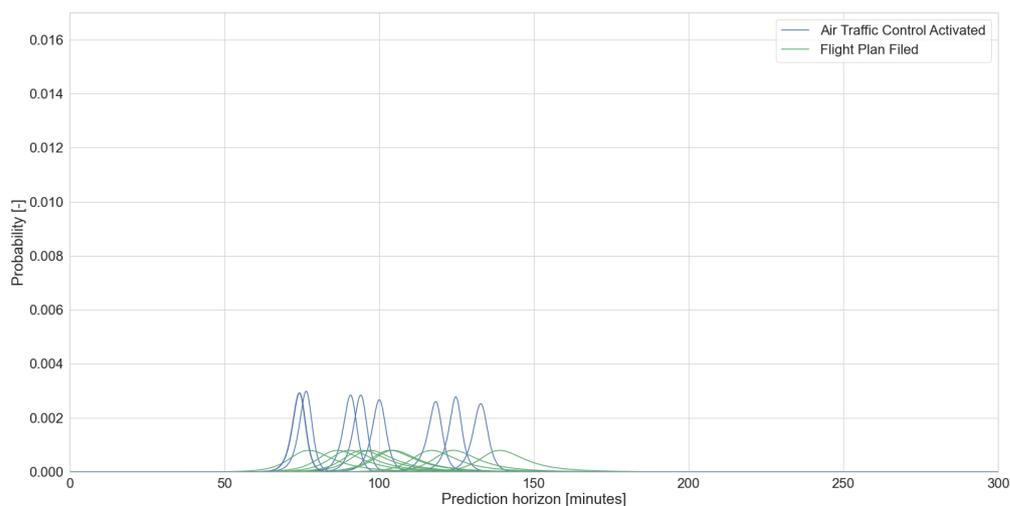


Figure 9.7: Probability Density Curves for aircraft flying over the RIVER and SUGOL stack, expected to arrive during a two-runway landing configuration

occurs. Since this output is given in probabilities, a definition for bunching could be given based on a threshold value: if the probability is higher than this threshold, the chances of bunching occurring are high, and thus debunching should be considered. The design of the debunching model will be discussed in the next section.

## 9.4. Solution Module

After detection of bunching, a solution for the debunching will be generated based on the aircraft instances, and the time with which the corresponding distributions should be shifted. Therefore, the independent variables are the delay each aircraft needs to absorb. The dependent variables then consist of the change in the

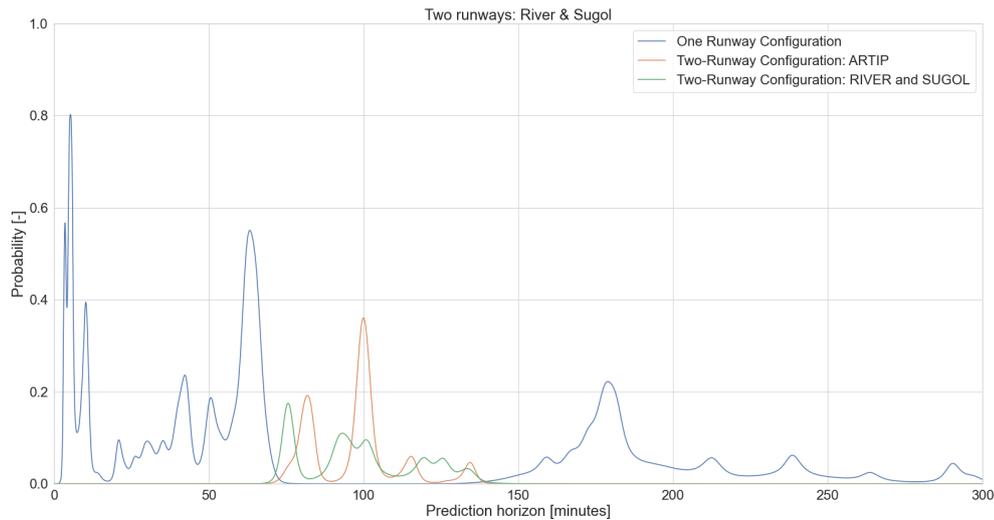


Figure 9.8: The bunching prediction timelines for the runway configurations, combined in one plot

bunching peak, and the delay necessary for an aircraft to achieve this decreased bunching peak. The solution to the problem can be found by stating it as an optimisation problem.

In order to gain an understanding of the type of optimisation problem that is to be solved, one has to examine the underlying equations. In section 9.3 it was explained that the Cumulative Density Function (CDF) was used to find the probability of an aircraft occupying the IAF. These CDFs were obtained by using Johnson Distributions as described in subsection 8.1.5. Note that the CDF is defined as the integral of the PDF, as shown in Equation 9.11 for the unbounded case, with the variables as defined in subsection 8.1.5.

$$CDF = F(x, a, b, loc) = \int_{-\infty}^{(x-loc)} \frac{b}{\sqrt{(x-loc)^2 + 1}} \cdot \phi(a + b \cdot \log((x-loc) + \sqrt{(x-loc)^2 + 1})) \, d(x-loc) \quad (9.11)$$

As stated before, the independent variables are the delay each aircraft needs to absorb, corresponding with a change in the *loc* parameter in the equations above. The main objectives for solving the bunching are to achieve a bunching percentage below a threshold value, to minimise the amount aircraft that need to absorb delay, and to minimise the time shift with which an aircraft is delayed.

### Solvers & Problem Formulation

Efficient solvers for such optimisation problems are gradient based optimisation algorithms where the search direction of the algorithm is defined by the gradient of the function at the current point. However, these methods do not suit problems with discrete-valued variables, non-convex objective functions and non-differentiable objective functions and constraints [46]. For example, the derivative information of the objective function might be unavailable, or impractical and time consuming to obtain. [9].

The optimisation problem at hand is identified to be non-linear, and non-convex due to the nature of the PDFs and the way the delay percentage values are calculated. Furthermore, it is a multivariate problem, where the amount of variables corresponds with the amount of aircraft, with the *loc* value for each aircraft being the optimisation variables. The problem should also be treated as bound-constrained, as each aircraft has a maximum amount of delay it can absorb.

The result of the algorithm that incorporates the equations in section 9.3 (Equation 9.8, Equation 9.9, and Equation 9.10) is an array consisting of the bunching probability at each time step on a given timeline. Therefore, a computationally efficient manner of achieving a bunching percentage below the threshold at each time step is to analyse the maximum value of the array, shown in Equation 9.12, and minimise that

function. However, the  $max()$  function is not available in gradient based optimisation solvers, as it is not continuously differentiable and can cause a gradient-based optimiser to fail to converge.

$$f(\vec{x}) = max(P_{>=2(t)}) \quad (9.12)$$

Instead, without the  $max()$  function the whole array should be optimised simultaneously, which adds an extra dimension to the problem, which common optimisation solvers cannot work with. Also, as the optimisation variables are defined as the locations of the probability density functions, differentiation can only be done at a single point (an  $x$  value) on the timeline, yielding gradient based optimisation not useful. Furthermore, it is assumed that most gradient based optimisation solvers require the algorithm to be written in closed form, which is complex and impractical due to the integral of the PDF and the summation and Pi-function in the probabilistic equations.

The benefit of the current optimisation problem is that the problem is in so-called 'white-box' form, meaning there is analytical knowledge of the equations (in contrast to so-called 'black-box' functions common in derivative free optimisation problems). This knowledge allows to define heuristics and to use previous knowledge of the problem to solve it. Concluding, the optimisation problem at hand can be defined as a multivariate mixed integer, non-convex, non-linear, bound-constrained optimisation problem, for which optimisation solvers will not suffice.

As no solvers could be used, the next step was to evaluate the use of gradient-based algorithms (which do not need the use of solvers) and derivative-free methods. Derivative-free methods typically limit performance of optimisation problems in terms of accuracy and expense, relatively to the use of gradient-based methods [35]. Therefore, the use of gradient-based methods should first be explored before exploring derivative-free methods. Although the algorithm to be optimised is not continuously differentiable, it is a white-box function, and therefore the derivative could be estimated using numerical differentiation. The benefit of using numerical differentiation is that only the function values are used, without requiring higher order information. A general downside of the method is the quality of the derivative estimate. However, as the delay values are discrete (measured in seconds) this might not be an issue. One method for numerical differentiation is using finite differences, which is based on the general equation for calculating the real derivative (shown in Equation 9.13). A commonly used formula for finite differences is the symmetric difference quotient, shown in Equation 9.14. The downside however, is that a large amount of function evaluations might be necessary.

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \quad (9.13)$$

$$\frac{f(x+h) - f(x-h)}{2h} \quad (9.14)$$

As stated before, the goal of the optimisation is to achieve a prediction timeline where the bunching percentage is below a given threshold. There are, however, several ways in which this can be achieved. The first method is to just minimise the maximum value of the delay (defined as  $f(\vec{x}) = max(P_{>=2(t)})$ ), constrained by the maximum delay bounds. This is shown in Equation 9.15, where  $\vec{x}$  is the vector of the delay values,  $n$  is the number of aircraft, and  $x_i^l$  and  $x_i^u$  are the upper and lower bounds of the delay respectively.

$$\begin{aligned} &\text{minimise} && f(\vec{x}) \\ &\text{subject to:} && x_i^l \leq x_i \leq x_i^u, \quad i = 1, \dots, n. \end{aligned} \quad (9.15)$$

The downside of this method however, is that this results in excessive delay as it disregards the threshold that needs to be reached. A way to mitigate this is using a multi-objective optimisation formulation, shown in Equation 9.16, where the sum of delay values vector is included in the objective function. The downside of this method is that it requires an a priori preference selection such that Weighted Sum or Penalty functions can be used [6]. The way an a priori preference selection is carried out will have influence on the outcome of the optimisation, and the mentioned multi-objective methods make the optimisation problem overly complicated.

A more elegant method is shown in Equation 9.17. There, the sum of the total delay values is minimised, while satisfying the bunching threshold constraint, which is defined using the bunching threshold ( $B_{\%}$ ) and the outcome of the maximum bunching probability ( $f(\vec{x})$ ). Formulating the problem in this way will ensure a

solution is found in which airlines will have to absorb the least amount of delay, while the bunching threshold is kept below the required value. However, no gradient based methods can be used with this formulation as the function  $f(\vec{x})$  is not in the objective function anymore.

$$\begin{aligned} &\text{minimise} && f(\vec{x}), \sum \vec{x} \\ &\text{subject to:} && x_i^l \leq x_i \leq x_i^u, \quad i = 1, \dots, n. \end{aligned} \quad (9.16)$$

$$\begin{aligned} &\text{minimise} && \sum \vec{x} \\ &\text{subject to:} && f(\vec{x}) \leq B_{\%} \\ &&& x_i^l \leq x_i \leq x_i^u, \quad i = 1, \dots, n. \end{aligned} \quad (9.17)$$

### Derivative Free methods & method selection

Even though the use of gradient-based methods might be beneficial in terms of performance compared to derivative-free methods, the use of gradient-based methods might result in worse solutions due to the local character of the algorithms. One way to overcome this is by using globalisation techniques such as Trust-Region, Line-Search or Multistart techniques [6] [35]. However, such problems might be better addressed by derivative-free optimisation techniques. Combined with the the optimisation problem formulation discussed above and shown in Equation 9.17, it was decided that derivative free methods should be implemented.

In general, derivative-free algorithms for optimisation can be divided between 'direct' and 'model-based' algorithms. In direct algorithms the search directions are determined by computing values from the function to be optimised directly. In model-based algorithms a surrogate model of the optimisation function is constructed from which values are obtained. A surrogate function is created by first sampling from the function within a given search space, and building a surrogate model using those values. The surrogate model is then optimised and evaluated instead of the original function, which is useful for black-box functions of high complexity or with high computational costs [48]. However, as previously stated, the optimisation problem at hand is in white-box form, and the computational complexity to evaluate function variables is relatively low. Therefore, it is chosen to use direct optimisation algorithms.

Another difference between algorithms is between stochastic and deterministic algorithms, which defines whether random search steps are used or not. In stochastic global search algorithms intermediate moves to points with a less efficient solution than the solution at hand are permitted, which mitigates local optimums and allows analysing more parts of the search-space. Stochastic algorithms are often used as the implementation is rather straightforward compared to deterministic algorithms [48].

Examples of direct search stochastic search algorithms with global characteristics are Hit-and-run algorithms, Simulated Annealing, Genetic Algorithms and Particle Swarm Optimisation. The downside of the first two methods are sensitivity to problem conditions and no guarantee of a good solution within a finite number of iterations [48]. Particle Swarm Optimisation (PSO) and Genetic Algorithms (GA) are similar in the sense that both algorithms are population-based search methods that move through solutions using a set of rules. Using PSO could lead to improvements in computational efficiency compared to GAs, specifically for unconstrained problems [26]. However, a GA is chosen due to it's intuitive nature and ease of implementation. Furthermore, with the use of a GA an efficient constraint handling method can be used (without the need for penalty functions), such that the problem formulation in Equation 9.17 can be handled. This constraint handling method will be elaborated on in the next subsection.

#### 9.4.1. Genetic Algorithm

In this chapter the general steps of the GA will be elaborated on, followed by a schematic overview of how the solver is used as a debunching module. A genetic algorithm mimics the principle of survival of the fittest found in natural selection. In general, a population of solutions is generated that is evaluated by a fitness function, and the best values are adapted and mutated such that new solutions can be found [48]. More specifically, the steps are [27]:

1. The algorithm starts by selecting random values for the delay of each aircraft, bounded by the maximum delay absorption for that aircraft. For the GA used in this research, binary encoding was used due

to the ease of implementation and since the operations in the GA are straightforward with this type of encoding. The random set of delay values is therefore encoded on a binary string of a given number of bits. Each binary string that represents a set of delay values is called a *chromosome*. For the start of the algorithm, a given number of chromosomes is created that together represent the *population*.

2. For each chromosome in the population the fitness is evaluated by the function to be minimised (the *fitness function*), which is the sum of the total delay, shown in Equation 9.17. For this, the chromosome is first decoded, such that it can be evaluated by the fitness function.
3. As the optimisation problem is constrained (shown in Equation 9.17), the fitness of each chromosome is also evaluated based on the constraint value. For each chromosome therefore, the delay values retrieved by decoding, and used to shift the CDF of the corresponding aircraft with the given amount of delay. Then, the maximum bunching probability is recalculated with the method presented in section 9.3. Note here that as the WTC separation has influence on the bunching probability, the sequence will inherently also be optimised for when optimising for the bunching probability.
4. The next step is *selection*, where certain chromosomes are randomly chosen for mating and reproduction. The general principle for selection is that chromosomes that have the best fitness value, will have a higher probability of becoming a *parent* for reproduction [26]. In the current research Tournament Selection is chosen because of its efficiency and ease of implementation. The principle of Tournament Selection is that multiple chromosomes are selected randomly from the population, and compared based on their fitness value. The chromosome with the best fitness value is selected for the generation of the new population. Commonly, the number of number of chromosomes selected for each tournament is 2 (called binary selection) or 3 [27]. The number of tournaments is equal to the pool size. As the current problem is constrained, a constraint handling method for GAs is used, presented in [12]. In this method, three extra rules are applicable to the tournament selection [12]:
  - (a) When two feasible (based on the constraint value) solutions are compared, the one with better fitness function value is chosen
  - (b) When one feasible and one infeasible chromosome are compared, the feasible solution is chosen
  - (c) when two infeasible chromosomes are compared, the one with smaller constraint violation is chosen

This means that a chromosome that satisfies the constraint (that says the bunching peak should stay below a given threshold) is always preferred over a solution that yields less delay (a lower fitness function value).

5. Now, with the population of chromosomes that result from the selection, a *crossover* is performed to produce new solutions. In a crossover, the chromosomes are selected in pairs, and a random point on the chromosomes is selected from which the binary values are exchanged between the chromosomes, creating new solutions. This process is shown in Figure 9.9. Whether the crossover happens between a pair is determined by a probability called the crossover rate, which will be tuned in subsection 9.4.2.
6. With the *offspring* that results from the crossover, the *mutation* operation will take place. In mutation for binary encoding, randomly chosen bits will be switched from 0 to 1 or vice versa. Whether this happens to a bit is determined by the crossover rate, which will be tuned in subsection 9.4.2. This operation is shown in ???. In case a mutation happens, the chromosome will yield a different value for the delay value for a given aircraft. This value will always be within the bound constraints as the bits are scaled between the bounds.
7. With the new population, the steps from Step 2 are repeated, until the maximum number of iterations is reached.

In Figure 9.11 a schematic overview of the debunching module is presented. It consists of the main steps of a GA given above, and of interdependencies with the other modules. These interdependencies are shown by the corresponding colors to those modules. It is shown that the solver is parallelised three times, from which the best solution is selected. The reason for this will be discussed in the section 9.5.

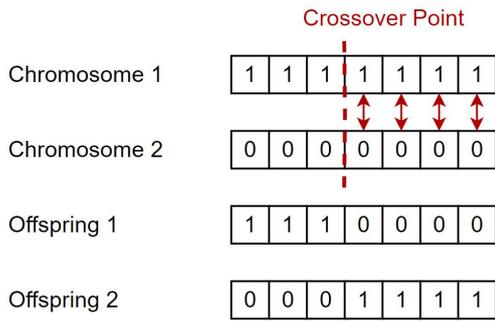


Figure 9.9: An example of Crossover in a GA

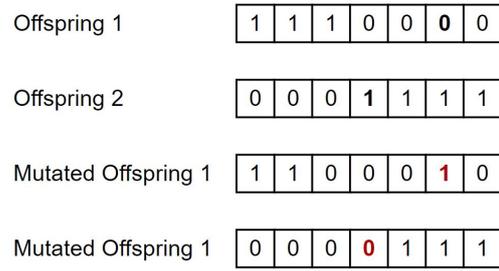


Figure 9.10: An example of mutation in a GA

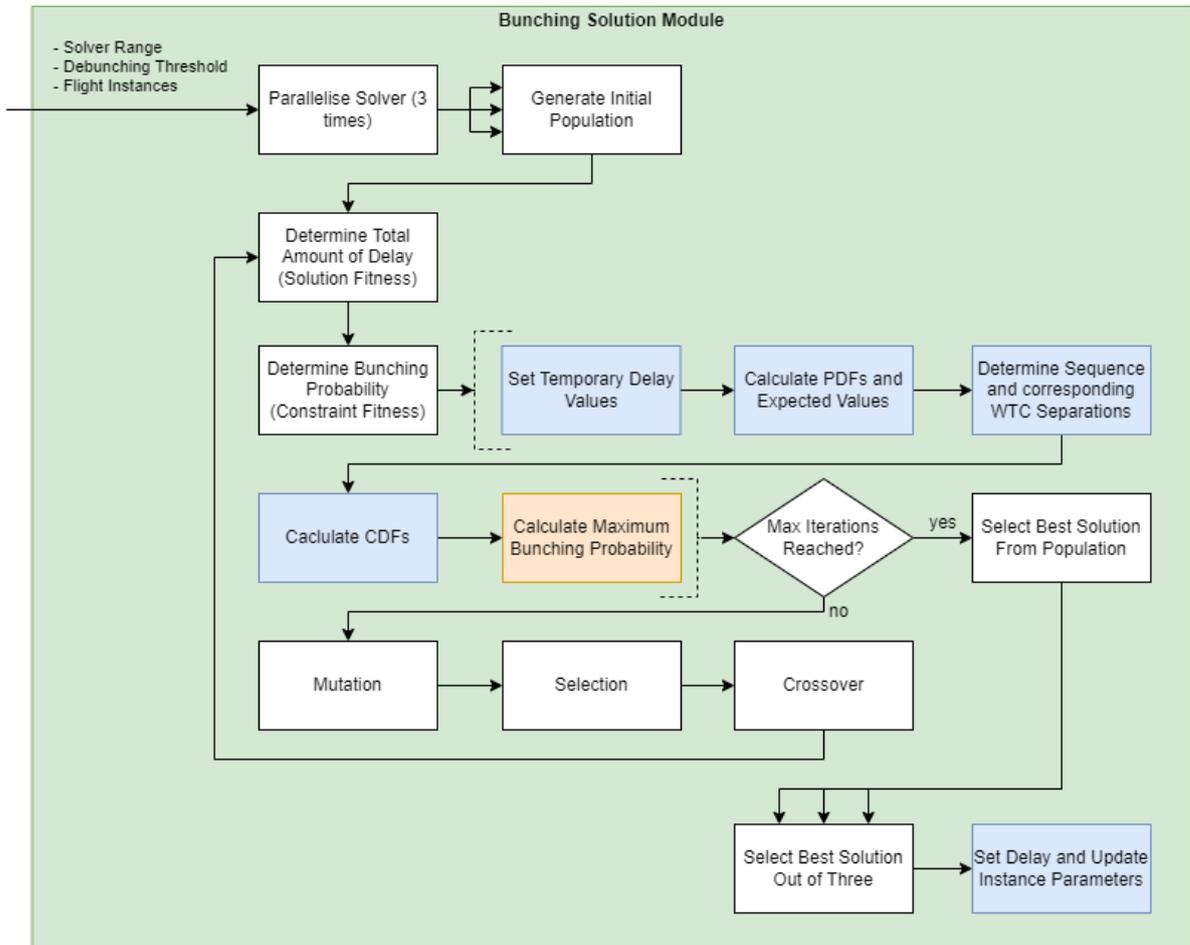


Figure 9.11: System architecture of the debunching module

### 9.4.2. Genetic Algorithm Parameter Selection

Setting the correct parameters for an GA is a complicated process which is problem specific, and requires tuning parameters which interact in a complex way. In [27] a review was performed for parameter setting in GAs as there are no definitive values for all types of problems. It was found that common population sizes ranged from 25 to 400, with some studies using lower or significantly higher values. For the crossover rate common rates ranged between 0.5 and 1, with most studies being at the higher end of the range (near 0.9). For the mutation rate common values ranged between 0.001 and 0.1.

For debunching, a preliminary analysis was conducted in order to understand the effect of the parameters above on the GA solution for debunching. Since a large range was found for the mutation rate in [27], a brief analysis was performed with mutation rates of 0.001, 0.01 and 0.1 (with: population size = 100, crossover rate = 0.9, number of iterations = 750, tournament selection size = 2). The preliminary findings suggested that a mutation rate of 0.1 stagnated improvement of the solution after a few iterations, as improved solutions were hardly found. The mutation rate of 0.01 resulted in improved iterations almost every iteration in the first iterations, but stagnated during later iterations (around 350 iterations). The mutation rate of 0.001 radically improved convergence and only resulted in stagnation when an apparent near-optimal solution was found. To check whether 0.001 would be the correct order of magnitude, a 0.0001 rate was also tested. This resulted in worse solutions than those provided by the 0.001 rate.

Based on the findings in [27] and the brief analysis mentioned in the paragraph above it was chosen to select a mutation rate range between 0.001 and 0.005 for further analysis. The crossover rate range will be varied between 0.85 and 0.95, and for the population size values of 50, 100 and 200 will be used based on the common values in [27]. Smaller values for the population size will not be used as the population size should commonly be higher than the dimensionality (the number of aircraft) [8]. In the preliminary analysis previously mentioned, it was found that the number of aircraft on the timeline often ranged between 50 and 100. Therefore, higher values for the population size will not be explored, as an increase in population results in an increase in computational complexity [8]. Since for debunching, the dimensionality continuously changes (as the number of aircraft on a given timeline is a dynamic process) a dynamic mutation rate of  $1/n$  is also explored, where  $n$  is the dimensionality times the number of bits. In [8] this rate was shown to work efficiently as it scales with the dimensionality of the problem at hand. Given the dimensionality above and a number of bits of 8, this mutation rate would yield values in the same order of magnitude as deemed effective in the brief analysis above (near 0.001).

Since tuning the mutation rate, crossover rate, population size and tournament selection size at the same time would result in too many experiments for the scope of this research, it was chosen to first only tune the crossover and mutation rate, with a fixed population size of 100 and a tournament selection size of 3. The maximum number of iterations was fixed at 750. For testing, the following independent variables and values used:

- Crossover rate: 0.85, 0.9, 0.95
- Mutation rate: 0.001, 0.002, 0.003, 0.004, 0.005,  $1/n$

Each experiment condition is repeated 10 times and the resulting values are averaged in order to average out some of the differences due to the iterative nature of the GA. With the crossover and mutation rates mentioned, the 10 repeats and the 4 samples, a total of 720 problems are solved with each 750 iterations of the GA. The results are shown in Figure 9.12, and in the Appendix for Figure B.1, Figure B.2 and Figure B.3.

From the graphs several things could be concluded. The first is that the mutation rate has more influence than the crossover rate after all iterations have been performed. It can also be seen that a mutation rate of 0.001 often converges slower, but then reaches a better minimum value than the other values. The mutation rate of 0.001 performs best in two of the 4 sub problems (Figure B.1, Figure B.3). In the other sub problems it performs second best once, and third best once. The mutation rate of  $1/n$  performs also best in two of the sub problems (Figure 9.12, Figure B.2), but it does not come second best in the others. Therefore, on average the 0.001 mutation rate performs best. The crossover rate of 0.9 performed best in three of the sub problems (Figure B.1, Figure B.2, Figure B.3) and second at the other (Figure 9.12). Therefore, this value was set as the default value for the crossover in the experiments in this research. The mutation rate was set to 0.001 because of the reasons mentioned above.

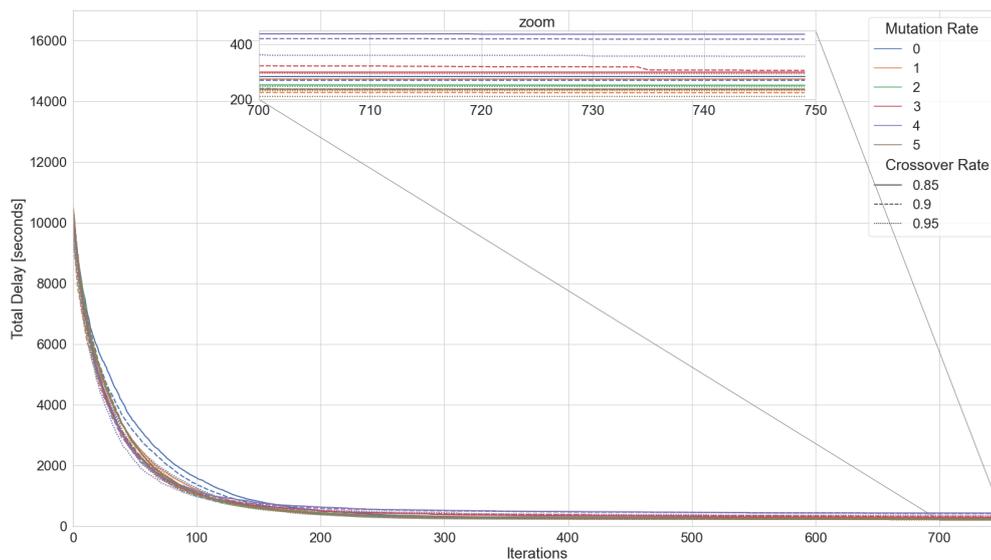


Figure 9.12: Sample problem 1: day 1, 35% threshold

With these parameters one more analysis was done to check the effect of the constraint on the convergence of the algorithm. In Figure 9.13 it was shown that although there were clear differences in the total seconds of delay, the convergence was not significantly impacted.

## 9.5. Verification and Validation of the Bunching Solver

For the verification of the solver a solution can easily be inspected both visually and by the data. The output of the bunching solver is shown in the Appendix, in Table B.1, where the threshold was set at 45%. This table shows the timestamp at which the solver was activated, which is at 5 different times during this day. Furthermore, it shows the ACID of the aircraft that were delayed, and the new delay that the solver imposed on the aircraft. The following columns provide information on the maximum delay that could be absorbed by the aircraft at that moment in time (in descent and cruise). The table also shows what the flight state is of the delayed aircraft at that point in time, and if the aircraft had been delayed in a previous iteration of the solver. In this particular example most aircraft were delayed in the early morning, on average 44 minutes before arrival at the IAF. Furthermore, for none of the aircraft the total of the maximum delay was necessary for debunching, and only one aircraft had to be delayed twice.

For verification, a bunching timeline for the single runway configuration was used at 2021-05-01 02:16:27. This means that the snapshot of the timeline is generated using real flights and corresponding EFD data, and therefore is a realistic scenario. The probability density curves for all flights are shown in Figure 9.14. The corresponding bunching prediction curve is shown by the blue line in Figure 9.16, where 41% is the maximum bunching probability. The solver was used for this situation by setting the bunching threshold requirement at 25%. The parameters of the algorithm were set as discussed in the section above. In Figure 9.15 the probability density curves after solving are shown, where it is visible that around the 150 minute mark the curves are more spread out as some curves were delayed (moved towards the right on the timeline). The red line in Figure 9.16 validates that the bunching prediction was decreased by the solver, and that the peak is now more spread out.

For the validation of the algorithm two methods were used. The first was to let the algorithm solve the same bunching timeline multiple times to check whether the model finds similar solutions in terms of total delay. Of course, it is not expected that the solutions are exactly the same in which aircraft get delayed, due to the random nature of the algorithm. This analysis was performed for the same timeline as used above,

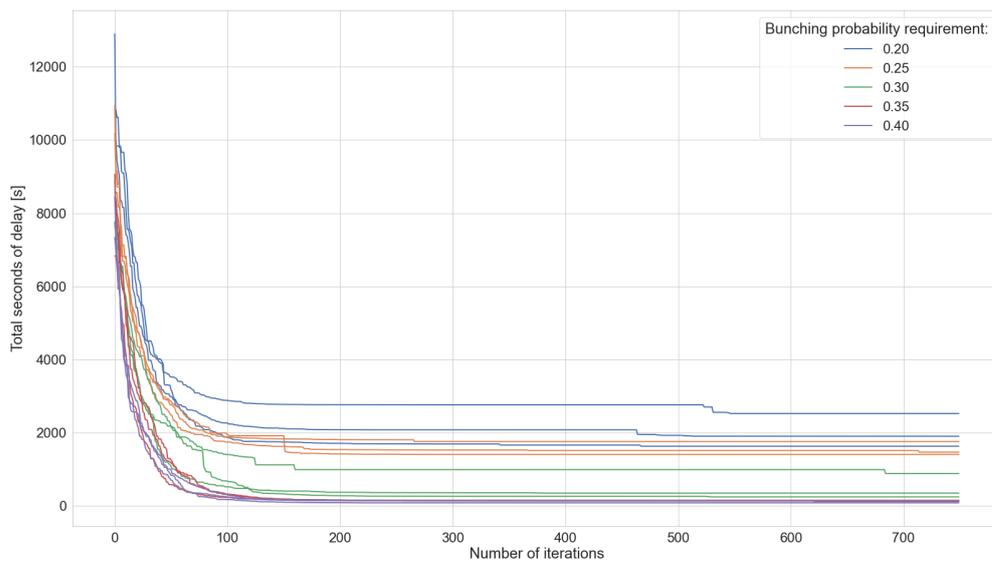


Figure 9.13: The effect of the bunching probability requirement on the Genetic Algorithm Convergence

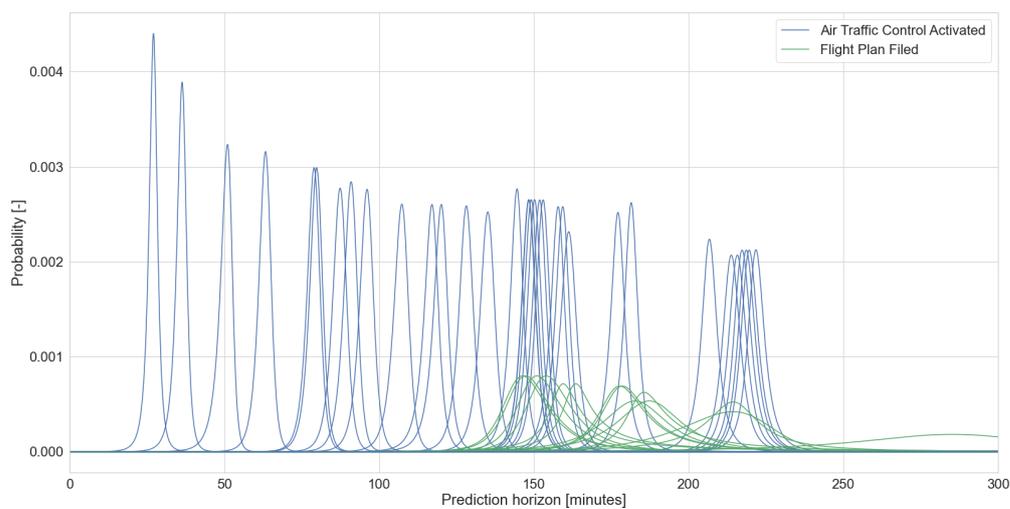


Figure 9.14: The Probability Density Curves for all flights on the prediction horizon at 2021-05-01 02:16:27

shown in Figure 9.14. In Table 9.2 the solutions were shown with the final achieved bunching probability. The requirement was set at 35%. It can be seen that in most of the solutions the requirement was met. But comparing these results to the results in which the percentage was slightly higher, this came at the expense of about 50 seconds extra delay. The conclusion that can be drawn from this is that due to the stochastic nature of the GA, the solutions might differ each time the solver is used, and therefore the solver will be activated 3 times in the experiments, from which the best solution is taken in terms of maximum delay. Note that this leads to an extra aircraft that is being delayed, which might not be optimal from an operational perspective. Therefore, in further developments, the solution in which less aircraft are delayed could be preferred at the cost of extra delay, which could easily be implemented by post processing the results of the 3 solvers.

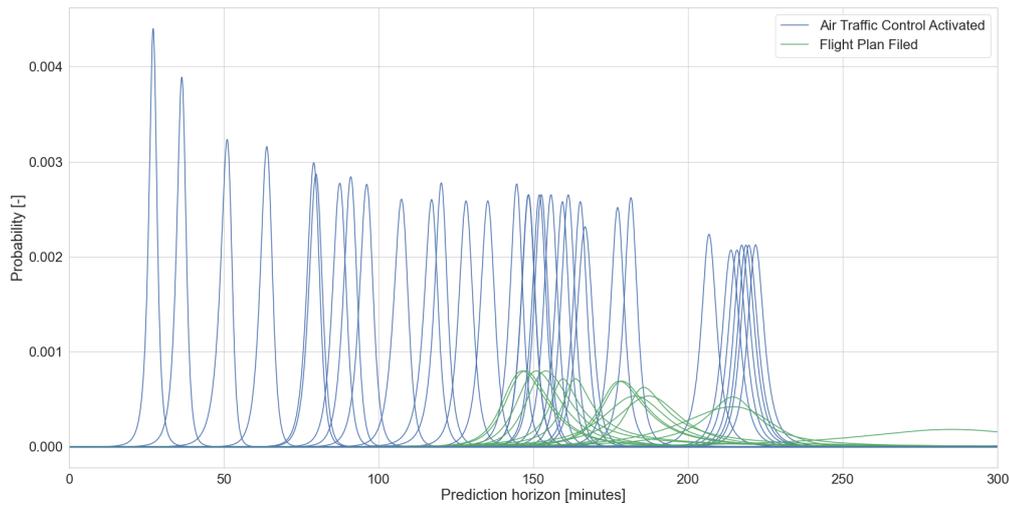


Figure 9.15: The Probability Density Curves for all flights, for the solved prediction horizon

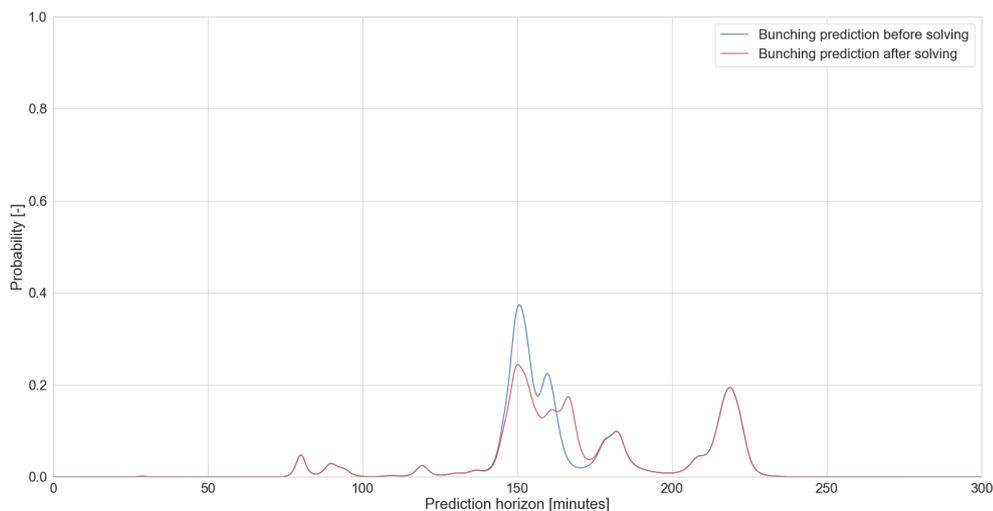


Figure 9.16: The bunching detection timeline before and after solving

The second way of validation was done by analysing the optimality of the solution by comparison with a grid search method. A grid search (also called exhaustive search or brute force) is a derivative free method in which possible values for all variables are equally spaced on a grid to find a (near) optimal solution to a problem. This means that every combination is tried, or almost all combinations in case a lower resolution is required for the solution. When trying every combination and thus finding the solution that leads to the minimum delay while satisfying the bunching threshold the optimal solution to the optimisation problem is found. With these values, the genetic algorithm solution can be validated and the algorithm can be analysed for optimality. The obvious downside of this method is that it is slow, as the resolution for the delay values that can be given is in seconds, and the maximum delay absorption for aircraft was found to be in the order of magnitude of hundreds of seconds.

As an example, trying every combination of delay with 40 aircraft with each 400 possible delay values,

	Bunching Percentage	AC ID Delayed	Imposed Delay [s]	Total Imposed Delay [s]
1	0.341 %	SIA324	[144]	144
2	0.351 %	SIA7957, SIA324	[20, 72]	92
3	0.341 %	SIA324	[144]	144
4	0.341 %	SIA324	[144]	144
5	0.346 %	SIA7957, SIA324	[41, 72]	113
6	0.341 %	SIA324	[144]	144
7	0.341 %	SIA324	[144]	144
8	0.350 %	SIA7957, SIA324	[20, 74]	94
9	0.351 %	SIA7957, SIA324	[10, 81]	91
10	0.341 %	SIA324	[144]	144

Table 9.2: Results of running the bunching solver 10 times for the same scenario

Solution Index	Delay Values [s]	Total Delay [s]	Constraint Value [-]
35201	[0, 220, 0, 10]	230	0.099291
36800	[0, 230, 0, 0]	230	0.099724
97601	[10, 210, 0, 10]	230	0.099203
99200	[10, 220, 0, 0]	230	0.09953
160001	[20, 200, 0, 10]	230	0.099021
161600	[20, 210, 0, 0]	230	0.099243
222401	[30, 190, 0, 10]	230	0.099281
224000	[30, 200, 0, 0]	230	0.098900
284801	[40, 180, 0, 10]	230	0.099429

Table 9.3: Solutions found by the Grid Search for the validation scenario

yields a permutation described as  $400P_{40}$ , which results in  $1000!/960!$  options. Therefore, a smaller sub problem is created which can be explored in a feasible time span. For this the resolution of the delay values is set to 10 seconds, as to decrease the number of options and to be able to give an indication of whether the bunching solver could lead to a near-optimum. The situation at 2021-05-01 02:16:27 was taken again, for which the PDFs were plotted in Figure 9.14. The maximum bunching peak was found to be near 150 minutes (shown in Figure 9.16). Therefore, the aircraft that were selected for the validation test problem were aircraft that were on the prediction horizon between 148 and 152 minutes. This resulted in four aircraft, for which the bounds on the maximum delay absorption are given by: [0, 491], [0, 407], [0, 560] and [0, 475]. These four aircraft resulted in a bunching peak of 0.17%. Therefore, the constraint for the debuncher was set at 10%. For each combination of values the effect on the bunching probability was calculated, and the solutions satisfying the bunching threshold were ordered based on the maximum delay that was imposed.

Hundreds of thousands of solutions were found that satisfied the constraint value. The minimum delay with which this is possible was found to be 230 seconds for which only tens of solutions were found, of which the first 10 options are shown in Table 9.4. Note that as the resolution is in 10 seconds, the optimum delay may be between 220 and 230 seconds.

The GA was activated 3 times, from which three different solutions were found. One solution yielded a total delay of 256 seconds. For the other two times the solution was found to be near-optimum. The first solution yielded a total delay of 231 seconds. The second one a total delay of 223, for which all iterations in which a new solution is found are shown in Table 9.4. From this it can be concluded that although the GA might be sensitive to initial conditions, it is capable of finding near-optimum solutions. The sensitivity to initial conditions can be mitigated by running several instances of the GA in parallel, and taking the best solution, as was done in this validation routine. In the experiments, the solver will be activated 3 times in parallel, from which the best solution is chosen.

GA Iteration	Delay Values [s]	Total Delay [s]	Constraint Value [-]
0	[319.3, 21.4, 202.3, 28.7]	571.9	0.099466
1	[239.7, 103.3, 26.2, 66.7]	436.1	0.099466
3	[35.4, 272.6, 13.1, 66.7]	388.0	0.094677
8	[35.4, 272.6, 4.3, 27.8]	340.3	0.087662
10	[35.4, 272.6, 1.1, 8.3]	317.5	0.084306
14	[35.4, 271.8, 1.1, 8.3]	316.7	0.084426
18	[35.4, 271.0, 1.1, 8.3]	315.9	0.084426
21	[35.4, 271.0, 1.1, 7.4]	315.0	0.084260
22	[35.4, 220.1, 1.1, 8.3]	265.1	0.091938
23	[35.4, 220.1, 1.1, 7.4]	264.1	0.091727
26	[34.5, 220.1, 1.1, 7.4]	263.2	0.091926
28	[35.4, 220.1, 0.0, 7.4]	263.0	0.091508
29	[34.5, 220.1, 0.0, 7.4]	262.1	0.091715
31	[34.5, 217.0, 1.1, 7.4]	260.0	0.092500
32	[3.8, 220.195, 1.1, 7.4]	232.5	0.098401
34	[3.8, 217.0, 1, 7.4]	229.3	0.098935
44	[3.8, 217.0, 0.0, 7.4]	228.2	0.098747
53	[0.0, 216.2, 0.0, 7.4]	223.6	0.099475

Table 9.4: Iterations of the GA for the validation scenario

## 9.6. AMAN module

In order to analyse what the effect of debunching is on the efficiency in the arrival process an AMAN module is used. As the AMAN is responsible for the planning of the inbound traffic (see Figure 3.1), the efficiency of the planning can be used as a proxy for the inbound traffic efficiency. The parameters of interest are the delay in the AMAN planning and the throughput per 20 minutes. For this an AMAN is created based on the AMAN currently in use at LVNL [37]. An important thing to note here is that an operational AMAN works with radar data, such that an accurate trajectory prediction can be performed to estimate the ETO's and the ETA. In the scope of this research however, no such radar data is used as the focus is on the uncertainty at higher prediction time frames. As discussed in section 5.6, the availability of radar data is limited outside the FIRAM. Furthermore, this research is not focused on the efficiency of the AMAN planning itself, as this would logically be better using radar data compared to EFD data. Instead, it focuses on the effect of debunching on the AMAN planning, which is achieved by comparing an AMAN planning fed by regular EFD predictions to an AMAN planning which is fed by EFD predictions influenced by the debuncher.

How the AMAN is modelled will be explained in the following sections, by first elaborating on the inputs and outputs in subsection 9.6.1 followed by the algorithmic steps explained in subsection 9.6.2

### 9.6.1. Data input and output

Before describing the way the AMAN is modelled, first the inputs and outputs are described. As described before, the AMAN was modelled in a way that it resembles the AMAN used at LVNL. Due to complexity however, the AMAN is not exactly replicated, as this would require construction of several extra modules (see section 3.1 for an overview of the modules). Instead, it was chosen to construct a data-based AMAN, which uses historical data and common-practice info acquired from the real AMAN at LVNL.

1. The first input is a table consisting of information on which IAF feeds which runway in case landing configuration with two runways is used. This table was acquired from LVNL and is also used in decision support tools, and is shown in Table 9.5
2. The second input is the historical recording of which runway was in use at what time, for 2020 and 2021. This configuration was retrieved from a historical data set of all landings, containing a field with the runway configuration at that point in time. This was filtered such that all runway configuration changes could be extracted with a timestamp. Using these runway configuration settings, and by cou-

Runway 1	Runway 2	ARTIP	RIVER & SUGOL
06	18R	18R	06
06	27	27	06
06	36R	36R	06
18C	22	22	18C
18C	36R	18C	36R
18R	18C	18C	18R
18R	22	22	18R
27	18C	27	18C
27	18R	27	18R
27	36C	27	36C
36R	36C	36R	36C

Table 9.5: Which IAF feeds which runway in a two runway landing configuration

Flight ID	Aircraft ID	ATA	Runway	IAF	SECTOR	ATO SECTOR	ATO IAF	Holding Count	Holding Time [s]
20775095	KLM856	2020-01-01 03:08:12	18R	ARTIP	1	2020-01-01 02:50:13	2020-01-01 02:56:52	0	0

Table 9.6: Example of data set

pling this with the simulation time (which uses the timestamps from the EFD messages) the AMAN planning becomes more realistic and the results from the experiments hold more value for true operations.

- The third input is a data set containing the average flying times from each IAF to a specific runway for aircraft in 2020. It is assumed that the estimates of flying times in the TMA from EFD data are inaccurate since they contain no airport specific procedure information, and AMAN planning information such as landing runways is not known by EUROCONTROL. Therefore, two post processed data sets from LVNL were merged that such that the flying time between a specific IAF-runway combination could be calculated. For each IAF-runway combination the average flying time was calculated, which serves as a proxy for the Trajectory Predictor module in a real AMAN. Of course, the exact flying times for all aircraft in the simulation are known, but using this would undermine the way uncertainty is used in the simulations of this research, as at the AMAN freeze horizon for a flight there is still uncertainty in the exact flying time between the IAF and a specific runway for a specific flight. The average flying times could have been divided per aircraft type as well, but this was left outside the scope of the research as for realistic values one would also have to take the controller influence into account.
- The fourth input is a table containing historical averages for landing speeds per aircraft type and operator. This table is also used in the current AMAN in use at LVNL. An example of the table was shown in Table 9.1. The correct value for the TAS is chosen based on the required WTC Separation distance (shown in Table 4.3). Based on this distance required between the two aircraft at final approach and the selected TAS for the follower, the WTC separation in time is calculated.

### 9.6.2. Algorithmic Steps

A schematic overview of the AMAN module is presented in Figure 9.17. The algorithmic steps are summarised below:

- Based on the current time and Input 1, get current runway configuration and IAF to runway configuration

Recording Timestamp	Aircraft ID	ADEP	Aircraft Type	WTC	Landing Runway	entry_sector	Landing Interval [s]	Landing Configuration	Take-off Configuration	Actual AMAN Slot [Unix]
2020-01-01 03:56:10	KLM894	ZSPD	B789	H	18C	2	201	18C—	24—	1577851002.0

Table 9.7: Example of the input data set for the AMAN containing the runway configurations

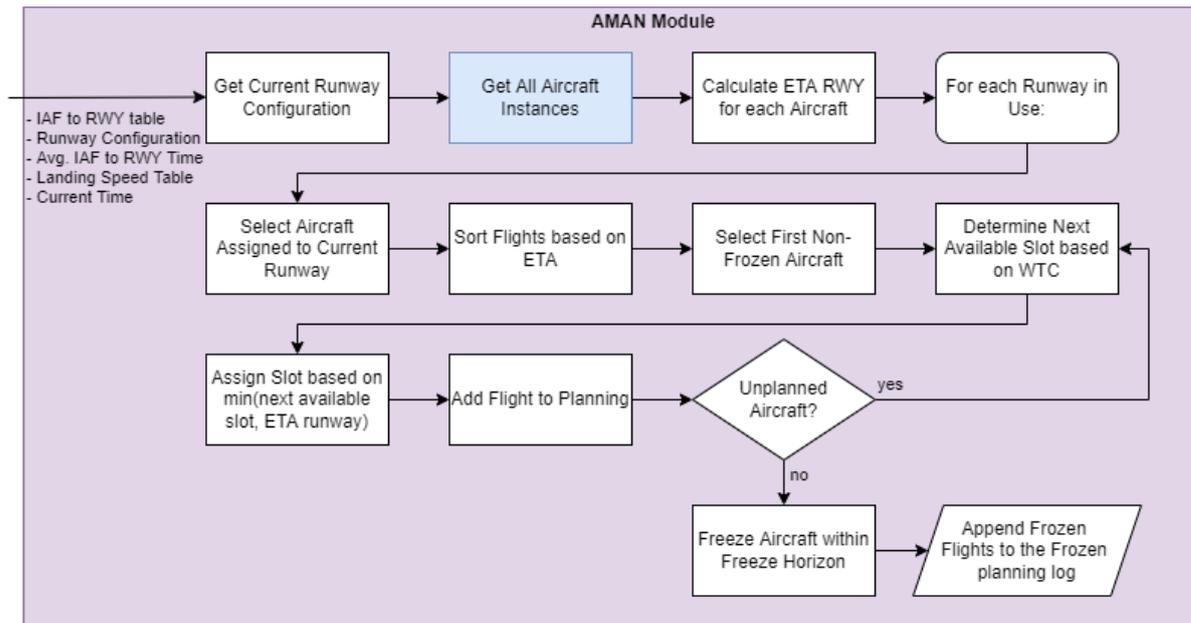


Figure 9.17: System architecture of the AMAN module

2. For each aircraft assign the flying time from its stack to the correct runway, calculate the new ETA based on the ETO stack and the flying time
- For each Runway in use, perform the following steps:
3. Get the information of all aircraft assigned to the current runway
  4. Sort the flights based on the ETA
  5. If there are no frozen flights, the first aircraft gets the slot equal to its ETA.
  6. Else, get second aircraft in the sequence, and determine the landing interval based on the WTC separation and the corresponding landing speed of the trailing aircraft.
  7. Calculate the next free slot based on this interval and the previous aircraft slot
  8. If the aircraft is scheduled later than the next free slot, the slot becomes the aircraft ETA, else the next free slot is assigned and delay is imposed
  9. The aircraft is added to the planning, and the steps from Step 6 are repeated for all other flights
  10. Freeze flights that are within freeze horizon. The freeze horizon is set at 18 minutes before IAF, as LVNL is actively researching changing the freeze horizon from 14 to 18 minutes before IAF [58].

### 9.6.3. AMAN Verification and Validation

The verification of the AMAN was performed for each step in the AMAN planning to check whether the results were as expected. An example is that the actual assigned landing interval should never be smaller than the required landing interval. A snippet of the AMAN output for day 5 is shown in Table 9.8. Note that the actual landing intervals are always larger than or equal to the required landing interval. Furthermore, the assigned slot should not be before the ETA Runway, and the required landing interval changes with the aircraft types.

As the AMAN module makes use of EFD data as input for the planning, it cannot be directly be compared to a real AMAN planning for validation. However, the working of the AMAN was validated by performing a sanity check on the calculated WTC separations and comparing these to the values shown in Table 9.7

Planning Order	Aircraft ID	Aircraft Type	AMAN Delay	AMAN Status	ETA Runway	Required Landing Interval	Actual Landing Interval	Assigned Runway	Assigned Slot
42	KLM1184	E75L	133	frozen	2021-05-01 05:33:48	00:01:17	00:01:17	6	2021-05-01 05:36:02
43	KLM74L	E75L	0	frozen	2021-05-01 05:42:29	00:01:17	00:06:27	36R	2021-05-01 05:42:29
44	KLM94B	E190	7	frozen	2021-05-01 05:43:41	00:01:19	00:01:19	36R	2021-05-01 05:43:48
45	KLM38D	E190	45	frozen	2021-05-01 05:44:15	00:01:19	00:01:19	36R	2021-05-01 05:45:00
46	KLM26B	E190	0	frozen	2021-05-01 05:49:41	00:01:19	00:04:40	36R	2021-05-01 05:49:41
47	KLM52F	E190	73	frozen	2021-05-01 05:49:47	00:01:19	00:01:19	36R	2021-05-01 05:51:00

Table 9.8: A snippet of the frozen planning of the AMAN Module

# 10

## Experimental Setup

In this chapter the experimental set-up will be explained. In the previous chapters the methodology for the construction of the separate modules was explained. In this chapter, those modules will be used together in order to perform experiments that will lead to answering the research questions. The experiments have been divided into three parts: the detection of bunching, debunching, and the effect of debunching on the Arrival Manager. These experiments will be elaborated on in section 10.1, section 10.2 and section 10.3.

### 10.1. Experiment 1: Bunching Detection

The research sub-question that needs to be answered by this experiment is the following: *What is the effect of the information accuracy on the debunching model?*. To answer this question, one has to look at the emergent patterns in the bunching detection module. For this, the sub-question mentioned above has been divided into the following questions, which need to be answered in this experiment:

- What is the average bunching probability over the whole prediction timeline
- What is the average bunching probability at a given prediction horizon
- How often is a given bunching probability reached at a given prediction horizon

#### Data

The data that will be used is 1 week of EFD messages, in May 2021. Due to computation increasing the number of weeks in the analysis could not be done within the scope of this analysis. The data is split up per day, so that 7 instances can run in parallel. This also means that at the start of a instance (the start of a day) there will be no information of an aircraft in the model yet, even though it should already be on the prediction timeline. It is tested that the effects of this are very limited as the traffic density is low during midnight and all aircraft get frequent updates, and therefore would be added to the timeline quickly.

The EFD messages are sent into the Message Class Module one by one. Each EFD message contains a timestamp, which indicates the order in which information is fed into the modules. The data contains information on each flight, so as the data is fed to the modules in the same way for each experiment the outcome is the same every time. Each time a minute has passed, an update from the bunching detection module is requested. Note that this means that multiple updates for flights might have occurred within this minute. Tests indicated that during peak hours this is up to 20 updates per minute. However, because of computational complexity an update frequency of 1 minute is chosen.

#### Independent Variables

In this experiment there are no independent variables.

#### Control Variables

All variables are changed with the input of the EFD message data in a deterministic manner. These variables are either fields in the EFD message, or changed based on the methodology presented in section 9.2.

### Dependent Variables

In order to answer all questions mentioned above the bunching probabilities for the whole timeline are stored at each iteration for all timelines (based on the runway configuration). The timestamps at which updates occurred are also saved as dependent variables.

## 10.2. Experiment 2: Debunching

In the debunching experiment the debuncher is tested to analyse the effects on the bunching prediction timeline. This experiment serves to answer the same sub-question as in the previous experiment, but this time the prediction timeline will be influenced by the debuncher.

### Data

The data that is used is the same as in Experiment 1.

### Independent Variables

The first independent variable is the timeline. In section 9.3 it was discussed that there are three timelines, based on the landing runway configuration and the Initial Approach Fixes. In this experiment, each timeline is solved separately with the other independent variables that will be discussed below. It should be noted that this can be done separately as there is no interference between the timelines. By separating the three timelines, the results specific for each timeline can be examined. Furthermore, the bunching behaviour at each timeline is different, requiring different settings for the percentage threshold, discussed below.

The second independent variable in this experiment is the percentage threshold values at which the debuncher will be activated. What percentage values for the thresholds will be used is dependent on Experiment 1, as a threshold that is too high leads to an experiment with no debunching. On the contrary, a threshold that is too low might result in solutions where there is excessive delay imposed on almost all aircraft since the maximum bunching peak has to be decreased too much. In section 11.1 the behaviour of the bunching peaks will be analysed, such that appropriate thresholds can be chosen section 11.2.

The third independent variable is the prediction horizon at which the timeline is solved. For this independent variable two values have been selected. The first value is a horizon of 40 minutes (before IAF). This option was selected based on an operational perspective of the debuncher and from the perspective of information uncertainty. As discussed in section 5.5, the radar is correlated at receipt of the ABI message at around 30 minutes before the ETO COP of an aircraft. In subsection 8.1.2 it was shown that the median flying time from COP to IAF was about 10 minutes. Therefore, at around 40 minutes before arrival at the IAF, an aircraft would be correlated with the radar such that trajectory predictions can be performed instead of relying on the EFD data (thus decreasing the uncertainty). The operational benefit of this debuncher therefore, would be before this horizon of 40 minutes before IAF. The second value that is selected for the prediction horizon is 18 minutes before IAF. This value was chosen as it is the minimum horizon at which the effect of the debuncher on the AMAN can be analysed, as the AMAN freeze horizon was set at 18 minutes as well (see section 9.6). Note that this horizon was not chosen from an operational standpoint, as it relies on estimates that are affected by the imposed delay immediately for the effect to be visible in the AMAN planning. Instead, it was chosen as at this prediction horizon there is less uncertainty, seen by the smaller width of the uncertainty probability density functions (see subsection 8.1.6). Therefore, the effects of debunching when uncertainty is lower can be compared to debunching where the uncertainty is higher (larger PDF width). From this, a conclusion can potentially be drawn on whether it is beneficial to do further research on the uncertainty analysis of the EFD estimates to construct better (and smaller width) PDF uncertainty curves, potentially increasing the debuncher effectiveness.

### Control Variables

The control variables again are the same variables as in the previous experiment. In addition, fixed solver parameters are used that have been discussed in subsection 9.4.2. These values consist of a population size of 100, a tournament selection size of 3, a 0.001 mutation rate and a 0.9 crossover rate. The maximum number of iterations was fixed at 750.

### Dependent Variables

The dependent variables are the same variables as in the previous experiment. In addition, information on the delayed aircraft are examined. For this, each delayed aircraft including all corresponding information (for example the flight state and time to fly) is stored with the time at which the aircraft was delayed, the amount of delay, and information on if the aircraft was delayed before.

### 10.3. Experiment 3: Arrival Manager

The third and final experiment serves to answer the following research question: *What is the effect of debunching on the efficiency and the stability of the arrival manager planning?*. As efficiency can be a broad term, two specific questions were created:

- What is the difference between ETO and EAT
- What is the throughput per hour

The difference between the ETO and the EAT is the delay that is assigned by the AMAN. This is the amount of seconds an ATCO needs to delay a flight in the Control Area such that it will be at the IAF at exactly the right time. Therefore, a decrease in delay results in an increase in arrival efficiency, as less work is required by the controller and less deviations are required (thus saving fuel).

The second question is a measure of efficiency in the arrival process that indicates the optimality of the metering and sequencing in an incoming arrival stream. As the sequence has influence on the WTC separation times at the runway (and IAF), an optimal sequence is a sequence in which the most aircraft can land within a period of time. The metering of the aircraft can be considered optimal when the aircraft arrive at exactly the required WTC separation that is required by the AMAN.

The stability in the arrival manager pre-planning is measured by analysing changes in sequence of the debunched traffic stream and the one without. As the data input is deterministic the change in order can be directly tied to the influence of the debunching. In operations, the pre-planning starts after radar correlation (see section 3.1). However, in the current research the radar range is not a limit, and the pre-planning is done throughout the whole timeline range. The changes in sequence due changes in the EFD estimates can easily be filtered out by comparing the debunched planning with the planning without.

In this experiment, no new simulations were performed. In Experiment 1 the output of the AMAN for each day was recorded as a dependent variable. For the simulations in Experiment 2 (with the given independent variables) the output AMAN was also recorded, where the effect of the debuncher should be noticeable. Therefore, the only change compared to the previous experiments is that the AMAN output is added as a dependent variable.



# 11

## Results and Discussion

This chapter contains the results of the experiments presented in chapter 10. The results of the uncertainty analysis have been presented in chapter 8, as the presented uncertainty curves served as one of the inputs for the experiments. The presentation of the results is structured in the same manner as the experiments in chapter 10. First, the results of the bunching detection are shown in section 11.1, followed by the debunching results in section 11.2, and concluded with the effects on the Arrival Manager in section 11.3.

### 11.1. Experiment 1: Bunching Detection

The main goal of this experiment was to analyse the effect of the information accuracy on the debunching model. In other words, one wants to know the emergent patterns on the bunching detection prediction timeline are. It should be noted that these patterns are in part due to the nature of the EFD input data at a given day (at what time a message is sent, how many aircraft there are, how the ETO estimates change over time and per message) and due to the uncertainty distributions that have been fitted with the months of EFD data. Therefore, one cannot simply attribute the emergent patterns on the bunching detection timeline to the information accuracy. Instead, a total analysis of these patterns is performed.

As discussed in section 10.1 the experiment has been conducted for 7 consecutive days. At each minute an update of the prediction timeline (of which an example is shown in ??) was performed. For all 7 days and all updates, the bunching probability over the whole timeline was averaged and plotted, shown in Figure 11.1. The same plot for a two-runway configuration is shown in Figure 11.2. It should be noted that the values in the second plot are an order of magnitude lower than those in the plot for a single-runway configuration. This is due to the fact that a double-runway configuration is less common and only used in peak traffic hours, which results in longer periods in which the bunching probability at a given prediction horizon is zero. What these plots show is that on average, the bunching probability is lower at higher prediction horizons than it is at shorter prediction horizons, mainly due to the increasing certainty in the curves.

The same effect is visible when plotting the maximum bunching probability that was reached during the 7 days, for each time bucket. This is shown in Figure 11.3 and Figure 11.4 for a double and a single runway configuration respectively. What these plots show, in other words, is the absolute maximum bunching probability that was reached at each time bucket during all simulation days. This is useful information as it shows similar patterns as in the average bunching probabilities: the maximum bunching probabilities that are reached increase with shorter prediction horizon, due to the fact that the overlap of curves with smaller curve width lead to higher bunching probabilities. It is shown that for a prediction horizon of 250 to 300 minutes before the IAF, the maximum bunching probability that is reached is 20%. From that point, it gradually builds towards a maximum value of about 90%. It is shown that the bunching probability never reaches 100%. Reaching this level would indicate two aircraft with exactly the same arrival time at the IAF, the same aircraft type and the same WTC separation, which is very unlikely.

In the analysis above the emergent patterns for the whole prediction timeline was analysed. However, to gain more insight in the bunching detection, the behaviour of the bunching prediction over the time span of the 7 days of simulation time is analysed. This does not only help with analysing the emergent patterns dur-

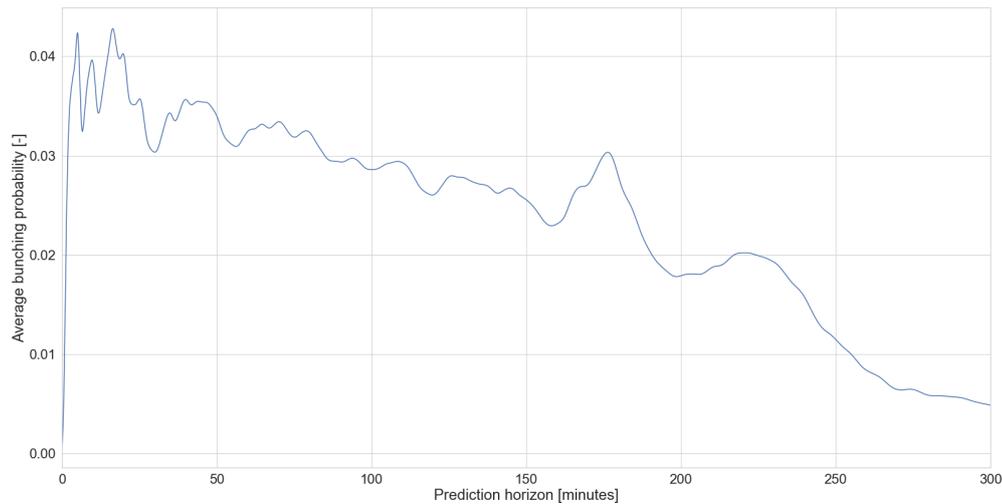


Figure 11.1: Average bunching probability at each prediction horizon for 7 consecutive days for a single runway configuration

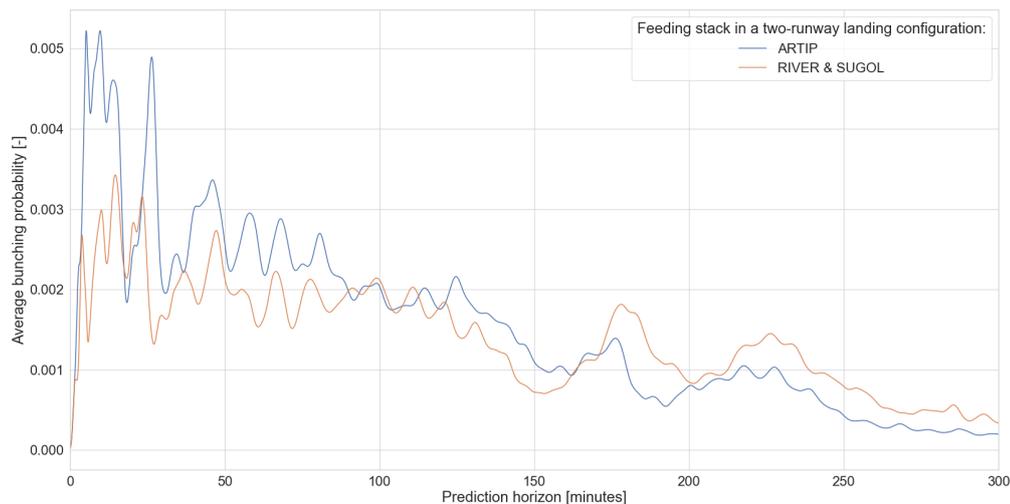


Figure 11.2: Average bunching probability at each prediction horizon for 7 consecutive days for a double runway configuration

ing a day, but it also gives insights in how often a given threshold is reached. For this, the maximum bunching probability at each time point for all 7 days was plotted in Figure 11.5 for a single runway configuration, and in Figure 11.6 for a double runway configuration. In other words, at each time point shown on the x-axis, a prediction timeline was constructed, of which the maximum value is plotted in this graph. The maximum value was determined for the prediction horizon range of 40 to 300 minutes, as discussed in section 10.2.

What can be seen is that commonly there is a buildup in the maximum bunching probability (which was confirmed by inspecting the prediction timelines visually each iteration). However, sometimes the sudden peaks appear in the bunching probability which drop quickly, which might be due to either due to a buildup just before the 40 minute cut-off mark, or due to a change in the EFD estimates that changes the sequence of the aircraft and the corresponding PDFs. The maximum bunching probability that is reached before the 40 minute cut-off is shown to be near 65%, also shown by the findings Figure 11.3. It is seen that the maximum

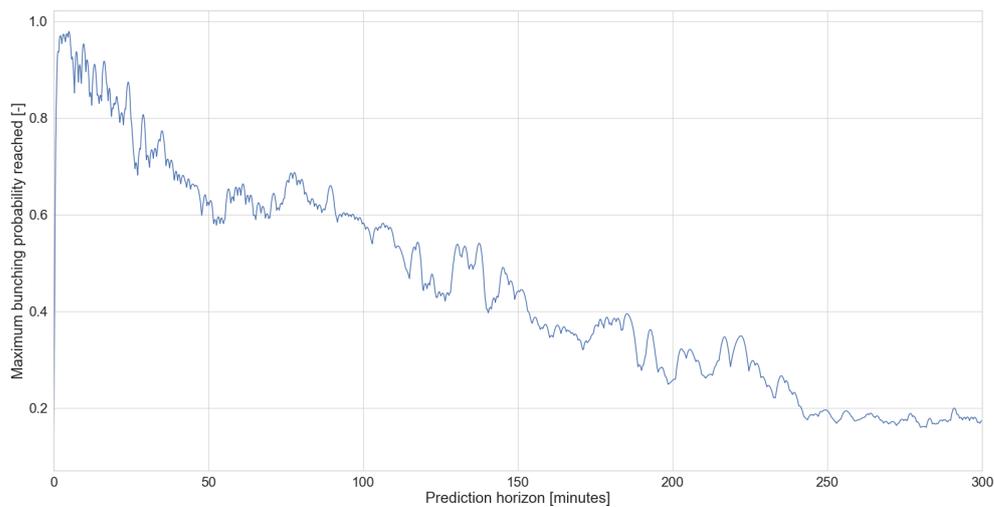


Figure 11.3: Maximum bunching probability that was reached at each prediction horizon for 7 consecutive days for a single runway configuration

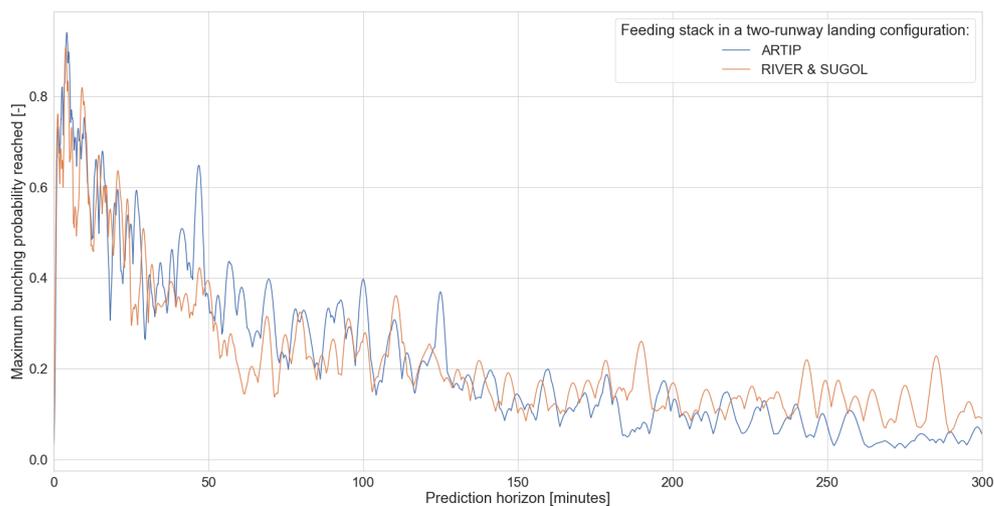


Figure 11.4: Maximum bunching probability that was reached at each prediction horizon for 7 consecutive days for a double runway configuration

bunching probability that is reached is highly dependent per day, but that often several peaks are present, which could be attributed to the inbound peaks during a day (as shown in Figure 4.2). In the beginning of almost each day the highest peaks with a wide base can be seen, which are due to the morning traffic peak.

Based on Figure 11.5 and Figure 11.6, values for the debunching thresholds can be set. It is evident that there is a large variability in the height of the peaks for each day. However, for the scope of this research and the experiments that are conducted in it, several fixed thresholds will be chosen such that the effects of decreasing the bunching percentage peaks can be analysed. Based on the peaks in Figure 11.5 for a single runway configuration, the thresholds that were chosen are: 55%, 50%, 45% and 40%. For the two-runway configuration the thresholds were set at 35%, 30%, 25%, and 20%. These thresholds are the values at which the debuncher is activated. The constraint value that is used in the debuncher itself is set 5% lower, meaning

that a trigger at 60% results in the debuncher decreasing the bunching to 55%.

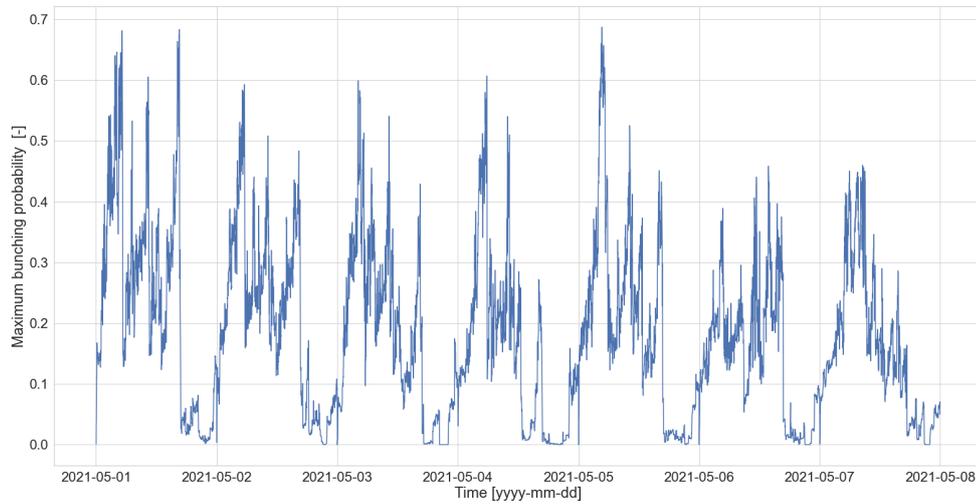


Figure 11.5: Maximum bunching probability at each point in time for 7 consecutive days for a single runway configuration

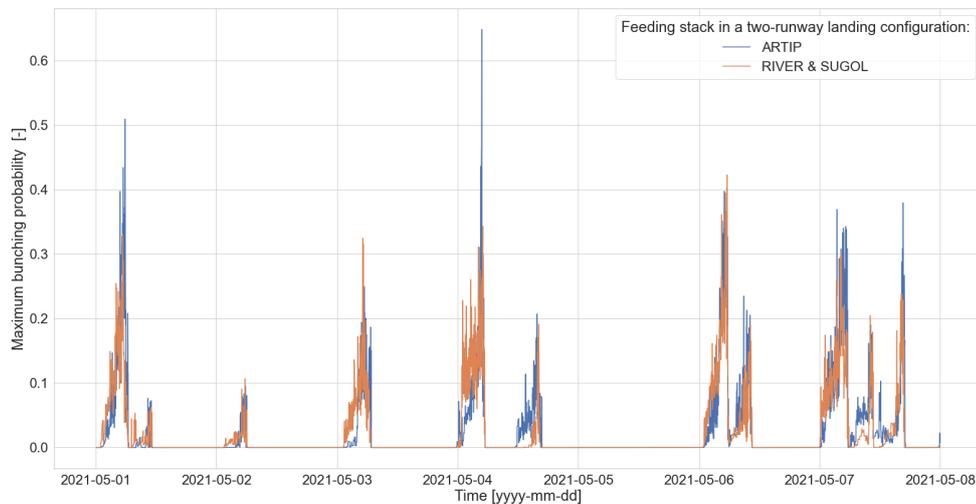


Figure 11.6: Maximum bunching probability at each point in time for 7 consecutive days for a double runway configuration

As discussed in section 11.2, there is a second independent variable for the debunching experiments that states the prediction horizon at which the bunching timeline is solved. The thresholds determined above were set for the 40 minute prediction horizon. However, in Figure 11.1 and Figure 11.3 it was shown that the average and maximum bunching probabilities increase with shorter prediction horizons. Therefore, using the same thresholds for the 18 minute horizon as for the 40 minute horizon would result in excessive delays, as the peaks near the 18 minute horizon would be too high to decrease. Therefore, new bunching threshold values have been determined. For this, the maximum bunching probability at each point for all 7 days was plotted again, but now with the same plot created with the 18 minute prediction horizon plotted in the background. This is shown in Figure 11.7. Note that each point on the x-axis of this plot represents the maximum value

of the bunching prediction timeline at that point in time, and that since now values up to 18 minutes are included, the peaks are higher than they were for the 40 minutes horizon. From this it was determined that the bunching thresholds for the 18 minute prediction horizon are 75%, 70%, 65%, and 60%.

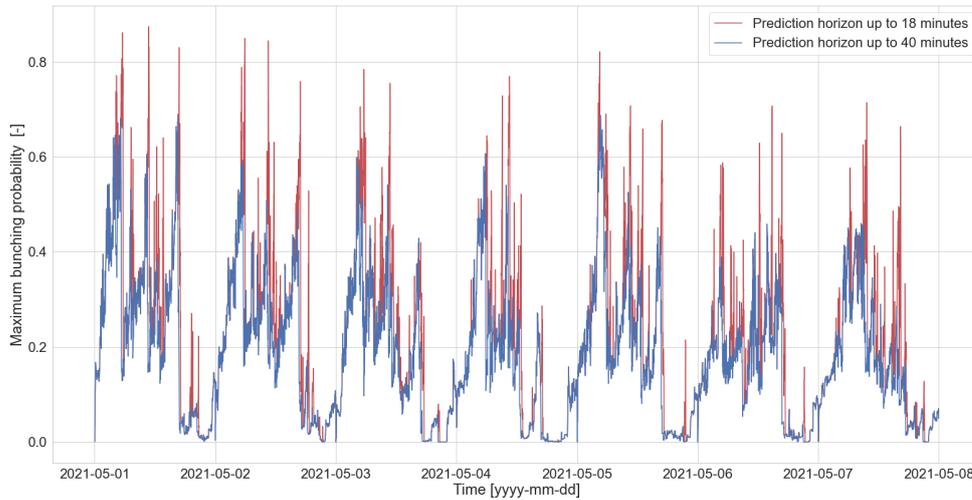


Figure 11.7: Maximum bunching probability at each point in time for 7 consecutive days for a double runway configuration, for prediction horizons of 18 and 40 minutes

## 11.2. Experiment 2: Debunching

In this experiment the effects of debunching the timeline are analysed. During the experiments it was found that because of the sharp bunching peaks that occur (shown in Figure 11.5), the debuncher was often activated, even though the peak could disappear in the update after. In order to cope with these peaks, and to not activate the solver when not necessary, the solver was only activated if the maximum bunching value was above the threshold value for 5 minutes. Therefore, peaks might still occur above the threshold value, but only for a maximum of 5 minutes. Using this, the results were drastically improved, and therefore this setting was used for all upcoming results.

For each day, summarising tables of the experiment outputs were constructed to show the effects of the debunching using different thresholds. These tables show the effect of the thresholds on the following things: the number of aircraft that are delayed, the total delay that is imposed by the solver during that day, the average time to fly to the IAF at the moment of the imposed delay, the average percentage of delay used of the total available delay for an aircraft, and the the number of aircraft that have been delayed a second time. The last two columns of the tables are the results of Experiment 3, and will be discussed in the next Section. For the first five days the results are shown in Table 11.1, Table 11.2, Table 11.3, Table 11.4 and Table 11.5. For day 6 and 7 the given thresholds were too high, resulting in no delayed aircraft for those days, and thus no table is presented.

The results show that the number of aircraft that get delayed and the amount of delay increases with the decreasing bunching threshold. It can be seen that as soon as the threshold becomes too low, both the delay and the number of aircraft increase exponentially, and on average a higher percentage of the maximum delay that can be absorbed by an aircraft is used. An intuitive explanation is that there is no room to space the probability density curves such that the bunching probability is decreased with 5%. Therefore, a lot of aircraft need to be delayed such that the curves can be spaced better. A conclusion that can be drawn from this is that the uncertainty is too high at this point. If the uncertainty would be lower, and the width of the curves smaller as a consequence, it would be easier to space the curves in such a way that the bunching percentage would be decreased without requiring high amounts of delay. This effect will be analysed in the next subsection.

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
55 %	3	110	45	13.3	0	-434	4.0
50 %	5	348	47	24.9	0	-613	5.7
45 %	13	1218	57	29.1	1	-427	4.0
40 %	217	30719	169	34.6	120	-770	7.2

Table 11.1: Output table for day 1, with a single runway configuration

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
55 %	0	-	-	-	-	-	-
50 %	0	-	-	-	-	-	-
45 %	5	616	48	31.1	0	-316	5.2
40 %	5	735	72	41.5	0	-399	5.6

Table 11.2: Output table for day 2, with a single runway configuration

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
55 %	0	-	-	-	-	-	-
50 %	8	268	43	12.9	0	-144	2.2
45 %	10	676	44	21.3	0	+48	0.7
40 %	31	3419	47	39.8	7	-157	2.4

Table 11.3: Output table for day 3, with a single runway configuration

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
55 %	0	-	-	-	-	-	-
50 %	0	-	-	-	-	-	-
45 %	3	181	58	20.8	0	-91	2.0
40 %	5	404	104	21.2	0	-274	6.0

Table 11.4: Output table for day 4, with a single runway configuration

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
55 %	5	360	84	19.3	1	-487	5.9
50 %	5	734	89	39.7	0	-133	1.6
45 %	9	1114	96	30.7	2	-136	1.7
40 %	28	2966	82	32.2	9	-977	11.9

Table 11.5: Output table for day 5, with a single runway configuration

From the results it can also be seen that the average flying time to IAF for the aircraft increases with decreasing bunching threshold. This is easily explained with Figure 11.1 and Figure 11.3, as the bunching threshold probability is reached earlier on the prediction horizon than higher bunching threshold probabilities. The results also show that aircraft commonly do not get delayed a second time after already being delayed. For low bunching thresholds with a large number of aircraft delayed this number increases. From an operational standpoint this might be unwanted, as a single shot solution might be preferred. This could of course easily be implemented as an constraint in the debuncher.

For the two-runway configuration the tables are shown in the Appendix. For ARTIP stack the tables are shown in Table C.1, Table C.2, Table C.3 and Table C.4. On day 2, 3 and 5 there were no aircraft delayed, so these tables were not shown. For the RIVER and SUGOL stack the tables are shown in Table C.5, Table C.6, Table C.7 and Table C.8. Again, day 2, 3 and 5 did not yield any delayed aircraft due to the low bunching probability shown in Figure 11.6. In the results, the same patterns as in single runway configuration are seen: decreasing the bunching probability leads to a high number of total delay and an increased number of aircraft delayed. However, a difference compared with the single runway configuration is that the buildup of these values seems to be more drastically. An explanation might lie in the fact that as the bunching threshold probabilities are lower on these timelines, decreasing the probability with 5% has a significant impact. In further research, a scaled decrease could be used, in which the bunching threshold would be decreased by a percentage relative to the percentage itself. For instance, a 20% decrease of a bunching threshold of 30% would lead to a 6% decrease in bunching threshold, to 24%. This would make the results more comparable to the results with higher threshold values.

The results given above were constructed using the 40 minute solver horizon. Now, the results of the 18 minute prediction horizon will be presented for the single-runway configuration, for which the thresholds were set at 75%, 70%, 65%, and 60%. For day 1 to 5 the results were presented in Table 11.6, Table 11.7, Table 11.8, Table 11.9 and Table 11.10 respectively. For day 6 and 7, the debuncher again was not triggered at these threshold values, as the bunching prediction did not reach those values for at least 5 minutes. For these results it was expected that higher certainty in the predictions (yielding smaller curve width) would yield to improved solutions in terms of total required delay and amount of aircraft delayed. As the bunching percentages are different (as shown in Figure 11.7), and as a consequence the threshold values are different than those used for the 40 minute prediction horizon, the results cannot be directly compared. However, similar patterns for the number of aircraft and amount of imposed delay can be seen as in the 40 minute solver horizon, with the most important one being that there is a sudden increase in amount of aircraft delayed as the bunching threshold becomes too low. Therefore, no clear conclusion can be drawn yet from the difference between the 18 and 40 minute solver horizon. In the next subsection the change in the AMAN will be considered, in which a difference is visible.

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
75 %	2	166	21	31.2	0	-152	1.4
70 %	3	228	21	31.1	0	-880	8.2
65 %	11	778	24	29.6	0	-953	8.9
60 %	71	11142	141	38.5	6	+139	1.3

Table 11.6: Output table for day 1, with a single runway configuration, with the 18 minute solver horizon

### 11.3. Experiment 3: Arrival Manager

As discussed in section 11.3, this experiment contained the same independent variables as Experiment 2, and the results were measured during the simulations of Experiment 2. In order to measure the effects on the AMAN planning, the planning for a given day without debunching (Experiment 1) was compared to the same day influenced by the debuncher in one of the simulations of Experiment 2. The results of this experiment are shown in the same tables in which the results of Experiment 2 are presented. the 'Change in AMAN' column represents the amount of seconds of change in the AMAN. A negative change means that

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
75 %	0	-	-	-	-	-	-
70 %	2	201	20	37.2	0	-245	4.0
65 %	3	299	20	39.3	0	-350	5.8
60 %	5	500	21	42.4	0	-393	6.5

Table 11.7: Output table for day 2, with a single runway configuration, with the 18 minute solver horizon

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
75 %	0	-	-	-	-	-	-
70 %	0	-	-	-	-	-	-
65 %	0	-	-	-	-	-	-
60 %	1	121	20	50 %	0	-66	1.0 %

Table 11.8: Output table for day 3, with a single runway configuration, with the 18 minute solver horizon

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
75 %	0	-	-	-	-	-	-
70 %	0	-	-	-	-	-	-
65 %	3	79	21	11.1	0	-4	0.1
60 %	4	371	24	39.3	0	-54	1.2

Table 11.9: Output table for day 4, with a single runway configuration, with the 18 minute solver horizon

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
75 %	0	-	-	-	-	-	-
70 %	0	-	-	-	-	-	-
65 %	3	306	55	39 %	0	-490	6.0
60 %	30	3789	54	43 %	9	-1024	12.5

Table 11.10: Output table for day 5, with a single runway configuration, with the 18 minute solver horizon

with the debuncher, the amount of delay imposed by the AMAN was decreased. A positive change means that the planning became less efficient, and more delay was imposed by the AMAN in the debunched simulation than in the one without. The '% of total AMAN Delay' shows by how much percent the delay was increased or decreased, relative to the total imposed delay by the AMAN for that given day to give an indication of the impact of the change of AMAN delay.

For the change in AMAN the two-runway configurations (shown in the tables in Appendix C) it can be seen that not all experiments yielded a decrease. The patterns in the results are less clear than they were for the single-runway configuration, and often the decrease in AMAN was significantly lower than the imposed delay in the solver, or the change in delay even became positive. The results were however verified by manually checking the changes in the AMAN planning. It was found that the delay of aircraft did have an effect on the planning, but often this yielded more AMAN delay for the flights after the delayed flight. It is assumed that effect might be due to the fixed decrease of 5% in the bunching threshold, as was discussed in the previous section.

However, for the results in Table 11.1, Table 11.2, Table 11.3, Table 11.4 and Table 11.5 it is shown that a decrease in the AMAN delay is often visible when a small number of aircraft (less than 10) is delayed by the debuncher. The total number of aircraft that were planned by the AMAN for each day were 294, 269, 246, 210,

237, 243 and 264 respectively, meaning that when 10 aircraft are delayed that means about 4% of the total aircraft for that day are influenced. For most of the scenarios with less than 10 aircraft delayed it was shown that the decrease in the AMAN delay is just over half the amount of seconds that was imposed on an aircraft en-route. For day 1 and day 5 (Table 11.1 and Table 11.5 respectively) some threshold values yielded AMAN decrease higher or similar to the imposed delay en-route. This decrease in AMAN delay is often a significant part of the total AMAN delay with changes in the order of 1%, up to 5.9% of the total AMAN delay for that day.

For the 18 minute solver horizon shown in Table 11.6, Table 11.7, Table 11.8, Table 11.9 and Table 11.10 it can be seen that the change in AMAN delay is improved compared to the change with the 40 minute solver horizon. For the 18 minute horizon, the total imposed delay is often similar to the decrease in delay that is found in the AMAN, and in some cases the decrease in AMAN delay is even larger than the imposed delay. What has changed in these situations compared to the 40 minute solver horizon is that the uncertainty is lower, and thus the probability density curve width is smaller. Therefore, it can be concluded that the impact the debuncher has on the delay in the AMAN is improved with decreasing uncertainty.

For situations where a large number of aircraft (more than 10) is delayed it can be seen that the change in AMAN delay is not comparable any more to the total imposed delay by the debuncher, with the imposed delay being significantly higher. One explanation is that the uncertainty is too high, such that delaying an aircraft has only a limited impact. Therefore, more aircraft need to be delayed, or higher amounts of delay need to be imposed such that the bunching peak decreases to the required limit. This spaces the incoming traffic stream more than would be required for the AMAN. On the contrary, in the hypothetical case where the uncertainty is very small, an aircraft would have to be delayed only a small amount to have impact on the bunching peak, or less aircraft would suffice as the change on the bunching peak is higher. To verify this concept, the situation with the 4 aircraft used in the debuncher validation (see section 9.5) was used. In Figure 11.8 two situations were constructed with the same four aircraft, but for Situation 1 the uncertainty was small (curve of AA at 20 minutes horizon), and for Situation 2 the uncertainty was high (curve of AA at 240 minutes prediction horizon). The bunching probability for Situation 1 was found to be 36.1%. For Situation 2 the bunching probability was 10.1%. In both situations, one aircraft was delayed with 200 seconds. In the Situation 1 this resulted in a new bunching probability of 25%, yielding a significant decrease. In Situation 2 this yielded a new bunching probability of 8.9%, yielding a much lower change in the bunching probability, which supports the analysis above.

This finding also indicates that if one wants to solve at different prediction horizons, the percentage at which the debunching is performed should be reconsidered. The width of the uncertainty curve has influence on the bunching percentage that is reached at a given threshold (as seen by Figure 11.1), and it also changes the effectiveness of the solver to reach a certain threshold. A Percentage too low could cause excessive delays which are not equal to similar to the delay decrease in the AMAN. Using a debuncher in an operational concept therefore, requires careful tuning of the solver horizon and the bunching thresholds before useful information can be provided to an Air Traffic Controller.

The imposed delay by the debuncher is often higher than the decrease in AMAN delay. In general, this means that decreasing the delay in the AMAN comes at a cost of extra delay en-route. Note however, that absorbing delay en-route is preferred as AMAN delay can only be imposed on an aircraft with inefficient tactical delay absorption techniques such as vectoring and airborne holdings (as discussed in section 3.2 and section 3.3). Furthermore, it should be noted that the perfect scenario in which the imposed delay is always equal to the delay decrease in the AMAN is not realistic, as for this to happen there should be no uncertainty, and as such a perfect Extended Arrival Manager would have been created. Instead, the debunching concept presented uses the fact that there is uncertainty, and leverages it to influence the incoming traffic stream. The results show that it is in fact possible to decrease the AMAN delay using this concept, although not consistently for a fixed threshold value.

In section 10.1 it was stated that another parameter of interest was the AMAN stability. One measure of the AMAN stability is the throughput, for which a common measure is the throughput per 20 minutes. For the results where a decrease in AMAN delay is visible it can be seen that the imposed delay for each aircraft is on average lower than 100 seconds. Therefore, the throughput is unlikely to change. As an example, in Figure C.1 the throughput for day 5 (Table 11.5) was plotted in 20 minute buckets. As this day had only a single-runway configuration (as seen in Figure 11.4), the AMAN throughput could be considered similar to

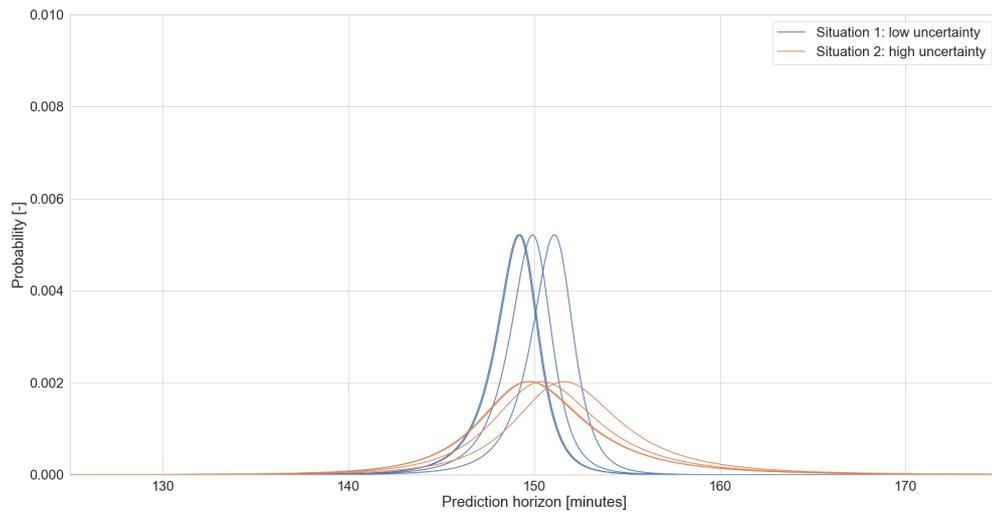


Figure 11.8: The probability density curves for two test cases illustrating the effectiveness of debunching with different uncertainty

the runway throughput for that day. It is shown that only the 40% threshold resulted in a change in throughput of 2 aircraft from one bucket to the next, and thus the effect on the AMAN throughput is low.

# 12

## Conclusion

The aim of this research is to analyse the effect of the inbound air traffic trajectory prediction accuracy on a proposed debunching concept. In the Literature Review and the Preliminary Findings it is concluded that the inbound arrival efficiency can be increased by extending the Arrival Manager freeze horizon. However, the lack of information at long ranges, the reliability of predicted arrival times, the division of the airspace and the influence of pop-up flights are identified as limiting factors for extending the freeze horizon. It is concluded that a probabilistic debunching concept can be introduced, which makes use of the available information at long ranges and uses the uncertainty in this information to increase the efficiency of the incoming traffic stream. The definition that is used for bunching in this research is a probabilistic one, where a situation yields a bunching probability if two or more aircraft have a probability of arriving at the IAF at the same time or within the required Wake Turbulence Category separation between the aircraft. The concept of debunching introduced in this research is to sequence the aircraft so that this probability decreases below a given threshold value.

In the uncertainty analysis of the EFD trajectory predictions it is concluded that relevant features for the division of the data are the time to arrival (in 5 minute buckets), the flight state and the Collaborative Decision Making state for the Tactical Activated flight state. By splitting the data with these features, clear differences are seen between the box plot statistics, yielding differences in the uncertainty distributions which can be used in the bunching prediction model. It is shown that a better representation of the uncertainty can be gained by increasing the relevant features. However, it is shown that using extra features than the ones mentioned is not possible given the amount of data available for this research.

A curve fitting routine is applied such that Probability Density Curves could be analysed. For most Probability Density Functions, there is a skew towards negative values, showing that ETO COP estimates in the EFD messages are later than the Actual Time Over the COP, which was retrieved from Radar data. This means there is a higher probability of an aircraft arriving earlier than the time provided in the EFD estimate than arriving later than that time. This negative skew is partly explained by aircraft flying direct routes towards the IAF due to the low traffic density, instead of following the standard prescribed route with which the trajectory predictions are performed. As this is a form of intent, it is hard to predict such patterns in a trajectory prediction. In EFD data for aircraft with an Air Traffic Control Activated flight state, the accuracy of the estimates is highest, with a width of 20 minutes centred around 0 for a prediction horizon between 60 and 65 minutes from the COP. This flight state is also shown to be the most common flight state, with around 85% of all EFD messages containing this flight state at the 60 minute prediction horizon. The second most common flight state is the Flight Plan Filed flight state, for which the width of the curve was shown to be 40 minutes at the 60 minute prediction horizon.

The first experiment shows that the average and maximum bunching probability increases with shorter prediction horizons. This shows that the uncertainty at long prediction horizons is too high to provide a bunching probability that can be debunched. For the 7 days that are analysed, it is shown that at prediction horizons of 250 to 300 minutes, the maximum bunching probability reached is 20%. However, at a 40 minute prediction horizon, this probability is around 65%. A bunching probability of 100% was never reached, as this

requires exactly the same predicted arrival time at the IAF, and the same WTC separation (which is a function of the aircraft type and the follower of the aircraft), which is shown to be unlikely. The emergent patterns of the bunching probability are shown to be different for each day in the experiment. For most days, it is shown that a morning peak is visible in which the bunching probability often reaches its maximum value for that day. This shows that the most operational potential lies in this morning peak.

In the second experiment the effects of debunching is analysed. It is shown that the number of aircraft that are delayed and the total amount of delay imposed on the aircraft increases with a decreasing bunching threshold. When the bunching threshold is too low compared to the bunching percentage that is reached on the timeline, the amount of imposed delay and the number of delayed aircraft increases exponentially. This is often seen for the bunching threshold of 40%. However, the threshold of 40% did not activate the debuncher for all days, as on days 6 and 7 of the experiment the bunching percentage stayed below this level for 5 consecutive minutes. The conclusion that can be drawn from this is that it is difficult to define fixed threshold values that result in a debunched timeline for each day, as the traffic patterns change from day to day. It is also shown that on average, only 20 to 40% of the total delay that aircraft could absorb is used, which is beneficial from an operational standpoint as aircraft do not have to be delayed by using minimum speeds. In the second experiment it is also shown that for a two-runway configuration similar patterns are observed in the data as for the single-runway configuration. However, the bunching probabilities are lower since there are fewer aircraft on these timelines than on the timeline for a single-runway configuration. Debunching, the bunching prediction for a fixed 5% decrease results in a high amount of delay as the relative decrease in the bunching probability, is higher. Therefore, a percentage decrease relative to the threshold value needs to be used to mitigate excessive delays.

The third experiment shows that using the probabilistic debunching technique, it is possible to have a positive effect on the Arrival Manager planning. However, this effect is not yet consistent as it is shown that for the double-runway configuration the decrease in the bunching percentage does not always lead to a decrease in the AMAN delay. Also, when the AMAN delay does decrease, the decrease is often lower than the imposed delay by the debuncher. For the single-runway configuration, it is shown that when less than 10 aircraft per day are delayed, a decrease in the AMAN delay is often seen. This shows that the debuncher is effective in solving high bunching peaks, where the debunching threshold is not set too low compared to the bunching prediction. Commonly, the decrease in delay in the AMAN is just over half the amount of seconds of delay that is imposed on the aircraft by the debuncher, and thus the decrease in AMAN delay often comes at the cost of extra delay en-route. However, this en-route delay can be absorbed more efficiently, as AMAN delay is commonly imposed on aircraft with inefficient tactical delay absorption techniques such as vectoring and airborne holdings. Therefore, a trade-off can be made between the amount of delay that an airliner is willing to absorb en-route compared to in the Control Area. It is also concluded that the stability of the AMAN planning is hardly affected as the debunching process has very little influence on the AMAN throughput due to the low amount of aircraft delayed, and the small amount of delay required.

Two solver horizons were used. The first solver horizon is set at 40 minutes before IAF, meaning that aircraft are eligible to be delayed by the debuncher before this point. This point in time is determined as this is the point from which more accurate information is available through the AMAN preview planning. Therefore, the operational benefit of the debuncher lies in the prediction horizon before this point. The second prediction horizon that is used is 18 minutes, as this is the point where the AMAN planning becomes frozen in the experiments. Analysing the debunching capabilities at this solver horizon allowed to analyse the effects of the lower uncertainty in the predictions on the debunching performance. Using this solver horizon it is found that the decrease in AMAN delay is improved compared to the 40 minute solver horizon, as the decrease in AMAN delay is often similar to the imposed delay by the debuncher, meaning that the decrease in AMAN delay could be achieved by a similar amount of delay en-route. It is concluded that when the uncertainty in the EFD estimates is decreased, the effectiveness of the debuncher increases, and the effect on the Arrival Manager can be improved. Therefore, currently the uncertainty at the operational range of 40 minute is too high for the proposed debunching concept, as a consistent decrease in the AMAN delay cannot yet be achieved. However, by decreasing the trajectory prediction uncertainty, the impact on the arrival manager efficiency can be increased, such that the proposed arrival concept can be used for increasing the inbound arrival efficiency.

# 13

## Operational Potential and Recommendations

Using the research throughout this report, several comments on the operational potential of the debuncher can be given. In this research a debunching concept was designed and tested, where the focus was on the dynamics of the bunching detection and the effects of debunching on these dynamics and the inbound traffic stream. In an operational context, the tool would be used at ranges where an Extended Arrival Manager concept would not be viable, and it could be used in conjunction with other decision support tools. One such decision support tool is a priority sequencer, in which aircraft that need to be prioritised due to transferring passengers could be exempt from being delayed. The current Genetic Algorithm that is used for solving the bunching optimisation problems could easily accommodate this feature by setting the bound-constraint of the given aircraft to zero. A final note on the Genetic Algorithm is that due to the algorithm's iterative nature, the convergence speed might not be suitable for live computations in operations. Further research could focus on (bound) constrained optimisation algorithms with global characteristics.

A further developed version of the tool could be used to assist the Air Traffic Controller by showing a measure for the bunching probability and provide options to solve it by proposing en-route delay. The results in this research showed that it is possible to decrease the AMAN delay using this concept. In other words, it was shown that the efficiency of the incoming traffic stream could be increased using the debuncher. However, this decrease in AMAN delay was not consistent for a fixed bunching threshold value. As the debunching concept is a probabilistic method the influence on the AMAN is not straightforward, it is complex or even impossible to guarantee a specific decrease in AMAN delay by absorbing delay en-route given the uncertainty that is inherent in the Air Traffic Management system.

It should also be noted that if one wants to debunch at different prediction horizons, the percentage at which the debunching is performed should be reconsidered. The width of the uncertainty curve was shown to have influence on the bunching percentage that was reached at a given threshold, and it also changed the effectiveness of the debuncher. A percentage too low caused excessive delays which were not equal to or similar to the delay decrease in the AMAN. Therefore, using a debuncher in an operational concept requires careful tuning of the solver horizon and the bunching thresholds before useful information can be provided to an Air Traffic Controller. In an operational tool, several levels for debunching could be proposed (using different thresholds), from which the Air Traffic Controller (or the Controller in the planning position) could select the appropriate solution based on the proposed delay and possibly the subjective workload at that time. The thresholds can also be tuned by selecting a bunching threshold level for a given time of day. For instance, it could be chosen to use the debuncher only for a part of the day, where the threshold is specifically tuned and tested for that part of the day. An example is the inbound peak in the morning, where it is seen that the bunching probability is often the highest for the day.

It was also shown that with decreasing uncertainty, the debunching results improved. Therefore, further development in the uncertainty analysis could increase the consistency of the tool. The uncertainty analysis in this research could be considered a baseline, or worst-case uncertainty analysis, as the amount of months

of data was limited, it included months of low traffic density (yielding low sample size), and the analysis was merely an experience-based uncertainty analysis. Improving the uncertainty analysis could be done in several ways. First, by increasing the amount of features considered for the grouping of data the uncertainty analysis could be improved. However, for this more data would be needed (especially from months with high traffic density). An ideal way to improve the uncertainty analysis would be to use machine learning techniques, as many features could be introduced by using multiple data sets as input (related to weather conditions instance). One such method is the Quantile Regression Forest Technique shown in [21] which, in contrast to a regular Random Forest approach, does not only provide a single improved estimate of arrival time, but generates a probability distribution of the arrival time. These probability distributions can easily be implemented within the current debunching concept. Another way of decreasing the uncertainty would be to use innovative data sources, such as down-linked data. Using this data would drastically improve uncertainty estimates as aircraft intent would be known.

# A

## Appendix: Uncertainty Analysis

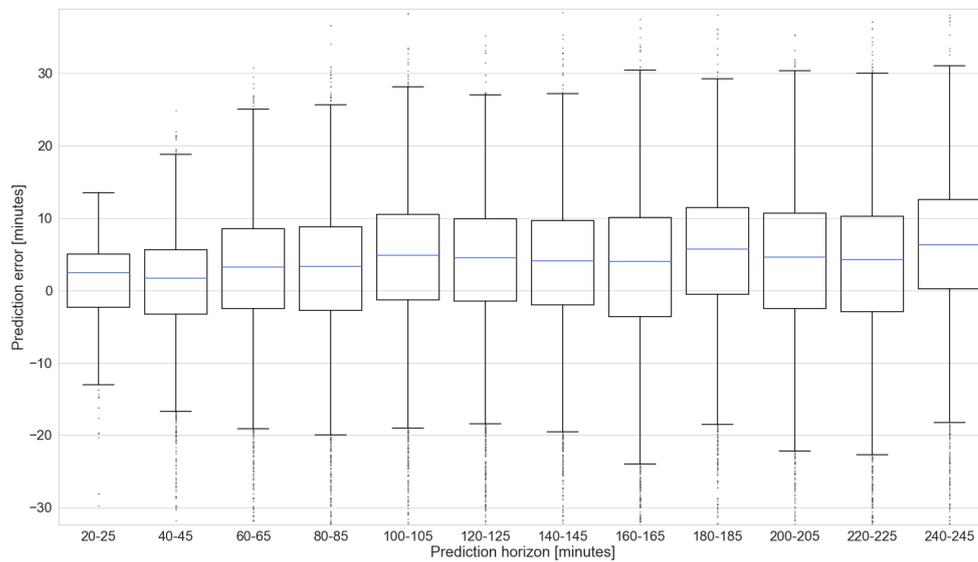


Figure A.1: Accuracy of the COP arrival predictions versus time to fly for the Flight Plan Filed flight state

Prediction Horizon [min]	AA		FI		TA (CDM)		TA (no CDM)		SI		FS		SU	
	Count	KS 95% CI	Count	KS 95% CI	Count	KS 95% CI	Count	KS 95% CI	Count	KS 95% CI	Count	KS 95% CI	Count	KS 95% CI
0 - 5	202319	[0.0125, 0.145]	4	-	3	-	23	[0.02, 0.0475]	0	-	0	-	2397	[0.0125, 0.1075]
5 - 10	158370	[0.0125, 0.1225]	5	-	117	[0.0375, 0.1975]	47	[0.0175, 0.2125]	0	-	0	-	1795	[0.015, 0.16]
10 - 15	133961	[0.0125, 0.115]	15	[0.05, 0.2675]	332	[0.0175, 0.29]	116	[0.0075, 0.265]	2	-	0	-	1973	[0.015, 0.1225]
15 - 20	116387	[0.01, 0.095]	88	[0.015, 0.09]	716	[0.01, 0.1525]	134	[0.0525, 0.35]	7	-	0	-	1066	[0.015, 0.15]
20 - 25	72492	[0.01, 0.09]	415	[0.005, 0.15]	1154	[0.01, 0.1025]	206	[0.0225, 0.21]	6	-	0	-	690	[0.0125, 0.24]
25 - 30	66106	[0.01, 0.0825]	602	[0.0075, 0.2]	1374	[0.0125, 0.1075]	810	[0.0125, 0.111]	11	[0.06, 0.06]	0	-	618	[0.01, 0.2225]
30 - 35	59055	[0.01, 0.075]	698	[0.0075, 0.1475]	2213	[0.01, 0.0775]	557	[0.0175, 0.24]	43	[0.0475, 0.1675]	0	-	608	[0.0125, 0.2375]
35 - 40	61050	[0.01, 0.07]	1967	[0.015, 0.105]	2769	[0.01, 0.0875]	428	[0.01, 0.105]	46	[0.03, 0.1275]	0	-	620	[0.02, 0.215]
40 - 45	59452	[0.0125, 0.07]	2767	[0.01, 0.12]	2294	[0.01, 0.145]	402	[0.01, 0.095]	80	[0.03, 0.21]	0	-	773	[0.0225, 0.1875]
45 - 50	59461	[0.01, 0.0875]	2317	[0.015, 0.1375]	2400	[0.01, 0.0875]	599	[0.015, 0.09]	83	[0.0025, 0.1625]	0	-	638	[0.01, 0.08]
50 - 55	50468	[0.01, 0.085]	2369	[0.0125, 0.1475]	1966	[0.01, 0.1075]	412	[0.01, 0.15]	101	[0.015, 0.2]	0	-	576	[0.01, 0.0825]
55 - 60	47739	[0.01, 0.07]	2511	[0.01, 0.1525]	1933	[0.01, 0.175]	300	[0.01, 0.1175]	133	[0.015, 0.2025]	0	-	603	[0.0125, 0.095]
60 - 65	48953	[0.01, 0.065]	3484	[0.015, 0.1425]	1790	[0.0125, 0.1825]	624	[0.015, 0.2575]	112	[0.005, 0.1925]	0	-	621	[0.0125, 0.09]
65 - 70	45020	[0.01, 0.0625]	3191	[0.0125, 0.165]	1638	[0.0125, 0.125]	681	[0.0175, 0.155]	169	[0.0125, 0.165]	0	-	557	[0.0125, 0.21]
70 - 75	42278	[0.01, 0.0675]	3679	[0.0125, 0.1575]	1717	[0.01, 0.105]	832	[0.0125, 0.11]	206	[0.0075, 0.2025]	0	-	588	[0.01, 0.2075]
75 - 80	40914	[0.01, 0.06]	4945	[0.01, 0.16]	1608	[0.01, 0.105]	865	[0.0125, 0.11]	269	[0.0125, 0.2225]	0	-	540	[0.0125, 0.2225]
80 - 85	36182	[0.01, 0.0625]	4894	[0.0125, 0.15]	1044	[0.0125, 0.1125]	656	[0.015, 0.1175]	247	[0.01, 0.2]	0	-	405	[0.02, 0.21]
85 - 90	33711	[0.0125, 0.06]	3806	[0.01, 0.1525]	783	[0.0125, 0.12]	405	[0.015, 0.36]	269	[0.005, 0.2375]	0	-	377	[0.0075, 0.2075]
90 - 95	32610	[0.01, 0.0625]	3373	[0.0125, 0.155]	566	[0.01, 0.095]	187	[0.005, 0.135]	231	[0.0325, 0.2825]	0	-	347	[0.04, 0.2225]
95 - 100	29781	[0.01, 0.06]	3821	[0.0125, 0.18]	698	[0.01, 0.0825]	283	[0.005, 0.1275]	262	[0.0075, 0.3075]	0	-	352	[0.0075, 0.095]
100 - 105	27633	[0.01, 0.06]	4353	[0.0125, 0.17]	879	[0.01, 0.105]	273	[0.0075, 0.145]	279	[0.0175, 0.2875]	0	-	325	[0.0125, 0.0875]
105 - 110	24983	[0.01, 0.0625]	4176	[0.0125, 0.16]	1043	[0.0125, 0.1275]	230	[0.0125, 0.1275]	319	[0.01, 0.2125]	0	-	315	[0.0125, 0.1075]
110 - 115	21102	[0.01, 0.0675]	4321	[0.0125, 0.1825]	1374	[0.01, 0.1125]	225	[0.01, 0.105]	336	[0.015, 0.2975]	1	-	373	[0.0075, 0.08]
115 - 120	18482	[0.01, 0.0625]	5263	[0.0125, 0.1425]	1534	[0.0075, 0.15]	270	[0.015, 0.1475]	366	[0.015, 0.26]	0	-	418	[0.0125, 0.11]
120 - 125	17577	[0.0125, 0.0625]	5244	[0.0125, 0.17]	1131	[0.0125, 0.1625]	294	[0.01, 0.1175]	354	[0.0075, 0.1725]	0	-	459	[0.01, 0.1075]
125 - 130	17015	[0.01, 0.06]	4229	[0.0125, 0.1725]	829	[0.0125, 0.1375]	426	[0.0225, 0.1375]	413	[0.0175, 0.215]	0	-	336	[0.01, 0.16]
130 - 135	16220	[0.01, 0.06]	3808	[0.01, 0.1775]	690	[0.01, 0.1175]	366	[0.01, 0.1225]	386	[0.02, 0.2425]	0	-	291	[0.0075, 0.105]
135 - 140	13866	[0.01, 0.06]	3959	[0.015, 0.1575]	619	[0.015, 0.1275]	466	[0.0225, 0.1375]	406	[0.01, 0.325]	0	-	214	[0.0175, 0.11]
140 - 145	12251	[0.01, 0.065]	3801	[0.0125, 0.1875]	602	[0.005, 0.1075]	371	[0.015, 0.205]	439	[0.0125, 0.2625]	1	-	139	[0.01, 0.0825]
145 - 150	10968	[0.01, 0.065]	3494	[0.015, 0.17]	384	[0.0125, 0.165]	369	[0.01, 0.1975]	425	[0.01, 0.2825]	1	-	150	[0.0125, 0.125]
150 - 155	10050	[0.01, 0.0625]	4118	[0.015, 0.16]	223	[0.0025, 0.1675]	355	[0.02, 0.175]	401	[0.01, 0.2325]	7	-	110	[0.005, 0.1275]
155 - 160	9174	[0.0125, 0.0625]	4608	[0.01, 0.1825]	182	[0.0225, 0.1925]	305	[0.0225, 0.1625]	439	[0.0175, 0.2125]	6	-	119	[0.0075, 0.06]
160 - 165	8527	[0.01, 0.06]	5695	[0.0175, 0.1975]	159	[0.01, 0.0775]	327	[0.0125, 0.1975]	423	[0.01, 0.2125]	13	[0.19, 0.4225]	121	[0.0025, 0.14]
165 - 170	7165	[0.01, 0.0675]	4007	[0.015, 0.17]	108	[0.0125, 0.2425]	285	[0.01, 0.15]	407	[0.015, 0.3825]	20	[0.22, 0.3575]	73	[0.0125, 0.205]
170 - 175	6460	[0.0125, 0.0675]	3913	[0.0125, 0.155]	84	[0.0125, 0.2]	256	[0.0225, 0.1775]	411	[0.0125, 0.25]	31	[0.0375, 0.555]	93	[0.005, 0.1525]
175 - 180	5799	[0.01, 0.07]	4021	[0.015, 0.185]	58	[0.0125, 0.1625]	233	[0.0125, 0.12]	387	[0.015, 0.3475]	36	[0.07, 0.1725]	57	[0.0125, 0.1575]
180 - 185	5599	[0.01, 0.0675]	3796	[0.015, 0.21]	37	[0.01, 0.1525]	191	[0.01, 0.27]	378	[0.01, 0.31]	28	[0.0225, 0.56]	49	[0.0375, 0.14]
185 - 190	5165	[0.01, 0.0725]	3452	[0.015, 0.2075]	11	[0.03, 0.08]	120	[0.0175, 0.385]	365	[0.0275, 0.4975]	25	[0.055, 0.245]	47	[0.0225, 0.28]
190 - 195	5171	[0.01, 0.075]	2796	[0.0125, 0.1825]	9	-	92	[0.0075, 0.1225]	364	[0.02, 0.5525]	50	[0.005, 0.11]	36	[0.0125, 0.2475]
195 - 200	4539	[0.0125, 0.0725]	3665	[0.0125, 0.1875]	1	-	78	[0.0875, 0.2825]	384	[0.0125, 0.57]	54	[0.0175, 0.2775]	33	[0.03, 0.065]
200 - 205	4624	[0.0125, 0.07]	3081	[0.015, 0.22]	9	[1.0, 1.0]	47	[0.0175, 0.1625]	371	[0.015, 0.37]	64	[0.0225, 0.315]	34	[0.15, 0.17]
205 - 210	4433	[0.01, 0.08]	2734	[0.0125, 0.175]	18	[0.3825, 0.52]	47	[0.0175, 0.2]	384	[0.0125, 0.375]	74	[0.04, 0.2625]	26	[0.06, 0.075]
210 - 215	4224	[0.01, 0.0675]	2983	[0.0075, 0.1975]	30	[0.055, 0.2075]	40	[0.05, 0.46]	308	[0.0075, 0.3775]	77	[0.04, 0.37]	28	[0.0025, 0.2525]
215 - 220	4144	[0.01, 0.075]	3673	[0.01, 0.1975]	36	[0.0525, 0.1225]	17	[0.1375, 0.36]	327	[0.0125, 0.3825]	79	[0.0075, 0.2175]	23	[0.0175, 0.0675]
220 - 225	3937	[0.01, 0.0675]	4864	[0.015, 0.1825]	38	[0.0275, 0.245]	43	[0.0775, 0.2975]	362	[0.0625, 0.795]	136	[0.0275, 0.2275]	18	[0.0175, 0.195]
225 - 230	3677	[0.01, 0.0775]	3919	[0.0125, 0.175]	73	[0.0125, 0.1375]	24	[0.015, 0.1725]	274	[0.0225, 0.415]	107	[0.09, 0.645]	17	[0.165, 0.2575]
230 - 235	3498	[0.01, 0.0725]	3919	[0.0125, 0.165]	74	[0.015, 0.145]	48	[0.065, 0.27]	282	[0.0425, 0.78]	116	[0.03, 0.29]	17	[0.0275, 0.0775]
235 - 240	3329	[0.01, 0.075]	3733	[0.01, 0.1725]	52	[0.025, 0.14]	62	[0.04, 0.4075]	291	[0.0075, 0.6325]	109	[0.005, 0.225]	16	[0.09, 0.3725]
240 - 245	3106	[0.01, 0.0775]	3305	[0.015, 0.1975]	39	[0.055, 0.24]	52	[0.0175, 0.195]	282	[0.0275, 0.7125]	113	[0.0075, 0.5125]	4	-
245 - 250	3072	[0.01, 0.0775]	2807	[0.0075, 0.185]	20	[0.025, 0.085]	60	[0.04, 0.18]	256	[0.01, 0.5525]	132	[0.005, 0.41]	6	-
250 - 255	3071	[0.01, 0.0775]	2533	[0.0125, 0.1975]	14	[0.085, 0.23]	44	[0.025, 0.115]	268	[0.0275, 1.0]	136	[0.0225, 0.2225]	7	-
255 - 260	3106	[0.01, 0.08]	2737	[0.015, 0.2]	9	-	56	[0.0075, 0.4475]	221	[0.025, 0.755]	149	[0.01, 0.295]	5	-
260 - 265	3179	[0.0125, 0.075]	2444	[0.0125, 0.1925]	2	-	88	[0.0675, 0.2475]	208	[0.0225, 1.0]	148	[0.01, 0.59]	9	-
265 - 270	3257	[0.01, 0.0725]	2084	[0.0125, 0.1725]	1	-	100	[0.005, 0.32]	235	[0.025, 1.0]	172	[0.02, 0.5225]	5	-
270 - 275	3556	[0.01, 0.0725]	1817	[0.015, 0.185]	0	-	112	[0.005, 0.4125]	179	[0.0125, 1.0]	142	[0.025, 0.65]	11	[0.1175, 0.1175]
275 - 280	3661	[0.0125, 0.075]	2008	[0.0175, 0.205]	0	-	101	[0.0425, 0.5725]	167	[0.2675, 1.0]	155	[0.035, 0.495]	3	-
280 - 285	3930	[0.0125, 0.0675]	1962	[0.01, 0.16]	0	-	60	[0.04, 0.2375]	169	[0.0125, 1.0]	150	[0.05, 0.375]	3	-
285 - 290	3710	[0.01, 0.0775]	2156	[0.0125, 0.1775]	0	-	40	[0.03, 0.52]	154	[0.02, 1.0]	156	[0.0425, 0.42]	5	-
290 - 300	3617	[0.01, 0.07]	2470	[0.0125, 0.1625]	0	-	33	[0.085, 0.2625]	154	[0.03, 1.0]	171	[0.0325, 0.6725]	6	-

Table A.1: The number of samples and goodness-of-fit statistics for the curve fitting routine

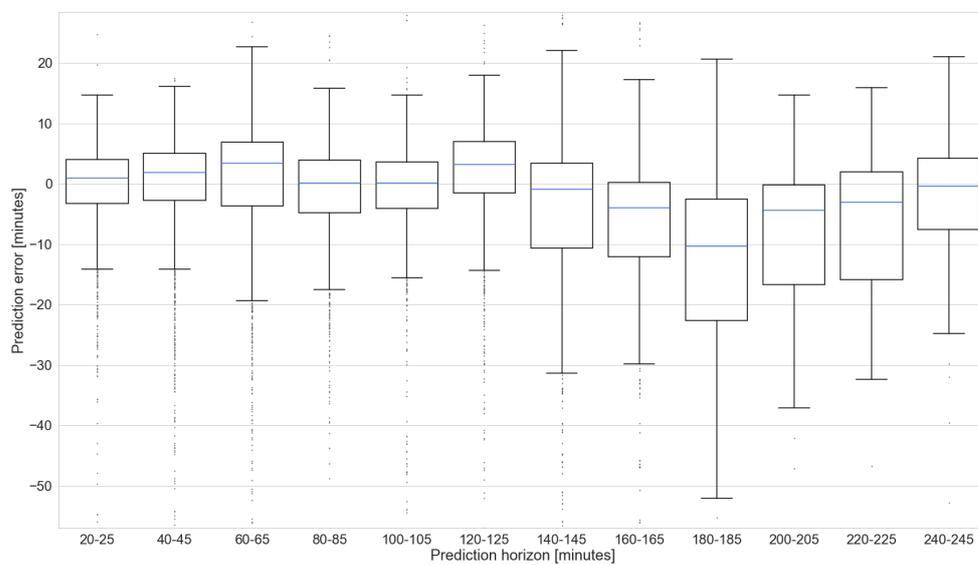


Figure A.2: Accuracy of the COP arrival predictions versus time to fly for the Tactical Activated flight state

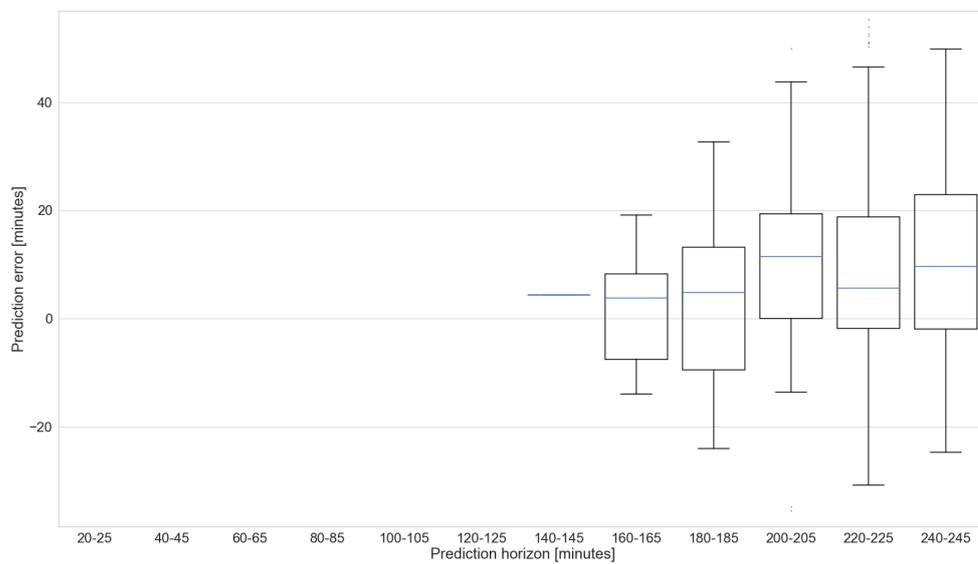


Figure A.3: Accuracy of the COP arrival predictions versus time to fly for the Filed Slot Allocated flight state

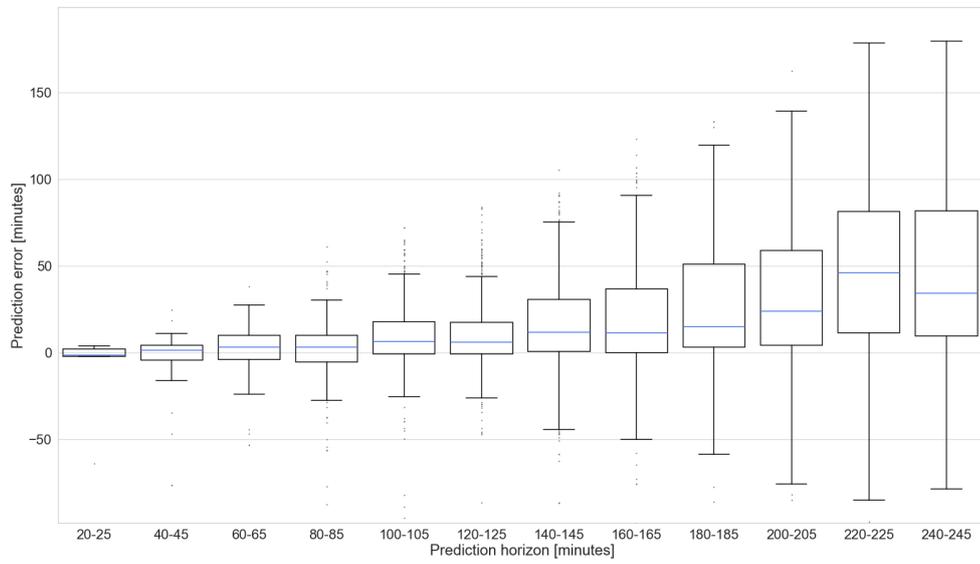


Figure A.4: Accuracy of the COP arrival predictions versus time to fly for the Slot Issued flight state

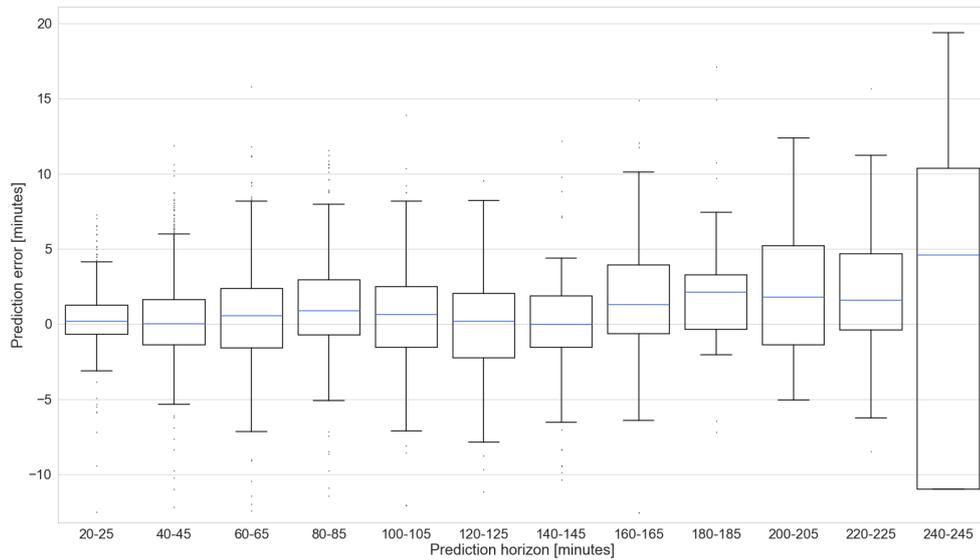


Figure A.5: Accuracy of the COP arrival predictions versus time to fly for the Suspended flight state

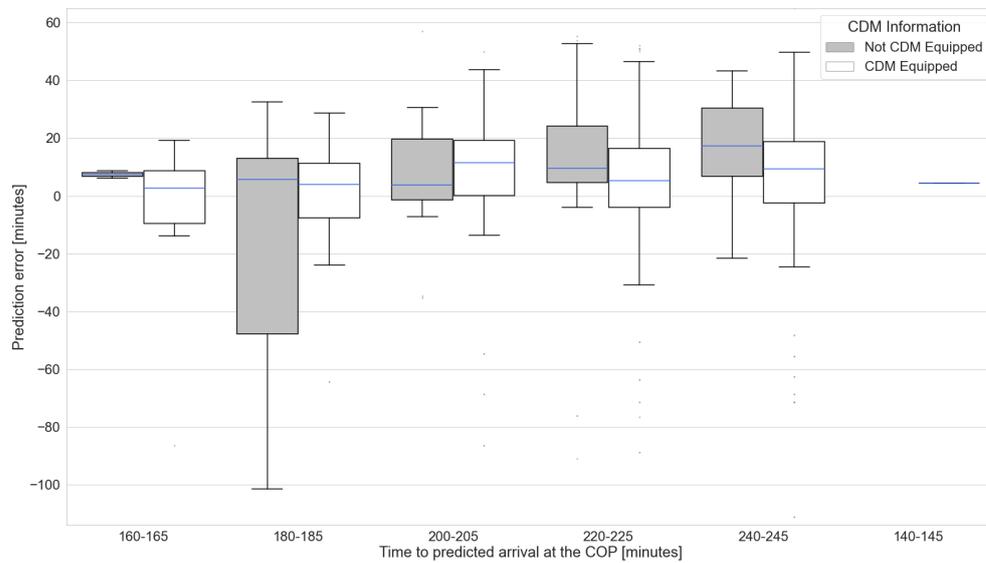


Figure A.6: Accuracy of the COP arrival predictions versus time to fly for the Filed Slot Allocated flight state, for CDM and non-CDM equipped messages

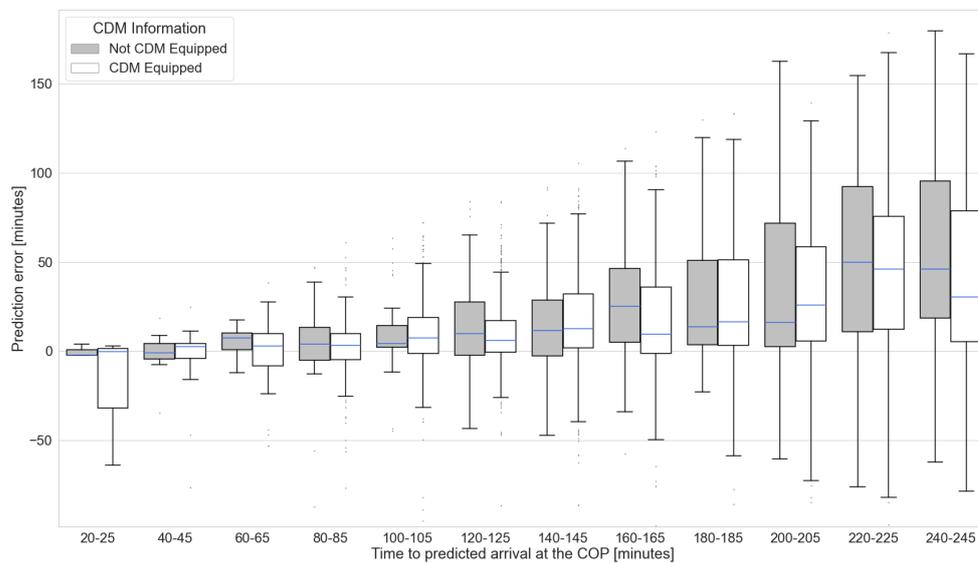


Figure A.7: Accuracy of the COP arrival predictions versus time to fly for the Slot Issued flight state, for CDM and non-CDM equipped messages

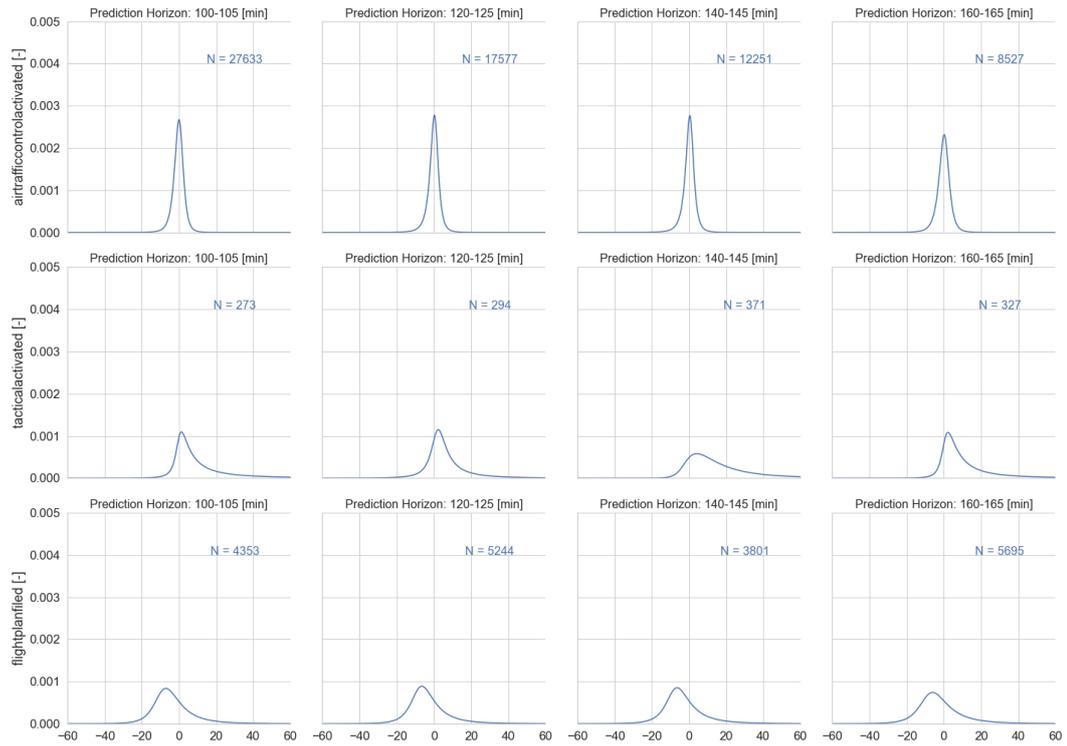


Figure A.8: Probability Density Curves for the AA, TA and FI flight state up to 165 minutes prediction horizon

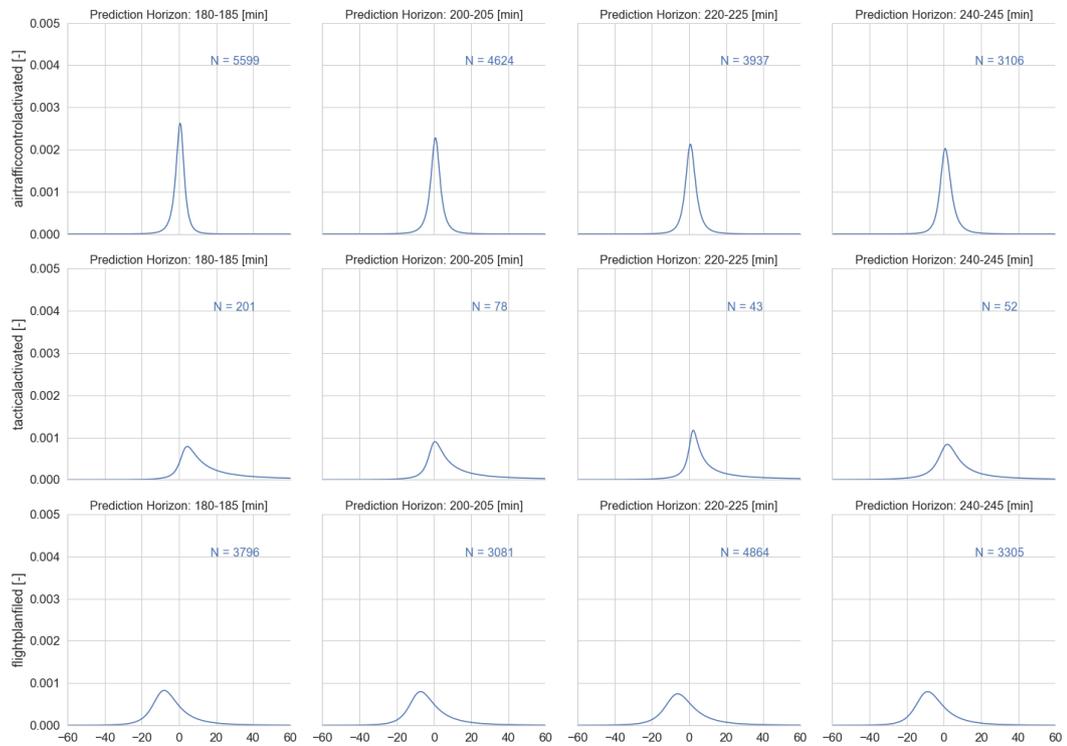


Figure A.9: Probability Density Curves for the AA, TA and FI flight state up to 245 minutes prediction horizon

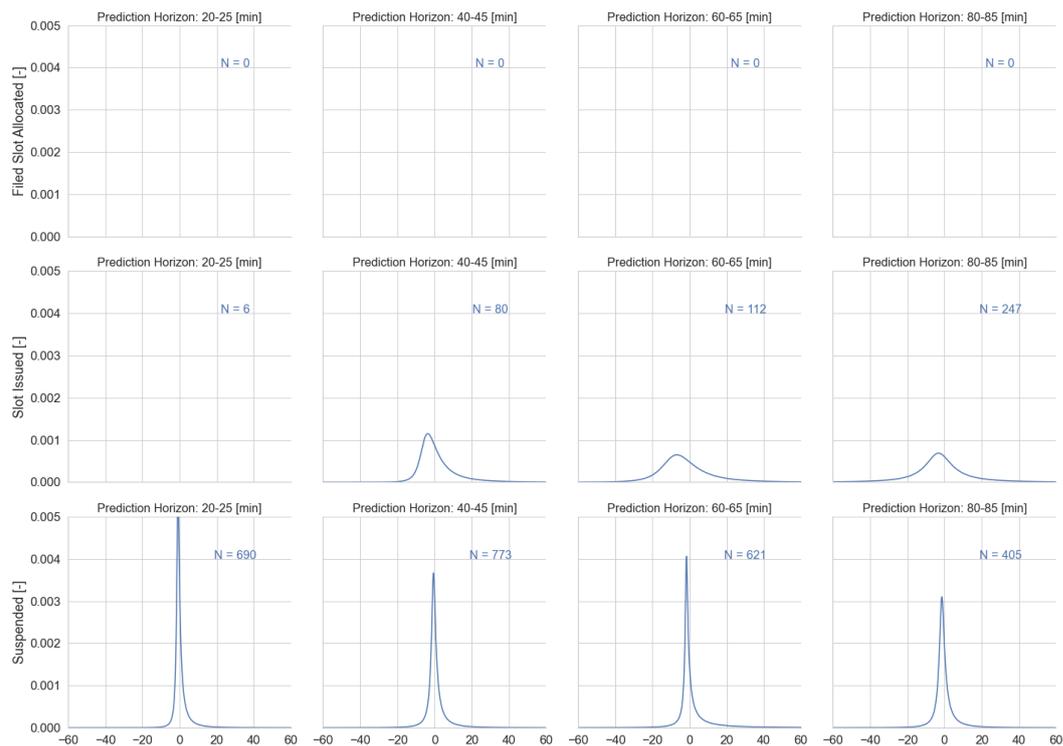


Figure A.10: Probability Density Curves for the FL, sI and SU flight state up to 85 minutes prediction horizon

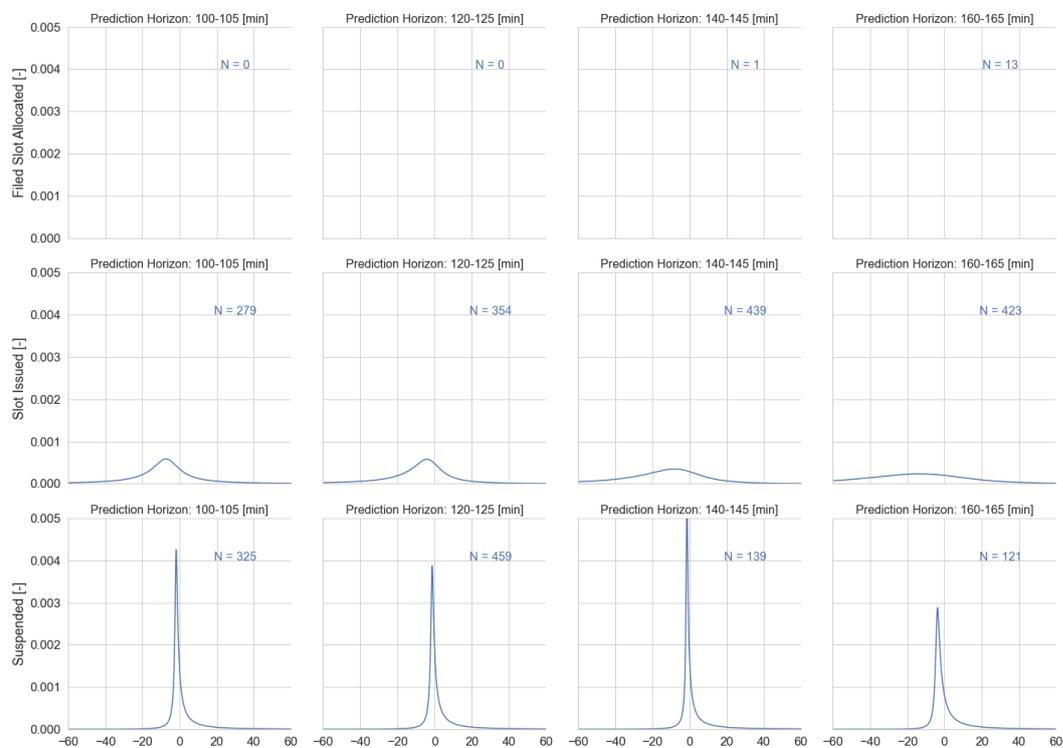


Figure A.11: Probability Density Curves for the FL, sI and SU flight state up to 165 minutes prediction horizon

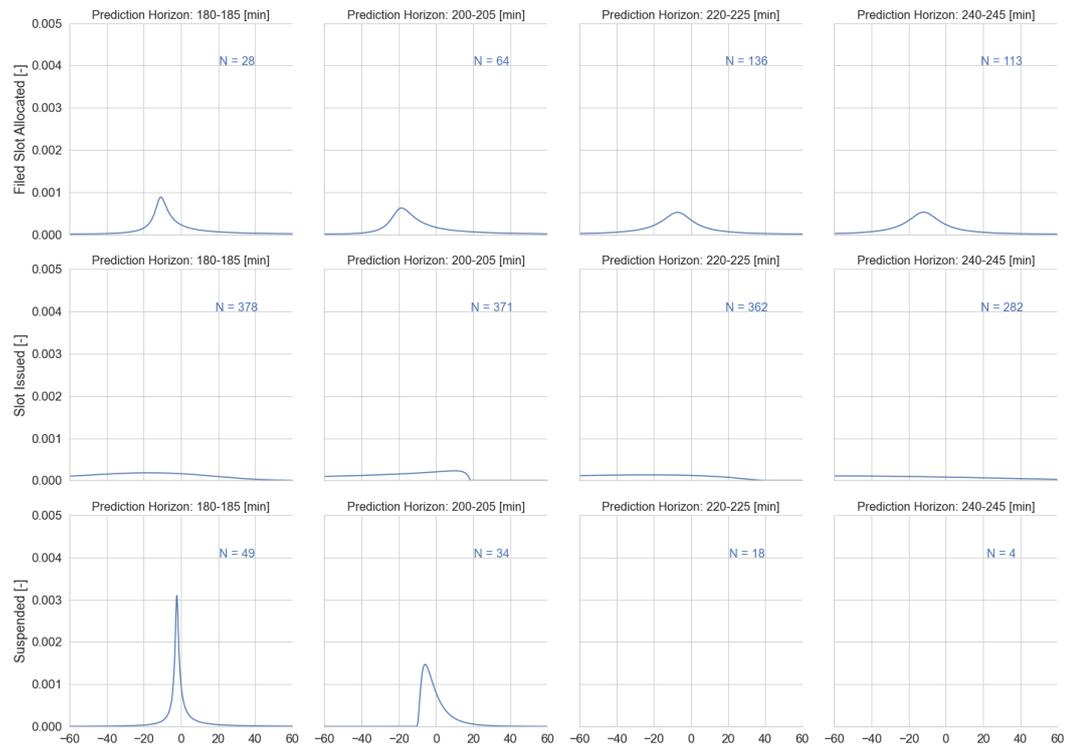


Figure A.12: Probability Density Curves for the FL, sl and SU flight state up to 245 minutes prediction horizon

# B

## Appendix: Methodology

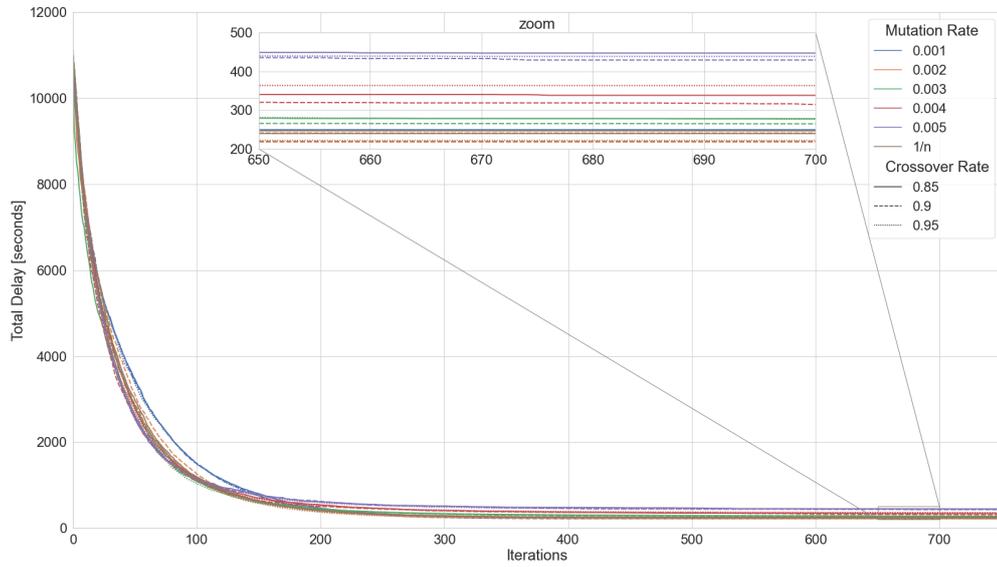


Figure B.1: Sample problem 2: day 1, 60% threshold

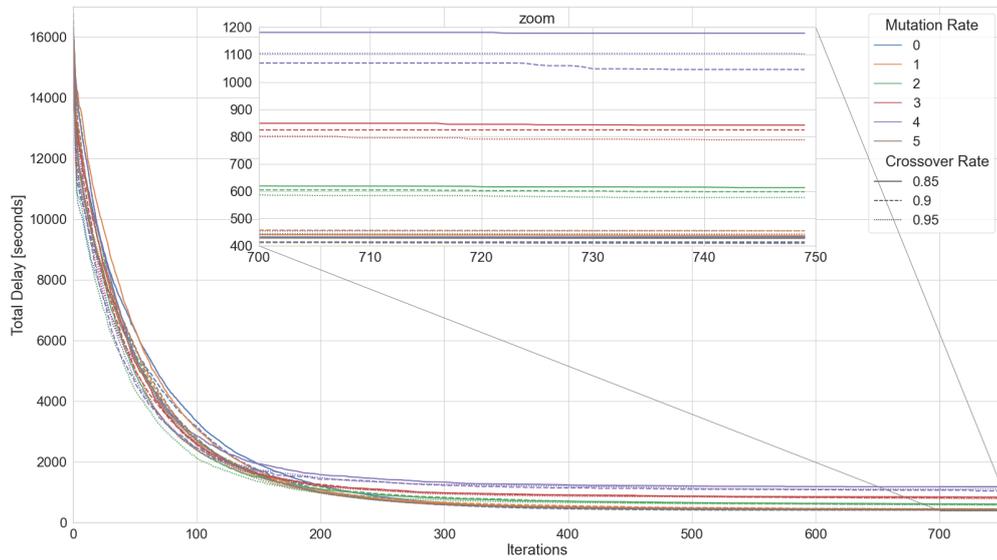


Figure B.2: Sample problem 3: day 4, 30% threshold

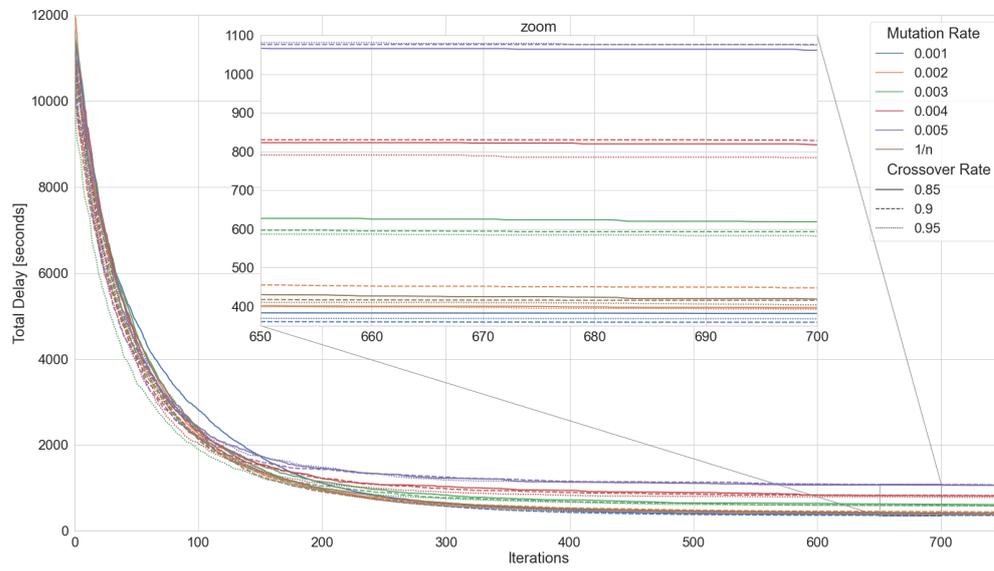


Figure B.3: Sample problem 4: day 4, 55% threshold

Timestamp	ACID	Flying Time to IAF [s]	New Imposed Delay [s]	Total Maximum Delay [s]	Maximum Delay Descent [s]	Maximum Delay Cruise [s]	Flight State [s]	Delay Imposed Before [s]
2021-05-03 04:11:09	SIA324	2486	255	341	286	55	AA	0
2021-05-03 04:11:09	AAI9702	5801	110	441	244	197	AA	0
2021-05-03 04:11:09	KLM96F	2557	123	220	220	0	AA	0
2021-05-03 04:34:40	DAL258	2268	53	286	225	61	AA	0
2021-05-03 04:34:40	DAL160	2269	71	278	234	44	AA	0
2021-05-03 04:34:40	AAI9702	4219	67	378	244	134	AA	110.0
2021-05-03 04:34:40	KLM52F	2602	140	280	213	67	AA	0
2021-05-03 04:34:40	NJE564E	2407	40	323	280	43	AA	0
2021-05-03 04:34:40	KLM84X	2199	53	213	213	0	AA	0
2021-05-03 04:34:40	KLM1382	2379	192	253	197	56	AA	0
2021-05-03 04:39:55	CAL073	2333	74	297	263	34	AA	0
2021-05-03 04:39:55	KLM74L	2298	54	219	206	13	AA	0
2021-05-03 05:12:48	KLM872	2321	109	369	319	50	AA	0
2021-05-03 05:12:48	KLM22A	2339	13	214	197	17	AA	0
2021-05-03 05:12:48	KLM62E	2540	108	217	203	14	AA	0
2021-05-03 05:12:48	KLM64J	2342	57	230	205	25	AA	0
2021-05-03 05:18:04	KLM431	2267	66	265	218	47	AA	0
2021-05-03 05:18:04	KLM1628	2355	59	203	197	6	AA	0
2021-05-03 05:18:04	KLM68T	2301	16	260	210	50	AA	0

Table B.1.: An example of the output table for the debunching module

# C

## Appendix: Results

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
35 %	3	395	56	40.7	0	-297	2.8
30 %	3	490	56	50.2	0	-54	0.5
25 %	6	1065	64	56.1	1	-179	1.7
20 %	27	3008	67	34.9	9	-916	8.5

Table C.1: Output table for day 1, with a double-runway configuration, fed by the ARTIP stack

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
35 %	0	-	-	-	-	-	-
30 %	0	-	-	-	-	-	-
25 %	7	735	52	39.2	0	-47	1.0
20 %	24	2358	61	35.9	10	+370	8.2

Table C.2: Output table for day 4, with a double-runway configuration, fed by the ARTIP stack

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
35 %	0	-	-	-	-	-	-
30 %	3	203	63	22.9	0	-100	1.9
25 %	15	2078	111	40.3	1	+5	0.1
20 %	41	4940	121	30.6	25	-616	11.9

Table C.3: Output table for day 6, with a double-runway configuration, fed by the ARTIP stack

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
35 %	0	-	-	-	-	-	-
30 %	0	-	-	-	-	-	-
25 %	0	-	-	-	-	-	-
20 %	15	1393	57	34.1	3	+232	4.2

Table C.4: Output table for day 7, with a double-runway configuration, fed by the ARTIP stack

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
35 %	0	-	-	-	-	-	-
30 %	2	128	74	18.6	0	+79	0.7
25 %	3	366	72	36.3	1	-278	2.6
20 %	17	2619	72	50.3	3	-529	4.9

Table C.5: Output table for day 1, with a double-runway configuration, fed by the RIVER and SUGOL stacks

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
35 %	0	-	-	-	-	-	-
30 %	0	-	-	-	-	-	-
25 %	0	-	-	-	-	-	-
20 %	8	1163	51	51.5	1	-144	3.2

Table C.6: Output table for day 4, with a double-runway configuration, fed by the RIVER and SUGOL stacks

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
35 %	0	-	-	-	-	-	-
30 %	2	174	74	26.3	0	+122	2.3
25 %	11	1585	54	57.3	1	-54	1.0
20 %	51	5280	421	32.7	29	-25	0.5

Table C.7: Output table for day 6, with a double-runway configuration, fed by the RIVER and SUGOL stacks

Bunching Threshold	# of AC Delayed	Total Delay [s]	Avg. Time to IAF [min]	% of Max Delay Avg	# of Multiple Delays	Change in AMAN [s]	% of total AMAN Delay
35 %	0	-	-	-	-	-	-
30 %	1	123	91	37.3	0	-72	1.3
25 %	2	264	106	28.1	0	-59	1.1
20 %	7	1493	82	62.2	1	-101	1.8

Table C.8: Output table for day 7, with a double-runway configuration, fed by the RIVER and SUGOL stacks

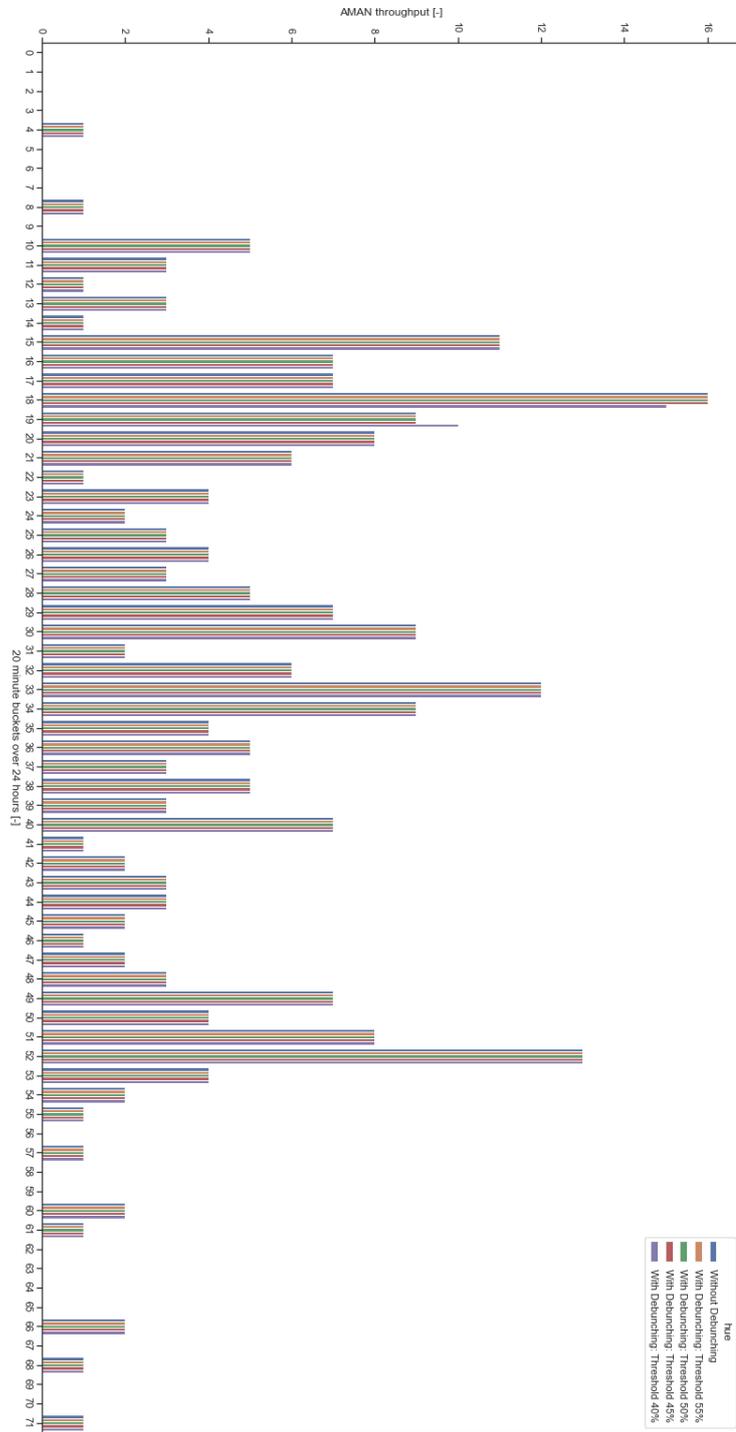


Figure C.1: The AMAN throughput per 20 minutes for each bunching threshold at day 5

# Bibliography

- [1] The Netherlands Amsterdam Airport Schiphol, Amsterdam. Traffic review 2019. Technical report, 2020. URL [https://assets.ctfassets.net/biom0eqyyi6b/2a3AhGB6ciY86vBhVbFrPW/da4d2bfd0cb0274834354ad5b7a5b925/Schiphol\\_Traffic\\_review\\_2019.pdf](https://assets.ctfassets.net/biom0eqyyi6b/2a3AhGB6ciY86vBhVbFrPW/da4d2bfd0cb0274834354ad5b7a5b925/Schiphol_Traffic_review_2019.pdf).
- [2] Hamsa Balakrishnan and Bala Chandran. Scheduling aircraft landings under constrained position shifting. In *AIAA guidance, navigation, and control conference and exhibit*, page 6320, 2006.
- [3] Jesper Bronsvoot. Contributions to trajectory prediction theory and its application to arrival management for air traffic control. *Ph. D. thesis*, 2014.
- [4] Jesper Bronsvoot, G McDonald, R Porteous, and E Gutt. Study of aircraft derived temporal prediction accuracy using fans. In *13th Air Transport Research Society (ATRS) World Conference*, 2009.
- [5] Sandrine Carlier, Ivan De Lépinay, Jean-Claude Hustache, and Frank Jelinek. Environmental impact of air traffic flow management delays. In *7th USA/Europe air traffic management research and development seminar (ATM2007)*, volume 2, page 16, 2007.
- [6] Coralia Cartis, Nicholas IM Gould, and Philippe L Toint. Evaluation complexity of adaptive cubic regularization methods for convex unconstrained optimization. *Optimization Methods and Software*, 27(2): 197–219, 2012.
- [7] Enrique Juan Casado Magaña. *Trajectory prediction uncertainty modelling for Air Traffic Management*. PhD thesis, University of Glasgow, 2016.
- [8] Stephen Chen, James Montgomery, and Antonio Bolufé-Röhler. Measuring the curse of dimensionality and its effects on particle swarm optimization and differential evolution. *Applied Intelligence*, 42(3): 514–526, 2015.
- [9] Andrew R Conn, Katya Scheinberg, and Luis N Vicente. *Introduction to derivative-free optimization*. 2009.
- [10] Richard Coppenbarger, Richard Lanier, Doug Sweet, and Susan Dorsky. Design and development of the en route descent advisor (eda) for conflict-free arrival metering. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*, page 4875, 2004.
- [11] Paula Leal De Matos and Richard Ormerod. The application of operational research to european air traffic flow management—understanding the context. *European Journal of Operational Research*, 123(1): 125–144, 2000.
- [12] Kalyanmoy Deb. An efficient constraint handling method for genetic algorithms. *Computer methods in applied mechanics and engineering*, 186(2-4):311–338, 2000.
- [13] Luis Delgado and Xavier Prats. En route speed reduction concept for absorbing air traffic flow management delays. *Journal of Aircraft*, 49(1):214–224, 2012.
- [14] Luis Delgado and Xavier Prats. Effect of wind on operating-cost-based cruise speed reduction for delay absorption. *IEEE transactions on intelligent transportation systems*, 14(2):918–927, 2013.
- [15] EUROCONTROL. European wake turbulence categorisation and separation minima on approach and departure. Technical report, 2015.
- [16] EUROCONTROL. Specification for on-line data interchange (oldi). Technical report, 2017.
- [17] EUROCONTROL. Aviation intelligence unit. atfm regulation : a power for good understanding how it works. Technical report, 2019.

- [18] Eurocontrol. Specification for trajectory prediction, edition 2.0. Technical report, March, 2017.
- [19] JU Eurocontrol-SESAR. European atm master plan-the roadmap for sustainable air traffic management, 2012.
- [20] Antony D Evans and Paul U Lee. Analyzing double delays at newark liberty international airport. In *16th AIAA Aviation Technology, Integration, and Operations Conference*, page 3456, 2016.
- [21] Stanley Förster, Michael Schultz, and Hartmut Fricke. Probabilistic prediction of separation buffer to compensate for the closing effect on final approach. *Aerospace*, 8(2):29, 2021.
- [22] Arno Gerretsen and Sip Swierstra. Sensitivity of aircraft performance to variability of input data. *EUROCONTROL Doc. CoE-TP-02005*, 2003.
- [23] I. Suci H. Koolen. Advanced atc twr implementation guide, edition 1.600. Technical report, July, 2020.
- [24] I. Suci H. Koolen. Dpi implementation guide, edition 2.600. Technical report, July, 2020.
- [25] Nathalie Hasevoets and Paul Conroy. Arrival manager: Implementation guidelines and lessons learned. *0.1, Brussels: EUROCONTROL*, 2010.
- [26] Rania Hassan, Babak Cohanin, Olivier De Weck, and Gerhard Venter. A comparison of particle swarm optimization and the genetic algorithm. In *46th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference*, page 1897, 2005.
- [27] Ahmad Hassanat, Khalid Almohammadi, Esra Alkafaween, Eman Abunawas, Awni Hammouri, and VB Prasath. Choosing mutation and crossover ratios for genetic algorithms a review with a new dynamic approach. *Information*, 10(12):390, 2019.
- [28] ID Hill, R Hill, and RL Holder. Fitting johnson curves by moments. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 25(2):180–189, 1976.
- [29] Icao. *Annex 2: Rules of the Air*. International Civil Aviation Organ, 1990.
- [30] International Civil Aviation Organization (ICAO). Nat doc 007. north atlantic operations and airspace manual. V.2018-1.
- [31] Michael R Jackson, Yiyuan J Zhao, and Rhonda A Slattery. Sensitivity of trajectory prediction in air traffic management. *Journal of Guidance, Control, and Dynamics*, 22(2):219–228, 1999.
- [32] H. Koolen. Fum implementation guide, edition 2.000. Technical report, July, 2020.
- [33] Hans Koolen and Ioana Coliban. Flight progress messages document, edition 2.700. Technical report, July, 2020.
- [34] James K Kuchar and Lee C Yang. A review of conflict detection and resolution modeling methods. *IEEE Transactions on intelligent transportation systems*, 1(4):179–189, 2000.
- [35] Jeffrey Larson, Matt Menickelly, and Stefan M Wild. Derivative-free optimization methods. *Acta Numerica*, 28:287–404, 2019.
- [36] Erich Leo Lehmann and Howard J D’Abrera. *Nonparametrics: statistical methods based on ranks*. Holden-day, 1975.
- [37] LVNL. System requirements document asap, unpublished. Technical report, 2019.
- [38] LVNL. *Integrated Aeronautical Information Package*, 2020. URL <https://www.lvn1.nl/eaip/2021-03-11-ATRAC/html/index-en-GB.html>.
- [39] Stéphane Mondoloni and Ibrahim Bayraktutar. Impact of factors, conditions and metrics on trajectory prediction accuracy. In *6th USA/Europe ATM R&D Seminar*, 2005.
- [40] A Montlaur and L Delgado. Flight and passenger delay assignment optimization strategies. *Transportation Research Part C: Emerging Technologies*, 81:99–117, 2017.

- [41] M. Tielrooij. *Arrival Management Support in the Presence of Prediction Uncertainty, PhD Thesis*. PhD thesis, TU Delft, 2022.
- [42] K Tysen Mueller, Rich Bortins, David Schleicher, Doug Sweet, and Rich Coppenbarger. Effect of uncertainty on en route descent adviser (eda) predictions. In *AIAA 4th Aviation Technology, Integration and Operations (ATIO) Forum*, page 6347, 2004.
- [43] Tasos Nikoleris, Gano B Chatterji, and Richard A Coppenbarger. Comparison of fuel consumption of descent trajectories under arrival metering. *Journal of Aircraft*, 53(6):1853–1864, 2016.
- [44] Angela Nuic. User manual for the base of aircraft data (bada) revision 3.10. *Atmosphere*, 2010:001, 2010.
- [45] Drs. B.I. Otto. *Capacity Declaration Amsterdam Airport Schiphol, Summer 2020*, 2019.
- [46] Alan R Parkinson, R Balling, and John D Hedengren. Optimization methods for engineering design. *Brigham Young University*, 5(11), 2013.
- [47] Victor Pereboom. A data-driven comparison of trajectory prediction methodologies using ads-b data. 2019.
- [48] Luis Miguel Rios and Nikolaos V Sahinidis. Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization*, 56(3):1247–1293, 2013.
- [49] Bill Roberson. Fuel conservation strategies: cost index explained. *Boeing Aero Quarterly*, 2(2007):26–28, 2007.
- [50] Julia Rudnyk, Joost Ellerbroek, and Jacco M Hoekstra. Trajectory prediction sensitivity analysis using monte carlo simulations based on inputs distributions. *Journal of Air Transportation*, 27(4):181–198, 2019.
- [51] Wilhelm Rumler, Thomas Günther, U Weißhaar, and H Fricke. Flight profile variations due to the spreading practice of cost index based flight planning. In *4th International Conference on Research in Air Transportation, Budapest*, 2010.
- [52] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.
- [53] SESAR. Extended aman osed, update 5.6.4 - step 1. Technical report, 2015.
- [54] Rhonda Slattery and Yiyuan Zhao. Trajectory synthesis for air traffic automation. *Journal of guidance, control, and dynamics*, 20(2):232–238, 1997.
- [55] James F Sliker and Samuel S Shapiro. The johnson system: selection and parameter estimation. *Technometrics*, 22(2):239–246, 1980.
- [56] Stéphanie Stoltz and Patrick Ky. Reducing traffic bunching through a more flexible air traffic flow management. In *Fourth International Air Traffic Management R&D Seminar ATM-2001, Santa Fe (New-Mexico, USA)*, 2001.
- [57] Maarten Tielrooij, Clark Borst, Marinus M Van Paassen, and Max Mulder. Predicting arrival time uncertainty from actual flight information. In *Proceedings of the 11th USA/Europe Air Traffic Management Research and Development Seminar*, pages 577–586, 2015.
- [58] Ferway To70. Business case extension aman radar horizon. work package 2: Location of current horizons, unpublished. Technical report, 2018.
- [59] Gonzalo Tobaruela, Peter Fransen, Wolfgang Schuster, Washington Y Ochieng, and Arnab Majumdar. Air traffic predictability framework—development, performance evaluation and application. *Journal of Air Transport Management*, 39:48–58, 2014.
- [60] Sergio Torres. Trajectory accuracy sensitivity to modeling factors. In *15th AIAA Aviation Technology, Integration, and Operations Conference*, page 2599, 2015.

- 
- [61] SESAR Joint Undertaking. European atm master plan. Technical report, 2019.
  - [62] Alexander Vanwelsenaere, J Ellerbroek, JM Hoekstra, and E Westerveld. Effect of popup flights on the extended arrival manager. *Journal of Air Transportation*, 26(2):60–69, 2018.
  - [63] Yan Xu, Ramon Dalmau, and Xavier Prats. Maximizing airborne delay at no extra fuel cost by means of linear holding. *Transportation Research Part C: Emerging Technologies*, 81:137–152, 2017.